STUDIES IN FUZZINESS AND SOFT COMPUTING

Eduardo Massad
Neli Regina Siqueira Ortega
Laécio Carvalho de Barros
Claudio José Struchiner

Fuzzy Logic in Action: Applications in Epidemiology and Beyond

Springer

Studies in Fuzziness and Soft Computing, Volume 232

Further volumes of this series can be found on our homepage: springer.com

Eduardo Massad, Neli Regina Siqueira Ortega,
Laécio Carvalho de Barros,
Claudio José Struchiner

Fuzzy Logic in Action: Applications in Epidemiology and Beyond

Springer

Authors

Prof. Eduardo Massad
School of Medicine
The University of São Paulo
Av.Dr. Arnaldo 455
São Paulo, CEP: 01246-903
SP-Brazil
E-mail: edmassad@usp.br

Neli Regina Siqueira Ortega
School of Medicine
The University of São Paulo
Av.Dr. Arnaldo 455
São Paulo, CEP: 01246-903
SP-Brazil
E-mail: neli@dim.fm.usp.br

Laécio Carvalho de Barros
Dept. of Applied Mathematics
State University of Campinas, Cidade
Universitária Zeferino Vaz
Barão Geraldo, Campinas,
CEP: 13083-970
SP-Brazil
E-mail: laeciocb@ime.unicamp.br

Claudio José Struchiner
PROCC, FIOCRUZ
Rua Leopoldo Bulhões 1480
Manguinhos, Rio de Janeiro,
CEP: 21041-210
RJ-Brazil
E-mail: stru@procc.fiocruz.br

ISBN 978-3-540-69092-4 e-ISBN 978-3-540-69094-8

DOI 10.1007/978-3-540-69094-8

Studies in Fuzziness and Soft Computing ISSN 1434-9922

Library of Congress Control Number: 2008928274

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed in acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

To those who contribute with their work
to the control of the transmission of
infectious diseases in the world

To Mara, Laila and Jamile

To my sweet love Eduardo Rissi and
our little flowers Ana Clara and Mariana

To Cristina, Otávio and Luiza

To Dita, Moyses, Noel, Ivan
Paula and Arthur

Foreword

Over four decades have passed since Lotfi Zadeh introduced the notion of fuzzy sets. During this period we have seen an impressive growth of conceptual, theoretical and methodological developments and a variety of applications in many fields. Current machinery in areas such as information storage and search, image processing and understanding, pattern recognition and control, computational biology and bioinformatics, to mention a few, have benefited considerably from the developments in fuzzy set theory. What is less visible is the paradigmatic change in the contemporary medical sciences, especially the changes that concern uncertainty, imprecision and vagueness. Despite the naturally imprecise character of the biological variables, only recently uncertainty, imprecision and vagueness have been viewed as essential, not only unavoidable, but in fact of a great practical utility. This is because uncertainty, imprecision and vagueness help to reduce complexity and to increase credibility and tractability, items that tend to maximize the usefulness of biological models.

Fuzzy Logic in Action: Applications in Epidemiology and Beyond, co-authored by Eduardo Massad, Neli Ortega, Laécio Barros, and Claudio Struchiner is a remarkable achievement. The book brings a major paradigm shift to medical sciences exploring the use of fuzzy sets in epidemiology and medical diagnosis arena. The volume addresses the most significant topics in the broad areas of epidemiology, mathematical modeling and uncertainty, embodying them within the framework of fuzzy set and dynamic systems theory. Written by leading contributors to the area of epidemiology, medical informatics and mathematics, the book combines a very lucid and authoritative exposition of the fundamentals of fuzzy sets with an insightful use of the fundamentals in the area of epidemiology and diagnosis. The content is clearly illustrated by numerous illustrative examples and several real world applications. Based on their profound knowledge of epidemiology and mathematical modeling, and on their keen understanding of the role played by uncertainty and fuzzy sets, the authors provide insights into the connections between biological phenomena and dynamic systems as a mean to predict, diagnose, and prescribe actions. An example is the use of

Bellman-Zadeh fuzzy decision making approach to develop a vaccination strategy to manage measles epidemics in São Paulo.

The book offers a comprehensive, systematic, fully updated and self-contained treatise of fuzzy sets in epidemiology and diagnosis. Its content covers material of vital interest to students, researchers and practitioners and is suitable both as a textbook and as a reference. The authors present new results of their own in most of the chapters. In doing so, they reflect the trend to view fuzzy sets, probability theory and statistics as an association of complementary and synergetic modeling methodologies.

Summing up, the authors have produced a book that presents a remarkably complete, well-organized, authoritative and reader-friendly exposition of the use of fuzzy sets in the area of epidemiology and diagnosis. The volume is a major contribution to a better understanding of how fuzzy set theory helps to manage complexity in biological phenomena and how to translate understanding into benefits for the human being.

Campinas, São Paulo, Brazil
April 2008

Fernando Gomide
University of Campinas

Contents

1 Introduction

Medicine is the science of intervention. Among human activities, medicine is rivaled only by agriculture (and more recently by modern engineering) in its capacity of changing the natural course of human fate. To intervene, however, one must predict. To predict the natural course of a system in the absence of intervention and to predict what is going to happen to such a system after the proposed intervention. Prediction, in turn, is strongly dependent on the scientific foundations of the subject in question. Although medicine men have been intervening in human diseases since the dawn of humanity (actually with rather questionable results) only very recently the scientific cornerstones of medicine have been laid down. After several centuries as a set of empirical practices, medicine gained some taints of natural history in the 19^{th} century and only in the first half of the 20^{th} century it overpassed the epistemological threshold. In the last 50 years, medicine started its approach to the formalism threshold, which is characterized by the application of formal logic and mathematics to the theoretical bases of diseases. Therefore, the marriage between mathematics and medicine is very recent and it started rather timidly with the works by the fathers of epidemiology. Scattered applications of mathematics in some biomedical areas was found in the literature until the early 1970s. Since then however, mathematical biology has changed a great deal. The magnificent 100 volumes of the series *Lectures Notes in Biomathematics* edited by Simon Levin is a respectable witness of those changes. In medicine specifically, mathematical models have been developed for some diagnostic problems and, to a great extent, to epidemiological questions.

1.1 Uncertainty *versus* Precision in Biomedicine and Epidemiology

Since the first attempts to develop mathematical and computer models in medicine, it was recognized that one of the significant obstacles was the inherent uncertainties, imprecisions, ambiguity and vagueness associated with medical concepts and with this type of applications. The study of physiological and

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 1–10, 2008.
springerlink.com

physiopathological problems is generally programmed by elaborating models which respond to the principles of formal logic (Guarini, 1994). This allows the transformation of the formal model into a mathematical model of reference which responds to the principles of set theory and its eventual implementation in computer systems. Medical problems, however, are not always amenable to straightforward algorithmic solutions by the lack of complete understanding of mechanisms of disease, inability to obtain complete information regarding the health (or ill) state of the organism, lack of precise ranges of normal (and abnormal) values of clinical parameters, and the inherent vagueness and ambiguity associated with medical concepts and terms. Disease diagnosis involves several levels of imprecision and uncertainty, a fact also true for epidemiological studies. A single disease may manifest itself quite differently in different patients and at different disease status. Furthermore, a single symptom may be indicative of different diseases, and the presence of several diseases in a single patient may disrupt the expected symptom pattern of any of them. This may cause a tremendous amount of imprecision and uncertainty in the interpretation of effect measures in the analysis. Also, the best and most useful descriptions of disease entities often use linguistic terms that are irreducibly vague. Indeed, as mentioned by Steimann (1996), the complexity of biological systems may force us to alter in radical ways our traditional approaches to the analysis of such systems. We therefore, may have to accept as unavoidable a substantial degree of fuzziness in the description of the behavior of biological systems as well as in their characterization.

Although traditional expert systems based on binary logic have been used successfully as diagnostic decision aids, their sequential processing of information and their almost universal use of probability theory (often Bayes' theorem) to represent those uncertainties in the medical context may be inappropriate and partly responsible for their limitations in certain applications. Therefore, nowhere in the field of science is the need for tools to deal with uncertainty more critical than in medicine and biology. In spite of its potential in dealing with uncertainties, very few works applying fuzzy logic concepts in epidemiological problems have been presented so far.

Over the past 30 years, the development of numerous mathematical methods and concepts have helped make the quantum leap in practical applications of fuzzy logic, in particular in devising controllers for complex industrial processes. However, biomedical applications of fuzzy logic concepts and methods have received less attention from experts, in spite of their tremendous potential as problem solving tools for those fields.

Notwithstanding, medicine is one field in which the applicability of fuzzy logic was quickly recognized in the mid-1970s. Within this field, it is the uncertainty found in the process of diagnosis of disease that has most frequently been the focus of applications, in particular in the design of expert systems. More recently, some workers have been trying to apply fuzzy logic concepts to populational biology with emphasis in epidemiological problems like causal studies, epidemic

models and designing of vaccination strategies. Some of those works are introduced along this book.

1.2 Epidemiological Modeling

The term epidemiology was apparently not used until the 1860s (Lilienfeld, 1979) and the profession only emerged in the early part of the 20th Century (Winslow, 1952). Certainly the work of Thomas Proudfoot (1833) and Henry Gaulter (1833) during the 1832 British cholera epidemic and of John Sutherland (1850), John Snow(1855) and William Budd (1849) during the mid-century epidemic was of a type that we would now recognize as epidemiological.

The last third of the twentieth century has seen rapid growth in the understanding and synthesis of epidemiological concepts (Rothman & Greenland, 1998). However, the fundamental concepts of epidemiology do not depend on empirical results but rather on the capacity of epidemiologists to formulate a theory of epidemiological concepts.

Physicians, the great majority of epidemiologists until the 1970s, have collaborate fruitfully with statisticians. Much of the theoretical development of modern epidemiology was contributed by statisticians (Rothman & Greenland, 1998), in particular in the epidemiological concepts related to chronic-degenerative diseases.

Epidemiology has been defined as the *study of disease distribution and its determinants in human populations* (Rothman & Greenland, 1998). The historical paradigm of epidemiological investigation still is the work by John Snow on cholera in the eighteenth century.

From the quantitative point of view, it is possible to identify two group of diseases amenable to be treated by different methods: the epidemiology of infectious diseases, and the epidemiology of non-infectious, also known as chronic-degenerative diseases (Massad, 1996). They differ in many aspects, the most important of which is the fact that infectious disease epidemiology is characterized by the presence of at least one other active player in addition to the human population, namely, the infectious agent or parasite (Halloran, 1998). As a consequence of *transmission*, a characteristic aspect of infectious disease epidemiology, is that, unlike non-infectious disease, the occurrence of infectious events depends on the number (or proportion) of susceptible AND infected individuals. This implies in non-linearities in the dynamical models of transmission.

Non-infectious disease epidemiology has dominated the stage of epidemiology. This is amply testified by the uneven number of authors working in either infectious or non-infectious disease epidemiology, characteristically biased in favor of the latter. If we take, for instance, one of the most recent textbook on epidemiology, the magnificent second edition of Rothman and Greenland's *Modern Epidemiology*, of its 32 chapters (and a total of 735 pages), only one chapter (25 pages) is specifically dedicated to infectious disease epidemiology. Table 1.1 summarizes the main differences between non-infectious and infectious diseases epidemiological approaches.

Table 1.1. Differences between non-infectious and infectious diseases epidemiologies

EPIDEMIOLOGY

Disease's character	Non-infectious	Infectious
Objective	*Causality*	*Control*
Cognitive approach	*Induction*	*Deduction*
Tools	*Statistics*	*Mathematics*
Models	*Functional*	*Structural*
Underlying aims	*Risk Factors*	*Mechanisms of the disease*

This bias towards non-infectious epidemiology has historical roots and is explained by the secondary role of infections as morbidity and mortality causes in industrialized countries (as a consequence of the development of hygienic habits, the discovery of vaccines and antibiotics), the main centers of theoretical developments of epidemiology. It explains the consequent theoretical bias of statistics, which permeates the epidemiological literature. As the main objective of non-infectious disease epidemiology is the establishment of *risk factors* (in a *causal* setting), heavy theoretical investment has been done in the development of statistical tools for dealing with risk analysis. In spite of the leading role of infectious diseases in the developing world, both in morbidity and mortality terms, for which *control* and risk analysis should be the target, only a small amount of the epidemiological editorial space has been dedicated to this sort of disease. However, history vindicated itself and the breathtaking number of emerging and re-emerging infections recorded in the last two decades, for which all the statistical armamentarium currently available is of limited use, is charging a heavy toll for both develop and developing countries. The hope lies in the development of quantitative tools able to deal with the strong non-linearities characteristic of infectious disease systems, in particular, mathematical models.

1.3 Mathematical Modeling

It is a now well established fact that mathematical modeling has demonstrated its value, playing a crucial role in the humanity's progress, both in describing and understanding the world and in the technological development. We can even say that physics and engineering were the main responsible for the progress of mathematical modeling, given the success of the models elaborated in those areas along the last four centuries, culminating with the technological revolution of the 20^{th} century. In addition, the power of mathematical models is spreading across all the theoretical and practical fields, from areas like health to social sciences, economy, management, business and finance. Due to its predictive capacity and its ability to provide a quantitative analysis of the phenomena, biomathematics is an area in expansion. Also, the provision of technological resources like computer power is allowing the development of more complex and bold models.

This tempting scenario is leading many researchers to propose models closed of the reality, making us to believe that the natural world could be described

through mathematical structures. However, the real world is much less precise and predictable than classical mathematical equations usually are, demanding appropriate tools to deal with uncertainties, imprecisions and vagueness. The desire to approximate the mathematical ideal to the real world has been impelling the development of *Artificial Intelligence* and *Soft Computing* techniques, areas of which fuzzy logic and fuzzy sets theory are part. *Artificial Intelligence* (AI) may be defined as "the branch of computer science that is concerned with the automation of intelligent behavior" (Luger & Stubblefield, 1998), and *Soft Computing* refers to AI areas that deal if soft/vague information.

Mathematical modeling requires creative and technical processes in the elaboration of the model as well as in the evaluation of the results. Building mathematical models involves several steps, such as: identification of the problem; definition of the assumptions focused in the particular aspects of the specific phenomenon; identification of the most important and relevant state variables and parameters, and the relationship between all of them; the choice of the mathematical structure to model, aiming to reach the best and more appropriated results; implementation of the computing with this mathematical structure, generating the results data; and evaluation of the model's performance. In all of these steps uncertainties and vagueness can be present, so soft computing techniques can be useful in different moments of the mathematical models building.

1.4 Fuzzy Logic in Biomedicine and Epidemiology

The first article considering the application of fuzzy sets theory in life sciences was proposed by L. A. Zadeh in 1968 titled *Biological applications of the theory of fuzzy sets and systems*, which was published in a book edited by L. D. Proctor in the following year (Zadeh, 1969). However, in spite of the proximity between fuzzy logic ideas and vagueness pertinent to the medical reasoning, few works applying this theory in medicine and correlated areas were published up to the 1980s. In fact, if we consider *MEDLINE*, the most important search database in biomedicine, we find that only 15 articles were published until 1991. Figure 1.1 presents the time evolution of fuzzy logic publications in *MEDLINE* database. We can note that the impact of fuzzy logic in biomedicine grew just after the 1990s, culminating with 265 articles published only in 2006.

Similar behavior is found if we consider the number of scientific journals that have published applications of fuzzy logic in biomedicine areas. Perhaps the most important of these journals is *Artificial Intelligence in Medicine* (AIMED), founded in 1989 by K. Sadegh-Zadeh. AIMED is a journal devoted to interdisciplinary publications, concerning the theory and practice of artificial intelligence in medicine, human biology and heath care. The first fuzzy logic publication in AIMED was in the volume 1, issue 2, in 1989. In this article the authors proposed a fuzzy decision-making for ECG diagnosis (Degani & Bortolan, 1989). Since then, AIMED had published more than 100 papers on fuzzy sets application in medicine and correlated areas, covering diagnosis systems, image processing, bio-engineering, decision making process and so on. A deeper analysis of the

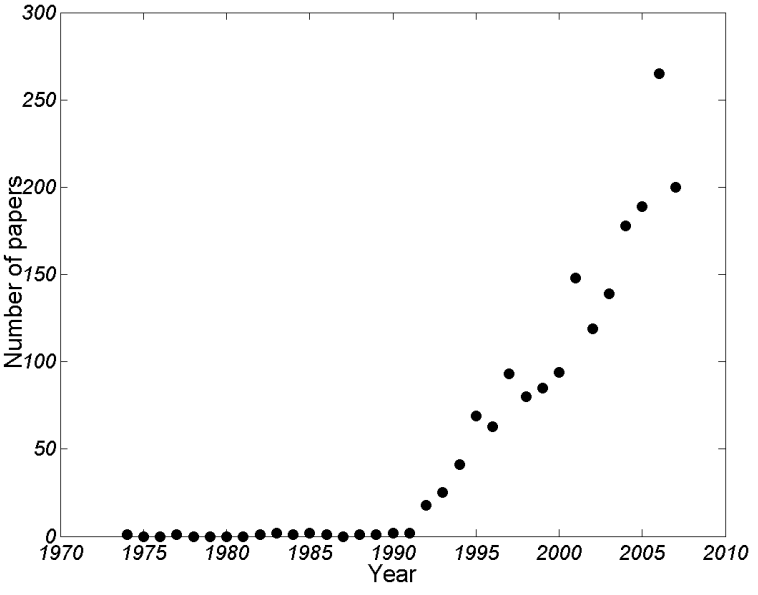

Fig. 1.1. Number of papers on fuzzy logic available in the *MEDLINE* database

development of fuzzy and neuro-fuzzy systems in medicine, from a historical point of view, can be found in Teodorescu *et al.* (1999b).

Despite the vagueness, imprecision and uncertainties also present in epidemiological problems, application of fuzzy logic in this area are recent and timid. One of the first articles that proposed the use of fuzzy sets theory in epidemiology was *Application of fuzzy numbers to assess opinion in epidemiologic studies* by Merilan and Roe, published in 1993 in the American Journal of Epidemiology. Two years later Bassanezi and Barros published the article titled *A simple model of life expectancy with subjective parameters*, in which the authors present an alternative structure to consider the fuzziness of parameters from a dynamical systems point of view (Bassanezi & Barros, 1995). In March 1997, Massad and collaborators presented two works that addressed the role that fuzzy logic can play in epidemiological studies and its interdisciplinary features: *Fuzzy logic in the analysis of vulnerability to HIV/AIDS infection in sexual partners of injecting drug users*, in the 8^{th} International Conference on the Reduction of Drug Related Harm, in Paris; and *Fuzzy Dynamic Systems in Epidemic Modeling*, in the 11^{th} International Conference on Mathematical and Computer Modeling and Scientific Computing, in Washington, USA (Massad *et al.*, 1997a and 1997b). In the same year Massad *et al.* presented two other works in the 2^{th} International Computer Science Conventions Symposium on Soft Computing (SOCO'97), in Nimes, France. In one of them they proposed the use of fuzzy linguistic models in AIDS clinical progression identification, and in the other they treated the uncertainties in risk estimations in epidemiology (Massad *et al.*, 1997c; Struchiner

et al., 1997). Also in 1997 Estrada-Peña proposed a fuzzy rule-based model to predicts the habitat availability for populations of a certain kind of tick, for the epidemiological surveillance of this tick populations in Spain (Estrada-Peña, 1997), which is probably the first fuzzy logic application in veterinary epidemiology.

The first practical application of fuzzy sets theory in public health, to the best of our knowledge, happened in São Paulo City, Brazil, when a system of fuzzy decision making was developed, in the consensus form with several types of health professionals, seeking to choose the best vaccination strategy against measles. The strategy proposed by the model was indeed implemented in the whole São Paulo State, denoting the capacity of adhesion of models based on fuzzy logic from the part of the health authorities (Massad *et al.*, 1999). Since then, few fuzzy epidemiological systems have been developed. In fact in the *MEDLINE* database we can found around 15 articles truly associated to epidemiological tasks (Ohayon, 1999; Pereira *et al.*, 2001; Campisi *et al.*, 2006; Drumond *et al.*, 2007). One of the main reasons for this absence is the fact that mathematical models in epidemiology are, in general, non-linear dynamical systems, which still consists in a challenge from the fuzzy systems point of view.

Thus, the future development of applications of fuzzy logic in epidemiology lies in the fields of modeling and system dynamics. A huge amount of editorial space has been dedicated to the development of mathematical models as applied to biomedical problems. However, uncertainty in the system structure and parameter values are still to be tackled with a formalism robust enough to circumvent some of the practical hurdles involved in the process of modeling real medical/biological problems. Also, the hybrid models combining fuzzy logic, neural networks and distributed processing (called soft computing) have a tremendous potential of practical applications. Indeed, the close association between artificial neural networks and fuzzy logic has lead to exciting developments in both fields, including the possibility of the extraction of "fuzzy" if-then rules from neural networks. Another important recent development has been the use of genetic algorithms for this purpose, and it seems likely that they will have a wider role to play. The mathematical problems involved in the process of fuzzy modeling, however, are still very important barriers to be overcome.

In addition, the modeling of uncertainty has recently attracted the attention of the wider mathematics community, and recent work suggests that alternative approaches may be valuable. Classical topology and ideas of partial metrics and filters have been found useful in developing alternative approaches. There has been a creative tension between the modeling of uncertainty using "fuzzy" methods, and those based on a probabilistic formalism. The emergence of a theory of possibility seems to be going some way toward clarifying the essential differences.

In summary, the epidemiological applications of fuzzy logic have been still isolated and/or insufficiently developed. As mentioned above, the prospective future of applications of fuzzy sets theory in biomedicine and epidemiology is very encouraging. We think that the foundations of fuzzy logic applications in

medicine and biology are already mature. However, some of the potential applications generally recognized as very important still involve mathematical difficulties which appear to be insurmountable. Only an interdisciplinary approach, a characteristic proposed by the very philosophy which motivated this book, can orient and motivate professionals from different areas to team up in order to circumvent the mathematical difficulties that still limit the applications of fuzzy logic to epidemiology and beyond.

1.5 Chapter Descriptions

This book is organized in three parts: basic concepts in fuzzy sets theory and epidemiology (chapters 2 to 7); fuzzy dynamical systems in epidemiology (chapters 8 to 11); and advanced techniques and overview in epidemiology and beyond (chapters 12 to 14).

In chapter 2 we present an introduction to fuzzy set theory ideas, showing the basic concepts of fuzzy sets and its relations, in a simple mathematical language. Aspects of uncertainties treatments of epidemiological variables is also discussed.

Chapter 3 is devoted to the modern epidemiology view, where aspects of modern epidemiology are briefly presented. Quantitative, statistic and other classical mathematical models in epidemiology are discussed, taking into account transmitted and non-transmitted diseases.

In chapter 4 we present the basic concepts of fuzzy probability and the probability of fuzzy events, discussing the differences between fuzzy and probability measures. The fuzzy expected value is also defined and the probabilities of fuzzy events are applied to answer epidemic questions.

In chapter 5 we discuss the role that fuzzy logic can play in the estimation of epidemiological risk, in a causal context. This is discussed taking into account the uncertainties and heterogeneity of individuals classification. The Odds Ratio concept is generalized in a fuzzy structure and a simulation is presented.

In chapter 6 we consider the fuzzy decision making in public health strategies design. The Bellman and Zadeh fuzzy decision making model is presented and it is discussed in an epidemiological context. A fuzzy model to design a vaccination strategy against measles performed inf São Paulo State is showed. Different control approaches in public health are discussed, considering their power in health authorities adhesion.

Fuzzy based-rule models are detailed in chapter 7. The main concepts and structures of fuzzy linguistic models is showed, and Mamdani and Sugeno approaches are presented and compared, from the epidemiological point of view. Three examples of fuzzy based-rule applied in the epidemiological context are presented: a fuzzy linguistic model for HIV Natural History; a fuzzy model to estimate the risk of neonatal death; and a fuzzy model to the quality of life evaluation.

Different approaches of fuzzy dynamical systems in epidemic modeling are presented in chapters 8, 9 and 10. Chapter 8 presents a dynamic structure based on linguistic fuzzy models, in which the dynamical process in found through the

iterative procedure on the fuzzy rules. This approach is applied in a SIS and in a SIR epidemic models to treat, respectively, canine rabies and measles spreading studies.

In chapter 9 two fuzzy dynamical approaches based on the differential equation structure and on the fuzzy differential inclusion are presented. Both the concepts of demographic and environmental fuzziness are analyzed with these approaches, and epidemiological studies are also done. Important epidemic concepts as the Basic Reproduction Number, R_0, are computed through SI ans SIS fuzzy models, comparing the results with those found by classical equations dynamic.

An alternative approach to elaborate fuzzy dynamical systems is presented in chapter 10. In this approach the fuzziness of the parameters of the classical differential equations is treated through a rule-based fuzzy model, resulting in a mixing of the dynamical structures exposed in the two previous chapters. This methodology is illustrated by two epidemiological examples: a HIV model for dynamical behavior between non-symptomatic and symptomatic seropositives individuals and a study of the influence of HIV epidemic in the expectancy of life in a group of seropositive individuals.

In chapter 11 we present a fuzzy Reed-Frost model, taking into account the heterogeneous infectivity in the individuals. The classical Reed-Frost model and a stochastical generalization of it are also presented and differences between probabilistic and possibilistic approaches are discussed. The dynamics of the models are simulated and the results are compared with real data on infectious disease.

Since epidemiological processes are non-linear and complex systems, sometimes it is necessary to aggregate complementary tools to treat the problem in a more realistic way. So, hybrid models in epidemiology are discussed in chapter 12, where different techniques, classical and fuzzy ones, are combined in a epidemiological context. To illustrate the power of this mixing we present two examples: the Bayesian statistic test with fuzzy hypothesis and a fuzzy linguistic model, in a decision making support context, for the optimal age for vaccination against measles; and a model that combines linguistic fuzzy models and numerical calculus to elaborate a predator-prey model to study the interaction between aphids and ladybugs in citriculture.

Finally, chapter 13 presents an overview of fuzzy logic in medical diagnosis, covering the diagnostic process, mathematical modeling of medical diagnostics, expert systems, Bayesian reasoning, belief networks, causal networks, and so on. Chapter 14 closes the book with the last remarks about the authors expectations of fuzzy logic applications in epidemiology and beyond, concerning about future perspectives of that theory.

1.6 The Authors

Eduardo Massad is Professor of Medical Informatics at the School of Medicine, the University of São Paulo, Brazil, is Honorary Professor of Tropical and Infectious Diseases at the London School of Hygiene and Tropical Medicine, UK and

is the Courage Fund Visiting Professor of Infectious Diseases at the National University of Singapore. He is co-author of "The Brain: From Fuzzy Arithmetic to Quantum Computing", Studies in Fuzziness and Soft Computing Vol. 165, 2005, Springer Verlag, ISBN 3-540-21858-0 (edmassad@usp.br).

Neli Regina Siqueira Ortega is bachelor and master in physics and Ph.D. in sciences by the University of São Paulo. Since 2001 she is researcher of Medical Informatics at the School of Medicine, the University of São Paulo, Brazil, where she teaches graduate courses of fuzzy sets theory in biomedicine and epidemiology, which consists in her main line of research (neli@dim.fm.usp.br).

Laécio Carvalho de Barros is Professor of Applied Mathematics at the Intitute of Mathematics, Statistics and Computational Sciences, the University of Campinas, and holds a Ph.D. degree in Applied Mathematics from the University of Campinas, São Paulo, Brazil, in 1997. His current research interests are modeling of biological phenomena and fuzzy sets theory (laeciocb@ime.unicamp.br).

Claudio José Struchiner is Professor of Epidemiology and Biostatistics at the Oswaldo Cruz Foundation in Rio de Janeiro, Brazil, and holds a DSc degree in epidemiology from the Harvard School of Public Health. Dr. Struchiner's research interests include the development of epidemiological methods for the study of infeccious diseases and the evaluation of vaccine efficacy (stru@fiocruz.br).

Acknowledgments

The authors would like to thank Dr. Harki Tanaka for his precious contribution in the final edition of this book, particularly in the figures edition, to Sandra Fabiana Almeida for the careful references revision and to the Brazilian institutions that supported the authors - CNPq, FAPESP, CAPES, FAPERJ, FINEP and LIM01/FMUSP.

2 Basic Concepts of Fuzzy Sets Theory

The concept of fuzzy sets was proposed by L. A. Zadeh in his paper published in 1965 (Zadeh, 1965). Since this seminal work, several researchers have importantly contributed for the development of fuzzy sets theory and its applications, resulting in the great success from the theoretical and technological points of view. Fuzzy sets theory proposes to deal with unclear boundaries, representing vague concepts and working with linguistic variables. In this sense, fuzzy sets emerged as an alternative way to deal with uncertainties.

The basic idea of a fuzzy set concerns the flexibility over the concept of belongingness. In classical sets theory one is able to classify (organize) objects in collections through a binary processes: accepting or rejecting the object as belonging to that collection. In fact, in a great number of situations it is possible to say if the element x belongs or not to the set A, that is, $x \in A$ or $x \notin A$. One can affirm without doubt, as an example, that the number 5 belongs to the natural numbers set and that -5 does not belong to this set. However, what to say about the fact that number 4 belongs to the group of numbers defined as "*around 5*"? Clearly, the answer in this case will depend on the context. The main difference between classical and fuzzy sets is that in the former there is a dichotomy notion that should necessarily be preserved.

There are many books available about fuzzy sets, many of them presenting this theory with high quality and detail (Negoita & Ralescu, 1975; Yager & Filev, 1994; Klir & Yuan,1995; Reznik, 1997; Pedrycz & Gomide, 1998 and 2007; Nguyen & Walker, 2000). So, the main objective here is rather to provide sufficient concepts for the reader to following the discussions presented, than to address in depth fuzzy sets theory.

2.1 Fuzzy Sets

Even though many sets present sharp boundaries (e.g., sex, matrimonial state, the set of students), there are several situations that a belongingness relation is not well-defined (e.g., the set of slim people, good students, high temperature). In these cases it is not easy to say if the element belongs or not to the given set.

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 11–40, 2008.
springerlink.com

To solve this problem Zadeh proposed a membership degree, according to which an element could belong partially to the set. In order to find a mathematical formalization to fuzzy sets Zadeh's approach based on the fact that any classical set could be represented by its characteristic function.

Definition 2.1. *Let U be a non-empty set and A a subset of U. The characteristic function of A is given by:*

$$A(x) = \begin{cases} 1 \; if \; x \in A \\ 0 \; if \; x \notin A \end{cases} .$$

Note that we are using the same symbol to designate the set, A, and its characteristic function, $A(x)$. So, $A(x)$ is a function whose domain is U and the image is contained in the set $\{0, 1\}$, with $A(x) = 1$ designating that the element x belongs to subset A, and $A(x) = 0$ designating that the element x does not belong to A. In this sense, the characteristic function $A : U \to \{0, 1\}$ describes completely the subset A in a well-defined boundary context. The fuzzy set concept consists in relaxing the belongingness constraint required by the function above, assuming intermediate membership values into the unit interval $[0, 1]$. The membership values can be understood as the degrees to which each object is compatible with the properties or features that characterizes the group (Pedrycz & Gomide, 1998 and 2007).

Definition 2.2. *Let U be a classical non-empty set. A fuzzy subset F of U is described by a function*

$$F : U \to [0, 1], \tag{2.1}$$

called membership function of fuzzy set F.

The value $F(x) \in [0, 1]$ indicates the membership degree of the element x of U in fuzzy set F, with $F(x) = 1$ and $F(x) = 0$ designating, respectively, the belongingness and not-belongingness of x in F, recovering the classical context[1]. In this sense, it is possible to say that a classical set is a particular case of a fuzzy set, whose membership function is its characteristic function. Note that the membership function of empty, $\emptyset$, and universe, U, sets are, respectively, $\emptyset(x) = 0$ and $U(x) = 1$ for all $x \in U$.

In the mathematical modeling context, the main contribution of fuzzy sets is its ability to deal with linguistic expressions, whose frontiers are not well-defined. As an example, consider the statement "*around 5*" in the natural numbers universe discussed above. In some contexts one could propose the following membership degrees for this set: $F(0) = 0$, $F(1) = 0.10$, $F(2) = 0.25$, $F(3) = 0.50$, $F(4) = 0.75$, $F(5) = 1$, $F(6) = 0.75$, $F(7) = 0.50$, $F(8) = 0.25$, $F(9) = 0.10$

[1] The membership degree of the element x of a fuzzy set F is also represented by $\mu_F(x)$, which is particularly useful in the membership functions context. In this book both notations are used, $F(x)$ and $\mu_F(x)$, depending on the purpose.

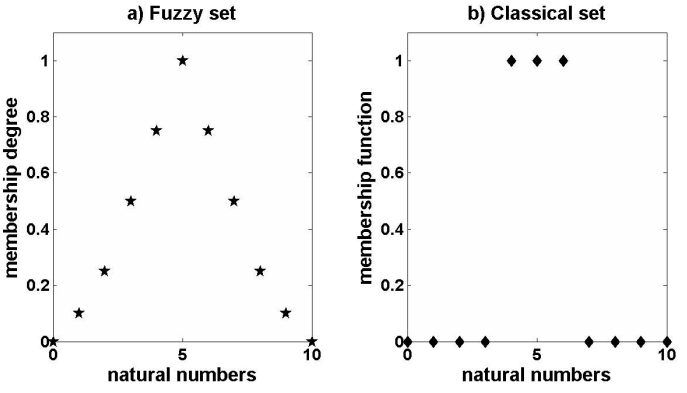

Fig. 2.1. (a) Fuzzy set to statement "around 5" in the Natural numbers; and (b) Classical set for the same statement

and $F(n) = 0$ if $n \geq 10$, which could be represented by a function as shown in figure 2.1(a). The value $F(4) = 0.75$, for instance, means that the compatibility degree of the number 4 with the statement "*around 5*" is 0.75, which reflects the uncertainty to classify this number as close to 5. Only for comparison, a classical set for the same statement, "*around 5*", could be the set $A = \{4, 5, 6\}$, whose membership function is presented in figure 2.1(b). In contrast, considering the classical set A the true value for number 4 is 1, highlighting the dichotomy situation.

When the universe set U is discrete or is a finite set with n elements, a fuzzy subset F of U is commonly represented by:

$$F = F(x_1)/x_1 + F(x_2)/x_2 + ... + F(x_n)/x_n$$

where the symbol "+" is only a mathematical notation, and it does not denote the standard algebraic summation. Usually, for simplicity, only elements of U with nonzero membership degree are listed. So, the fuzzy subset of the natural numbers universe "$F = around\ 5$" may be represented by:

$$F = 0.10/1 + 0.25/2 + 0.50/3 + 0.75/4 + 1/5 + 0.75/6 + 0.50/7 + 0.25/8 + 0.10/9. \tag{2.2}$$

In many applications the fuzzy set cannot be a discrete set. In those cases, the set could be represented by a continuous function. Consider, for instance, the case in which one wishes to classify the individuals of some population as exposed or not exposed to cigarette smoke. Basically, the level of "*cigarette smoke exposition*" depends on two factors: 1) the level of cigarette smoke inhaled by his/her own consume, if the person is a smoker; and 2) the level of cigarette smoke inhaled from his/her environment; both factors considering a period of time.

According to the first factor, a person that usually smokes 40 cigarettes per day is more exposed than another that consumes, say, 3 cigarettes per day.

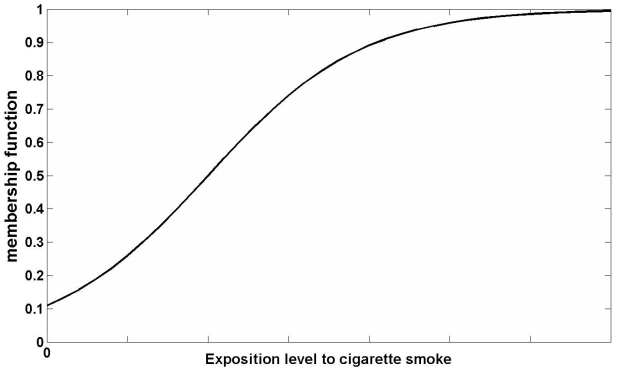

Fig. 2.2. Fuzzy set to describe the exposition to cigarettes smoke

Naturally, there are many intermediate cigarettes consumers between 3 and 40 cigarettes and, consequently, the exposition intensity varies continually. Analogous reasoning is valid if we consider the quantity of cigarette smoke present in the environment, for the second factor. In this sense, even people that do not smoke, but nevertheless are exposed to cigarette smoke, either in work places or because their partners are smokers, has a certain level of exposition. For simplicity we can consider that there are a basal level of cigarette smoke exposition for everybody.

Thus, this uncertainty to classify the individuals as exposed to cigarette smoke can be represented by a fuzzy set of "*exposition level to cigarette smoke*" shown in figure 2.2. Note that in this figure the membership degree for the non exposed, $F(0)$, is not zero, representing the situation described above. So, if even the person does not smoke he/she is considered exposed to cigarette smoke with a basal membership degree equal to 0.1, in this example. Also, it is worth noting that the shape of the membership function is an arbitrary choice, depending on the context.

There are infinite functions capable of describing a membership function associated with a linguistic predicate. The choice depends on the concept to be represented and on the context in which this is used. However, independently of the fuzzy sets shape it is possible to operate with them by means of the applications of mathematical operators over their membership functions, as it is commonly done in classical sets theory. In next section we present the operators to aggregate fuzzy sets most used in the majority of practical applications, highlighting their most important properties.

2.2 Operations with Fuzzy Sets

Before discussing the operations over fuzzy sets it is interesting to review the classical sets operators from a new point of view, that is, from their characteristic

functions. The fundamental operations over classical sets are union, intersection and complement. Consider the A and B subsets of the universe set, U. Then, the union of the sets A and B (denoted by $A \cup B$) is the collection of those objects that belong either to A or B. The intersection of A and B (denoted by $A \cap B$) is the collection of those objects that belong both to A and B. The complement of a set A (denoted by A^cor $\overline{A}$) is the collection of those objects that belong to U but do not belong to A. Although it is not common, all of these operations could be described through the characteristic function defined in 2.1. In terms of their characteristic functions, the union, intersection and complement of sets A and B could be defined, respectively, by:

$$(A \cup B)(x) = \max[A(x), B(x)], \tag{2.3}$$

$$(A \cap B)(x) = \min[A(x), B(x)], \tag{2.4}$$

and

$$A^c(x) = 1 - A(x), \tag{2.5}$$

for all $x \in U$, where $A(x)$, $B(x)$, $(A \cup B)(x)$, $(A \cap B)(x)$ and $A^c(x)$ are the results of the characteristic function of the element x to the sets A, B, $A \cup B$, $A \cup B$ and A^c, respectively. Note that the most important requirement to the intersection operator is that it must return 1 when both arguments are 1, and 0 otherwise. Likewise, the union operator must return 1 if at least one argument is equal to 1, and to 0 if both arguments are 0.

Example 2.3. Suppose that the universe set U is a collection of patients labeled 1, 2, 3, 4 and 5. Let A and B be the classical sets composed by patients that present fever and cough, respectively. So, table 2.1 illustrates the union, intersection and complement of the sets A and B by application of the expressions (2.3-2.5).

Therefore, the set of intersection is composed only by patients 2 and 5, who have simultaneously fever and cough.

Table 2.1. Illustration of the union, intersection and complement of the sets A and B, from the classical point of view

Patient (x)	**Fever A(x)**	**Cough B(x)**	**(A ∪ B)(x)**	**(A ∩ B)(x)**	**A^c(x)**	**B^c(x)**
1	1	0	1	0	0	1
2	1	1	1	1	0	0
3	0	1	1	0	1	0
4	0	0	0	0	1	1
5	1	1	1	1	0	0

As in the classical sets operations, their main properties may also be described through their characteristic functions. So, for all $x \in U$ we have:

1. Commutativity:

$$(A \cup B)(x) = A(x) \cup B(x) = B(x) \cup A(x) = (B \cup A)(x)$$
$$(A \cap B)(x) = A(x) \cap B(x) = B(x) \cap A(x) = (B \cap A)(x)$$

2. Associativity:

$$(A \cup (B \cup C))(x) = (A(x) \cup B(x)) \cup C(x) = ((A \cup B) \cup C)(x)$$
$$(A \cap (B \cap C)) = (A(x) \cap B(x)) \cap C(x) = ((A \cap B) \cap C)(x)$$

3. Idempotency:

$$(A \cup A)(x) = A(x) \cup A(x) = A(x)$$
$$(A \cap A)(x) = A(x) \cap A(x) = A(x)$$

4. Distributivity:

$$A(x) \cap (B(x) \cup C(x)) = (A(x) \cap B(x)) \cup (A(x) \cap C(x))$$
$$A(x) \cup (B(x) \cap C(x)) = (A(x) \cup B(x)) \cap (A(x) \cup C(x))$$

5. Involution:

$$(A^c)^c(x) = 1 - A^c(x) = 1 - (1 - A(x)) = A(x).$$

The operations over fuzzy sets are generalized from the classical ones, by changing the characteristic function into the membership function. In this sense, the *standard fuzzy operators* for union, intersection and complement are defined as in (2.3-2.5), where A and B are fuzzy subsets of U, and $A(x)$, $B(x)$, $(A \cup B)(x)$, $(A \cap B)(x)$ and $A^c(x)$ are the *membership degrees* of the element x to the sets A, B, $A \cup B$, $A \cap B$ and A^c, respectively. Figures 2.3, 2.4 and 2.5 show these operations on fuzzy sets.

Example 2.4. Consider the situation described in example 2.3, where fever and cough are now fuzzy sets, that is, the patients may belong partially to these sets. Thereby, a patient with body temperature equal to 39 Celsius degrees is considered certainly with fever ($A(39) = 1$), and another with body temperature equal to 37 Celsius degrees is considered as belonging to the fever set with 0.3 membership degree. Table 2.2 presents a possible fuzzification of the situation described in this example.

Observe that in the fuzzy context it is possible to perform a better discrimination of the patients condition, since patients 1 and 2, for instance, are equally members of the fever set in a classical context but this could be different in a fuzzy approach.

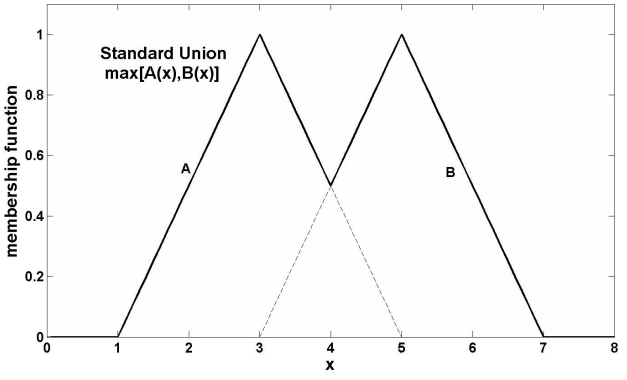

Fig. 2.3. Standard union of fuzzy sets

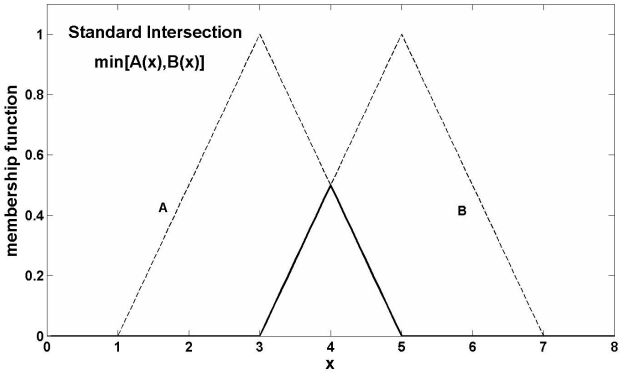

Fig. 2.4. Standard intersection of fuzzy sets

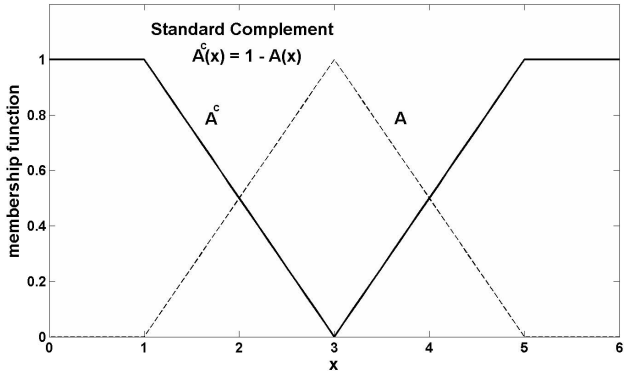

Fig. 2.5. Standard complement (negation) of fuzzy sets

Table 2.2. Illustration of the union, intersection and complement of the sets A and B, from the fuzzy sets point of view

Patient (x)	Fever A(x)	Cough B(x)	(A ∪ B)(x)	(A ∩ B)(x)	$A^c(x)$	$B^c(x)$
1	0.7	0.4	0.7	0.4	0.3	0.6
2	1.0	1.0	1.0	1.0	0.0	0.0
3	0.2	0.7	0.7	0.2	0.8	0.3
4	0.5	0.5	0.5	0.5	0.5	0.5
5	1.0	0.8	1.0	0.8	0.0	0.2

One of the most interesting consequences of the fuzzy set definition, in contrast with its classical counterpart, is the failing of the Law of Excluded Middle and the Law of Contradiction. In the classical approach we have that $A \cup A^c = U$, the Law of Excluded Middle, and $A \cap A^c = \emptyset$, the Law of Contradiction. However, due to the membership function flexibility this does not occur in fuzzy sets theory. In other words, it is possible for an element to belong partially both to fuzzy sets A and its complement A^c. We may observe in figures 2.6 and 2.7 that $A \cup A^c \neq U$ and $A \cap A^c \neq \emptyset$, as a result of the fuzziness involved.

Another interesting relation, and useful in the modeling context, of fuzzy sets is the *inclusion*. Let A and B be fuzzy sets. So it is possible to say that A is a

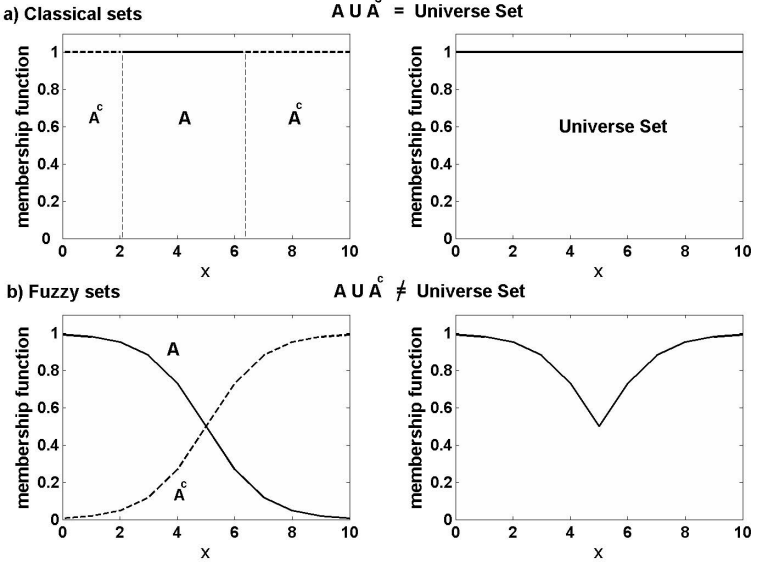

Fig. 2.6. Law of Excluded Middle in the classical and fuzzy approaches

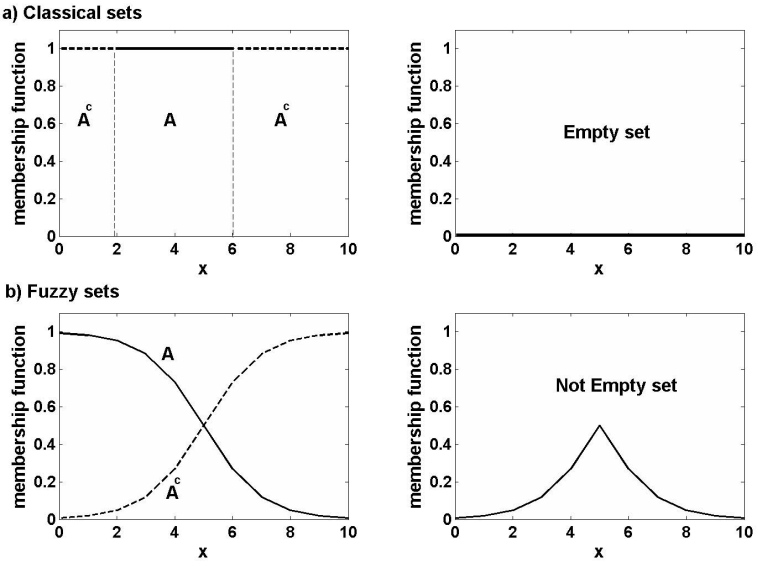

Fig. 2.7. Law of Contradiction in the classical and fuzzy approaches

subset of B, $A \subseteq B$, iff $A(x) \leq B(x)$ for all $x \in U$. Note that, since the empty set $\emptyset$ has membership function $\emptyset(x) = 0$ and the universe set has membership function $U(x) = 1$, for all $x \in U$, then it is possible to affirm that $\emptyset \subseteq A$ and $A \subseteq U$ for all A.

Example 2.5. To illustrate the inclusion concept among fuzzy sets, consider a study whose main objective is to evaluate the impact of poverty on life expectancy of a group of individuals, as proposed by Bassanezi and Barros (1995). Assuming that poverty may be characterized by the individuals' income, a possibility for the "poor" fuzzy set (P_k) is the membership function:

$$P_k(r) = \begin{cases} \left[1 - (\frac{r}{r_0})^2\right]^k & if\ 0 \leq r \leq r_0 \\ 0 & if \quad r > r_0 \end{cases}, \tag{2.6}$$

where r is the income, r_0 is the income threshold above which life expectancy of the group is not affected, and k is an environmental parameter that characterize the group (for instance, rural, urban, Indians, etc). Figure 2.8 illustrates the poverty fuzzy sets for two values of k.

It is easy to see in the figure that if $k_1 \geq k_2$ then $P_{k_1} \subset P_{k_2}$, since given an income r, the degree of poverty obeys $P_{k_1}(r) \leq P_{k_2}(r)$, for all r. So, it is possible to say that P_{k_1} is the fuzzy set of poor individuals, while P_{k_2} is the fuzzy set of very poor individuals.

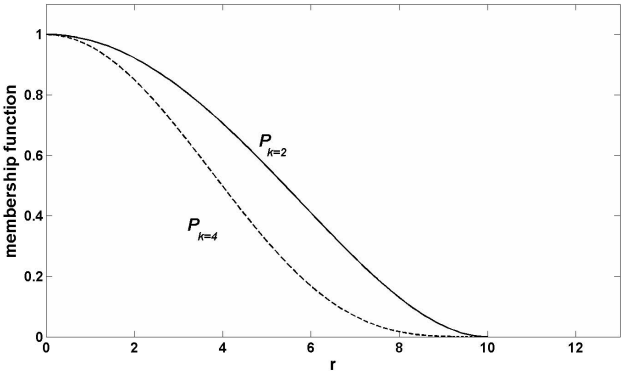

Fig. 2.8. Fuzzy set for "poor", $P_{k=2}$, and for "very poor", $P_{k=4}$. The set "very poor" may be seen as a fuzzy subset of "poor", since a person that is "very poor" is also "poor" (Bassanezi & Barros, 1995).

The most important classical sets operations properties could be generalized and verified for the *standard fuzzy operations*, particularly the properties of commutativity, idempotency, associativity, distributivity, and involution. As in classic logic, the operations of union and intersection correspond, in the fuzzy sets context, to the disjunction (*or*) and conjunction (*and*) operators respectively. In fuzzy sets theory these disjunction and conjunction operators are generalized into the so-called *triangular conorms* and *triangular norms*, respectively.

2.2.1 Triangular Norms and Conorms

In fuzzy set theory the choice of the operations to disjunction and conjunction between fuzzy sets is arbitrary. The operators *max* and *min* chosen previously to designate the union and the intersection between fuzzy sets, called *standard fuzzy operators* chosen to generate expressions (2.3-2.5). However, in the fuzzy context, there are many ways to define these operators.

The conjunction operator could be generalized for any *triangular norm*, also called $t-norm$ and denoted by $t(x, y)$, which is a mapping $[0, 1] \times [0, 1]$ into $[0, 1]$, that satisfies the following set of axioms, for all $x, y, z, w \in [0, 1]$:

i. $t(x, y) = t(y, x)$ (commutativity)
ii. $t(x, t(y, z)) = t(t(x, y), z)$ (associativity)
iii. $t(x, y) \leq t(z, w)$ if $x \leq z$ and $y \leq w$ (monotonicity)
iv. $t(x, 1) = t(1, x) = x$ (boundary conditions).

In the same way, it is possible to define fuzzy disjunction operators as a *triangular conorms*, also called $t-conorm$ and denoted by $s(x, y)$, as a mapping

$[0,1]\times[0,1]$ into $[0,1]$, that satisfies the following set of axioms, for all $x, y, z, w \in [0,1]$:

i. $s(x,y) = s(y,x)$ (commutativity)
ii. $s(x, s(y,z)) = s(s(x,y), z)$ (associativity)
iii. $s(x,y) \leq s(z,w)$ if $x \leq z$ and $y \leq w$ (monotonicity)
iv. $s(x,0) = s(0,x) = x$ (boundary conditions).

There is a duality between *t-norms* (conjunction) and *t-conorms* (disjunction) operators mediated by the *negation* operator, $x' = 1 - x$. The fuzzy conjunction $t(x,y)$ and disjunction $s(x,y)$ operators form a dual pair if they satisfy the following condition:

$$1 - t(x,y) = s(1-x, 1-y)$$

and

$$1 - s(x,y) = t(1-x, 1-y),$$

which in the classical sets theory corresponds to the DeMorgan's laws:

$$(A \cup B)^c = A^c \cap B^c$$

and

$$(A \cap B)^c = A^c \cup B^c.$$

So, the choice of a fuzzy disjunction operator determines the choice of the fuzzy conjunction operator, and vice versa, if it is wished that the operators are dual.

There are many dual $t-norm$ and $t-conorm$ operators over the negation $x' = 1 - x$, some of them are illustrated below. Due to their simplicity and properties, the most commonly used are the *max* and *min* operators. However, it is common the use of other disjunction/conjunction pairs, like *Algebraic product/algebraic sum, Drastic product/drastic sum* and *Bounded difference/bounded sum*, defined as:

Algebraic product/algebraic sum:

$$t(x,y) = xy \tag{2.7}$$

and

$$s(x,y) = x + y - xy. \tag{2.8}$$

Drastic product/drastic sum:

$$t(x,y) = \begin{cases} \min[x,y] & if\ \max[x,y] = 1 \\ 0 & if\ \max[x,y] < 1 \end{cases} \tag{2.9}$$

and

$$s(x,y) = \begin{cases} \max[x,y] \; if \; \min[x,y] = 0 \\ 1 \qquad if \; \min[x,y] > 0 \end{cases} . \tag{2.10}$$

Bounded difference/bounded sum:

$$t(x,y) = \max[0, x+y-1] \tag{2.11}$$

and

$$s(x,y) = \min[1, x+y]. \tag{2.12}$$

An important property about disjunction operators is that all $t-norms$ are bounded above by the *min* and bellow by the *drastic product*. In the same way, all $t-conorm$ are bounded above by the *drastic sum* and below by the *max* operator. In addition, the *min* and *max* operators are the only $t-norm$ and $t-conorm$ that satisfy the idempotency condition, that is, $t(x,x) = x$ and $s(x,x) = x$, for all $x \in U$. To know more about $t-norms$ and $t-conorms$ we recommend the book by Nguyen and Walker (2000).

2.2.2 Cartesian Product

From the inference and decision making points of view, another important concept in fuzzy set theory is the Cartesian Product. Technically, such operation is similar to the intersection operator. The difference is that in the former the universes sets involved could be distinct, while in the intersection the fuzzy subsets involved are necessarily from the same universe set. In fact, this difference becomes a great advantage of Cartesian product in the applications of fuzzy sets, as we will see in the case of linguistic models and in fuzzy control theory.

Definition 2.6. *Suppose that A and B are fuzzy subsets of the universe sets U and V, respectively. The Cartesian product of A by B is a fuzzy subset of the classic Cartesian Product set $U \times V$ whose membership function is given by:*

$$(A \times B)(x,y) = \min[A(x), B(y)], \tag{2.13}$$

with $x \in U$ and $y \in V$.

That is, if x belongs to A with $A(x)$ degree and y belongs to B with $B(y)$ degree, the pair (x,y) belongs to fuzzy Cartesian product, $A \times B$, with the minimum degree between $A(x)$ and $B(y)$. The Cartesian product concept plays a fundamental role in the *Fuzzy Relations* and, consequently, in linguistic fuzzy modeling, as it will be discussed in the chapter 7.

2.3 Fuzzy Sets and Membership Functions

In principle, any function as $F : U \rightarrow [0, 1]$ is a candidate to membership function, independently of its shape. Clearly, each family of functions presents specific properties and the choice will depend on the application. However, the most commonly used in practice are triangular, trapezoidal, Gaussian and sigmoidal functions. All of them are defined in the universe of the real numbers, that is, the universe set is $U \equiv \mathbb{R}$.

A very special class of fuzzy sets is the so-called "fuzzy numbers", which are defined in section 2.4.2. This importance is due to the fundamental role that their play in fuzzy modeling, particularly in fuzzy linguistic models. In this sense, the majority of the fuzzy sets belongs to the fuzzy numbers class.

A fuzzy subset F of $\mathbb{R}$ is called *Triangular* if its membership function is a triangular function. This function is specified by three parameters, $F(x : a, b, c)$, such as:

$$F(x : a, b, c) = \begin{cases} 0 & if \quad x < a \\ \dfrac{x-a}{b-a} & if \ a \leq x < b \\ \dfrac{c-x}{c-b} & if \ b \leq x < c \\ 0 & if \quad x > c \end{cases} .$$

A fuzzy subset F is called *Trapezoidal* if its membership function is a trapezoidal function, which is specified by four parameters as $F(x : a, b, c, d)$:

$$F(x : a, b, c, d) = \begin{cases} 0 & if \quad x < a \\ \dfrac{x-a}{b-a} & if \ a \leq x < b \\ 1 & if \ b \leq x < c \\ \dfrac{d-x}{d-c} & if \ c \leq x < d \\ 0 & if \quad x > d \end{cases} .$$

Triangular and trapezoidal functions are widely applied, particularly in fuzzy control systems, due to their simple expressions and computational efficiency. Figure 2.9 presents examples of these membership functions. It can be noted that triangular membership function is a particular case of the trapezoidal membership function, when $b = c$.

A *Gaussian* fuzzy set is a set whose membership function is Gaussian. This function is specified by two parameters $F(x : m, \sigma)$, as follows:

$$F(x : m, \sigma) = \exp\left(-\frac{(x-m)^2}{\sigma^2}\right), \tag{2.14}$$

where m and σ denote the center and width of the function. The function shape could be controlled by the parameter σ. If σ is small a "thin" function is

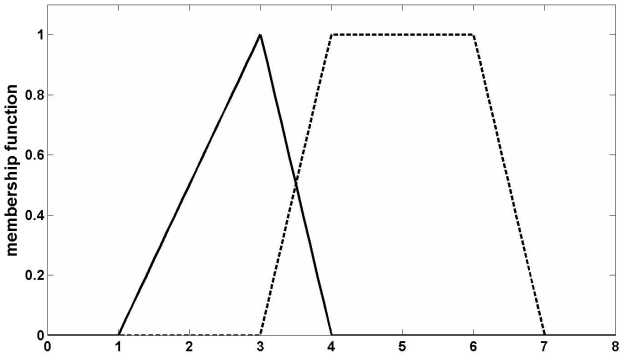

Fig. 2.9. Fuzzy sets represented by triangular and trapezoidal membership functions

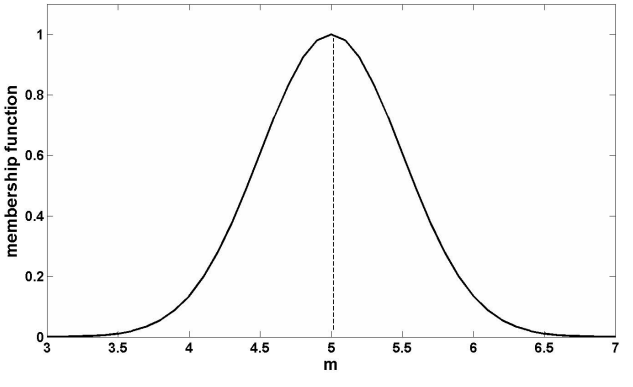

Fig. 2.10. Fuzzy set represented by a Gaussian membership function

generated, while a large σ will generate a "flat" membership function. Figure 2.10 shows an example of Gaussian membership.

A fuzzy subset F is called *Sigmoidal* if its membership function is a sigmoidal function. This function is specified by two parameters $F(x : a, b)$ as follows:

$$F(x : a, b) = \frac{1}{1 + \exp[-a(x - b)]} \quad . \tag{2.15}$$

As can be noted from figure 2.11, if parameter a increases, the transition from 0 to 1 becomes sharper, and $F(x = b) = 0.5$ for all a, indicating the point where transition from membership to non membership occurs. The fuzzy set "*exposition level to cigarette smoke*" above is an epidemiological concept that may be described through the sigmoidal membership function with $a > 0$.

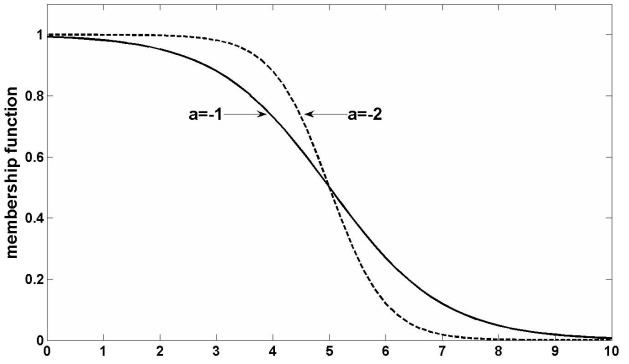

Fig. 2.11. Fuzzy set represented by a sigmoidal membership function

Independently of the shape, the membership function classifies and organizes the elements in a different way than that of classical logic.

As well discussed in Pedrycz and Gomide (1998 and 2007), there are two views associated with fuzzy sets theory, concerning *fuzziness* and *uncertainty* concepts. In the fuzziness of the information the membership degree $F(x)$ quantifies how x is compatible with the statement "x *is* F", or with the characteristics that describe the set F. This concept is more strongly linked to the ability to use *linguistic variables*, which is the great contribution of fuzzy set theory in several areas of applications, for example in the modeling of medical questions. From the uncertainty point of view, $F(x)$ is concerning on how likely x is to occur given that a group of constraints is present, F, represents the available knowledge about the variable in study. From both points of view, fuzzy sets theory could contribute to model epidemiological and public heath problems.

Fuzzy sets theory incentives us to revisit several classical medical and epidemiological concepts, expanding our insights of the problems. In a classical standard approach, for instance, the diagnostic process should be ideally a classificatory process able to determine the crisp divide between healthy and non-healthy individual. Let us imagine that a reliable measure of health is available and that a "normal" threshold, below which an individual is classified as non-healthy exists. A good clinician should then be able to classify individuals below (diseased) or above (healthy) such a threshold. Figure 2.12 illustrates this classical approach. So, an individual with a certain score 3.8 would then be considered as not healthy, according to this approach. Individuals in the borderline, however, would present additional classification problems.

Since the most fundamental aspects of fuzzy set theory is the idea of graded membership, it is possible to classify the individuals as healthy and non-healthy by the use of this sets. In the fuzzy approach the healthy and non-healthy sets could be represented by overlapping triangles with the y-axis indicating the grade of membership in the fuzzy set, like in figure 2.13 (Bellamy, 1997). In this case,

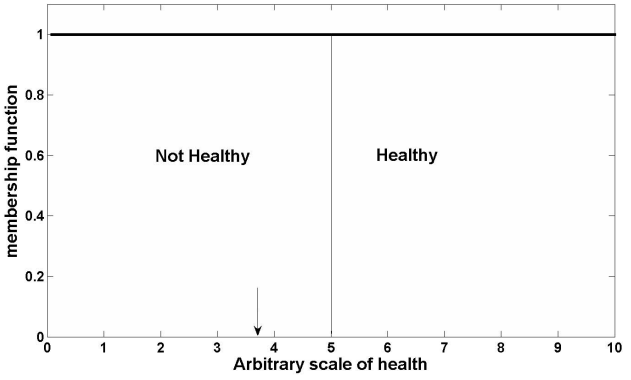

Fig. 2.12. Sets to illustrate the health classificatory process in the classical approach (modified from Bellamy, 1997)

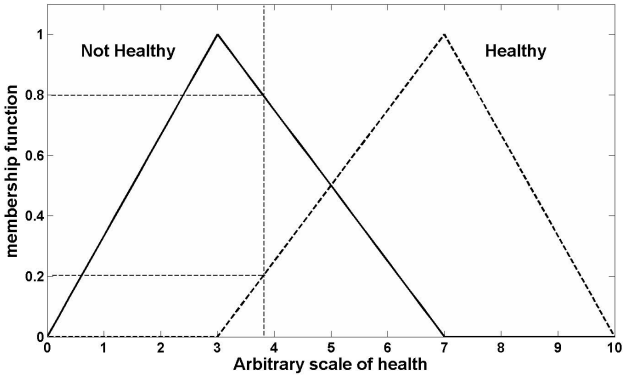

Fig. 2.13. Sets to illustrate the health classificatory process in the fuzzy approach (modified from Bellamy, 1997)

the same individual who scored 3.8 in the health scale now has 0.2 degree of membership to the set of healthy individuals and 0.8 degree of membership to the set of non-healthy individuals.

For a long time the concepts of disease and health are seen as opposed by the medical community, that is, disease is the absence of health and vice-versa. However, in the fuzzy approach, the concepts of disease and health are rather complementary than contradictory. Thereby, a new concept of health and disease can be established, which can provoke transformations in other conceptual constructions of medicine as, for instance, nosology (Sadegh-Zadeh, 1994, 1998 and 1999). Along this book several medical and epidemiological situations that could be modeled through fuzzy set theory will be presented.

2.3.1 Properties and Characteristics of Fuzzy Sets

Fuzzy sets have several characteristics and properties whose explanation facilitates the understanding and the development of fuzzy models. We list bellow those most important and common, recommending to the interested reader the books by Klir and Yuan (1995) and, Pedrycz and Gomide (1998 and 2007).

Fuzzy Singleton

A particular case of a finite fuzzy set is a *fuzzy singleton*. A fuzzy singleton is a set whose membership function is defined by:

$$A(x) = \begin{cases} 1 & if \quad x = x^* \\ 0 & if \quad x \neq x^* \quad for \quad all \quad x \in U \end{cases} .$$

This fuzzy singleton play an important role in fuzzy control systems development due to its simplicity and fast computation, commonly used in linguistic models as consequent of fuzzy rules.

Normality

Consider a fuzzy set A in the universe of discourse U. The set A is called *normal* if its membership function attains 1, that is, if there is x such that $A(x) = 1$. Otherwise, A is called subnormal. The maximum value of $A(x)$ is also referred to as the height of A, $hgt(A)$. Figure 2.14 shows an example of normal and subnormal fuzzy sets.

A subnormal fuzzy set is a set with no full members, containing only partial members. This notion contrasts with classical situation where a set is either nonempty or empty (black or white situation). A subnormal fuzzy set introduces a gray area between these extremes. In general, the human reasoning corresponds

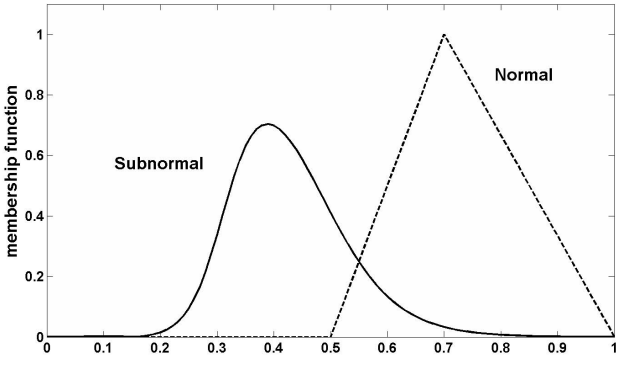

Fig. 2.14. Normal fuzzy set and subnormal fuzzy set

to normal fuzzy sets. However, there are situations, particularly in biomedicine and epidemiology where vagueness permeates many issues, whose reasoning is better described by subnormal fuzzy sets. Subnormal fuzzy sets play an important role in the modeling context, where they are usually generated during the inference process, as it will be seen in chapter 7.

Cardinality

The cardinality of a classical set is the number of all elements belonging to it. This concept is generalized to fuzzy sets through the definition:

$$Card(A) = \sum_{x \in U} A(x), \tag{2.16}$$

where the universe U is a finite set and the summation symbol referred to the standard algebraic sum. As an example, consider the fuzzy set "$F = around\ 5$", discussed above and defined in the universe $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$,

$$F = 0.10/1 + 0.25/2 + 0.50/3 + 0.75/4 + 1/5 + 0.75/6 + 0.50/7 + 0.25/8 + 0.10/9. \tag{2.17}$$

So, the cardinality of F is given by:

$$Card(F) = 0.10 + 0.25 + 0.50 + 0.75 + 1 + 0.75 + 0.50 + 0.25 + 0.10 = 4.2. \tag{2.18}$$

Note that in fuzzy approach the cardinality may assume noninteger number. So, it generalizes the concept of the classic counting. In chapter 9 and 11 this account processes will be discussed in the epidemiological context.

Normalization

In addition to the definitions presented above there are several transformations that may be performed on fuzzy sets. An important transformation of one-argument mapping, due to its applications in control systems, is the *normalization* of a fuzzy set. This operation consists in converting a subnormal fuzzy set (nonempty) into its normal version, dividing the original membership function by the height of A, that is,

$$Norm\_A(x) = \frac{A(x)}{hgt(A)}. \tag{2.19}$$

Support and Core

The support of a fuzzy set A, denoted by $Supp(A)$, is the set of all elements of U that belong to A with nonzero degree, that is,

$$Supp(A) = \{x \in U | A(x) > 0\}. \tag{2.20}$$

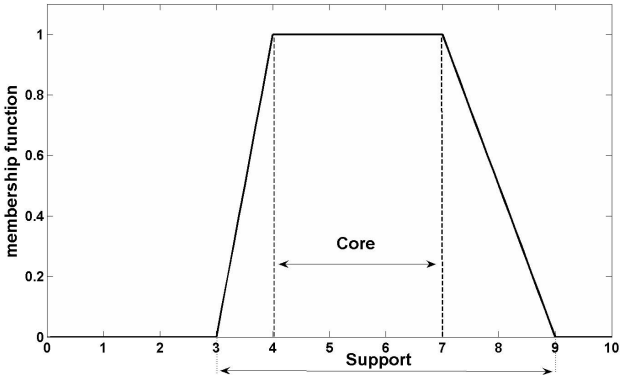

Fig. 2.15. Illustration of *core* and *support* sets. Note that both are crisp sets.

The core of a fuzzy set A is the set of all elements of U that belongs to A with membership degree equal to 1, that is,

$$Core(A) = \{x \in U | A(x) = 1\}. \tag{2.21}$$

So, if A is a subnormal fuzzy set, then its core is the empty set, $Core(A) = \emptyset$.

Core and support definitions are related concepts, since they identify elements belonging to the fuzzy set and $Core(A) \subseteq Supp(A)$. Note that both core and support are crisp sets in U. Figure 2.15 illustrates the support and core concepts of a fuzzy set.

2.3.2 $\alpha - cut$ or $\alpha - level$

The set of all elements that belong to a fuzzy set A with at least α degree is called $\alpha - cut$, or $\alpha - level$, and denoted by $[A]^\alpha$, that is,

$$[A]^\alpha = \{x \in U | A(x) \geq \alpha\}. \tag{2.22}$$

So, the set $[A]^\alpha$ consists of those elements of U whose membership degree is larger than α. The largest level is $\alpha = 1$, and it determines a set of U belonging completely to A. It is easy to see that $[A]^\beta \subset [A]^\alpha$ if $\beta > \alpha$. An example of $\alpha - cut$ is shown in figure 2.16.

The $\alpha - cut$ set definition is important since it provides a different way to consider a fuzzy set. According to the representation theorem, every fuzzy set may be represented by the aggregation of their $\alpha - cuts$ sets (Negoita & Ralescu, 1975). In this sense, all fuzzy sets could be decomposed into a family of their $\alpha - cuts$.

The $\alpha - cut$ concept assumes great importance in the fuzzy set theory analysis and, in a certain sense, turn the *core* and *support* sets into particular cases of the $\alpha - level$ sets (Negoita & Ralescu, 1975). On the other hand, the importance of

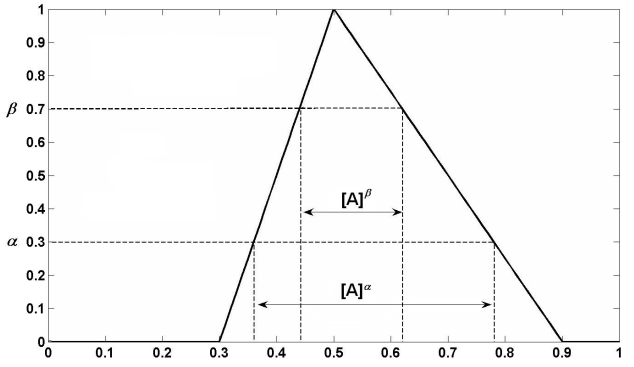

Fig. 2.16. Illustration of $\alpha - level$ sets. Note that these $\alpha - level$ sets are crisp sets.

the representation theorem relates to the fact that problems formulated in the fuzzy sets structure may be solved using non-fuzzy techniques, taking advantage of classical formalism and results available. Several important concepts in fuzzy analysis, as differentiability and integrability, are built from the $\alpha - cut$ concepts (see Dubois & Prade, 1982a, 1982b and 1982c; Puri & Ralescu, 1983; Kaleva, 1987; Seikkala, 1987).

2.4 Extension Principle and Fuzzy Numbers

One of the most useful transformation, with important consequences from the theoretical and applied points of view, is the so-called *extension principle* (Dubois *et al.*, 1995; Ortega *et al.*, 2003; Barros & Bassanezi, 2006). Essentially, the extension principle is a tool used to extend some classical mathematical concepts, as a mathematical functional, for instance. There are many ways to apply this extension, but the most famous and used is the extension principle proposed by Zadeh (1975a, 1975b and 1975c) and later further characterized by Yager (1986).

Another outstanding concept in fuzzy sets theory is that defined as *fuzzy number*. As discussed previously, fuzzy numbers apply a really important role in fuzzy modeling because it allows the quantification of qualitative predicates and to compute with them.

Extension principle and fuzzy numbers are intimately linked because it is through the extension operations that the computations with fuzzy numbers becomes possible.

2.4.1 Zadeh Extension Principle

Extending concepts of classical sets theory to fuzzy sets theory is a necessity, particularly in the theoretical developments. The extension method proposed by Zadeh, usually called *Extension Principle* only, is one of the basic ideas to

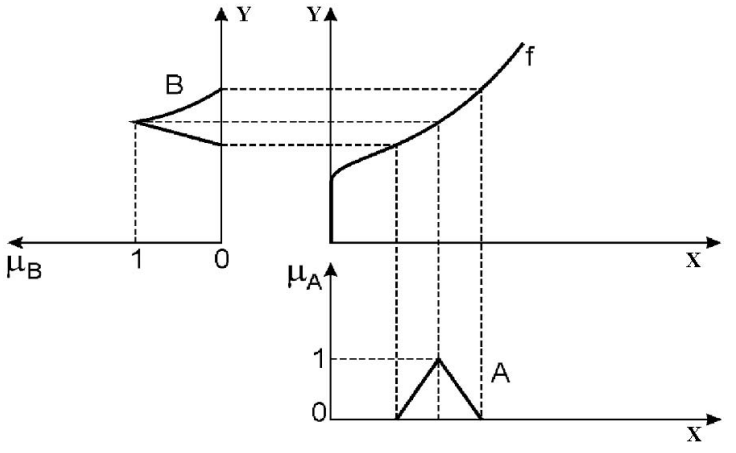

Fig. 2.17. Illustration of the Extension Principle application (modified from Klir and Yuan, 1995)

process the extension of the classical mathematical concepts into fuzzy ones (Zadeh, 1975a, 1975b and 1975c).

Let us consider two crisp sets X and Y and f a mapping from X to Y, $f : X \rightarrow Y$. Let A be a fuzzy subset of X, $A \in X$. So, the Extension Principle allows to build the image of A under the crisp mapping f as a fuzzy set $B = f(A)$ in Y, whose membership function is given by:

$$B(y) = \begin{cases} \sup_{x \in f^{-1}(y)} A(x) & if \quad f^{-1}(y) \neq \emptyset \\ 0 & if \quad f^{-1}(y) = \emptyset \end{cases} \quad for \quad all \quad y \in Y, \tag{2.23}$$

where $f^{-1}(y)$ denotes the set of all points $x \in X$ such that $f(x) = y$. Figure 2.17 illustrates the Extension Principle action considering the mapping $f : X \rightarrow Y$ and A a triangular fuzzy set.

Example 2.7. Let a mapping $f : I\!R \rightarrow I\!R$ and f given by $f(x, y) = x^2 - 5x + 8$ with $x \in X$. Consider the finite fuzzy subset F in $I\!R$ chosen in the beginning of this chapter to describe a statement "around 5" and illustrated in figure 2.1:

$$F = 0.10/1 + 0.25/2 + 0.50/3 + 0.75/4 + 1/5 + 0.75/6 + 0.50/7 + 0.25/8 + 0.10/9.$$

Remember that the symbol + is not the standard algebraic summation, but only a mathematical notation denoting the aggregation of the elements. So, applying the extension principle (2.23) we find the fuzzy set of $B(y) = f(around\ 5)$:

$$\begin{aligned} B(y) = f(around\ 5) &= 0.10/f(1) + 0.25/f(2) + 0.50/f(3) + 0.75/f(4) \\ &+1/f(5) + 0.75/f(6) + 0.50/f(7) + 0.25/f(8) + 0.10/f(9), \end{aligned} \tag{2.24}$$

resulting in:

$$B(y) = 0.10/4 + 0.25/2 + 0.50/2 + 0.75/4 + 1/8 + 0.75/14 + 0.50/22 + 0.25/32 + 0.10/44. \tag{2.25}$$

Note that $f(1) = f(4)$ and $f(2) = f(3)$. Thus, the possibility of the images $y = 4$ and $y = 2$, with $y \in Y$, are found by means of *sup* operator, such as:

$$B(y) = \sup[0.10, 0.75]/4 + \sup[0.25, 0.50]/2 + 1/8 + 0.75/14 + 0.50/22 + 0.25/32 + 0.10/44. \tag{2.26}$$

Therefore,

$$B(y) = 0.50/2 + 0.75/4 + 1/8 + 0.75/14 + 0.50/22 + 0.25/32 + 0.10/44.$$

This procedure and the final result are shown in figure 2.18.

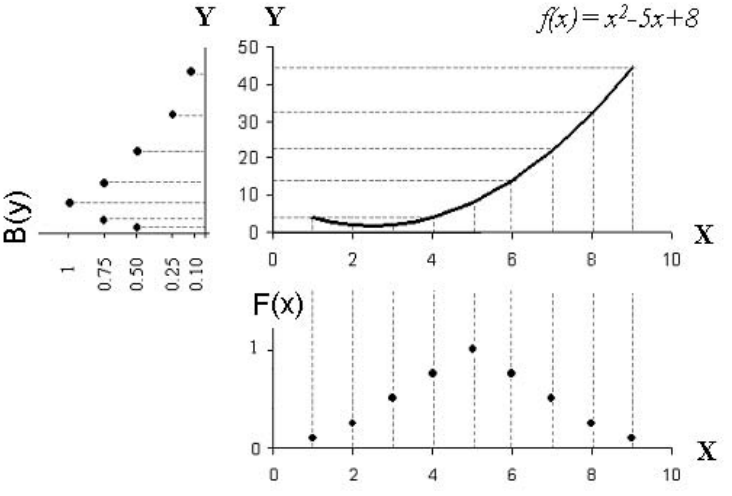

Fig. 2.18. Example of the extension principle application

Although we have defined the Extension Principle just for a mapping of a single variable, it is easily generalized to functions of many variables. For more details about transformations on fuzzy sets and the Extension Principle see Klir and Yuan (1995) and, Pedrycz and Gomide (1998, 2007). In chapter 8 we will present an application of this concept in an epidemic issue where a multivariated function is considered (Ortega *et al.*, 2003).

2.4.2 Fuzzy Numbers

Fuzzy numbers are entities useful to quantify fuzzy concepts. They are defined on the set of real numbers, $\mathbb{R}$, and their membership functions are mapping such

as $\mu : I\!R \rightarrow [0,1]$. Usually fuzzy numbers represent statements like "around of", "small", "very large" and so on.

Definition 2.8. *A fuzzy subset A in $I\!R$ is called fuzzy number when:*

- *All $\alpha-$levels of A are non-empty, with $0 \leq \alpha \leq 1$, that is, A must be normal;*
- *All $\alpha-$levels of A are closed intervals of $I\!R$;*
- *The support of A, that is, $SuppA = \{x \in I\!R : A(x) > 0\}$.*

Note that all real number r is a particular fuzzy number with the given membership function (Barros & Bassanezi, 2006):

$$\mu_r(x) = \begin{cases} 1 \;\; if \;\; x = r \\ 0 \;\; if \;\; x \neq r \end{cases} . \tag{2.27}$$

The membership functions most used to represent fuzzy numbers are triangular and Gaussian shapes, illustrated in figures 2.9 and 2.10, when it is considered only a non-null Gaussian membership function in a limited interval. Note that a fuzzy number must not have a symmetrical shape. The fuzzy set F in example 2.7 is a symmetrical fuzzy number.

If the fuzzy set satisfy all conditions above and besides its Core set is an interval, so it is called *fuzzy interval*. The most used membership function to represent fuzzy intervals is trapezoidal (see figure 2.9). Figure 2.19 illustrates:

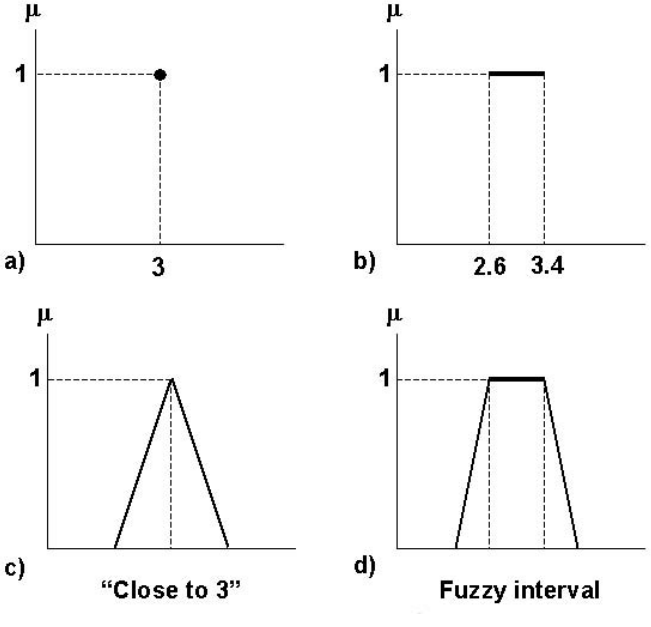

Fig. 2.19. Illustration of classical and fuzzy numbers and intervals, respectively

a) the classical number 3; b) the classical interval [2.6, 3.4]; c) a fuzzy number that represent a statement "close to 3"; and d) a fuzzy interval that represents a statement "around [2.6, 3.6]". It is not the objective here to detail fuzzy intervals and their properties, and for the interested reader we recommend the book of Pedrycz and Gomide (2007).

The great advantage of fuzzy numbers is that it is possible to compute with them. Thus, we can define arithmetic operations on fuzzy numbers.

2.4.3 Arithmetic Operations on Fuzzy Numbers

Fuzzy arithmetics is developed by means of two methods basically: a method based on interval arithmetics, using the $\alpha - levels$ concepts; and a method that applies the extension principle, by which operations on real numbers are extended to fuzzy numbers (Klir & Yuan, 1995). In this section we define the fuzzy operations basing on the extension principle. In this case, these definitions can be seen as a particular case of the extension principle, both for functions of one, $f(x)$, and two variables, $f(x, y)$.

Definition 2.9. *Let A and B be fuzzy numbers and δ a real number.*

a) The sum and the difference of two fuzzy numbers A and B are also fuzzy numbers, $(A + B)$ and $(A - B)$ respectively, whose membership functions are given by:

$$\mu_{(A+B)}(z) = \sup_{\{(x,y):x+y=z\}} \min\left[\mu_A(x), \mu_B(y)\right], \tag{2.28}$$

and

$$\mu_{(A-B)}(z) = \sup_{\{(x,y):x-y=z\}} \min\left[\mu_A(x), \mu_B(y)\right], \tag{2.29}$$

where $x \in X$, $y \in Y$ and (x, y) is an ordinated pair in the Cartesian product space $X \times Y$.

b) The multiplication of the fuzzy number A by a real number δ is a fuzzy number (δA), whose membership function is given by:

$$\mu_{(\delta A)}(z) = \sup_{\{x:\delta x=z\}} \left[\mu_A(x)\right] = \begin{cases} \mu_A(\delta^{-1}z) & if \quad \delta \neq 0 \\ 0 & if \quad \delta = 0 \end{cases}. \tag{2.30}$$

c) The multiplication of a fuzzy number A by a fuzzy number B is a fuzzy number $(A \cdot B)$, whose membership function is given by:

$$\mu_{(A\cdot B)}(z) = \sup_{\{(x,y):x\cdot y=z\}} \min\left[\mu_A(x), \mu_B(y)\right]. \tag{2.31}$$

d) The division of a fuzzy number A by a fuzzy number B, if $0 \notin Supp(B)$, is a fuzzy number (A/B), whose membership function is given by:

$$\mu_{(A/B)}(z) = \sup_{\{(x,y):x/y=z\}} \min\left[\mu_A(x), \mu_B(y)\right]. \tag{2.32}$$

Clearly, the condition imposed over the support of fuzzy number B in the definition of the fuzzy division is necessary to avoid division by zero. The application of definition 2.9 is illustrated in an example below, considering the summation operator.

Example 2.10. Let A and B be discrete fuzzy numbers, describing the statements "around 4" and "around 5", respectively, such that:

$$A(x) = 0.4/1 + 0.6/2 + 0.8/3 + 1/4 + 0.7/5 + 0.1/6$$

and

$$B(y) = 0.5/3 + 0.8/4 + 1/5 + 0.2/6.$$

Then, the sum of A and B is a fuzzy number $(A + B)$ where:

$$\begin{aligned}\mu_{(A+B)}(z) \; &= 0.4/4 + 0.4/5 + 0.4/6 + 0.2/7 + 0.5/5 + 0.6/6 + 0.6/7 + 0.2/8 \\ &+ 0.5/6 + 0.8/7 + 0.8/8 + 0.2/9 + 0.5/7 + 0.8/8 + 1/9 + 0.2/10 + 0.5/8 \\ &+ 0.7/9 + 0.7/10 + 0.2/11 + 0.1/9 + 0.1/10 + 0.1/11 + 0.1/12,\end{aligned} \tag{2.33}$$

where $z = x + y$ with $x \in A$ and $y \in B$.

As in example 2.7, since there are ordinated pairs (x, y) that map the same image, we must apply the operator sup to find the final membership degrees. So, by means of the definition above we find:

$$\begin{aligned}\mu_{(A+B)}(z) = \; &0.4/4 + 0.5/5 + 0.6/6 + 0.8/7 + 0.8/8 + 1/9 + 0.7/10 \\ &+0.2/11 + 0.1/12 + 1/9 + 0.7/10 + 0.2/11 + 0.1/12.\end{aligned} \tag{2.34}$$

Note that the support of the fuzzy number $(A+B)$ is greater than the support of A and than the support of B. Occurs that from the definition to summation operator on fuzzy numbers results a larger fuzzy number. In some sense, it means that the fuzziness increases when the fuzzy arithmetic sum operation is applied. It is also observed with difference and multiplication operations.

An interesting property of the fuzzy arithmetic sum and difference operastion is that, if A and B are triangular fuzzy numbers, so the fuzzy number found through the application of these operators is also a triangular fuzzy number, that is, the membership functions of $(A + B)$ and $(A - B)$ have triangular shapes.

All arithmetic operations on fuzzy numbers in definition 2.9 are described based on binary operations of their membership degrees. Nevertheless, these

arithmetic operations can also be described by means of the $\alpha - levels$ of the fuzzy numbers, which in some situations allow easier calculations. This fact is a consequence of a theoretical result that provides the $\alpha - levels$ for the Zadeh extension (Pedrycz & Gomide, 2007). This result restricted to arithmetic operations between fuzzy numbers provides the properties bellow expressed in (2.35) - (2.39).

Properties of $\alpha -$ levels for fuzzy numbers

Let A and B be fuzzy numbers, with $\alpha - levels$ given respectively by $[A]^\alpha = [a_1^\alpha, a_2^\alpha]$ and $[B]^\alpha = [b_1^\alpha, b_2^\alpha]$; and δ a real number.

a) The sum and the difference between A and B are fuzzy numbers, $(A + B)$ and $(A - B)$ respectively, whose $\alpha - levels$ are given by:

$$[A + B]^\alpha = [A]^\alpha + [B]^\alpha = [a_1^\alpha + b_1^\alpha, a_2^\alpha + b_2^\alpha], \tag{2.35}$$

and

$$[A - B]^\alpha = [A]^\alpha - [B]^\alpha = [a_1^\alpha - b_2^\alpha, a_2^\alpha - b_1^\alpha]. \tag{2.36}$$

b) The multiplication of the fuzzy number A by a real number δ is a fuzzy number (δA), whose $\alpha - levels$ are given by:

$$[\delta A]^\alpha = \delta[A]^\alpha = \begin{cases} [\delta a_1^\alpha, \delta a_2^\alpha] & if \quad \delta \geq 0 \\ [\delta a_2^\alpha, \delta a_1^\alpha] & if \quad \delta < 0 \end{cases} . \tag{2.37}$$

c) The multiplication of a fuzzy number A by a fuzzy number B is a fuzzy number $(A \cdot B)$, whose $\alpha - levels$ are given by:

$$[A \cdot B]^\alpha = [A]^\alpha [B]^\alpha = [\min P, \max P], \tag{2.38}$$

where $P = \{a_1^\alpha b_1^\alpha, a_1^\alpha b_2^\alpha, a_2^\alpha b_1^\alpha, a_2^\alpha b_2^\alpha\}$.

d) The division of a fuzzy number A by a fuzzy number B, if $0 \notin Supp(B)$, is a fuzzy number (A/B), whose $\alpha - levels$ are given by:

$$\left[\frac{A}{B}\right]^\alpha = \frac{[A]^\alpha}{[B]^\alpha} = [a_1^\alpha, a_2^\alpha]\left[\frac{1}{b_2^\alpha}, \frac{1}{b_1^\alpha}\right]. \tag{2.39}$$

Arithmetic operations on fuzzy numbers based on their $\alpha - levels$ are widely applied in the analytical calculus carried out in chapters 9, 10 and some applications presented in chapter 12.

To finish this brief review about fuzzy sets and their properties we wish to remark that the definition presented in 2.1 refer to so called *ordinary fuzzy set* or *Type-1 fuzzy set*. However, there are several generalizations of Type-1 fuzzy sets

based on the generic concept of partial membership, which the most common is Type-2 fuzzy set. These generalizations are not the subject of this book, but for the interested reader we recommend the book by Mendel (2001), where these issues are well detailed and a rich discussion about their practical applications is presented.

2.5 Fuzzy Relations

Like fuzzy sets, fuzzy relations are generalizations of the classical relation. A classical relation describes the relationship between two or more objects. A relationship between two objects is called a binary relationship, between three objects is called a ternary relationship, and so on. For instance, the relationship between father and son could be characterized by the binary relation: $(father, son)$.

A classical relationship obeys the classical characteristic function. So, the relationship of friendship among two persons, designated as "*friends*", considers that in human relationships a person either is or is not your friend, which certainly is a simplification of the reality. On the other hand, a fuzzy relationship of friendship between two persons considers the friendship degree between them. So, two or more individuals may be linked with different degrees of friendship, from 1 (they are certainly friends) down to 0 (they are not friends).

A binary classical relation on variables x and y, whose domains are X and Y respectively, may be defined as a set of ordered pairs in the Cartesian Product space $X \times Y$ (see definition 2.6). For instance, the relation "*more than*", between two real numbers, could be formally defined as:

$$R = \{(x, y) | x > y;\ x, y \in I\!R\}. \tag{2.40}$$

In general, a relation between n objects $x_1, x_2, ..., x_n$, that is a *n-ary* relation, whose domains are $X_1, X_2, ..., X_n$, is a subset of the Cartesian space $X_1 \times X_2 \times ... \times X_n$. For instance, someone can use a *n-ary* relation to describe a patient's state p, that presents signs s_1, s_2 and s_3, with family antecedents f_1 and f_2, taking medicines m_1, m_2 and m_3, resulting in a relation having nine arguments $\{p, s_1, s_2, s_3, f_1, f_2, m_1, m_2, m_3\}$.

A binary relation in $X \times Y$, for instance, could be described by its characteristic function that maps an ordered pair (x, y) in $X \times Y$ to 0 (if the relation does not hold) or 1 (if the relation holds), that is, $R : X \times Y \rightarrow \{0, 1\}$. So, a fuzzy binary relation is defined by its membership function that maps the ordered pairs of $X \times Y$ to the relationship degree, which is itself a number in the interval $[0, 1]$, that is, $R : X \times Y \rightarrow [0, 1]$. More generally, a fuzzy relation of n objects $(x_1, x_2, ..., x_n)$, whose domains are $X_1, X_2, ..., X_n$ is defined by its membership function that maps $R : X_1 \times X_2 \times ... \times X_n \rightarrow [0, 1]$. If the possible values of the relation are discrete, the fuzzy relation could be expressed in a matrix form.

Let us see a practical example of fuzzy relation. We know, from hundreds of years of medical practice, that three clinical findings, *Headache* and *Cough*, (two symptoms) and *Fever* (a sign) are associated, at different levels, with

several possible diagnostics. Let us take, for instance, *Endocarditis* (End.), *Pneumonia* (Pn.), *Pertussis* (Pt.), *Tuberculosis* (Tb.) and *Common Cold* (C.C.). Consider, also, that one wishes to express the fuzzy relation of the diagnostic system in terms of signs/symptoms, *Headache, Fever, Cough*, and diseases, $\{End., Pn., Pt., Tb., C.C.\}$. So, a fuzzy relation R of medical knowledge that relates those symptoms and signs to the set of possible diseases may be like the matrix below:

$$R = \begin{array}{c} \\ Headache \\ Fever \\ Cough \end{array} \begin{array}{c} \begin{array}{ccccc} End. & Pn. & Pt. & Tb & C.C. \end{array} \\ \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.0 & 0.8 \\ 0.9 & 1.0 & 0.3 & 1.0 & 0.2 \\ 0.2 & 0.4 & 0.7 & 1.0 & 0.1 \end{bmatrix} \end{array}.$$

Then, the relationship degree between *Headache* and *Pneumonia* is zero, that is, from the relationship matrix proposed there is no direct relation between *Headache* and *Pneumonia*. On the other hand, the relationship between Fever and Pneumonia is one characterizing the complete relationship. One could suggest that the relationship between symptoms/signs and diseases depends on the age, which is reasonable for several pathologies. To aggregate the age information it will be necessary one dimension more in the relationship matrix, resulting in a ternary fuzzy relation *age,symptons/signs,diseases*.

Summarizing the last example, consider a specific patient that presents a *persistent* fever, *intense* and *constant* cough and no complain of headache (the italicized words are meant to emphasize the vagueness of clinical findings and complains). We may assign to this specific patient the following fuzzy set A:

$$\begin{array}{c} Head. \quad Fev. \quad Cou. \\ A = \begin{bmatrix} 0.0 & 0.7 & 1.0 \end{bmatrix}. \end{array}$$

Suppose now that a doctor wants to know what is the possibility that this patient is affected by any of the diseases considered above. The question consists in how to join the characteristics of that patient with the information contained at the symptom-sign/disease relationship matrix. To answer to the doctor we can use one of the most precious resources of fuzzy relationships, that is, the *Composition of the Fuzzy Relations*.

2.5.1 Composition of Fuzzy Relations

Let X and Y be the universe of discourse of the variables x and y, respectively, and x_i and y_j the elements of X and Y. Consider R the fuzzy relation that maps $X \times Y$ in $[0, 1]$ and a possibility distribution in X, $\pi_x(x_i)$. Then, the possibility distribution in Y (see chapter 4) can be given by:

$$\pi_Y(y_j) = \oplus_{x_i}[\pi_X(x_i) \otimes \pi_R(x_i, y_j)], \tag{2.41}$$

where the symbol $\oplus$ denotes some fuzzy disjunction operator, the symbol $\otimes$ denotes some fuzzy conjunction operator, and $\pi_R(x_i, y_j)$ is a possibility distribution of the relationship between x and y.

If X and Y are finite sets, the procedure to calculate the composition rule is similar to that used in multiplication of matrix, where the fuzzy conjunction and disjunction operators correspond, respectively, to the multiplication and addition steps. The rule of composition is not unique, since different choices can be made with regard to the conjunction and disjunction operators. In practice the two most used are the composition $max - min$ and the composition $max - product$. Considering the two possibilities equation (2.41) gives:

$$\pi_Y(y_j) = \max_{x_i}[\min(\pi_X(x_i), \pi_R(x_i, y_j))], \tag{2.42}$$

to the $max - min$ composition, and

$$\pi_Y(y_j) = \max_{x_i}[\pi_X(x_i) \cdot \pi_R(x_i, y_j)], \tag{2.43}$$

to $max - product$ composition.

A fuzzy set A (e.g. the set of the symptoms related to a patient) could be composed with the fuzzy relation R (e.g. representing the medical knowledge that relates the symptoms in the set S to the diseases in the set D) by the compositional rule of inference, inferring the set B (e.g. the set of the possible diseases on the patient) as:

$$B = A \circ R. \tag{2.44}$$

So, we can evaluate the possibility distribution of the diseases for the patient A through the $max - min$ fuzzy composition between the status of the patient and the relationship matrix of symptoms/diseases. Using equation (2.42) we find:

$$\pi_B(d) = \max[\min(\pi_A(s), \pi_R(s, d)], \tag{2.45}$$

where d is the disease set, *End., Pn., Pt., Tb., C.C.*, $\pi_B(d)$ is the possibility distribution of d, s is the symptoms/signs set, $\{$*Headache, Fever, Cough*$\}$, and $\pi_R(s, d)$ is the possibility distribution of the relation $R(s, d)$. The possible disease of this patient can be calculated by operating the matrices

$$\begin{bmatrix} 0.0 & 0.7 & 1.0 \end{bmatrix} \circ \begin{bmatrix} 0.0 & 0.0 & 0.3 & 0.0 & 0.8 \\ 0.9 & 1.0 & 0.3 & 1.0 & 0.2 \\ 0.2 & 0.4 & 0.7 & 1.0 & 0.1 \end{bmatrix}, \tag{2.46}$$

which results in the following possibilities of diagnosis for this patient:

$$\begin{matrix} End. & Pn. & Pt. & Tb & C.C. \end{matrix}$$

$$\begin{bmatrix} 0.7 & 0.7 & 0.7 & 1.0 & 0.2 \end{bmatrix}.$$

that is, the highest diagnostic possibility for this hypothetical patient is tuberculosis, although a final decision for discriminating this diagnosis from endocarditis, pneumonia and pertussis should require further investigations.

Notice that the answer of the composition is also a fuzzy set. In other words, it does not answer which disease the patient has, but supplies the distribution of the patient's possibilities in the diseases set. However, based on the answer obtained the doctor can take decisions, choosing, for instance, more detailed laboratory exams, investigating with more insistence the possibility of a Tuberculosis diagnostic and discarding the possibility of a Common Cold.

The fuzzy logic framework has been utilized in several different approaches to modeling the diagnostic process. In one of the first attempt to apply fuzzy sets theory to medical diagnosis, Sanchez (1979) proposes a model in which the medical knowledge is represented as a fuzzy relation between symptoms and diseases, as illustrated by example above.

More recently, Reis *et al.* applied the $max-min$ composition of fuzzy relations to predict the need of neonatal resuscitation. Although the simplicity of this approach, one of the most advantage of its application is its capacity to deal with several variables. In Reis' work 61 clinical factors were considered and the fuzzy expert system presented a sensitivity of 82.4% and a specificity of 93.0% in the identification of newborn's life-threatening situation in the delivery room. In this work 303 deliveries were followed up and the area under ROC curve was 0.93 (Reis *et al.*, 2004 and 2005).

This fuzzy relation approach was also applied by Lopes and collaborators to develop a fuzzy expert system to diagnosis differentiation of urinary incontinence. In this work 6 possible diagnostics and 35 clinical factors were considered, and when compared with a panel of experts in the analysis of 195 clinical cases the system presents excellent agreement (kappa's value equal to 0.98 and $p < 0,001$) for the most optimist conditions, and it presents substantial agreement (kappa's value equal to 0.69 and $p < 0,001$) for the most pessimist conditions (Lopes *et al.*, 2006).

The importance that fuzzy relation plays in the systems modeling field will be more evident along this book. In chapter 11 the $max - min$ fuzzy composition is also applied in a dynamical epidemiological system.

3 Modern Epidemiology

In his masterpiece *The Growth of Biological Thought* Ernest Mayr (1982) postulated that the change of the typological to populational thinking is the most important conceptual revolution in biology. This conjecture of Mayr was based on his interest in the evolution and variation of birds, and his observations of the variety within populations led him to conclude that it is individuality that is the chief characteristic within populations rather than any criterion of sameness (Childs, 1999). Variation, therefore, is the biological substrate upon which evolution by natural selection and other stochastic mechanism constructs life and humans beings are not exception. In typological thinking the **type** (the equivalent to the statistical mean in populational thinking) is **real** and the **deviation from the type** (the equivalent to the statistical variance in populational thinking) is an **abstraction**. In contrast, in population thinking the **variance** is **real** and the **mean**is an **abstraction.**

The human variability is reasonably well known nowadays and this variation poses some important challenges as far as medicine is concerned. On the one hand, it is the human variation that makes medicine such an interesting and exciting subject. On the other hand, quoting Sir Willian Osler, "if it were not for the great variability among individuals medicine might as well be a science and not an art", that is, there is an urgent need of scientific tools to deal with the observed variability. It is an irony that the above quotation of Osler emphasizes variability but that the so-called Oslerian medical thinking, that is, the concept of the body as a machine that when broken must be repaired, implies in a typological thinking.

Epidemiology, in contrast, by dealing with populations of sick and healthy individuals is strongly based on the emphasis on heterogeneity and the epidemiological studies are designed to characterize those populations, to extract their essence, and to use the information to propose risk factors and control measures. So the development of tools to deal with variability within and between populations is central to modern epidemiology, from statistical tools, to mathematical tools to non-binary logic tools, such as the theory of fuzzy sets.

This chapter is intended to be an introduction to the main theme of this book, that is, modern epidemiology and mathematical tools usually applied in it. Many

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 41–57, 2008.
springerlink.com

of these tools are revisited along the book, considering the fuzzy logic approach. However, before we start to describe fuzzy logic applications in epidemiology, let us first briefly describe the classical quantitative models as applied to epidemic problems.

3.1 Statistical Models in Epidemiology

Statistical models in epidemiology assume that epidemiological studies generate data in which the response measurement for each subject may take one of only two possible values, the so-called *binary* response (Clayton & Hills, 1993). This kind of data are usually generated by two different types of study, namely, *cohort* studies, in which a group of people are followed through some period of time in order to study the occurrence (or not) of a certain event of interest; and *cross-sectional* study, in which the prevalence of an event is determined in a certain instant (or limited period) of time.

The *risk* parameter of the binary model is the *probability of failure*, π, which corresponds to the probability that the event under observation does occur. The complement of risk, $1-\pi$, is the *Probability of survival*, that is, the probability that the event does not occur. An important alternative way of parametrizing the probability model is in terms of the *odds* of failure versus survival, that is,

$$\frac{\pi}{1-\pi}. \tag{3.1}$$

Another central concept of statistical epidemiology is the *conditional probability*, expressed as the Bayes theorem. This approach incorporates the potential causes of the event under observation, which is called *exposures*. So, the aim of such an approach is to determine the conditional probability of developing the event (disease) given that one is exposed to the potential cause. More appropriated still is to determined the *risk ratio*, defined as the ratio of the conditional probability of developing a disease given one is exposed to a certain cause, $p(D|E)$, to the conditional probability of developing the disease given one is not exposed to the cause, $p(D|\ \bar{E})$, such as:

$$RR = \frac{p(D|E)}{p(D|\ \bar{E})}. \tag{3.2}$$

Assuming that the purpose of models is to allow predictions, we need a way of choosing the values of the parameters of the model. In modern statistics the central concept to the process of parameter estimation is *likelihood* (Clayton & Hills, 1993). In general if we observe f failures in n subjects, the likelihood for π is

$$(\pi)^f\,(1-\pi)^{n-f}\,. \tag{3.3}$$

As the likelihood, being the serial product of probabilities (numbers less than one) is usually a very small number, it is, therefore, convenient to use logarithms of likelihood, the *log likelihood*.

Also important in epidemiological studies is the concept of *rate*, always defined as the variation of some measure as related to other measure (usually time). So, if we have as a general example in which the measure, M, varies of ΔM in an interval of Δt, the rate, λ is equal to

$$\lambda = \lim_{\Delta t \to 0} \frac{\Delta M}{\Delta t} = \frac{dM}{dt}. \tag{3.4}$$

When the measure is the occurrence of a given disease, the rate of interest is called *incidence*, or the number of new cases per unit of time.

In risk analysis the rate at which an event may occur is called *hazard* and is the variation in the number of events over the total number of individuals at risk. The hazard is also called *force of morbidity*, or *force of mortality*.

In cohort studies it is common to estimate the cumulative probability of survival up to the end of the n^{th} time interval. This procedure, called *Kaplan-Meier* method, yields the most likely value of the survival curve.

If, on the one hand in cohort studies the relationship between exposure and disease incidence is investigated by following the entire cohort and measuring the rate of occurrence of new cases in the different exposure groups, on the other hand, in *case-control* studies the subjects who develop the disease (cases) are registered by some other mechanism than follow-up, and a group of healthy subjects (controls) is used to represent the subjects who do not develop the disease (Clayton & Hills, 1993). In this sort of study we calculate the ratio of the case/control ratio among exposed ($\pi_1/(1-\pi_1)$) to the case/control ratio of unexposed group ($\pi_0/(1-\pi_0)$), also called *odds ratio*:

$$OR = \frac{\pi_1/(1-\pi_1)}{\pi_0/(1-\pi_0)}. \tag{3.5}$$

Other kind of causal studies in epidemiology include multivariate models in which a set of potential causes is taken into account simultaneously, in order to estimate their composed risk to exposed, as related to unexposed individuals. The most popular multivariate model currently is the *logistic model*. In this model the Odds of being a case is given by (Clayton & Hills, 1993):

$$OR = K\frac{\pi}{1-\pi}, \tag{3.6}$$

where K is the ratio between the probability that a failure is sampled as a case and the probability that a survivor is sampled as a control. On a log scale we have

$$\log(Odds) = \log(K) + \log\left(\frac{\pi}{1-\pi}\right). \tag{3.7}$$

When the disease is rare (which is often the case in real epidemiological studies) the probability of failure in the study base is small and the odds of failure are related to the rate λ by ;

$$\frac{\pi}{1-\pi} \approx \lambda T, \tag{3.8}$$

where T is the duration of the study. Therefore,

$$\log(Odds) = \log(T) + \log(\lambda) \tag{3.9}$$

and it can be demonstrated that effects estimated from a logistic regression model are also estimates of effects on the log rate in the study base.

Obviously it is far beyond the scope of this book an exhaustive review of statistical models in epidemiology. We just sampled some central concepts from the book by Clayton and Hills (1993), an authoritative work to which the reader may refer to, in order to illustrate the basic differences between this sort of approach to epidemic problems and the ones described in the next section.

3.2 Mathematical Models in Epidemiology

Several have been the classifications of models in infectious disease. So, for instance, Anderson and May (1991) use to classify models in microparasites (viruses, bacteria and protozoa) as *prevalence*-based models, and models in macroparasites (helminths, flat-worms, etc.) as *density*-based models. Other authors use a mathematical approach to classify models in *deterministic* and *stochastic* models (Bailey, 1975).

Current models take several forms, although most fall into two broad categories, which will be detailed further on: analytical and computer simulations. Compared to computer simulation models, analytical models tend to be relatively simple, usually sets of differential equations that keep track of a few important variables. In contrast, computer simulation models try to incorporate many more of the variables influencing transmission.

3.2.1 The Reproduction of an Infection

The central parameter related to the intensity of transmission of infections is the so called basic reproduction number (R_0), defined by Macdonald (1952) as the number of secondary infections produced by a single infective in an entirely susceptible population (see next section). Originally applied in the context of malaria, R_0 is a function of the vector population density as related to the host population, m, the average daily biting rate of the vector, a, the host susceptibility, b, the vector mortality rate, μ, the parasite extrinsic incubation period in days, n, and the parasitemia recovery rate, r, according with the (now) historical equation:

$$R_0 = \frac{ma^2 b \exp[-\mu n]}{r\mu}. \tag{3.10}$$

From the definition of the basic reproduction number it can be demonstrated that if R_0 is not greater than one, that is, when an index case (the first infective individual) is not able to generate at least one new infection, the disease dies out. Hence, in the original Macdonald analysis, R_0 coincides with the threshold for the infection persistence. For an interesting historical account of R_0 see Dietz (1993).

In the context of directly transmitted infections, R_0 is defined in terms of the contact parameter, β, which comprises the probability of a potentially infectious contact and the probability that this contact generates a new infection, the total size of the involved population, N, and the time an infectious individuals remain infective, T, such that

$$R_0 = \beta NT. \tag{3.11}$$

Table 3.1 illustrates some of the estimated values of R_0 for some infections and some distinct populations in different moments in time:

Table 3.1. Estimated values of R_0 for some infections, distinct populations and in different moments in time (modified from Anderson and May, 1991)

Infection	**Geographical Location**	**Time Period**	$\mathbf{R_0}$
Measles	England	1947 - 1950	13 - 14
	USA	1918 - 1921	5 - 6
Pertussis	England	1944 - 1978	16 - 18
Chicken Pox	USA	1912 - 1921	7 - 8
Diphtheria	USA	1918 - 1919	4 - 5
Mumps	England	1960 - 1980	7 - 8
Rubella	England	1960 - 1970	6 - 7
Poliomyelitis	USA	1955	5 - 6
Malaria	Nigeria	1972	80 - 200
HIV	England	1981 - 1985	2 - 5
	Kenya	1981 - 1985	11 - 12
	USA	1981 - 1984	5 - 6
	Brazil	1991	90
Dengue	Brazil	1996	1 - 2
	Brazil	2001	5 - 12

In his 1952 seminal paper, Macdonald (1952) addressed the problem of a system involving one vector (*Anopheles* mosquitoes) and one host (men). As mentioned above, his definition of R_0 is the number of secondary infections in the first generation, that is, produced by a single infectee along its entire infectiousness period. We shall deduce an explicit expression for R_0 from an intuitive perspective to show that it coincides with the threshold for the establishment of the disease. We do this because, as shown in the next section, for more complex systems this approach does not work in such a simple way.

Let us begin by assuming that the index case is a human host. The question to be answered is how many human secondary infections this index case produces in his/her entire infectiousness period.

Let N_m be the number of female mosquitoes. Let a be the average daily biting rate female anophelines inflict in the human population. The number of bites in the human population per units of time is, therefore, $N_m a$. Let N_h be the number of humans and r be the rate of recovery from parasitemia in the human cases. Therefore, the index case produces $\frac{N_m a}{N_h r} c_{h \to m}$ infected mosquitoes, where $c_{h \to m}$ is

the probability that a mosquito gets the infection after biting an infective human. Those $\frac{N_m a}{N_h r} c_{h\to m}$ infected mosquitoes, in turn, produce $a\frac{N_m a}{N_h r} c_{h\to m} \frac{1}{\mu} b_{m\to h} e^{-\mu\tau}$ new human cases in the first generation, where $\frac{1}{\mu}$ is the average life expectancy of mosquitoes, $b_{m\to h}$ is the probability that a human gets the infection after being bitten by an infective mosquito and $e^{-\mu\tau}$ is the fraction of the infected mosquito population that survives through the extrinsic incubation period τ of the parasite. Note that, once infective a mosquito is assumed to remain so for life. Therefore, the expression for R_0 is (Macdonald, 1952):

$$R_0 = a\frac{N_m a}{N_h r} c_{h\to m} \frac{1}{\mu} b_{m\to h} e^{-\mu\tau}. \tag{3.12}$$

Similarly, if we begin with an infective mosquito as an index case, and compute the number of infected mosquitoes this index case produces in the first generation we get the same expression. Let us now see how this deduction can be performed by a dynamical system approach (Lopez *et al.*, 2002).

Let Y_h be the number of infected humans, and Y_v the number of infected vectors. We can write

$$\begin{aligned} \frac{dY_h}{dt} &= \frac{Y_v a}{N_h} b_{v\to h} S_h - rY_h \\ \frac{dY_v}{dt} &= \frac{S_v(t-\tau)a}{N_h} c_{h\to m} e^{-\mu\tau} Y_h(t-\tau) - \mu Y_v \end{aligned}, \tag{3.13}$$

where S_h and S_v are the number of susceptible humans and vectors, respectively.

To deduce the threshold for the disease to establish in the human population we analyze the stability of the trivial solution $S_h = N_h$, $S_v = N_v$, $Y_v = Y_h = 0$, that is, the solution representing the absence of the infection. Linearizing the system (3.13) around the trivial solution we get

$$\begin{aligned} \frac{dy_h}{dt} &= y_v a b_{v\to h} - ry_h \\ \frac{dy_v}{dt} &= \frac{N_v a}{N_h} c_{h\to m} e^{-\mu\tau} y_h(t-\tau) - \mu y_v \end{aligned}, \tag{3.14}$$

where y_v and y_h are small deviations from zero. From the system (3.14) we get the following characteristic equation

$$\begin{vmatrix} -(\lambda + r) & ab_{v\to h} \\ \frac{N_v a}{N_h} c_{h\to m} e^{-\mu\tau} e^{-\lambda\tau} & -(\lambda+\mu) \end{vmatrix} = 0 \tag{3.15}$$

or

$$\lambda^2 + (\mu + r)\lambda + \mu r - \frac{N_v a}{N_h} c_{h\to m} e^{-\mu\tau} e^{-\lambda\tau} a b_{v\to h} = 0. \tag{3.16}$$

It follows that the roots of equation (3.15) or (3.16) have negative real parts if

$$\mu r - \frac{N_v a}{N_h} a b_{v\to h} c_{h\to m} e^{-\mu\tau} > 0. \tag{3.17}$$

The above result is the same as that obtained by the intuitive McDonald's approach.

This still holds true for slightly more complex systems, like those with one vector and two hosts populations or two vectors with one host populations. In these cases, the expression for R_0 is partitioned in a sum with the individual terms of each component of the transmission chain (Burattini *et al.*, 1998).

In a classical paper Diekmann *et al.*(1990) propose a new definition of the basic reproduction number for infections which we now study how it compares with the classical Macdonald definition described above.

Those authors define R_0 as being the greatest eigenvalue of an operator which they call "the next generation operator (NGO)". The case of vector-transmitted infections was analyzed in a recent book by Diekmann and Heesterbeek (2000).

In this section we give the next generation operator for the case of one-vector/one-host, exemplified by malaria. In this case, the next generation operator reduces to a two-by-two matrix

$$NGO = \begin{pmatrix} A_{v\to v} & A_{v\to h} \\ \\ A_{h\to v} & A_{h\to h} \end{pmatrix}. \tag{3.18}$$

The elements have the following interpretation. The element $A_{v\to h}$, for instance, means the number of infected humans generated by a single infected vector during its infectious period. Therefore, we have

$$\begin{aligned} A_{v\to v} &= 0 \\ A_{v\to h} &= a\frac{1}{\mu}b_{m\to h}e^{-\mu\tau} \\ A_{h\to h} &= 0 \\ A_{h\to v} &= \frac{N_m a}{N_h r}c_{h\to m} \end{aligned} .$$

In this case the greatest eigenvalue of the NGO matrix, that is, R_0^{NGO}, is

$$R_0^{NGO} = \sqrt{a\frac{N_m a}{N_h r}c_{h\to m}\frac{1}{\mu}b_{m\to h}e^{-\mu\tau}}, \tag{3.19}$$

which is the square root of the Macdonald R_0. It follows from the general theory of the Next Generation Operator (Diekmann *et al.*, 1990) that if $R_0^{NGO} < 1$ ($R_0^{NGO} > 1$) the disease cannot (can) invade the host population.

3.2.2 Analytical Models

Analytical models are those which involve the association of a set of equations to each step individuals from the community take with the development of the natural history of the infection. Also called dynamical models, they capture the structure of the disease as this take their natural course. They can be of either deterministic or stochastic nature. Deterministic models are those which use

difference, differential, integral or functional differential equations to describe the changes in time of the sizes of the epidemiological classes. So, consider, for instance, the picture described in figure 3.1, which describes the progress of a viral infection, such as measles, through a host.

The curves in figure 3.1 illustrate the growth of the virus population, the immune response to the virus, and the timing of acute disease, a^*. The block diagram in figure 3.2 below represents the flow of the transmission between infection categories. In this simple situation, the total population is considered to be constant, and therefore the birth and death rates are equal (μ). People are born susceptible X, are infected at a rate λ, passing by a latent state, H, before developing the acute disease with a rate σ. Those infected and infectious, Y, recover from the infection at a rate γ, remaining immune, Z, for life.

The system illustrated by figure 3.1 is extremely simple but it captures the essence of the transmission chain between the involved categories. In general, systems like this, although still very simple in its biological assumptions are too complex to have analytical solution, i.e., the solution of the associated system of differential equations without the help of numerical simulations. This is the first problem posed by modeling biological phenomena: the highest the biological realism the lower the probability of analytical solution. Figure 3.3 illustrates, through a flow chart, the algorithm to solve infectious diseases problem by mathematical and computer modeling.

As can be noted from figure, everything starts, as in any scientific approach, by defining the problem to be modeled. The next step is to design the model, in this case, the set of differential equations associated to the disease categories. If the system is simple enough then the analytical solution (by pencil and paper) is feasible. If so, then is just to interpret the results and the problem is finished. If, on the other hand, the problem is too complex, then it is necessary the

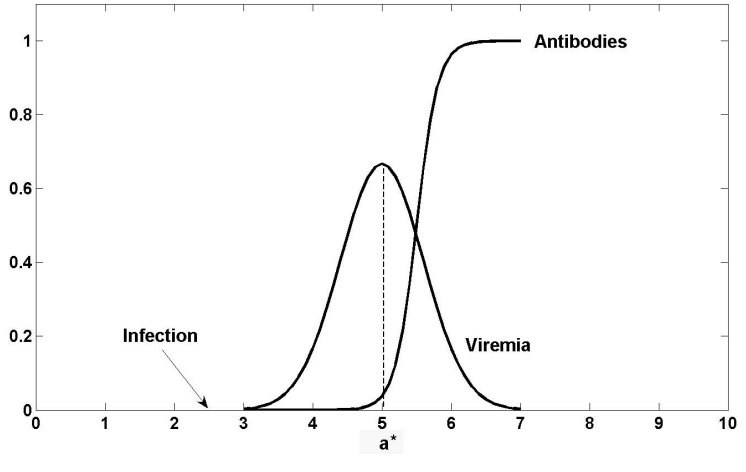

Fig. 3.1. Typical progress of a viral infection, such as measles, in a SIR (susceptible-Infected-Recovered) dynamical model

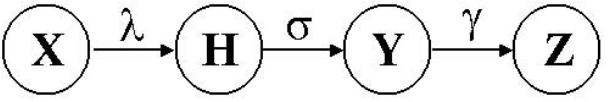

Fig. 3.2. Flow diagram for a epidemic system type Susceptible-Latent-Infected-Recovered

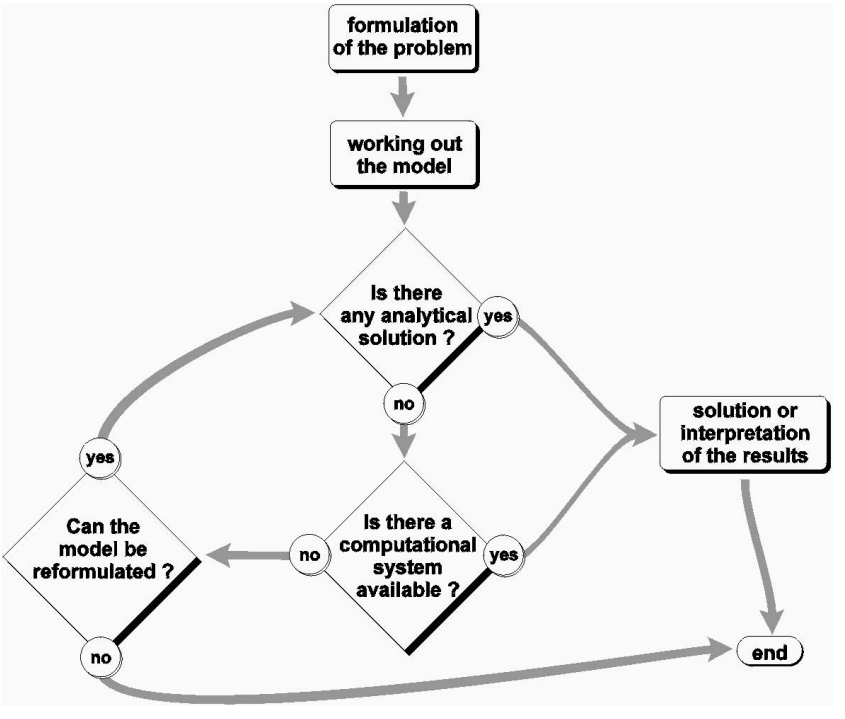

Fig. 3.3. Algorithm to solve infectious diseases problems by mathematical and computer modeling

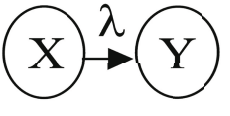

Fig. 3.4. The simplest epidemic model, Susceptible-Infected type

application of numerical methods for the computational solution. If the system to solve the model numerically is available, then is just to interpret the results and the problem is finished. If the computational system is not available, then the model needs to be reformulated, whenever possible, and the problem starts

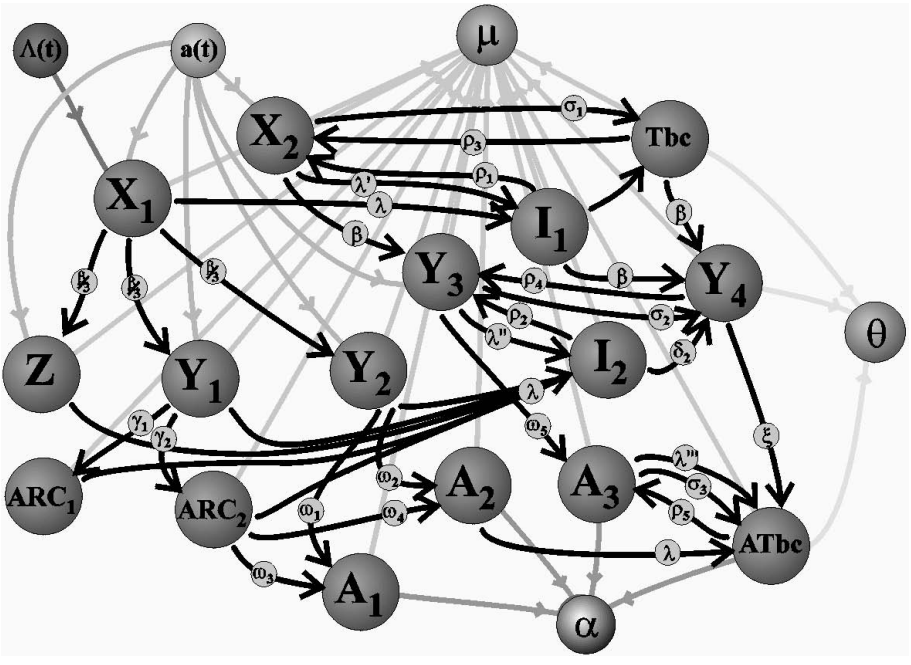

Fig. 3.5. Compartmental model for the interaction between AIDS and crack abusers (Burattini *et al.*, 1998)

again. In case the model cannot be reformulated, then it is not possible to solve it and the system stop. Let us see some examples of the situation described above.

In figure 3.4, it is illustrated the simplest epidemic model as possible, involving only two categories, namely, susceptibles, X, and infected, Y. Individuals acquire the infection with a rate λ, normally denoted the *force of infection*. This latter parameter is related to the incidence rate, *i.r.*, by equation:

$$i.r. = \lambda X. \tag{3.20}$$

In this simple, and unrealistic model, it is associated the following system of ordinary differential equation (ODE):

$$\begin{aligned} \frac{dX(t)}{dt} &= -\lambda X(t) \\ \frac{dY(t)}{dt} &= \lambda X(t) \end{aligned} \quad . \tag{3.21}$$

If we assume the total population as a constant, we can work with proportions, in the sense that $X(t) + Y(t) = 1$. The analytical solution of equations 3.21 is then straightforward:

$$X(t) = X(0) \exp[-\lambda t]$$

and

$$Y(t) = 1 - \{X(0)\exp[-\lambda t]\}\,.$$

In the other end of the spectrum we have a model which, by incorporating almost all the biological realities, is so complex that no analytical solution is possible. Figure 3.5 illustrates a compartmental model for the interaction between AIDS and crack abusers (Burattini *et al.*, 1998).

The associate system of ODE is then:

$$\begin{aligned}
\frac{dS(t)}{dt} &= \delta S_0 - \lambda_{S_1} S(t)\,(*) - \lambda_{S_2} S(t)\,(**) + a_5 C(t) + a_6 D(t) + \\
&\quad a_9 U(t) - \ (a_1 + a_2 + \mu)\,S(t) \\
\frac{dU(t)}{dt} &= \delta U_0 - \lambda_{U_1} U(t)\,(*) - \lambda_{U_2} U(t)\,(**) + b_2 S_Y(t) \\
&\quad - \ (b_3 + b_8 + b_9 + \mu + \mu_U + \omega_U)\,C_Y(t) \\
\frac{dC(t)}{dt} &= \delta C_0 + \lambda_{C_1} C(t)\,(*) + b_1 S_Y(t) + b_4 D_Y(t) \\
&\quad - \ (b_5 + b_7 + \mu + \mu_c + \omega_c)\,C_Y(t) \\
\frac{dD(t)}{dt} &= \delta D_0 + \lambda_{D_1} D(t)\,(*) + \lambda_{D_2} D(t)\,(**) + b_3 U_Y(t) + b_7 C_Y(t) \\
&\quad - \ (b_4 + b_6 + \mu + \mu_D + \omega_D)\,D_Y(t) \\
\frac{dS_Y(t)}{dt} &= \delta S_{Y_0} + \lambda_{S_1} S(t)\,(*) + b_5 C_Y(t) + b_6 D_Y(t) + b_9 U_Y(t) \\
&\quad - \ (b_1 + b_2 + \mu + \omega_S)\,S_Y(t) \\
\frac{dU_Y(t)}{dt} &= \delta_Y U_Y + \lambda_{S_2} S(t)\,(**) + \lambda_{U_1} U(t)\,(*) + \lambda_{U_2} U(t)\,(**) + b_2 S_Y(t) \\
&\quad - \ (b_3 + b_8 + b_9 + \mu + \mu_U + \omega_U)\,C_Y(t) \\
\frac{dC_Y(t)}{dt} &= \delta_Y C_Y + \lambda_{C_1} C(t)\,(*) + b_1 S_Y(t) + b_4 D_Y(t) \\
&\quad - \ (b_5 + b_7 + \mu + \mu_c + \omega_c)\,C_Y(t) \\
\frac{dD_Y(t)}{dt} &= \delta_Y D_Y + \lambda_{D_1} D(t)\,(*) + \lambda_{D_2} D(t)\,(**) + b_3 U_Y(t) + b_7 C_Y(t) \\
&\quad - \ (b_4 + b_6 + \mu + \mu_D + \omega_D)\,D_Y(t) \\
\frac{dS_A(t)}{dt} &= c_9 U_A(t) + c_6 D_A(t) + c_5 C_A(t) - (c_1 + c_2 + \mu + \alpha)\,S_A(t) \\
\frac{dU_A(t)}{dt} &= \omega_U U_Y(t) + c_2 S_A(t) - (c_3 + c_8 + c_9 + \mu + \mu_U + \alpha_U)\,U_A(t) \\
\frac{dC_A(t)}{dt} &= c_1 S_A(t) + c_8 U_A(t) + c_4 D_A(t) + \omega_C C_Y(t) \\
&\quad - \ (c_5 + c_7 + \mu + \mu_C + \alpha_C)\,C_A(t) \\
\frac{dD_A(t)}{dt} &= c_3 U_A(t) + c_7 C_A(t) + \omega_D D_Y(t) - (c_4 + c_6 + \mu + \mu_D + \alpha_D)\,D_A(t),
\end{aligned}$$

where $(*)$, $(**)$ and δ are given by:

$$(*) \quad = S_Y(t) + U_Y(t) + C_Y(t) + D_Y(t) + S_A(t) + U_A(t) + C_A(t) + D_A(t)$$

$$(**) = U_Y(t) + D_Y(t) + U_A(t) + D_A(t)$$

$$\delta \quad = \mu S + \mu_U U + \mu_C C + \mu_D D + \mu S_Y + \mu_U U_Y + \mu_C C_Y + \mu_D D_Y$$

$$+(\mu + \alpha)S_A + (\mu_U + \alpha_U)U_A + (\mu_C + \alpha_C)C_A + (\mu_D + \alpha_D)D_A.$$

As can be noted, this model has no analytical solution whatsoever.

This examples illustrate how to deal with deterministic models, and the difficulties involved in their solutions.

In stochastic models, there are probabilities at each time step of moving from one epidemiological class to another. When these models are simulated with the probabilities calculated using random number generators, the outcomes of different runs are different so that this approach is often called Monte Carlo simulations; conclusions are obtained by averaging the results of many computer simulations.

Stochastic epidemiological models incorporate chance, but it is usually harder to get analytic results for these models. Moreover, computational results are also harder since Monte Carlo simulations require many computer runs (25 to 100 or more) in order to detect patterns and get quantitative results. Even for stochastic epidemiological models where parameters are estimated by fitting the mean of the simulation to data, it may not be possible to find confidence intervals on these parameters estimates (Hethcote, 2000). As an example of this kind of approach we may consider the model by Haas (personal communication) who considered the individuals from a given population as occupying the sites of a net. The model then generates, at each Monte Carlo step, a certain number of sites to be checked. If the drawn site is occupied by a susceptible individual then the model generates a random number. If the random number is lower than the *a priori* probability that a susceptible gets the infection, then this individual gets the infection. Otherwise the site remains susceptible. If, on the other hand, the drawn site is occupied by an infected individual, then the model generates another random number. If this random number is lower than the *a priori* probability that an infected individual recovers from the infection, then this individual recovers from the infection. Otherwise, the site remains occupied by an infected individual, and so on. Each Monte Carlo step is considered to be a time step so the temporal evolution of an epidemic can be followed up. This model has shown a good retrieving capacity when tested against a real epidemic.

If it is harder to get analytic results of stochastic models, they are not impossible to be analyzed. Let us see, as an example of an analytical solution of a stochastic approach, the model by Massad *et al.* (1994a), who considered the transmission of HIV, the causative agent of AIDS, among injecting drug users, by applying the Mcdonald's (1952) model originally designed for vector-borne infections, to the injecting apparatus. One of the problems faced by those authors was the estimation of the probability, δ, that a needle which pricked an infected individual would get infected.

The authors considered that this chance, δ, depends on the statistical distribution of HIV serum title in the host population. The mean titers of HIV in the blood of contaminated individuals is quantified in the literature as tissue culture infective doses (also called by some authors as "infectious particles") (Ho *et al.*, 1995, Levy, 1988). Let us then suppose that the infective inoculum for the needle is of i Units Infective for Tissue Culture (UITC). So, the probability, $P(i)$, of finding i UITC in a needle with a residual volume of ν ml of blood, after 'biting' an infected individual, is:

$$P(i) = \sum_{n=0}^{\infty} \frac{e^{-n\nu}(n\nu)^i}{i!} P(n \mid \varphi), \tag{3.22}$$

where n corresponds to the plasmatic concentration of HIV (in UITCs) in a randomly selected individual from an infective population with an average φ UITCs per ml of blood.

Assuming that one UITC is *sufficient* to infect a needle, then the probability of having *at least* one infective inoculum, δ, is:

$$\delta = P(i \geq 1) = [1 - P(0)]\,. \tag{3.23}$$

In order to illustrate the above analysis let us consider two possible situations:

a) the infective inocula are homogeneously distributed among the infected population. Then,

$$P(n \mid \varphi) = \begin{cases} 1 \; if \; n = \varphi \\ 0 \; if \; n \neq \varphi \end{cases} \tag{3.24}$$

and so, $P(i)$ reduces to the Poisson distribution:

$$P(i) = \frac{e^{-\varphi\nu}(\varphi\nu)^i}{i!}. \tag{3.25}$$

Therefore,

$$\delta = 1 - e^{-\varphi\nu}; \quad and \tag{3.26}$$

b) the infective inocula are heterogeneously distributed among the infected population. Then:

$$\delta = P(i \geq 1) = 1 - \sum_{n=0}^{\infty} e^{-n\nu} P(n \mid \varphi). \tag{3.27}$$

Now, assuming that the distribution of infective inocula among the infected population, $P(n \mid \varphi)$, is a Negative Binomial Distribution with parameter κ (Spiegel, 1992), as described previously for several epidemiological studies (Anderson, 1982), with form:

$$P(n \mid \varphi) \equiv P(n \mid \varphi, \kappa) = (1 - \epsilon)^{\kappa} \frac{\Gamma(\kappa + n)}{\Gamma(\kappa)} \frac{\epsilon^n}{n}, \tag{3.28}$$

where

$$\epsilon = \frac{\varphi}{\varphi + \kappa}. \tag{3.29}$$

Considering that the Probability Generating Function, G, for the Negative Binomial Distribution is given by (Feller, 1968):

$$G(z) = (1-\epsilon)^{\kappa}(1-\epsilon z)^{-\kappa} \tag{3.30}$$

we finally obtain:

$$\delta = P(i \geq 1) = 1 - (1-\epsilon)^{\kappa}\left(1-\epsilon e^{-\nu}\right)^{-\kappa}, \tag{3.31}$$

where ν, as mentioned above, is the residual volume of blood in the needle and κ can be obtained from the distribution of inocula in the population, being defined as:

$$\kappa = \frac{\frac{\varphi^2}{\sigma^2}}{1-\left(\frac{\varphi}{\sigma^2}\right)}, \tag{3.32}$$

where φ is the mean and σ^2 is the variance of the inoculum distribution among the population.

The parameter δ,was estimated from the data described by Ho *et al.*(1995). From these data we get the average concentration of infective inoculum per unit of blood volume ($\varphi = 30$ UITC/ml) and its respective variance ($\sigma^2 = 1.0 \times 10^3$). Therefore, by taking the residual volume of blood in the lumen of the needles as equal to 6.6 μl (a typical '30 $\times$ 7' needle), we obtain the value of $\delta = 0.18$ assuming a homogeneous, and $\delta = 0.24$ assuming a heterogeneous distribution of infective inocula, respectively.

3.2.3 Computer Simulation Models

Computer simulation models consist of interactions, in a computational environment, of individuals from a given population. If a certain number of those individuals are infected by a certain pathogen, the potentially infective contacts between susceptible and infected individual may generate new infections with a certain probability. The already well-known ONCHOSIM model (Plaisier *et al.*, 1990) for the study of onchocerciasis, described as a computer simulation is in reality a stochastic model.

Computer simulation approach can be seen as an alternative for analytic models is which it is created an environment in the computer screen, where artificial organisms "live" and can mimic the epidemic situations (Silveira *et al.*, 1995).

Since the creation of the Game of Life by Conway in 1970, probably the first computer model employing cellular automata (CA) principles, CA have been used in a large range of applications, including evolutionary biology, predator-prey interactions, airship design, robotics, epidemics, etc.

Another advantage of computer models is that they allow researchers to observe events in a much shorter time span than they can do in nature. Hypotheses can be tested in this kind of model by recreating the reality using rules as simple as possible. In this context, this kind of model is a 'tool to think', a machine designed to execute ideal experiments, where the simple conditions of interactions among organisms are re-created and the consequences can be analyzed. Let us

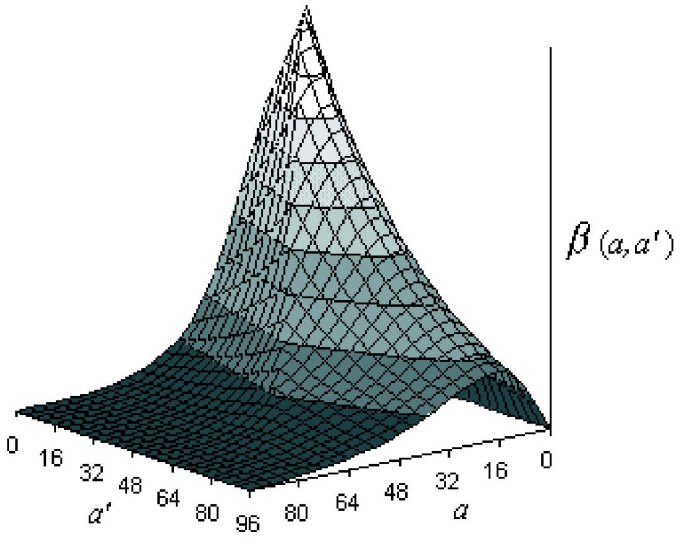

Fig. 3.6. Graphical shape of the function for $\beta(a, a')$ for the epidemiology of rubella in São Paulo, Brazil (Massad *et al.*, 1995), from analytical model.

see an example that illustrates the concept of computer simulation models in epidemiology, due to Silveira *et al.* (1996).

In mathematical epidemiology of infections a central problem is the heterogeneity in the potentially infective contact rates between susceptible and infected individuals. These heterogeneities can arise in several cases, like the distribution of susceptibilities, heterogeneities in infectiousness, in behavior or, simply the age distribution of those contacts. So, the potentially infective contact rate $\beta(a, a')$, describes those contacts between individuals susceptible of age a with infective individuals of age a'. These rates have been treated in the specialized literature as a matrix, described by the first time by Anderson and May (1991), as the Who-Acquired-the-Infection-from-Whom (WAIFW) matrix. Although of wide use, this matrix has the disadvantage of using a discrete rate $\beta(a, a')$. We have been dealing with this problem for some time and have proposed a continuous equivalent for the WAIFW matrix (Massad *et al.* 1994b), which has the form:

$$\beta(a,a') = \frac{\kappa_1}{\kappa_2}\frac{1}{\Gamma(\kappa_1+1)}\frac{\left(\frac{a}{\kappa_2}\right)^{\kappa_1}\exp\left(-\frac{a}{\kappa_2}\right)}{[2-\exp(-\kappa_3 a)]}\exp\left(-\kappa_3\left|a'-a\right|\right), \tag{3.33}$$

where Γ is the Gamma function and κ_i are parameters that can be fitted by Maximum Likelihood techniques to real epidemic situations. By doing this we managed to find a suitable function for $\beta(a, a')$ for the epidemiology of rubella in São Paulo, Brazil (Massad *et al.*, 1995). Figure 3.6 illustrates the graphical shape of such a function.

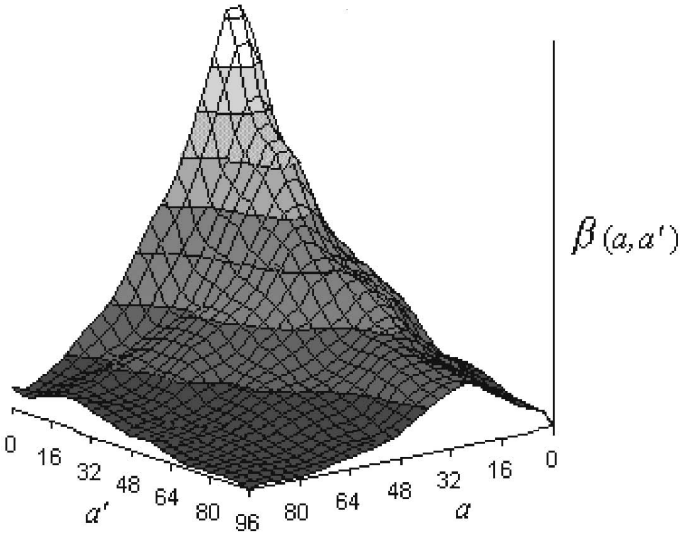

Fig. 3.7. The age-dependent contact rates as a function of ages a and a' (Massad *et al.*, 1995), from computer simulation

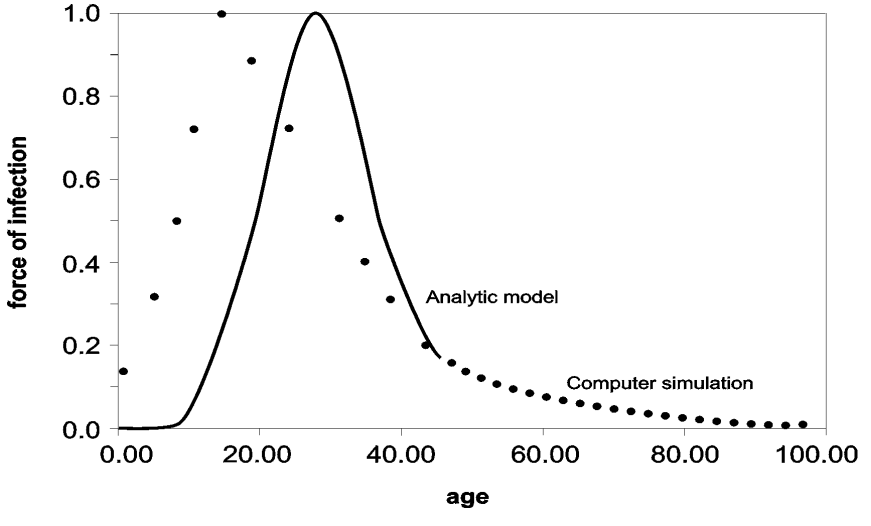

Fig. 3.8. The age-dependent force of infection calculated by both analytic and computer simulation approaches

An alternative approach is to use computer simulation systems, applying cellular automata and genetic algorithm principles (Silveira *et al.*, 1995). So we performed a set of simulation in a computer-based environment which mimics

the interactions, or contacts, described by the maximum searching distance for individuals, represented by pixels in a computer screen, to get in contact with other individuals, presence or absence of disease, presence or absence of immune response to the infection, duration of infectiousness, etc.

After reaching the equilibrium, we can plot the age-dependent contact rates as a function of ages a and a'. Results can be seen in figure 3.7.

It noteworthy the striking similarities between the analytical (figure 3.6) and the computer simulation (figure 3.7) models.

Another derived parameter related to transmission, is the force of infection, defined above as the rate at which susceptible individuals acquire the infection, and has the form:

$$\lambda(a) = \int_0^\infty \beta(a, a')y(a')da', \tag{3.34}$$

where $y(a')$ is the proportion of infectious individuals. Equation (3.34) can be estimated by both analytic and computer simulation approaches. Figure 3.8 shows the age-dependent force of infection calculated by both methods.

Again, it is noteworthy the striking similarities between both approaches.

This model exemplifies the computer simulation method and also demonstrates how it compares with analytic models.

Another interesting computer model is the Cybermouse, an on-line virtual laboratory animal for exploring the immune system (Cybermouse, IPC, Freemont, CA, U.S.A). This model is similar to the model described above in the sense that it applies CA principles to study the interaction of immune cells with antigens. One of its main applications have been to provide testable predictions about how the AIDS virus spreads through the immune system and how the systems "remembers" a pathogen years after the original infection and respond with a quick counterattack.

3.3 What Lies in the Future?

How should epidemiology face up to the challenge of the post-genome sequencing world? The high profile of genetic research will surely increasingly influence epidemiology, indeed it is already noticeable that relative risks only a little above unity get treated with considerably more excitement than if similar relative risks were associated with other exposures (Smith & Ebrahim 2001). This demonstrates the central role of individuality in medical and epidemiological thinking of the 21^{st} century.

The steady shift from the typological Oslerian "body-as-a-machine" medical approach to the populational Garrodian (from Archibald Garrod, Osler's successor in the Regius Professor of Medicine chair in Oxford) "chemical individuality" approach is gaining momentum with the huge amount of genetical information. The need for quantitative tools able to deal with this trend is impossible to be superestimated. All quantitative community is, therefore, invited to participate in this effort.

4 Probability, Possibility and Fuzzy Events

It is very common to confuse fuzzy sets theory with the theory of probabilities. Frequently membership degree is misinterpreted as probability value and the membership function as a statistic distribution function. This confusion emerges due to the fact that there is a narrow relationship among the two theories and, under certain aspects, the fuzzy theory presents characteristics very similar to the probability theory. In fact, its possible to see from the theory of measures that the probability measure is a particular case of fuzzy measure (Klir & Yuan, 1995; Pedrycz & Gomide, 1998 and 2007; Barros & Bassanezi, 2006). Nevertheless, it is important to understand the differences between these theories for best comprehend how they can be complementary, and how they can work together, as interesting and powerful mathematical tools.

The theory of probabilities and the theory of fuzzy sets work, in general, with different types of uncertainties, as discussed shortly in the introductory chapter (see chapter 2). In the theory of probabilities the event is well defined and the doubt hovers on the occurrence of the event. However, once the event happened there will be no more doubt. It is possible to calculate how is the probability of, in an urn containing n_w white balls and n_r red balls, one raffle a red ball. Nevertheless, once the ball is raffled there is nothing else to do, the ball will be white or red, and the uncertainty disappears. On the other hand, suppose that there is inside the urn several balls of red tones, varying of the intense red to the white one. In this case it is not possible to ask "*How is the chance of raffling a white ball*", simply because it is difficult to decide about the pink balls. Actually, it is not possible to make a typical probabilistic question because in this situation the event is not well defined. There are "almost red" balls, "almost white" balls and balls with several tones of rose, configuring an imprecise situation. It is important to know in these cases what is the appropriate question to be asked and how to answer it.

Another aspect that differentiates the two theories is the fact of the theory of probabilities do not consider, in general, subjectivities as the individuals' psychological characteristics, for instance, a very common phenomenon with the people that gamble in the Lottery, including professionals of the statistic areas.

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 59–78, 2008.
springerlink.com

It is known that people rarely bet a set of numbers in sequence as $1, 2, 3, 4, 5$ and 6. They have, in general, the *sensation* that this set is much less probable of happening than a set of assorted numbers. However, it is not surprise that they feel like this, once this event hardly ever is verified. In fact, it hardly ever happens that a set of numbers in sequence to be raffled. Nevertheless, it is known that the probability of happening any combination of numbers is the same, but if one statistician is asked about what kind of sequence he/she bets in the Lottery rarely the answer is a sequence of numbers. In other words, even knowing that the probability of this bet is the same as any other, he/she does not *feel comfortable* in betting it, although if it bet he/she would have greater chance of winning alone if raffled. This example illustrates the fact that people do not assume a merely statistical approach in decision making. So, depending on the subject the theory of probabilities may not be enough to analyze the problem and others approaches may be needed to treat the uncertainties involved in the process. It is worth to point out that Bayesian statistics could model situations as in examples mentioned above. However, the main difference consists in the ability of fuzzy theory to treat uncertainties with non stochastic behavior.

The objective of this chapter is to discuss subjects related with types of subjectivities and the appropriate mathematics to describe them. We discuss and try to clarify the basic differences and likeness between probability and possibility measures, the last being an exclusive concept of fuzzy theory. In this sense, some concepts related to *Fuzzy Analysis* are presented, as fuzzy measure and fuzzy integral, both used to define the *Fuzzy Expected Value* of an uncertain variable. Such concepts will be applied in the epidemiological context presented the following chapters, particularly in chapters 9 and 12.

4.1 Probability and Fuzzy Measures

The notion of measure generalizes the usual concepts of length, area, volume and so on. Probability is a typical case of measure, with the restriction that its codomain must be the interval [0,1]. In this chapter we will study the measure of probability aiming to compare it with the so called fuzzy measure. The interested reader in General Theory of the Measure may consult several available books, among them we recommended the one by Bartle (1995).

4.1.1 Probability Measure

In the Probability Theory the main interest consists of knowing how to measure the chance of the occurrence of an event, or in other words, its probability, which is represented by a number between 0 and 1.

The typical way to define the probability of an event $A \subset \Omega$ is given by:

$$P(A) = \frac{n(A)}{n(\Omega)} \tag{4.1}$$

which is the ratio between the number of favorable cases to A and the total number of the all possible cases. Ω is also called space of events.

The central idea in the definition of probability is to identify each event as a subset of the sample space and to associate to this set a number, indicating the chance of occurrence of that event. In this sense, the probability is simply a real function of sets. The axiomatic definition of probability was presented in 1933 by Kolmogorov and it can be summarized in three properties (Kolmogorov, 1933):

- P_1. For all $A \in \mathbf{A} \Longrightarrow 0 \leq P(A) \leq 1$;
- P_2. $P(\Omega) = 1$;
- P_3. If A_1, A_2,...,A_i are disjunct two by two, then

$$P(\cup_{i \in N} A_i) = \sum_{i \in N} P(A_i).$$

where $\mathbf{A}$ represents a class of events of Ω called $\sigma - algebra$.

In the theory of probability, or usually in Theory of Measure, P should be $\sigma - additive$, that is, the property P_3 must be satisfied. However, this property is very strong, excluding concepts that may be measurable. The following example illustrates a situation where the $\sigma - additive$ condition seems not to hold.

Example 4.1. Suppose that we want to measure the productivity of a group of workers in a given factory. Let $\mu(A)$ be the productivity of a subset A of these workers. In that case, it is not reasonable that μ be necessarily additive, that is, $\mu(A \cup B) = \mu(A) + \mu(B)$, although $A \cap B = \phi$.

If the groups A and B work separated, it is reasonable that $\mu(A \cup B) = \mu(A) + \mu(B)$. However, if the groups A and B interact, then the equality may not be verified. If there is cooperation among the groups, then it may happen that $\mu(A \cup B) > \mu(A) + \mu(B)$. On the other hand, it can happen $\mu(A \cup B) < \mu(A) + \mu(B)$ if there is incompatibility among the operations of A and B.

Summarizing, if additive functions are acceptable to measure uncertainties, it is reasonable to suppose that non-additive functions can also play this role, if the problem in question demands it. In fact, the number of professionals that try to deal with uncertainties using mathematical tools to quantify them are increasing, particularly among health professionals. Fuzzy measure is one of the available tools for this task.

4.1.2 Fuzzy Measure

Aiming to relax the rigid property of $\sigma - additive$ of classic measure, Sugeno (1974) suggested the concept of the fuzzy measure. He proposed the substitution of the axiom P_3 from the probability measure by properties that indicate the continuity of a measure, which will be defined below. Here the measures with this characteristic will be called *Measures of Sugeno*.

Definition 4.2. *A mapping of subsets of* Ω, $\mu_S : P(\Omega) \rightarrow [0, 1]$, *is called a Sugeno Measure if:*

1. $\mu_S(\phi) = 0$ *and* $\mu_S(\Omega) = 1$*;*
2. $\mu_S(A) \leq \mu_S(B)$ *if* $A \subseteq B$*;*
3. If $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_i \subseteq \ldots$ *then* $\mu_S(\cup_{i \in N} A_i) = \lim_{i \to \infty} \mu_S(A_i)$*;*
4. If $A_1 \supseteq A_2 \supseteq ... \supseteq A_i \supseteq ...$ *then* $\mu_S(\cap_{i \in N} A_i) = \lim_{i \to \infty} \mu_S(A_i)$*.*

Another class of measures that plays a prominent role in fuzzy sets theory are the so-called *fuzzy measure*. The fundamental property that a measure must have in order that an integration theory could be developed is the monotonicity. With this tool it is possible to broaden the reaching of mathematical modeling of problems close to reality.

The definition of fuzzy measure varies in the literature. For instance, it is quite common to require just that this measure should be monotonous and positive. Nevertheless, all of these different definitions require that the measure of the emptiness is *zero* and that the monotonicity hold, as the capacity measure of Choquet satisfy (Nguyen & Walker, 2000), for instance. In this chapter the definition that we will adopt for fuzzy measure is the following:

Definition 4.3. *A mapping* $\mu : P(\Omega) \to [0, 1]$*, is called a fuzzy measure if*

1. $\mu(\phi) = 0$ *and* $\mu(\Omega) = 1$*;*
2. $\mu(A) \leq \mu(B)$ *always that* $A \subseteq B$*.*

Under this definition the Sugeno measure is a particular case of the fuzzy measure.

4.1.3 Possibility Measure

In 1978 Zadeh published the first article addressing the measure of possibility (Zadeh, 1978). This article brings important discussions. One of those discussions is regarding the statement, apparently naïve but very frequent in the day by day; "*such fact is possible but unlikely*". This suggests that, independently of the possibility notion adopted, (π), it is expected that $\pi(A) \geq P(A)$. Nowadays, such inequality has been widely discussed and it is usually called the *Principle of Consistence* (Dubois & Prade, 1980). Before the introduction of the concept of possibility measure it is interesting to ponder on different ways to taking into account the uncertainties treatment.

Suppose that in a certain problem we want to obtain the value of a parameter ω_0. However, the only information available is that such value belongs to a space Ω. This partial knowledge with regard to ω_0 indicates that some uncertainty model should be designed to estimate ω_0. If, in the attempt to estimate ω_0 it is found that there are more plausible elements than others in the Ω space, how to treat this information? The Bayesian approach suggests the assumption of a probability distribution and how to find this distribution is a subject to be discussed. On the other hand, in fuzzy sets theory the expert's gradual information is translated into membership functions, which indicate how elements of Ω space should have bigger or smaller "weight", in agreement with the expert knowledge. Therefore, the crucial difference between probability and possibility

theories consists in choosing a mathematical treatment of uncertainties based on a model conceived *a priori* or a model built hand to hand with experts.

In summary, in the stochastic theory the available information are treated through probability density functions. In fuzzy sets theory such informations are modeled by membership functions $\varphi : \Omega \to [0,1]$, which may be elaborated according to an expert knowledge. In the situation exposed above, $\varphi(\omega)$ indicates the possibility of ω to be ω_0. As we will see, the function φ is seen as a *distribution of possibilities* on Ω.

Definition 4.4. *A possibility distribution on the set $\Omega \neq \phi$ is a function $\varphi : \Omega \to [0,1]$ satisfying $\sup_{\omega\in\Omega} \varphi(\omega) = 1$ for some $\omega \in \Omega$.*

Note that any normal fuzzy subset of Ω define a possibility distribution on Ω (see chapter 2, section 1.2).

Definition 4.5. *A measure of possibility on Ω is a function of sets $\pi : \mathbf{P}(\Omega) \to [0,1]$, which satisfies:*

- π_1. *$\pi(\phi) = 0$ and $\pi(\Omega) = 1$;*
- π_2. *For all family $A_i \in I$ of subsets of Ω it is verified*

$$\pi\left(\cup_{i\in I} A_i\right) = sup\left\{\pi(A_i) : i \in I\right\}, \tag{4.2}$$

where $\mathbf{P}(\Omega)$ is the power sets of Ω, that is, the set of all crisp (classical) subsets of Ω.

In a narrow sense, a possibility measure is nothing more that a fuzzy subset of the power sets of Ω, since its codomain is the interval $[0,1]$. Note that, given the measure of possibility π, this induces a possibilities distribution function, φ_π, on Ω through its restriction to the elements of Ω. In other words, $\varphi_\pi(\omega) = \pi(\omega)$. On the other hand, given a function of possibility distribution $\varphi : \Omega \to [0,1]$, this induces a possibility measure on Ω given by:

$$\pi(A) = \begin{cases} \sup_{\omega\in A} \varphi(\omega) \; if \; A \neq \phi \\ 0 \qquad\quad if \; A = \phi \end{cases},$$

for all $A \subset \Omega$.

A consequence of the possibility measure definition is the fact that, for any crisp subset A and B of Ω, $\pi(A \cup B) = \max[\pi(A), \pi(B)]$.

In the following example we emphasize the difference between probability and possibility measures, highlighting however that not all membership functions can be interpreted as a possibility distribution.

Example 4.6. In epidemiology it is very common the attempt to evaluating how strong is the infectiousness of a given agent. This in general, is estimated through the number of people that are infected by an infectious individual. Therefore, such infectiousness is narrowly related with the rate $\beta \in [0,1]$, that susceptible

individuals are infected. The parameter β, called *Coefficient of Transmission*, is directly linked to the chance of happening the transmission of the disease when there is a contact between a susceptible individual and an infected one. In the classical deterministic models this parameter could be evaluated as an average over the population, assuming that all infectious individuals present the same 'power' to infect the susceptible ones. However, this homogeneity hypothesis consists in an oversimplification of reality, since it is expected that there are individuals with larger power to transmit the disease then others (Sadegh-Zadeh, 1999). This heterogeneity depends on several factors, among them the viral or parasite load of infected individual. So, a mathematical model that takes into account this heterogeneity should describe the transmission coefficient β as $\beta = \beta(v)$, where v is a new parameter of the model, representing the viral or parasite load of the infected individuals.

Considering the above, it is reasonable to suppose that the parameter β could be modeled as a membership function of some fuzzy subset of the viral or parasite load domain. Since v could be translated by a number, its domain could be the real numbers set. So, the transmission coefficient could be the map $\beta : I\!R \to [0, 1]$.

In this case, the number $\beta(v)$ reflects the degree with which the values β and v are associated, that is, the effect of v on the parameter β. In this sense, if the effective value (v_0) of the viral or parasite load is known, the β value is inferred as a consequence.

When building a membership function for the parameter $\beta(v)$ we are, in some way, taking into account the uncertainties in the description of the values assumed by it. However, no information about the viral or parasite load was considered in order to find v_0, and nothing indicates that the membership function β corresponds to a possibility distribution. However, an expert could offer informations about v_0 as, for instance, that $v_0 \in [v_{\min}, v_{\max}]$. In that case, we may design a mathematical model to evaluate v_0, since a space for this parameter was selected. In addition, if it is known that there are values of $v \in [v_{\min}, v_{\max}]$ more plausible than others, and with "weights" $\rho(v) \in [0, 1]$, then, it is possible to propose a possibility distribution for v to evaluate v_0, given by:

$$\rho : [v_{\min}, v_{\max}] \to [0, 1]. \tag{4.3}$$

The function ρ may be directly built by an expert, in agreement with his/her empiric knowledge. In addition, ρ does not need to be a probability density distribution, that is, its integral does not need to be necessarily equal to 1, which reinforce the difference between the fuzzy and the stochastic approaches to obtain an estimate. This example illustrates that the probability density distribution is for the probabilities measure what the possibility distribution is for the possibility measure.

Consider $A \subset I\!R$. If for some $v \in [v_{\min}, v_{\max}]$, $\rho(v) = 1$, then the possibility distribution ρ induces a measure of possibility on $I\!R$ given by:

$$\pi(A) = \begin{cases} \sup_{v \in A} \rho(v) \ if \ \ A \neq \phi \\ 0 \qquad\quad if \ A = \phi. \end{cases}$$

It is not difficult to verify that π is also a fuzzy measure. It is interesting to observe that this kind of measure could be adopted by an expert of the public health area to estimate how much a group of individuals are infected. Clearly, this choice consists in an upper-estimate, since the group is being 'measured' by the individual with larger viral or parasite load.

Finally, if $\int_{I\!R} \rho(v)dv = 1$, then the probability density distribution ρ induces a probability measure in $I\!R$ given by $P(A) = \int_A \rho(v)dv$.

We finish this session revisiting the *Principle of Consistence* and indicating some mathematical models to transform probability in possibility and vice-versa, which may be important in many practical problems. Consider, for instance, that we want to build the possibility distribution (membership function) of a fuzzy set from a collection of statistical data. Also, assume that we want to elaborate a probability density function from a possibility distribution. In these cases, operations that allow to transform a measure in other would be certainly useful. These transformations are also interesting to compare the information obtained from the two approaches when applied both in the same phenomenon. Although there are in the literature several ways to transform probability into possibility, or vice-versa, all of them in accord to the Principle of Consistence (Dubois & Prade, 1980), that is,

$$P(A) \leq \pi(A) \; for \; all \; A \subseteq \Omega.$$

We concentrate on the transformation method between probability and possibility for the case in which Ω space is finite. Suppose that $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ and that

$$1 = \pi(\omega_1) \geq \pi(\omega_2) \geq \ldots \geq \pi(\omega_n)$$

and

$$P(\omega_1) \geq P(\omega_2) \geq \ldots \geq P(\omega_n).$$

Then, the simplest transformations is given by:

$$\pi(\omega_i) = \frac{P(\omega_i)}{P(\omega_1)}$$

or

$$P(\omega_i) = \frac{\pi(\omega_i)}{\sum_{i=1}^{n} \pi(\omega_i)}.$$

One practical example of the utility of the probability/possibility transformations it is found in (Castanho, 2005; Castanho *et al.*, 2007), where a complete study of diagnosis of prostate cancer is presented, and summarized below.

4.1.4 Probability/Possibility Transformations Applied to the Prostate Cancer Analysis

When diagnosing a prostate cancer, doctors evaluate its stage to indicate the appropriate treatment. It is known that treatments as surgery or radiotherapy have great chance of cure when the tumor is confined to the organ.

To make that evaluation doctors have information given by the clinical exam (rectal touch and/or image modalities), blood test that measures the level of prostatic specific antigen (PSA) - that is a substance that increases when the tumor increase - and by biopsy. In the biopsy the tumor is classified by the score of Gleason in agreement with the degree of differentiation of the cells (aggressiveness of the tumor) and the proportion of the affected gland. Combining those three variables and using statistical data, there are in the urological literature tables that indicate the patient's probability to be in certain stage of the disease development (Partin *et al.*, 1990, 1997 and 2001): with involvement of the prostatic capsule, involvement of seminal vesicles, or with involvement of pelvic linfonodes. Those classifications are clearly imprecise and the borders among those stages are not well defined. Therefore, the idea of treating those classifications as linguistic variables are quite reasonable.

Thus, with the variables used to predict the stage of the prostate cancer, Castanho and collaborators developed a system based on fuzzy rules, with the purpose of obtaining the stage of the disease (Castanho *et al.*, 2007). In this system the output variable (stage of the disease) is modeled through fuzzy sets. For each real value that represents the system output, it corresponds a membership degree to the fuzzy set that describes the disease stage. The proposition "stage of the disease is confined", for instance, allows to see this degree as the possibility that the disease is confined to the organ. In that way, that proposition defines a possibility distribution in the set of individuals. The systems based on fuzzy rules assume an important role in fuzzy theory and it will be presented in detail in chapter 7.

In this way, information of the same phenomenon may be given either in probabilistic terms (tables of probability) or possibilistic terms (system based on fuzzy rules). To verify if those information are consistent, a probability/possibility transformation should be used.

Consider, for instance, a patient with the following pre-surgical data: the clinical status was classified as *palpable*; confined in the *less of half of a lobe* (*lobe* is each one of the two parts in that the prostate is subdivided); PSA level equal to $5.3ng/ml$ (ng means nanogram) and degree of Gleason in the biopsy equal to 7.0. According to the system based on fuzzy rules proposed by Castanho and collaborators, the possibility of that patient having cancer confined to the organ is 0.60; the possibility that he has capsular involvement is 0.93 and the possibility that he has involvement of vesicles and/or linfonodes is 0.11. Using the transformation

$$P(\omega_i) = \frac{\pi(\omega_i)}{\sum_{i=1}^{n} \pi(\omega_i)}$$

the following probabilities are obtained: 0.36 of the cancer to be confined to the organ; 0.57 that he has capsular involvement and 0.07 of having involvement of vesicles and/or linfonodes. In tables of probability available (Partin *et al.*, 1990, 1997 and 2001) it is found 0.33; 0.52 and 0.14, respectively. Therefore, at least for this case, it is possible to affirm that the results supplied in terms of probability and possibility indicate the same stage of the disease development.

Other transformations could be found in several fuzzy logic articles and books as Civanlar and Trussel (1986 and 1990), Dubois and collaborators (1993), Sudkamp (1992) and, Klir and Yuan (1995).

The idea now is to build a fuzzy integral from the fuzzy measure for, finally, finding a fuzzy expected value, as it is usually made on the classical stochastic theory (Barros & Bassanezi, 2006). In the next section we study two kinds of fuzzy integrals.

4.2 Fuzzy Integrals

At the end of the previous section we presented a justification of the importance of studying the concept of integral. However, such concept in mathematics, and in exact sciences in general, dispenses any motivation. Its importance for the related theoretical disciplines is not smaller than that due to the applications in several types of problems, as calculus of volumes, areas, energy, work etc. Here the integral calculation will be applied to the study of the expected value of uncertain variables. Thereby, in this section will be presented some concepts and properties of integrals with respect to the classic and fuzzy measures.

Nowadays there are many fuzzy integrals available in the literature of fuzzy mathematics, all of them starting from a fuzzy measure and, therefore, it does not demand the $\sigma - additive$ condition. Here we will only present two kinds of fuzzy integrals: the integral of Choquet and the integral of Sugeno.

4.2.1 The Integral of Choquet

The integral of Choquet of the function $f : \Omega \to [0, \infty)$ with respect to the measure μ, not necessarily additive, is given by:

$$(C)\int_{\Omega} f d\mu = \int_0^{\infty} \mu\{\omega \in \Omega : f(\omega) > \alpha\} d\alpha \tag{4.4}$$

where the last is the Riemann integral. Intuitively, to find the Choquet integral of the function $f : \Omega \to [0, \infty)$, it is simply to use the concept of Riemann integral in the levels of f, which is defined in the codomain of that function. Choquet used this integral concept with regard to the measure of capacity, which is not additive, in mechanics (Nguyen & Walker, 2000).

4.2.2 The Integral of Sugeno

The integral of Sugeno (1974) was proposed in the decade of 1970 and it was one of the first fuzzy integrals definition. Such integral was defined in order to

defuzzify a fuzzy number from a measure that is not necessarily $\sigma - additive$. The definition of the Sugeno integral, given below, applies to functions whose codomain is the interval $[0,1]$, that is, to membership functions of fuzzy sets.

Definition 4.7. *Let be $f : \Omega \to [0,1]$ a function and μ a fuzzy measure on Ω. The integral of Sugeno of f, on Ω with regard to measure μ, is the number:*

$$(S)\int_{\Omega} f d\mu = \sup_{0\leq\alpha\leq 1}[\alpha \wedge \mu\{\omega \in \Omega : f(\omega) \geq \alpha\}] = \sup_{0\leq\alpha\leq 1}[\alpha \wedge \mu\{\omega \in \Omega : f(\omega) > \alpha\}]. \tag{4.5}$$

If A is a classical subset of Ω, then

$$(S)\int_{A} f d\mu = \sup_{0\leq\alpha\leq 1}[\alpha \wedge \mu(A \cap H(\alpha)], \tag{4.6}$$

where $H(\alpha) = \mu\{\omega \in \Omega : f(\omega) \geq \alpha\}$.

It is interesting to notice that $H : [0,1] \to [0,1]$ is a decreasing and continuous function in almost all points of the domain (Barros & Bassanezi, 2006). This will be important in the understanding of the next results.

Some theorems will be enunciated. They have great usefulness in some examples and applications of this book, particularly in chapters 9 and 12.

Theorem 4.8. *Let $f : \Omega \to [0,1]$ be a typical membership function and μ a fuzzy measure on Ω. If the function $H(\alpha) = \mu\{\omega \in \Omega : f(\omega) \geq \alpha\}$ has a fixed point $\overline{\alpha}$ then*

$$(S)\int_{\Omega} f d\mu = \overline{\alpha} = H(\overline{\alpha}). \tag{4.7}$$

In other words, the value of the Sugeno integral of f coincides with the fixed point of H, if it exists.

The proof of this theorem is not simple, involving some specific concepts of mathematical analysis and it is beyond the goal of this text, however, it can be found in Kandel (1986). Figure 4.1 presents the illustration of it.

The boldface part of the curve in figure 4.1 indicates the value of $[\alpha \wedge H(\alpha)]$ for $\alpha \in [0,1]$, and has as *sup* the value $\overline{\alpha}$, which is the intersection of the bisector with the graph of $H(\alpha)$. That is, it coincides with the fixed point of $H(\alpha)$.

Theorem 4.9. *Let $f : \Omega \to [0,1]$ be a typical membership function and μ a classical measure on Ω. Then,*

$$|(S)\int_{\Omega} f d\mu - \int_0^1 H(\alpha)d\alpha| \leq \frac{1}{4}. \tag{4.8}$$

Note that $\int_0^1 H(\alpha)d\alpha = \int_0^1 \mu\{\omega \in \Omega : f(\omega) \geq \alpha\}d\alpha$ is the Choquet integral of the function f.

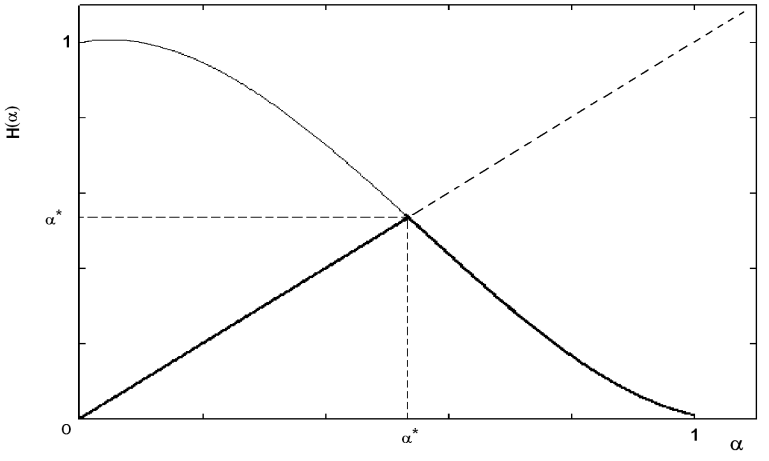

Fig. 4.1. Illustration of the Theorem that the value of the Sugeno integral of f is equal to the fixed point of H (Barros & Bassanezi, 2006)

Proof: Let be $\overline{\alpha} = H(\overline{\alpha})$ the fixed point of $H(\alpha)$. So,

$$\left|(S)\int_{\Omega} f d\mu - \int_0^1 H(\alpha)d\alpha\right| = \left|\overline{\alpha} - \int_0^1 H(\alpha)d\alpha\right| =$$

$$\left|\int_0^{\overline{\alpha}} 1 d\alpha - \left[\int_0^{\overline{\alpha}} H(\alpha)d\alpha + \int_{\overline{\alpha}}^1 H(\alpha)d\alpha\right]\right| = \left|\int_0^{\overline{\alpha}}(1 - H(\alpha))d\alpha - \int_{\overline{\alpha}}^1 H(\alpha)d\alpha\right|$$

$\leq \int_0^{\overline{\alpha}}(1 - H(\alpha))d\alpha \leq \int_0^{\overline{\alpha}}(1 - H(\overline{\alpha}))d\alpha = \int_0^{\overline{\alpha}}(1 - \overline{\alpha})d\alpha = (1 - \overline{\alpha})\overline{\alpha} \leq \frac{1}{4}$ since $\alpha \in [0, 1]$.

The penultimate inequality above is valid because H is a decreasing function. This theorem was originally proposed by Sugeno in 1974 for inexact variable, that is, any measurable function $X : \Omega \to [0, 1]$. At the end of this section it will be enunciated in such a context.

Properties of the Sugeno integral

Like the Choquet integral, the Sugeno integral is not linear. However, both of them are monotonous, that is,

$$(S)\int_{\Omega} f d\mu \leq (S)\int_{\Omega} g d\mu \quad if \quad f \leq g.$$

Bellow are other properties of the Sugeno integral, whose proofs can be found in (Barros, 1992; Barros & Bassanezi, 2006).

Let $f : \Omega \to [0, 1]$ and $g : \Omega \to [0, 1]$ be functions, μ a fuzzy measure on Ω and A and B subsets of Ω. Then, the following properties hold:

1. If $f(x) = k$, then $(S)\int_A f d\mu = k \wedge \mu(A)$;

2. If $f \leq g$, then $(S)\int_A f d\mu \leq (S)\int_A g d\mu$;
3. $(S)\int_A (f \vee g) d\mu \geq (S)\int_A f d\mu \vee (S)\int_A g d\mu$;
4. $(S)\int_A (f \wedge g) d\mu \leq (S)\int_A f d\mu \wedge (S)\int_A g d\mu$;
5. If $A \subset B$, then $(S)\int_A f d\mu \leq (S)\int_B f d\mu$;
6. $(S)\int_{A\cup B} f d\mu \geq (S)\int_A f d\mu \vee (S)\int_B f d\mu$; and
7. $(S)\int_{A\cap B} f d\mu \leq (S)\int_A f d\mu \wedge (S)\int_B f d\mu$;

where $f \vee g$ and $f \wedge g$ are, respectively, the maximum and minimum functions between f and g.

Like the integral of Lebesgue, that is used to obtain the expected value of an random variable with regard to a measure of probability, the integral of Sugeno has been used to obtain the fuzzy expected value of an uncertain variable with regard to a fuzzy measure.

Definition 4.10. *Let $X : \Omega \to [0,1]$ be an uncertain variable (typically a membership function) and μ a fuzzy measure on Ω. The Fuzzy Expected Value (FEV) of X is the real number:*

$$FEV(X) = (S)\int_{\Omega} X d\mu = \sup_{0\leq\alpha\leq 1} [\alpha \wedge \mu\{\omega \in \Omega : X(\omega) \geq \alpha\}]. \tag{4.9}$$

The following result was enunciated by Sugeno in 1974 and it is a consequence of theorem 4.9.

Corollary 4.11. *Let $X : \Omega \to [0,1]$ be a normalized random variable, which is typically a membership function, and P a probability measure on Ω. Then, it follows that:*

$$|FEV(X) - E(X)| \leq \frac{1}{4}. \tag{4.10}$$

where $E(X)$ is the expectation of the random variable X.

To proof this corollary it is enough to remind that the fuzzy measure in this case is a probability measure and to use theorem 4.9 (Sugeno, 1974).

This corollary legitimates the use of the fuzzy expectation in substitution to the classic one, when the uncertainty involved on the phenomenon under study is not originated from randomness but rather from the different possibilities for the variable in question.

It is important to point out that it is not possible to reduce the maximum difference of $\frac{1}{4}$ in theorem 4.9 above. In other words, there are functions f that this difference is attained (Ralescu & Adams, 1980). However, by choosing certain categories of functions, such differences are substantially reduced (Bassanezi & Barros, 1995). Actually, we verified that, for random variables X with symmetrical distribution the following theorem holds:

Theorem 4.12. *Let $X : \Omega \to [0,1]$ be a random variable with a probability density function $f : [0,1] \to [0,1]$, symmetrical in relation to $x = \frac{1}{2}$, that is, $f(x) = f(1-x)$ for all $x \in [0,1]$. Then,*

$$FEV(X) = E(X).$$

Proof: Since $f(x) = f(1-x)$, then

$$\int_0^{\frac{1}{2}} f(x)dx = \int_0^{\frac{1}{2}} f(1-x)dx = \int_1^{\frac{1}{2}} -f(x)dx = \int_{\frac{1}{2}}^1 f(x)dx.$$

On the other hand,

$$\int_0^1 f(x)dx = 1 = \int_0^{\frac{1}{2}} f(x)dx + \int_{\frac{1}{2}}^1 f(x)dx \Longrightarrow \int_0^{\frac{1}{2}} f(x)dx = \int_{\frac{1}{2}}^1 f(x)dx = \frac{1}{2}$$

and $H(\alpha) = P\{\omega \in \Omega : X(\omega) \geq \alpha\} = \int_\alpha^1 f(x)dx$.

Thus,

$$H(\frac{1}{2}) = \int_{\frac{1}{2}}^1 f(x)dx = \frac{1}{2}$$

and, consequently, $FEV(X) = \frac{1}{2}$.

In addition, it is known that the stochastic expectation $E(X)$ of any symmetrical random variable coincides with the median, and this is equal to $\frac{1}{2}$, what proves the theorem.

To finish this section we will enunciate a method to obtain $FEV(X)$ for the case in which X is an uncertain variable that assumes a finite number of values.

Theorem 4.13. *Suppose that the variable $X : \Omega \longrightarrow [0,1]$ assumes only $n+1$ values, $\{a_i\}_{1 \leq i \leq n+1}$, and let be $\{\mu_i\}_{i=1}^n$ the distinct values of $\mu\{\omega \in \Omega : \omega \geq a_i\}$, excluding the values $\mu = 1$ and $\mu = 0$. Assuming, without loss of generality, that $0 \leq a_i \leq a_j \leq 1$ if $i \leq j$, then:*

$$FEV(X) = \int_\Omega X d\mu = \textit{median of } A,$$

where $A = \{a_1, a_2, \ldots, a_{n+1}, \mu_1, \ldots, \mu_n\}$ is an increasing ordered set.

It is interesting to remember that the median of an increasing ordered sequence $\{a_n\}_{n \in N}$ is defined by;

$$med(\{a_n\}) = \begin{cases} a_{(n+1)/2} & if \quad n \; is \; odd \\ \dfrac{a_{\frac{n}{2}} + a_{(\frac{n}{2}+1)}}{2} & if \quad n \; is \; even \end{cases} . \tag{4.11}$$

The proof of the theorem 4.13 could be found in the book by Kandel (1986).

4.2.3 Fuzzy Expected Value Applied to the Accidents of Traffic Analysis, in the São Paulo City, Brazil

São Paulo is one of the most populous cities of the world, counting now about 17 million inhabitants in the called *Great São Paulo* (the town and main municipal districts of the surroundings). As expected, the traffic is one of the principal problems of the city, representing a great challenge for the local authorities. Only in the city of São Paulo it is registered two vehicles per inhabitant. Among the problems associated to the traffic we have the jam of the principal roads, resulting in hundreds of kilometers of traffic jam a day in the rush hours, and in the great number of accidents.

The accidents of traffic in the city corresponded in 1999 to 14% of the distribution of all deaths for external causes. In São Paulo is registered one traffic accident every 3 minutes, resulting in about 180.000 per year. Of these accidents, one death is registered every 6 hours, or 4 deaths a day, and about 1.500 per year.

The accidents of traffic vary from serious accidents, with death occurrence in the place, to light accidents, small crashes without victims. Seeking a better planning of resources and strategies definition that leads to reduction of the number of the accidents and deaths, information about the distribution of the types of accidents in time and in space are routinely collected and organized by the Company of Engineering of Traffic (CET), resulting in an extensive database on the accidents in the city.

Since the information about the gravity of the accident is as important as their numbers, from the decision making point of view, the objective in this study was to analyze the distribution of the accidents weighted by their gravity. The idea was to compare the analysis of the absolute data with that weighted by the gravity through the Fuzzy Expected Value. We analyzed the April data of 1997 and 1999.

In 1997 it was registered a total of 17,762 occurrences. From that, 12,304 (82.7%) without victims and 2,570 (17,3%) with victms. In 1999 it was registered a total of 14,874 occurrences, from that 15,227 (85.8%) without victims and 2,535 (14,3%) with victims. Table 4.1 below presents the weekly distribution of accidentes considering only the accidents with victims.

It is possible to observe in table that the day with the greater occurrence number was Saturday, in 1997, and Friday, in 1999, as expected. The gravity of an accident is classified into four categories: cases without victims, cases with light victims, cases with serious victims and, cases with fatal victims. In order to calculate the Fuzzy Expected Value for the number of accidents with victims we elaborated the following fuzzy set to the *gravity of accident*:

$$gravity = 0.3/light + 0.7/serious + 1.0/fatal.$$

The fuzzy measure considered was $\mu(S) = \#S/\#\Omega$, given by:

$$\mu\{\omega \in \Omega : X \geq \alpha\} = \begin{cases} 1 & if \;\; 0 \leq \alpha \leq 0.3 \\ (total - light)/total & if \; 0.3 < \alpha \leq 0.7 \\ fatal/total & if \; 0.7 < \alpha \leq 1.0 \end{cases} . \qquad (4.12)$$

Table 4.1. Weekly distribution of the absolute number of accidents with victims in April 1997 and in April 1999, from CET of São Paulo City

Day	**1997**	**1999**
Sunday	405	348
Monday	260	334
Tuesday	358	339
Wednesday	362	324
Thursday	327	426
Friday	363	438
Saturday	460	361
Total	2,535	2,570

As an example consider the number of accidents on Saturday, which recorded 237 light cases, 197 serious cases and 26 fatal cases. For this day the fuzzy measure above is:

$$\mu\{\omega \in \Omega : X \geq \alpha\} = \begin{cases} 1 \quad if \;\; 0 \leq \alpha \leq 0.3 \\ 0.485 \; if \; 0.3 < \alpha \leq 0.7 \\ 0.057 \; if \; 0.7 < \alpha \leq 1.0 \end{cases} . \tag{4.13}$$

The calculus of $[\alpha \wedge \mu\{\omega \in \Omega : X(\omega) \geq \alpha\}]$ in (4.9) for Saturday provide:

$$\{0.057,\ 0.300,\ 0.485,\ 0.700,\ 1.000\}$$

and, therefore, the FEV is equal to 0.485, since this is the maximum value of the sequence. Table 4.2 presents the results of the FEV for everyday of the week.

If we rank the days of the week in a crescent way for the absolute number of accidents we will have that, in April of 1997, Saturday occupies the first place, followed by Sunday and Friday, and in 1999 the first place in number of accidents is Friday, followed by Thursday and Saturday. However, when we considered the *gravity* of the accidents we obtain a different sequence. In April of 1997 Saturday continues occupying the first place, however, Friday occupies the second place now and Sunday is in third. As for 1999, a larger modification is observed. In

Table 4.2. FEV results for the number of accidents with victims weighted by the *gravity* of the accident, for April 1997 and April 1999 (Data from CET of São Paulo City)

Day	**1997**	**1999**
Sunday	0.469	0.454
Monday	0.381	0.347
Tuesday	0.366	0.345
Wednesday	0.384	0.435
Thursday	0.370	0.399
Friday	0.474	0.402
Saturday	0.485	0.421

this case, Sunday is the most worrying day, followed by the Wednesday and Saturday.

The main objective of this example is to illustrate that the results can be different when analyzed under a fuzzy perspective. The advantage of the use of FEV in this example was to allow a subjective evaluation for the absolute number of accidents, taking into account the gravity of these.

4.3 Probability of Fuzzy Events

As discussed before, fuzzy logic and probability theory, although similar in certain perspectives, were designed to different tasks. The values of probability measure has been classically defined as a number between 0 and 1 that preserves the additive property. However, in the real world there are cases where such a precise probability measure is difficult to obtain. This motivated the development of interval probability theories, one of the best known being the Dempster-Shafer theory of evidence (see Shafer,1987, for details), which is also called the theory of belief functions. Fuzzy probability in turn is a generalization of interval probability in which the probability value is bounded by a fuzzy set (Klir & Yuan, 1995).

The notion of fuzzy events and their probability measures were first introduced by Zadeh (1978). Since then, several other authors have contributed to comprehension of the relationships between fuzzy logic and probability theory. A remarkable account of the histories about fuzzy logic and probability is given by Yen and Langari in their book (1999).

A way of uniting the theory of probabilities and fuzzy logic is to consider the *probability of a fuzzy event*. An event is a subset of the sample space that shares a common characteristic of interest. However, there are situations when the event has not well-defined sharp boundaries. These events are called fuzzy events. Formally, a fuzzy event is a fuzzy subset of the sampling space and the probability of a fuzzy event may be seen as a generalization of the probability theory (Yen & Langari, 1999; Pedrycz & Gomide, 2007).

As a classical event is a crisp subset of the sample space, a fuzzy event is simply a fuzzy subset of the sample space.

Consider a probability space $(\Omega, \mathbf{A}, P)$, in which Ω is the space of events, $\mathbf{A}$ is a $\sigma - algebra$, and P is a probability measure in Ω. A fuzzy event in Ω is a fuzzy subset A of Ω whose membership function $\mu_A : \Omega \rightarrow [0, 1]$ is measurable. That is, μ_A is a random variable, in a classical sense.

Thus, the probability of A is given by:

$$P(A) = E(\mu_A),$$

where $E(X)$ represents the mathematical expectancy of the random variable X. That is,

$$E(X) = \int_{\Omega} X dP. \tag{4.14}$$

in which the integral is the Lebesgue integral in relation to the measure of probability P.

Note that if A is a classical event, then its membership function is $\chi_A : \Omega \to \{0, 1\}$, which is a discrete random variable. So the expression (4.14) has a form:

$$E(\chi_A) = 0 \cdot P(A^c) + 1 \cdot P(A) = P(A). \tag{4.15}$$

Thus, the equation (4.14) generalize the classical case.

In the fuzzy case μ_A can be not discrete, but from (4.14) and remembering that $\chi_\emptyset \equiv 0$ and $\chi_\Omega \equiv 1$, we find:

$$P(\emptyset) = 0 \quad and \quad P(\Omega) = 1,$$

which result in $0 \leq P(A) \leq 1$, since $0 \equiv \chi_\emptyset \leq \mu_A \leq \chi_\Omega \equiv 1$. This prove that, in fact, equation (4.14) defines a probability measure.

It is interesting to point out that if the probability measure is originated from a function $f : I\!R \to [0, \infty]$ with $\int_{-\infty}^{\infty} f(x)dx = 1$, then

$$P(A) = E(\mu_A) = \int_\Omega \mu_A dP = \int_0^1 \mu_A f(\mu_A) d\mu_A, \tag{4.16}$$

in which the last integral is in the Riemman sense.

It is interesting to note that expression (4.16) considers both fuzzy and stochastic uncertainties involved in the problem. While the membership function μ_A takes account of the fuzzy uncertainty, the density function $f(x)$ takes account of the stochastic uncertainty. In addition, the expression (4.16) can be interpreted as the average of μ_A weighted by $f(\mu_A)$.

In relation to the probability of fuzzy events, it is worthwhile to observe that in the classical case, under the Borel $\sigma - algebra$ $\mathbf{B}$, the probability measure P on $\mathbf{A}$ defines a probability P_X on $\mathbf{B}$ in the following way:

$$P_X(X \in B) = \int_B f(x)dx, \quad for \quad all \quad B \in \mathbf{B}. \tag{4.17}$$

In the fuzzy case, that is, when B is a fuzzy subset of $I\!R$, with μ_B measurable, Zadeh (1978) generalized the expression (4.17) to:

$$P_X(X \ is \ B) = \int_B \mu_B f(\mu_B) d\mu_B, \quad for \quad all \quad B \in \mathbf{B}. \tag{4.18}$$

For discrete variable the expression (4.18) can be rewrite as:

$$P_X(X \ is \ B) = \sum_i \mu_B(x_i) P(x_i). \tag{4.19}$$

Recently, Massad and collaborators presented two examples of applications of fuzzy probabilities to hypothetical epidemic situations, in order to answering two epidemic questions, the second of which involving Bayesian aspects central to causal studies (Massad *et al.*, 2003). This work will be discussed in the next section.

4.3.1 Fuzzy Probabilities of Epidemic Events

Certainly the most important application of fuzzy logic in epidemiology and health areas consists in its linguistic variables approach. This powerful tool is able to deal with the huge number of uncertainties characteristic of epidemic problems but also has enormous potential to solve problems that the complementary field of probability theory fails to do in any reasonable way.

Questions like "What is the *probability* that an individual (a population) *very* exposed to a certain cause will develop a *severe* disability?" are very difficult to deal with classical probability tools. The hedges emphasized **very** and **severe** are quantifiers of epidemic situations which are prone to be appropriately included in epidemic models by the use of fuzzy sets theory.

Suppose that X is a random variable that counts the total number of HIV (the causative agent of AIDS) positive individuals in a sample of 100 individuals taken from a population of intravenous drug abusers with seroprevalence of 50% (that is half of this population has the AIDS virus, or at least its specific antibodies). The interesting fuzzy event can be expressed by the question "What is the probability of the fuzzy event that this sample of 100 individuals has *s*everal more positive than negative individuals?". This fuzzy event may be characterized by the following membership function:

$$\mu_A(x) = \begin{cases} 0 & if\ 0 \leq x \leq 50 \\ (x-50)/30 & if\ 50 \leq x \leq 80. \\ 1 & if\ x \geq 80 \end{cases} \tag{4.20}$$

In a population with 50% seroprevalence to HIV, the probability distribution function is described by a Binomial distribution with $p = q = 0.5$ and for a sample of 100 individuals we have:

$$P(x_i) = \frac{100!}{i!\,(100-i)!} 0.5^i 0.5^{100-i}. \tag{4.21}$$

Therefore, the fuzzy probability of getting *s*everal more positive than negative individuals in a sample of 100 individuals is given by equation (4.19) and for this example results in 0.067 or 6.7%.

Let us now suppose that we are interested in answering the question "What is the fuzzy probability that an individual *v*ery active sexually will develop AIDS *v*ery quickly?". We are now dealing with the conditional probability that an individual will develop AIDS with a certain speed given that he/she is subject to a certain risk due to his/her sexual activity level. In this case we need to define the conditional probability of the fuzzy event. For this, suppose that A and B are any two events in a sample space Ω. The probability of both fuzzy events occurring is defined as (Yen & Langari, 1999):

$$P\left(A \cap B\right) = \sum_{x \in S} \mu_A(x) \mu_B(x) P(x). \tag{4.22}$$

Therefore, the conditional probability of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{4.23}$$

where $P(B) \neq 0$ and is given by Eq. (4.19).

Suppose now that the time in years taken to develop full-blown AIDS disease is a fuzzy variable and that the fuzzy event "to develop AIDS *very* quickly" is described by the following membership function:

$$\mu_A(x) = \begin{cases} 1 & if \quad x \leq 5 \quad years \\ (10-x)/5 & if \ 5 \leq x \leq 10 \ years \\ 0 & if \quad x \geq 10 \quad years \end{cases} \tag{4.24}$$

and that the incubation period is described by the prevalence profile shown in figure 4.2:

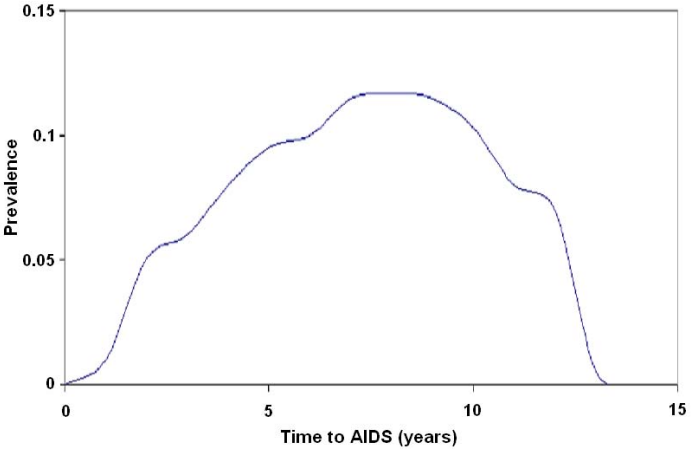

Fig. 4.2. Prevalence profile of HIV, in time to AIDS (Massad *et al.*, 2003)

Suppose also that the sexual activity, expressed by the annual rate is a fuzzy variable and that the event "sexually very active" is described by the following membership function:

$$\mu_B(x) = \begin{cases} 0 & if \quad x \leq 4 \quad years^{-1} \\ (x-3)/8 & if \ 4 \leq x \leq 10 \ years^{-1} \\ 1 & if \quad x \geq 10 \quad years^{-1} \end{cases} \tag{4.25}$$

and that the distribution of new sexual partners per year is given by figure 4.3.

So, the probability that an individual *very* active sexually will develop AIDS *very* quickly is obtained by applying equations (4.19), (4.22) and (4.23) and results in 0.734 or 73.4%. In these calculations it was considered the term $P(x)$

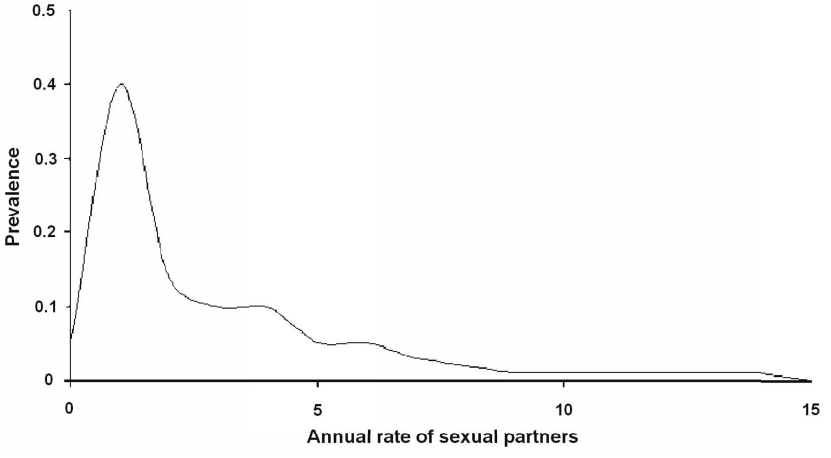

Fig. 4.3. Distribution of new sexual partners per year (Massad *et al.*, 2003)

of equation (4.22) as the fuzzy probability calculated by equation (4.19), that is, the product of the prevalence by its membership degree (Massad *et al.*, 2003).

4.4 Final Considerations

In this chapter we present some mathematical concepts of fuzzy theory, as fuzzy measure, fuzzy integral, and probability of fuzzy event. The main goal was compare the fuzzy and probabilistic approaches, highlighting the differences and similarities among them. Despite of the possible mathematical difficulties for a not familiarized reader, the presented examples can serve as possibilities of applications in epidemiology problems. We hoped these can act as seeds for future works. In the next chapter we will be devoted to the called fuzzy linguistic models, addressing static and dynamical systems.

5 Fuzzy Logic and Risk Estimators

Epidemiology is concerned with the identification of risk factors for diseases. This process relies on the definition and estimation of measures of association between a putative risk factor and its putative outcome. In order to achieve this end, several measures of association (Kleinbaum *et al.*, 1982) are commonly used, namely the risk ratio (RR), the risk difference (RD), the attributable risk (RA), the odds ratio (OR) and the hazard rate ratio (HR). Some of these measures were introduced in chapter 3 and we revisit their definitions at the beginning of this chapter. Often, the causal pathway between risk factors and disease outcome is translated into statistical models and the measures of association of interest can be estimated as functions of the model parameters.

In all steps above, sources of uncertainties can be present. Uncertainty in epidemiology is not restricted to sampling variations or individual heterogeneity. Uncertainties also arise from ignorance on how best assign individual subjects to the exposure and disease categories as well as ignorance on the very definition of the putative causal pathway. Several levels of imprecision and uncertainty, particularly in epidemiological studies, permeates the process leading to the establishment of a risk factor-disease association. As discussed in chapter 1, this may cause a tremendous amount of imprecision and uncertainty in the interpretation of effect measures of covariates of interest.

Nowhere in the field of biosciences is the need for tools to deal with uncertainty more critical than in medicine and epidemiology. A rigorous treatment of the various dimensions involved is still lacking, and recent contributions to this topic can be found in Almond (1995) and Pearl (2000). In what follows, we review the various dimensions contributing to the definition of measures of association in epidemiology. We further indicate the initial steps in using fuzzy logic to extend the definitions of epidemiologic measures of association in the presence of uncertainties about the membership of study subjects to epidemiologic categories of interest.

5.1 Risk and Common Measures of Association Used in Epidemiology

In chapter 3, section 3.1, *risk* was defined in the context of a binomial model of disease occurrence as the probability of failure, i.e., the probability that the

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 79–95, 2008.
springerlink.com

event under observation does occur within a certain specified period of time and conditional on not being lost to follow-up from any other causes during that period. Along with the definition of risk, π, the reader was also introduced to definition of *odds* ($\frac{\pi}{1-\pi}$).

The concept of risk plays a central role in modern Epidemiology (Morgenstern *et al.*, 1980; Rothman *et al.*, 2008). In causal studies individuals at risk are supposedly exposed or non-exposed to a certain cause and are then categorized into diseased and non-diseased. The definition of the concept of risk is given at the individual level while its estimator is defined at the population level. Most epidemiological measures relating the association between disease (or other event of interest for that matter) and its putative "cause" are derived from the concepts of outcome risk or odds.

Suppose that we intend to compare two groups subject to treatments A and B. Let π_A and π_B denote the risk of experiencing the event of interest in each group. Common measures of association between treatment and the outcome of interest can be derived from the risk or the odds. They are:

RD - risk difference $= \pi_A - \pi_B$;
RR - relative risk $= \dfrac{\pi_A}{\pi_B}$;
RA - attributable risk $= \dfrac{\pi_A - \pi_B}{\pi_B}$; and
OR - odds ratio $= \dfrac{\frac{\pi_A}{1-\pi_A}}{\frac{\pi_B}{1-\pi_B}}$.

Section 3.1 also introduces the concept of *hazard rate* denoted by λ. Common measures of association derived from this concept are:

HD - rate difference $= \lambda_A - \lambda_B$, and
HR - rate ratio $= \dfrac{\lambda_A}{\lambda_B}$.

5.1.1 Potential Outcomes

The framework of potential outcomes (Splawa-Neyman, 1923; Holland, 1986; Rubin, 1974) adds a causal interpretation to slightly modified versions of the measures described above. Under this approach, the effect of an intervention (e.g., treatment or exposure) is defined by the comparison of what would happen to some individual or group under different treatment or exposure conditions. In order to introduce the main ideas, we require a notation for observed and latent quantities (Joffe *et al.*, 2004).

Let the subscript i denote the unit of observation or subject. We denote the observed level of the treatment whose effect we wish to measure by A_i. By Y, we denote the continuous or discrete outcome of interest. We denote by Y_i^a potential outcomes that we would see in subject i were that subject to receive a hypothetical level of treatment a. Until the time a decision is made about treatment, the outcome Y_i^a observed under any given treatment a is potentially observable. Y_i^a is

observed if subject i receives treatment level $A_i = a$. If $A_i \neq a$, Y_i^a is not observed and counterfactual reasoning is required to assign values to it. The vector $\mathbf{Y} \equiv Y^a$ denote the all potential outcomes Y^a for a given subject i.

Under the potential outcome framework, causal effects are comparisons of different potential outcomes Y_i^a for the same units i under different treatment levels a and a'. However, the outcomes Y_i^a and $Y_i^{a'}$ are not simultaneously observable in subject i and, therefore, one cannot directly compare these potential outcomes for any individual subject. Instead, one can estimate the distributions or expectations of potential outcomes in a group defined by pretreatment covariates ($\mathbf{V}$) under common assumptions that are often made and detailed next.

We denote the joint density of the potential outcomes $\mathbf{Y}$ in a subset of the data defined by covariates $\mathbf{V}$ by $f(\mathbf{Y}|\mathbf{V})$. Analogously, we denote the marginal density of the potential outcome Y^a in the subset $\mathbf{V}$ by $f(Y^a|\mathbf{V})$. The joint density is not generally estimable and the models used to estimate $f(Y^a|\mathbf{V})$ (densities) or $E(Y^a|\mathbf{V})$ (expectations) require additional assumptions.

The first common observed assumption implies no interference and the stability of treatment among units of observation. It requires that the potential outcomes in the i-th individual are independent of the treatment assignment and outcome in another unit of observation. Cox (1958) called this assumption the assumption of no interference between units. Rubin (1980) dubbed this SUTVA for the stable unit treatment value assumption, but more recently, he has adopted the term stability assumption (Rubin, 1990). Under the stability assumption, if there are two treatments, then there are only two possible outcomes for each person, one for each treatment, and the two completely exhaust the possibilities.

Just as it is not possible to observe both potential outcomes in one individual, the average causal effect in the population is also not observable. To make inferences based on this information, we need additional assumptions regarding the mechanism of assignment of treatments to the individuals. Once a treatment is assigned, then, under this model, the potential outcome that will be observed for that individual is determined, i.e., the potential outcomes $\mathbf{Y}$ do not predict treatment assignment once one accounts for baseline covariates $\mathbf{X}$. It is also assumed that there is a nonzero probability that subjects with any covariate level $\mathbf{X}$ receive treatment level a, that is,

$$p(A = a|\mathbf{X}, \mathbf{Y}) = \mathbf{p}(\mathbf{A} = \mathbf{a}|\mathbf{X}) > \mathbf{0}$$

for all a, $\mathbf{X}$, and the assumption is known as strong ignorability of treatment assignment (Rosenbaum & Rubin, 1983).

Randomization, such as by the flip of a fair coin, is an assignment mechanism that generally ensures that the treatment assigned to an individual is independent of the potential outcomes in that individual. If the treatment is allocated randomly, then under the stability assumption, the expected value if everyone received one treatment would be equal to the expected value of those who actually receive that treatment. More generally, it is assumed that within strata defined by covariates $\mathbf{X}$, groups assigned to various levels of treatment are

comparable except for the effect of treatment (Greenland & Robins, 1986). Marginal structural models (MSMs) provide a way to model these aggregate causal effects (Joffe *et al.*, 2004).

5.2 Infectious Diseases and Violation of the Stability Assumption

The development up to now has relied on the stability assumption, that is, the assumption that there is no interference between the units. However, SUTVA is often violated in infectious diseases (Halloran & Struchiner, 1995). A fundamental aspect of infectious diseases is transmission from one host to another, however, so that whether a person becomes infected depends upon who else is infected. Sir Ronald Ross called this phenomenon "dependent happenings" (Halloran & Struchiner, 1991) and it has consequences for the applicability of the model for causal inference to infectious diseases. Let's examine the model of causal inference used by Rubin and others in light of questions posed in evaluating interventions against infectious diseases, say a vaccine.

Assume that there are two people, and that person 2 would become infected through a contact with person 1 under the situation that neither is vaccinated. Assume further that if person 1 is vaccinated, he will not become infected. Then person 2 will not become infected, even if he is not vaccinated. The infection outcome in person 2 depends on the vaccination status of person 1. The assumption that the outcome in any individual is independent of the treatment assignment in other persons, the assumption of no interference between units, is violated.

The dependence in happenings in infectious diseases differs in its essence from the interference discussed by Cox (1958). The interference in agricultural experiments, for example, is a nuisance that we try to be rid of by leaving guard rows between the different plots. The contacts and mixing patterns among members of a population, however, take place even without the presence of the infectious agent. The exposure to infection provided by the other members of the population in infectious diseases, either directly or via vectors, is essential to transmission as well as for evaluating the effects of the intervention. If no one were exposed to infection, no one would become infected, all observed outcomes would be zero, and the difference between all observed outcomes would be zero. The nature of transmission itself makes it a *sine qua non* of infectious diseases. This essential difference between dependent happenings and interference must be kept in mind in further examination of the model of causal inference. The attempt to prevent the exposure to infection, or the interference, can itself be viewed as a type of cause, or manipulation, to be evaluated using a different type of causal effect, the indirect effect (Halloran & Struchiner, 1991).

Consequences of the violation of the stability assumption include the need for an expanded representation of outcomes, and the existence of different kinds of effects, such as direct and indirect effects. Effects of interest include changes in susceptibility as well as changes in infectiousness. As an alternative, we can

define the transmission probability formally as an average causal parameter of effect in a population by conditioning on exposure to infection. We address this topic next.

5.2.1 Vaccine Efficacy

Given the structure of dependent happenings in infectious diseases which breaks the commonly made SUTVA assumption, effects of a vaccine, either unconditional indirect or total effects are difficult to define formally using the model for causal inference based on potential outcomes. The assignment mechanism can influence the sampling mechanism when it determines who is exposed to infection raising problems that require further inquiry. Here, we review the role of differential exposure to infection in defining direct and indirect effects of a vaccine as a particular application in infections diseases of the model of causal inference based on potential outcomes (Halloran & Struchiner, 1995; Halloran *et al.*, 1999).

What does it mean exactly when we say that a vaccine is 80 percent efficacious? Due to the structure of dependent happenings in infectious diseases, vaccination can produce several different kinds of effects, at both the individual and the population levels, not all of them described by a single estimate. Therefore, one should make explicit the mechanism of vaccine action under consideration. For example, vaccination can induce a biologically protective response in a vaccinated individual or reduce the degree or duration of infectiousness for other individuals. Moreover, widespread vaccination in a population can reduce transmission and produce indirect effects, even in individuals who were not vaccinated. In addition, vaccination may result in change in behavior such that vaccinated people might change their rate of making contacts with potentially infectious sources altering the exposure to infection and the overall benefits of vaccination.

The biologic aspects of vaccination that are of interest or that need to be taken into account when estimating vaccine efficacy determine the necessary components of the epidemiologic study design. In planning a study to evaluate the effects of vaccination, one has to contemplate the choice of unit of observation, comparison groups, parameter of effect, and level of information required.

Vaccine efficacy and effectiveness (VE) are generally defined as one minus some measure of relative risk (RR) in the vaccinated group compared with the unvaccinated group:

$$\mathrm{VE} = 1 - \mathrm{RR}.$$

This basic expression can accommodate different dimensions involved in the interpretation of this concept. For example, the groups being compared could be composed of individuals or of populations or communities. In addition, particular study designs allow for the interpretation of VE as the direct protective effect of vaccination. Under alternative design options VE estimates how vaccination alters the infectiousness of a person who becomes infected.

The primary effect of interest of vaccination is how well it protects the vaccinated individual. Biologically, the protective immune response can reduce the probability that a person becomes infected given a specified exposure to or inoculum of an infectious agent. That is, it can reduce the *transmission probability.* If a vaccinated person becomes infected, the immune response might reduce the degree or duration of disease or the probability of dying from the disease. It may also alter the rate of disease progression. Studies with either infection or disease as outcomes are sometimes used to measure vaccine efficacy for susceptibility (VE_S), though the distinction between infection and disease should always be kept in mind.

As another example of effect of interest, a vaccinated person who becomes infected may also be less infectious to other susceptibles or be infectious for a shorter period of time. The vaccine efficacy for infectiousness (VE_I) is of interest because a vaccine that reduces infectiousness could have important public health consequences. Other effects are possible (Halloran *et al.*, 1999).

Let's consider first the problem of estimating VE_S based on information on transmission probabilities (p_{ij}), the probability that, conditional upon a contact between an infective source with covariate status i and a susceptible host with covariate status j, successful transfer and establishment of the infectious agent will occur. The transmission probability could also be defined conditional on a specified level of inoculum. A related concept is the secondary attack rate, (SAR_{ij}) defined as the proportion of susceptibles with covariate status j making contact with an infectious person of covariate status i who become infected. Let 0 and 1 denote being unvaccinated and vaccinated, respectively. Then, for example, p_{01} denotes the transmission probability per contact from an unvaccinated infective person to a vaccinated uninfected person. Let $p_{\cdot 0}$ and $p_{\cdot 1}$ denote the transmission probability to unvaccinated and vaccinated susceptibles, respectively, where the dot in the subscript can denote any vaccine status or an average across the population. Then $\mathrm{VE}_{s,p}$ based on the transmission probability or secondary attack rate is estimated from

$$\mathrm{VE}_{S,p} = 1 - \frac{p_{\cdot 1}}{p_{\cdot 0}} = 1 - \frac{\mathrm{SAR}_{\cdot 1}}{\mathrm{SAR}_{\cdot 0}} = 1 - \frac{\frac{\text{vaccinated infections}}{\text{vaccinated exposures}}}{\frac{\text{unvaccinated infections}}{\text{unvaccinated exposures}}}.$$

Estimating vaccine efficacy from the transmission probability ratios requires information on who is infectious and when, and whom they contact and how. The type of contact and the infectiousness of the infective source will determine the inoculum level per contact. If it were possible to measure the different types of contacts, then the transmission probability for each type of contact could be estimated, and the $\mathrm{VE}_{S,p}$ estimates could be stratified by type of contact. If it is not possible to measure the levels of infectiousness, the inoculum level, or the different types of contacts, then the estimates will reflect the unmeasured heterogeneities.

As a second example, let's turn our attention now to the problem of estimating vaccine effect on infectiousness, VE_I. The efficacy of a vaccine in reducing infectiousness, VE_I, can be estimated epidemiologically by comparing the per-contact

transmission probability from vaccinated people who become infected with the transmission probability from unvaccinated people who become infected. The relative risk comparison groups are defined according to the vaccination status of the infectious person contacting the susceptible person. In contrast to VE_S, which can be estimated using either conditional or unconditional parameters, VE_I can generally be estimated using only conditional measures such as the transmission probability or secondary attack rate,

$$\mathrm{VE}_{I,p} = 1 - \frac{p_{10}}{p_{00}} = 1 - \frac{\mathrm{SAR}_{10}}{\mathrm{SAR}_{00}}.$$

Information on exposure to infection is often difficult or impossible to collect. More commonly, studies are designed to estimate VE_S from events per person-time of potential rather than actual exposure or simply from the proportion of people who become infected in the vaccinated compared with the unvaccinated groups. Studies for estimating VE_I can be incorporated into those for estimating VE_S, p based on the transmission probability, if the vaccination status of the infectious person in a contact is known. The analysis can then simply stratify on the vaccination status of both the infectious and susceptible persons in the contact to get estimates of VE_S and VE_I. One can specify additional study designs for estimating specific measures of indirect, direct, total and overall effectiveness of widespread vaccination in populations based on the choice of comparison groups, the unit of observation, the choice of parameter, and the amount of information about the transmission system required for estimation.

5.3 Measures of Association for Non-observable Subsets of the Target Population

Actual study populations are often heterogeneous in biological, social or environmental characteristics relevant to the validity of field trials conducted to evaluate interventions against infectious diseases. These heterogeneities result in differences in susceptibility, exposure to infection, outcome assessment and propensity to loss to follow up. Sometimes a few of these factors can be identified and measured and are represented explicitly by the vector of covariates $\mathbf{X}$ in analytical models. Therefore, such measured differences can be controlled for in the analysis.

However, most sources of heterogeneity, or lack of comparability among study subjects remain unknown. Individuals are heterogeneous regarding their susceptibility to infection, development and duration of natural immunity, etc. They also react differently once vaccinated, and immune responses, ranging from total lack of protection to protection that is partial or complete, are found for the different vaccines. Other sources of heterogeneity include vector behavior or competency, seasonal variations in climate, spatial clustering, and age-related inoculation rates or immunity. As we saw above, in the context of estimating vaccine efficacy, the definition of treatment effects when estimating the efficacy

of control measures against infectious diseases requires the specification of covariates not observable at the time a treatment decision is made.

Under heterogeneity one has the choice to report stratum-specific or summary measures of effect, the latter representing weighted means of stratum specific measures. It is well known that both measures might differ depending on the distribution pattern of vaccine coverage among the various strata (Halloran *et al.*, 1992). The idea of stratum specific measures of effect can be taken one step further by defining treatment effects for strata of the target population characterized by variables not observable at the time a treatment decision is made (Joffe *et al.*, 2007). Among several possible frameworks for doing this, Frangakis and Rubin (2002) introduced the concept of principal stratification in causal inference. The idea was later applied in the context of assessing vaccine efficacy by Hudgens, Hoering, and Self (2003). These authors postulated a principal stratum of volunteers doomed to become infected whether randomized to placebo or vaccine and developed hypotheses tests of vaccine effect specific to this stratum. We briefly review the rationale of this approach. In order to do that we need to expand on the notation introduced to describe the model of potential outcomes above.

Let, in addition to the notation already introduced for the model of potential outcomes, $\mathbf{S}$ be the vector of potential auxiliary outcomes inducing a stratification in the target population within which one is interested in characterizing the causal effects of infectious diseases control measures. Therefore, S^a denotes the level of the auxiliary variable were a subject given treatment a.

Based on the idea of Principle Stratification, Gilbert, Bosch, and Hudgens (2003) noted that for a vaccine trial, each subject enrolled in the trial must be one of four unknowable types. This stratification of the target population is motivated by the type of immune response elicited by the vaccine. In the simplest case, the immune system when in contact with an antigen increases the concentration of circulating antibodies against this antigen in the blood. These antibodies are expected to increase the protection of the individuals against the pathogen in question. Therefore, the subgroup of the "protected" individuals consists of those who would mount an immune response if given a vaccine and thus have a lower chance of becoming infected in this case, but who would remain susceptible to becoming infected if given the placebo.

A vaccine does not always confer protection and might "cause" infections by a variety of mechanisms. Some vaccines are actually live but weakened pathogens. If the weakened pathogen is used in a weakened host, disease producing infection might result, as is the example of the attenuated yellow fever (Massad *et al.*, 2005a; Struchiner *et al.*, 2004) and polio vaccines (de Oliveira & Struchiner, 2000). As a second mechanism, vaccine induced antibodies produced by the immune system can "enhance" rather than reduce the chance a disease producing infection occurs. Additionally, vaccines might induce an auto-immune reaction which could hamper the ability of the immune system to fight infection or disease. As a third mechanism, a vaccine might induce a perception of protection and encourage risky behavior that results in more exposure to pathogens, thus

Table 5.1. Principal Strata

Principal Strata	**Potential Outcome treatment A (V=1; C=0)**	**Population Distribution**
Naturally Resistant $(S^0 = 1; S^1 = 1)$	$Y^0 = 0$ $Y^1 = 0$	θ_{00}
Harmed $(S^0 = 0; S^1 = 0)$	$Y^0 = 0$ $Y^1 = 1$	θ_{10}
Protected $(S^0 = 0; S^1 = 1)$	$Y^0 = 1$ $Y^1 = 0$	θ_{01}
Doomed $(S^0 = 0; S^1 = 0)$	$Y^0 = 1$ $Y^1 = 1$	θ_{11}

leading to an increased probability of infection. These mechanisms lead to the possibility of a second principal stratum, the subgroup of the "harmed" consisting of those who would have a higher chance of becoming infected if given the vaccine than if given the placebo.

The final two strata comprise the subgroups of the "naturally resistant" individuals and those uncapable of responding to the antigen challenge and therefore "doomed" to becoming infected since their susceptibility remains the same irrespective of their vaccination status. Let S^a denote the auxiliary variable, level of protection (high = 1, low = 0) conferred by treatment a. In general, the principal strata are not fully identified from the data, because one cannot simultaneously observe both potential auxiliary outcomes S^0 and S^1. Table 5.1 describes the potential outcomes in the four possible strata.

In this context, the strata comprised by the subgroups of "naturally resistant", "harmed", and "doomed" can be be regarded as non-informative in what concerns the estimation of vaccine efficacy, and we may be interested in the effect of this vaccine among subjects that could respond to the vaccination, i.e., the "protected" stratum. If accepting this rationale, we should concentrate our attention on comparing the proportion in this subgroup that would have become infected if not vaccinated, $p(Y^1 = 1|\text{protected})$, and the proportion in the same group who would have become infected if not vaccinated, $p(Y^0 = 1|\text{protected})$.

5.3.1 Randomization and Baseline Transmission

We have argued that evaluation of intervention measures against infectious diseases shares the same general principles of validity with epidemiologic causal inference, i.e., the process of drawing inferences from epidemiologic data aiming at the identification of causes of diseases. Judicious exercise of these principles indicates that, for meaningful interpretation, measures of efficacy of disease control interventions require definitions based upon arguments conditional on the amount of exposure to infection, and specification of the initial and final states in which one believes the effect of interest takes place (Struchiner *et al.*, 1994; Struchiner & Halloran, 2007). The paradigm of the randomized, double-blinded, placebo controlled field trials is usually regarded as the gold standard in

order to achieve this end. This view is supported by the fact that most sources of heterogeneity, or lack of comparability among study subjects remain unknown and cannot be controlled for in the analysis. Randomization and double-blinding are two strategies designed to distribute these unmeasured heterogeneities approximately equally between the comparison groups. We review some misconceptions about randomization and what is actually achieved by this treatment assignment mechanism in the context of infectious diseases epidemiology. Our primary concerns here are concepts of study design and interpretation of the efficacy estimates.

In any epidemiologic study conceived to assess the effect of a certain treatment on an outcome of interest, comparison groups must be, in all material respects, alike except for their treatment status. The statement can be interpreted as if the same results would be expected if treatment status had been exchanged between the two groups. In other words, exchangeability assures comparability between treatment groups and is an important requirement for valid epidemiologic inference on the effects or causal role of the treatment of interest. Conversely, inherent differences in risk between treated and untreated individuals imply lack of comparability between treatment groups which could potentially bias the estimation of the effects of said treatment on disease risk, a condition known as confounding in epidemiology.

When the outcomes under study are independent, exchangeability guarantees that it would be possible to describe the occurrence of the outcome of interest among the treated individuals, had they not been treated, from the observed data on the untreated. The latter sentence describes only partially the concept of exchangeability but is sufficient to assure identifiability of causative parameters (Greenland & Robins, 1986) in chronic disease epidemiology, where happenings are independent. Complete exchangeability must also guarantee that it would be possible to describe the occurrence of the outcome of interest among the untreated individuals, had they been under treatment, from the observed data on the treatment group.

In the context of vaccine evaluation, the requirement that the vaccinated and unvaccinated be exchangeable was noted as early as 1915 by Greenwood and Yule in their criteria for valid efficacy or effectiveness studies (Greenwood & Yule, 1915). Field trials that comply with this requirement are believed to yield unconfounded estimates of vaccine efficacy. This belief stems from the analogy one could make between a vaccine and the treatment factor in epidemiologic studies. However, direct application to vaccine field trials of the concepts briefly described in the previous two paragraphs is not possible without further qualification.

Halloran and Struchiner (1995) separate evaluation of vaccines on the one hand, conditional on exposure to infection, and on the other hand, not conditioning on exposure to infection. Thus, we must be aware that the concept of vaccine efficacy is not unique and be explicit about our intents. In addition, since exchangeability within both pairs of comparison groups does not necessarily hold simultaneously, field trials that yield valid measures of vaccine

efficacy of one kind can potentially lead to biased estimates of efficacy of a different kind.

By the same token, partial and complete exchangeability must also be further qualified. Partial exchangeability is expressed as the counterfactual reasoning requiring that the ideal unvaccinated control group describe the potential outcome in the vaccinated group in the absence of vaccination. In actual field trials, however, due to the indirect protection of the unvaccinated group which is brought about by the presence of the individuals who became immune by the vaccine, a mechanism known as herd immunity (Fine, 1993), even partial exchangeability might not be achieved giving rise to different concepts of vaccine efficacy (Halloran & Struchiner, 1991; Halloran *et al.*, 1991). Complete exchangeability requires, in addition, that the outcome observed in the vaccinated subjects describe the potential outcome in the unvaccinated group had it been vaccinated, or, phrasing it in a different way, if vaccination states were exchanged, the value observed for the incidence among vaccinated and unvaccinated subjects would have been the same. Again, due to the indirect effects of a vaccine, the latter statement gives rise to different interpretations. This translates into different concepts of measures of vaccine efficacy that are discussed below.

The principle of exchangeability in actual vaccine field trials thus involves at least two dimensions: (i) where in the sequence of pathogenic processes comparisons between vaccinated and unvaccinated is being sought; and (ii) how we interpret the counterfactual reasoning implicit in the principle of exchangeability. Dimension (i) leads to the concept of biological efficacy and dimension (ii) to the concepts of direct and indirect effects of a vaccine. It then becomes a challenge to epidemiologists to design studies where comparability is ensured and to data analysts to develop methods to control for departures from the exchangeability principle.

Perhaps, the single most difficult study design requirement an epidemiologist must fulfill is to assure that effective exposure to the infectious challenge must be identical in the case of the treated and untreated study participants. While the necessity of comparability of personal attributes in the two groups is common to epidemiologic studies in chronic and infectious diseases, the requirement of comparability of exposure to infection is specific to epidemiologic studies in infectious diseases and more subtle to fulfill.

Notice that exposure to infection might be the same within any trial but the study population participating in trials taking place at other locations and time may be subject to different baseline inoculation rates rendering comparison of measures of vaccine efficacy across geographic locations or time more difficult. Equal amount of exposure to infection in the vaccinated and unvaccinated groups is an important requirement for the assumption of exchangeability to hold. Valid comparisons, however, must be further qualified by making explicit reference to the underlying level of exposure to infection in order to be appropriately interpreted. It is well known that the background level of transmission is a function of seasonal factors (density of mosquitoes, climate, etc.), other concomitant

Table 5.2. Expanded potential outcome framework presented in table 1 to accommodate infectious challenge S^a

Principal Strata	**Challenge**	**A (V = 1 / C = 0)**	**Population Distribution**
Naturally Resistant	$S^0 = 0$ $S^1 = 0$	$Y^0 = 0$ $Y^1 = 0$	θ_{000}
	$S^0 = 1$ $S^1 = 1$	$Y^0 = 0$ $Y^1 = 0$	θ_{100}
Harmed	$S^0 = 0$ $S^1 = 0$	$Y^0 = 0$ $Y^1 = 1$	θ_{010}
	$S^0 = 1$ $S^1 = 1$	$Y^0 = 0$ $Y^1 = 0$	θ_{110}
Protected	$S^0 = 0$ $S^1 = 0$	$Y^0 = 0$ $Y^1 = 0$	θ_{001}
	$S^0 = 1$ $S^1 = 1$	$Y^0 = 1$ $Y^1 = 0$	θ_{101}
Doomed	$S^0 = 0$ $S^1 = 0$	$Y^0 = 0$ $Y^1 = 0$	θ_{011}
	$S^0 = 1$ $S^1 = 1$	$Y^0 = 1$ $Y^1 = 1$	θ_{011}

control measures besides the vaccine, and changes in transmission brought by the vaccine itself.

In table 5.2, we expand the previous potential outcome framework presented in table 5.1 to accommodate this new dimension, i.e., infectious challenge. Notice that in table 5.2 we treat the infectious challenge as a binary discrete variable. In this case, one could argue that the informative principal stratum of interest is given by the "protected" subgroup that were also challenged by natural exposure to infection. Struchiner *et al.* (1994) and Struchiner and Halloran (2007) discuss the cases in which S^a can be continuous or integer valued random variable.

Data on actual exposure to infection are scarce, but could indirectly be approximated by surrogate variables that are easier to collect. Variables that could help in controlling for differences in exposure to infection include time since arrival in the endemic area, reported number of previous morbidity episodes, clinical signs (splenomegaly), and, possibly, serology at the start of follow up period. The same objective could be achieved by taking advantage of the observed clustering of cases within households in the same village. It seems that individuals living in the same household are more homogeneously exposed to infection, therefore, trials that compare vaccinated to unvaccinated persons matched on household are less prone to bias from differences in exposure to infection. Finally, one could reconstitute exchangeability by the appropriate use of mathematical models. Data collected prior to the intervention helps to project the baseline transmission level into the post intervention period, allowing for the construction of a comparison standard estimating what exposure to infection would have been in the absence of intervention.

In summary, in order to assure validity of comparison one must guarantee, through appropriate mechanisms of vaccine assignment, exchangeability according to the various aspects of the transmission process, ie., the infectivity of the infectious source, the susceptibility of the susceptible, and the type of contact of the susceptible with the infectious source.

5.4 Fuzzy Logic and Risk Estimators

In the previous sections we introduced a causal inference approach to the study of infections diseases epidemiology based on a model of potential outcomes. In addition to a counterfactual reasoning, this model also makes use of the notions of risk, based on conditional probability, and causal effects, based on the comparison of epidemiologic risk among treatment groups. This approach made explicit the main sources of uncertainties in epidemiologic reasoning. The presence of sources of uncertainties in this formulation provides the opportunity to address uncertainty, imprecision and vagueness in epidemiology from a fuzzy logic perspective. We identify two key steps in this direction: the formulation of fuzzy conditional probabilities and the definition of fuzzy measures of association.

5.4.1 Fuzzy Conditional Probability

The following paragraphs, adapted from *Example 1* in Coletti and Scozzafava (2004), serves as motivation for our discussion. When faced with the statement E={"Mary is sick"}, it is natural to think that the statement was made in the presence of some information about Mary's morbidity status. This information allows one to refer to a suitable membership function of the fuzzy subset of "diseased people". The suitability of interpreting the statement about Mary's health condition as an event, and the values of the membership function corresponding to the relevant fuzzy set as probabilities is usually challenged in the literature. The description of one's health condition, including Mary's, can be regarded as vague and "vagueness is looked on as referring to the intended meaning (i.e., a sort of *linguistic* uncertainty) and not as an uncertainty about facts" (Coletti & Scozzafava, 2004), the latter being the very subject of probabilities but not the former.

Suppose now that Mary's condition can be defined by a laboratory exam and we describe the linguistic uncertainty via a membership function, say, equal to 0 (healthy) for values of this laboratory exam below a certain cutoff point z^1, equal to 1 (diseased) for values of the exam above the cutoff point $z_2 > z_1$, and increasing from 0 to 1 in the interval from z^1 to z_2. The assignment of such a membership function is pretty general, the only restriction being the range of values from 0 to 1.

On the other hand, the assignment of a subjective probability p to the statement E={"Mary is sick"} has to obey certain rules such as the axiom of additivity. Implicitly, and even in the absence of any other justification, the complementary statement E^c="Mary is not sick" has also been assigned the subjective

probability $1-p$ in order to fulfill the consistency argument represented by the additivity rule. Notice, however, that the assignment of a subjective probability to the conditional statement $E|A_x = \{$"Mary is sick"|"Mary's lab exam is x"$\}$ does not imply any secondary assignment to the conditional statement $E|A_x^c$ where A_x^c denotes the complementary statement $\{$"Mary's lab exam is not x"$\}$.

Coletti and Scozzafava (2004) propose, then, to identify the values of the membership function with suitable conditional probabilities, which, in this case could be expressed as:

$$H_0 = \text{Mary's lab exam is} < z_1$$
$$H_1 = \text{Mary's lab exam is} > z_2$$

and, further assume that $P(E|H_0) = 0$ and $P(E|H_1) = 1$.

5.4.2 Fuzzy Measures of Association

Suppose now, as a second complementary approach to the fuzzification of the ideas expressed in this chapter, that individuals are assumed to be exposed to some risk factor according to a certain fuzzy set membership functions and their response is also categorized according to other fuzzy set membership functions. Risk analysis is then performed by applying maximum likelihood and fuzzy set theory. These procedures allow us to calculate *Fuzzy Relative Risks* and *Fuzzy Odds Ratios* under individual heterogeneity (Massad *et al.*, 2003), which are more realistic estimators of risk assessment than their classical crisp counterparts.

We begin by defining a *fuzzy risk* estimator. In classical epidemiology, one kind of risk estimator is the so-called *risk ratio*, RR. Under *crisp logic* setting, the risk ratio is defined as the ratio of the conditional probability of developing a disease given one is exposed to a certain cause, $p(D|E)$, to the conditional probability of developing the disease given one is not exposed to the cause, $p(D|\ \bar{E})$, such as (see chapter 3):

$$RR = \frac{p(D|E)}{p(D|\ \bar{E})}. \tag{5.1}$$

However, if under crisp logic setting the risk ratio is defined in terms of conditional probabilities, under fuzzy logic setting it must be defined in terms of *conditional possibilities*. A *possibility* distribution function, r, associated with a fuzzy subset A is numerically equal to its grade of membership function μ_A (Zadeh, 1978), such that:

$$r(x) = \mu_A(x) \text{ for all } x \in X$$

with measure π given by:

$$\pi(A) = \max_{x \in A} r(x). \tag{5.2}$$

We can now define a *Fuzzy Risk Ratio*, FRR, as the ratio of the conditional possibility of developing a given disease severity, d, given that one is exposed to

Table 5.3. Heterogeneity in the risk classes

		$\bar{E}$	
		$D \mid \bar{E}$	$\bar{D} \mid \bar{E}$
E	$D \mid E$	*doomed*	*at risk*
	$\bar{D} \mid E$	*protected*	*resistant*

a certain level of a causal factor, e, to the conditional possibility of developing disease severity, d, given that one is not exposed to the causal factor, $\bar{e}$.

One basic difference between crisp and fuzzy setting approaches is the fact that in fuzzy logic the intersection of a given fuzzy set, A and its complement, A^c is different from the empty set, such that

$$A \cap A^c \neq \emptyset. \tag{5.3}$$

We, therefore, have to consider the uncertainty generated by the fuzziness due to the possibility that the sets of *exposed* and *non-exposed* intersect and create a subset with diseased and non-diseased individuals. The conditional possibility of developing the disease, given that one belongs to the subset defined by the intersection $E \cap E^c$ is given by:

$$Poss(D \,|E \cap E^c). \tag{5.4}$$

The *fuzzy risk ratio* estimator, FRR, should then be defined in terms of *conditional possibilities*, and it is expected that it should be proportional to the ratio between the conditional possibility of developing disease given that one is exposed to a suspected factor, $Poss(D \,|E)$, to the conditional possibility of developing disease given that one is not exposed to that factor, $Poss(D \,|\bar{E})$, such that:

$$FRR \propto \frac{Poss(D \,|E)}{Poss(D \,|\bar{E})}. \tag{5.5}$$

Let us now examine a more general situation, due to Greenland (1987), who considered the theoretical possibility of four types of individuals (see section 5.3); those *doomed*, who would develop a certain disease, independently of being exposed or not to the suspected cause; those *resistant*, who would never develop the disease, either exposed or not; those *protected*, for whom the exposure is a protection factor, that is, they develop the disease if *not* exposed to the suspected cause; and those *at risk*, who develop the disease only if exposed to the suspected cause. This classification assumes a high degree of heterogeneity in the population and involves several uncertainties in the definition of each class. Table 5.3 shows the four types of individuals and their respective risk categories.

A fuzzy logic approach should then consider the degree of membership of individuals to each of the above sets. The next step should be the definition of the *conditional possibilities* associated to each of the fuzzy subsets. A *Fuzzy*

Odds Ratio (*FOR*) estimator could then be defined as the ratio of the following conditional possibilities:

$$FOR = \frac{Poss(\bar{D} \mid \bar{E}) \wedge Poss(D|E)}{Poss(D|\ \bar{E}) \wedge Poss(\bar{D}\ |E)}. \tag{5.6}$$

where the symbol $\wedge$ is the *min* operator.

The rationale behind this expression is the following: individuals who do not develop the disease if not exposed *and* develop the disease if exposed, are classified as individuals at risk. Those who develop the disease if not-exposed *and* do not develop the disease if exposed are classified as protected. Hence, the ratio between individuals at risk and protected should provide a good measure of association. Considering the uncertainties related to the classification criteria and the consequent heterogeneity in the population, a fuzzy approach generating the conditional possibilities above would define the association between cause and effect depending on the ratio expressed by equation (5.6) being greater (positive association), or lower (negative association) than one.

We simulated the theoretical scenarios above in order to assess the fuzzy approach. The characteristic function for the *exposed* fuzzy subset, $\mu_E(x)$, was generated as a uniform distribution. The characteristic function for the *diseased* fuzzy subset, $\mu_D(x)$, was associated to the exposed by the following logistic model:

$$\mu_D(x) = \frac{\exp(\beta\mu_E(x))}{1 + \exp(\beta\mu_E(x))}, \tag{5.7}$$

where β is a parameter whose signal and magnitude indicate the direction and intensity of the association. The conditional possibilities were defined according

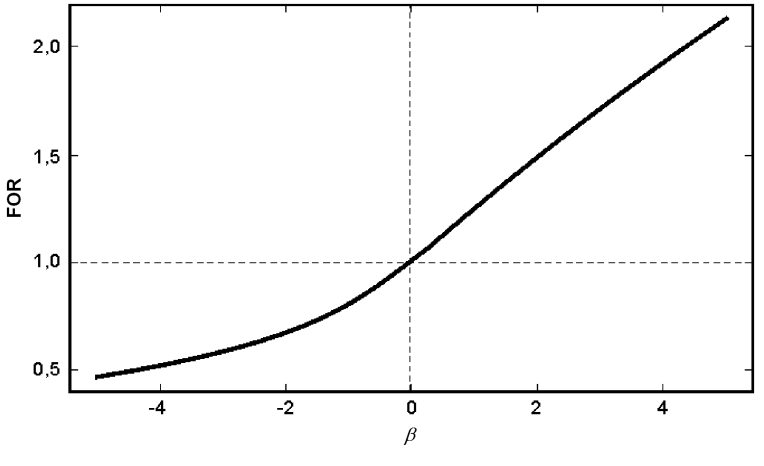

Fig. 5.1. Simulation of the *fuzzy odds ratio* (FOR) according to equation (5.6), for each value of β (Massad *et al.*, 2003)

to the rule $Poss(\bar{D} \mid \bar{E}) = \max_{x \in X}[\min(\mu_{\bar{D}}(x), \mu_{\bar{E}}(x))]$, where the operator min was applied at individual level and the operator max was applied at the populational level. Figure 5.1 shows the results of the simulation.

It can be noted from figure that when β is lower than zero the fuzzy odds ratio, *FOR*, is lower than one, indicating a *protecting* effect of the exposure. On the other hand, when β is greater than zero the FOR is greater than one, indicating a *risk* effect of the exposure. When β is equal to zero, the FOR is equal to one, indicating the absence of association between exposure and disease. It can also be noted from figure 5.1 that the FOR is non-linearly related to the association parameter β (Massad *et al.*, 2003).

The results of the simulations above have demonstrated a good plausibility of the fuzzy logic approach since they confirm what should be expected to find under the conditions imposed by the simulations assumptions. We could even conjecture that the degree of nonlinearities and other relations between the first derivative of the function presented in figure 5.1 are, to a certain level, related to the uncertainties in the process of classifying individuals in the several subsets of the model. However, future research is still needed to clarify this point.

6 Fuzzy Decision Making in Public Health Strategies

Making decision is one of the most fundamental activities of human beings (Klir & Yuan, 1995; Yager & Filev, 1994; Zadeh, 1973). This is particularly true in Public Health where decisions usually have relevance for millions of people. In the field of vaccination strategies design, decision making concerning the target population for the immunization program, the proportion of susceptibles to be vaccinated, the optimal age to immunize children and the nature of the strategy, e.g. selective or indiscriminate, are examples of the variables to be optimized, subject to a set of constraints. As an example, we present in this chapter a fuzzy model to decision making applied to the design of the vaccination campaign against measles in São Paulo, Brazil (Massad *et al.*, 1999)

Decision making comprises the study of how decisions are actually made and how they can be made better or more successfully (Klir & Yuan, 1995). Models of human decision making generally include the aggregation criteria or criteria of constraints (Zimmermann, 1996). For the case that criteria and/or constraints cannot be modeled crisply but as fuzzy sets a decision has been defined by Bellman and Zadeh (1970) as the intersection of fuzzy sets representing either objectives or constraints. The grade of membership of an object in the intersection of two fuzzy sets, that is, the "fuzzy set decision" was determined by the use of both the *min* operator or the *product* operator (Zimmermann, 1996).

While decision making under conditions of risk have been modeled by probabilistic decision theories and game theories, fuzzy decision theories attempt to deal with vagueness and monospecificity inherent in human formulation of preferences, constraints, and goals (Klir & Yuan, 1995).

In the first paper on fuzzy decision making Bellman and Zadeh (1970) suggest a fuzzy model of decision making in which relevant goals and constraints are expressed in terms of fuzzy sets, and a decision is determined by an appropriate aggregation of these fuzzy sets. The decision models have the following components (Klir & Yuan, 1995):

- a set A of *possible actions;*
- a set of *goals*, $G_i (i \in \mathbf{N})$, each of which is expressed in terms of a fuzzy set defined on A;

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 97–110, 2008.
springerlink.com

- a set of *constraints,* $C_j (j \in \mathbf{M})$, each of which is also expressed in terms of a fuzzy set defined on A.

The fuzzy set of decision, D, is that which simultaneously satisfies the given goals G_i and constraints C_j, and is:

$$D(a) = \min \left[\inf_{i \in \mathbf{N}} G_i(a), \inf_{j \in \mathbf{M}} C_j(a) \right] \tag{6.1}$$

for all $a \in A$.

6.1 Designing a Vaccination Strategy

Let us assume that the objective of a vaccination campaign is the reduction of the incidence of an infection like measles in children below 14 years of age, the age interval where viral infections are most likely to be circulating. This assumption is based on previous works which demonstrated that the force of infection of the measles virus has a strong age-dependence, peaking around 2 years of age in the absence of vaccination (Anderson & May, 1991). Therefore, in spite of the high proportion of cases in the age interval between 20 and 39 years of age, the highest incidence rate (normalized per 100,000 inhabitants) observed during the epidemic occurred in children below 5 years old. In addition, contact patterns suggest that adult cases are the product of infective contacts of susceptible individuals in that age interval with children below 14 years old (Massad *et al.*, 1994b), the target age interval of the vaccination campaign. All the subsequent analysis in this work are based on the assumptions above.

We begin by considering 8 possible vaccination strategies, composed by combinations of *Selective* vaccination,S_i, meaning vaccinating only children without vaccination record in the past, and *Indiscriminate* vaccination, I_j, that is, vaccinating children irrespective of previous immunization history (i and j stands for the age intervals). Besides, we considered the use of *Mobile Units*, *M.U.*, meaning those vaccination sites that are not part of the Primary Care Network, as opposed to *Fixed Units, F.U.*, those belonging to the network. Table 6.1 shows the various vaccination strategies considered.

The number of children, as well as the estimated proportion and number of susceptible children (assuming the seroepidemiological profile of 1994 and the drop in the routine measles vaccine coverage discussed above) in each age interval of São Paulo State is shown in table 6.2.

The last column of table 6.2 is the maximum theoretical number of children to be vaccinated in each age interval in order to stop the progression of the current epidemics. The optimal strategy, therefore, would be that which would maximize the number of susceptible children vaccinated in the target age interval, without wasting resources by over-vaccinating children in any specific age interval.

The next step was to invite a number of experts from the Health Secretary of São Paulo with great experience in vaccination campaigns in order to provide a

Table 6.1. Possible vaccination strategies (modified from Massad *et al.*, 1999)

Strategy	Age intervals and immunization history	Units Type
1	S_{9m-6y} and I_{6y-14y}	M.U.+F.U.
2	S_{9m-6y} and I_{6y-14y}	F.U.
3	S_{9m-14y}	M.U.+F.U.
4	S_{6y-14y} and I_{9m-6y}	M.U.+F.U.
5	I_{9m-14y}	M.U.+F.U.
6	S_{9m-6y}	F.U.
7	S_{9m-6y}	M.U.+F.U.
8	I_{9m-6y}	M.U.+F.U.

Table 6.2. Number, proportion of susceptible and number of susceptible children in the target age-interval (modified from Massad *el al.*, 1999)

Age	Number of* children	Proportion of+ susceptible	Number of susceptible
9m	49,500	0.65	32,175
10m	49,500	0.50	24,750
11m	49,500	0.50	24,750
12m	49,500	0.50	24,750
1-2y	640,609	0.10	64,061
3-5y	2,515,711	0.05	125,786
6-14y	5,920,000	0.05	296,000
Total	9,274,331	-	592,272

* Estimated from official data.
\+ Estimated by dynamical modeling (Massad *et al.*, 1994b).

scale of efficacy and/or constraints of each of the possible strategies considered. The variables chosen by this experts team were:

- *compliance* by the population, that is, the proportion of the target population expected to attend the campaign convocation of each possible strategy;
- *human resources*, a relative scale of the staff required (including the training) for the implementation of each possible strategy;
- *transportation*, a relative scale of the difficulties in transport of people and material of each possible strategy;
- *communication*, a relative scale of the difficulties in explain to the population each possible campaign.

The minimum value of each of the variables will be that which determine the success of the strategy. The result of such a consultation to the experts is presented in table 6.3.

Values provided by the experts can be considered either as a proportion of expected success of each strategy or as degrees of membership to the fuzzy sets of

Table 6.3. Variables determinants of strategy success (Massad *el al.*, 1999)

Strategy	**Compliance**	**Human Resources**	**Transp.**	**Communic.**	min
1	0.30	0.30	0.20	0.30	0.20
2	0.45	0.60	1.00	0.50	0.45
3	0.70	0.50	0.30	0.40	0.30
4	0.40	0.40	0.30	0.40	0.30
5	0.80	0.20	0.20	0.80	0.20
6	0.60	1.00	1.00	0.70	0.60
7	0.50	0.60	0.60	0.60	0.50
8	1.00	0.70	0.40	1.00	0.40

successful strategies. In both views the *min* operator is the one which determine the expected results of each strategy. In addition, the *max* operator could be applied in this stage of the analysis if we consider the variables presented in table 6.3 as the only constraint of the strategies. According to this method, the strategy which maximizes the success of the campaign would be the strategy number 6.

The *min* values of the variables presented in table 6.3 allowed us to estimated the expected number of children, in each age class, that would be vaccinated in each of the possible strategies. So, for instance, strategy number one has as limitation the transport of people and materials and would, therefore, cover only 20% of the target population. As that strategy proposed to vaccinate children selectively from 9 months to 6 years of age and indiscriminately from 6 to 14 years of age, only 20% of the susceptibles below 6 years and 20% of all children from 6 to 14 years old would receive the vaccine. The minimum square of the difference between the number of children desired to receive the vaccine and the number of children that the strategy would actually vaccinate in each age class should determine the efficacy of each possible strategy, according to the definition of optimal strategy, as presented above.

A normalized scale of the efficacy of each strategy is shown in table 6.4. This was obtained by assuming that the most efficacious strategy is the one with the minimum square difference, assigned value 1. The others are obtained as a relative scale basing on multiples of the minimum square difference. Table 6.4 shows also the result of the economic costs of each strategy. This was calculated assuming a unit cost of US$0.25 for the single measles vaccine, US$1.40 for the measles-mumps-rubella (MMR) vaccine (applied only in children older than one year of age) and a unit cost of US$0.75 for the application of the vaccines. So, the economic cost of each strategy is obtained by the sum of the vaccine and application unit costs times the total number of doses of each vaccine used (measles and MMR).

The next step in the analysis is to compare the two constraints to the success of each strategy, namely, those relative to the technical constraints (*adhesion, human resources, transportation* and *communication*) and those relative to costs.

Table 6.4. A comparative scale of relative efficacies and economic costs for each strategy (Massad *et al.*, 1999)

Strategy	Number of vaccinated	Relative efficacy	Economic costs (US$)	Relative costs
1	1,243,254	0.049	3,178,223	0.533
2	2,797,322	0.098	5,959,168	1.000
3	177,682	1.000	414,359	0.070
4	1,095,099	0.127	2,743,384	0.460
5	1,854,866	0.045	4,730,907	0.794
6	177,763	0.770	308,758	0.052
7	148,136	0.761	370,509	0.062
8	1,341,732	0.147	3,352,374	0.563

Table 6.5. Degree of memberships of technical and costs constraints for each strategy (Massad *et al.*, 1999)

Strategy	Technical constraints	Costs# constraints	min
1	0.20	0.467	0.20
2	0.45	0.000	0.00
3	0.30	0.930	0.30
4	0.30	0.540	0.30
5	0.20	0.206	0.20
6	0.60	0.948	0.60
7	0.50	0.938	0.50
8	0.40	0.437	0.40

Complement of column 5 (*Relative costs*) of table 6.4

For this we took the minimum between the minimum of the variables presented by the experts (last column of table 6.3) and the complement to the relative costs scale (1-relative cost), so that both scales are in the same constraint direction, such that their minimum values represent the maximum constraint, as shown in table 6.5:

Now we have all the components of the decision model:

- a set A of *possible actions*: the eight possible strategies;
- a set of *goals*, G_i ($i \in \mathbf{N}$) defined on A: the relative efficacy of each possible strategy (third column of table 6.4) [1]; and
- a set of *constraints* C_j ($j \in \mathbf{M}$), defined on A: the minimum between the technical and costs constraints (last column of table 6.5).

[1] Remarking that by "*goal*" (this is the jargon in fuzzy optimal control theory) we mean the achievable efficacy of each possible strategy and not the major goal of controlling the epidemic.

Table 6.6. Fuzzy decision setting (Massad *et al.*, 1999)

Strategy	$G_i(a)$	$C_j(a)$	$D(a)$
1	0.049	0.200	0.049
2	0.098	0.000	0.000
3	1.000	0.300	0.300
4	0.127	0.300	0.127
5	0.045	0.200	0.045
6	0.770	0.600	0.600
7	0.761	0.500	0.500
8	0.147	0.400	0.147

The fuzzy decision, D, that simultaneously satisfies the given goals G_i and constraints C_j, is then:

$$D(a) = \min\left[G_i(a), C_j(a)\right] \tag{6.2}$$

for all $a \in A$, that is:

Therefore, the strategy that has the maximum degree of membership in the set of decision is strategy number 6, which selectively vaccinate children aged from 9 months to 6 years, using only Fixed Units of the health system. This strategy was then recommended to São Paulo public health authorities.

6.2 The Measles Epidemic in São Paulo

In São Paulo State, routine measles vaccination started in 1973. In spite of this, recurrent epidemics continue to occur until 1987, when the first mass vaccination campaign against measles was carried out, lessening the average incidence rate to something around 0.1 per 100,000 inhabitants.

By the end of September, 1996, the number of measles cases notified to São Paulo health authorities started to raise, interrupting a stability verified since the last major epidemic, in 1987. After March, 1997, the number of new cases started an exponential trend, characterizing the beginning of a new epidemic, which reached a total of 23,915 confirmed cases after one year, with 23 deaths. Regarding the age profile of the epidemic, it is noteworthy that 47% of the cases occurred in young adults, aged 20-29 years. The second age interval in number of cases, 15%, was that of children bellow one year old. However, the highest incidence rate, normalized per 100,000 inhabitants occurred among that latter age class. In what follows we briefly describe this episode, presented in details by Massad *et al.* (1999).

Table 6.7 describes the age profile of the epidemic, expressed as annual incidence rates, normalized by 100,000 inhabitants:

As can be seen from table 6.7, the highest incidence rates occurred in infants below one year of age, seconded by young adults in the age interval which corresponds to the expected age adults have greatest contact with young children.

Table 6.7. Age-related incidence rates per 100,000 inhabitants (Massad et al., 1999)

Age (years)	São Paulo city	State countryside	Total
< 1	871.50	94.17	482.84
1-4	115.99	15.32	65.65
5-9	61.21	13.13	37.17
10-14	36.17	5.93	21.05
15-19	67.27	11.34	39.31
20-29	314.30	29.85	172.08
30-44	56.52	7.54	32.03

Those adults belong to the reproductive age stratus and probably represent the parents of the children under the highest attack rates.

Figure 6.1 shows the epidemic wave in São Paulo State (bold continuous line), in the interior of the State (broken line) and in the City of São Paulo (dashed line), during the year of 1997.

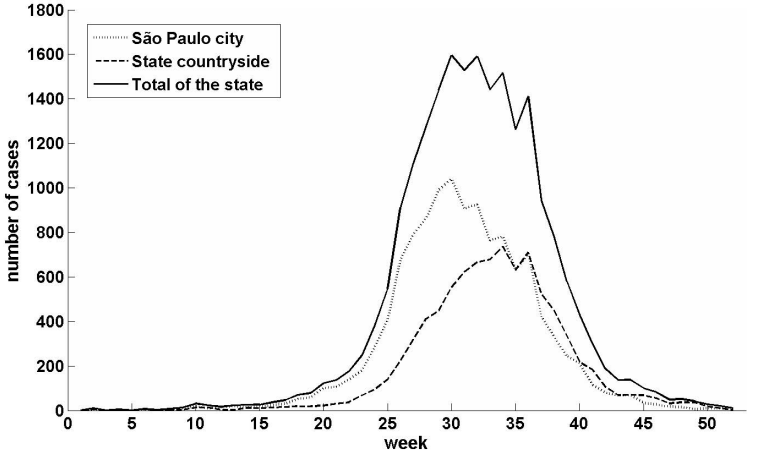

Fig. 6.1. Epidemic wave of measles in São Paulo, Brazil, in 1997. The two vertical doted lines mark the moments of the two campaigns (Massad *et al.*, 1999).

6.3 The Impact of the Vaccination

Health impact assessment (HIA) is a developing approach that assesses the health impacts of a proposal on a population, and produces a practical set of recommendations to inform the decision-making process of the proposal. The purpose is to influence decision makers to increase positive health impacts of a proposal and decrease any identified negative impacts (Quigley & Taylor, 2004; Health Development Agency - UK, 2002). It is not an academic exercise. HIA

aims to provide a practical public health approach that can be used to address health concerns about a proposal and to reduce health inequalities (Department of Health - UK, 1999).

6.3.1 Forecasting and Projection Models

As mentioned in chapter 3, three major aims of mathematical models in epidemiology can be identified: the first centers on the need for scientific understanding and precision in the expression of current theories and concepts; a second aim, linked to the first, is the role of theory in identifying areas in which better epidemiological data is required to refine prediction and improve understanding; and the third, and in many instances, the most difficult objective is that of prediction (Anderson, 1988). In addition to these three aims of modeling we propose a fourth objective: the generation of testable hypotheses by providing a theoretical framework on which plausible scenarios can be simulated in a computer environment (in silicon experiments).

Prediction in general science can be divided into two components: forecasting and projections (Keyfitz, 1972). A forecast is an attempt to predict what will happen. A projection is an attempt to describe what would happen, given certain hypotheses (Caswell, 2000). Among the tools available to the modern epidemiologists for both forecasting and projection are the mathematical (or dynamical) models, which, when well structured, can provide predictive capacity to the public health professional, helping in the design, and assessment of the impact of control strategies (Amaku *et al.*, 2003; Burattini *et al.*, 1998; Massad *et al.*, 1995; Burattini *et al.*, 1993). For instance, by projecting what would happen with a given population if individuals were not vaccinated, it is possible to quantify the relative impact of a specific vaccination program.

In what follows we illustrate the application of a projective model do the Severe Acute Respiratory Syndrome (SARS), describe in details in Massad *et al.* (2005b).

Severe Acute Respiratory Syndrome (SARS) is a recently discovered infectious disease with high potential for transmission (WER, 2003), transmitted by droplet and direct contact and caused by a new strain of corona virus (CDC, 2003). On 5 July 2003, World Health Organization (WHO, 2003) announced that the last known chain of human-to-human transmission of the SARS corona virus had been broken. A cumulative number of 8422 cases have been reported worldwide to the WHO, with 908 deaths, as of August, 2003.

In the end of 2002, reports from China suggested that a new, highly contagious, and very severe atypical pneumonia of unknown cause was occurring in the Guangdong province. As it reached southeastern Asian countries, the condition appeared to be particularly prevalent among health care workers and their household members. In response to that threat, on March 13, 2003, WHO issued a global alert, for the first time on more than a decade, and instituted worldwide surveillance. On March 27, scientists in the WHO laboratory network reported major progress in the identification of the causative agent, a new member of the corona virus family.

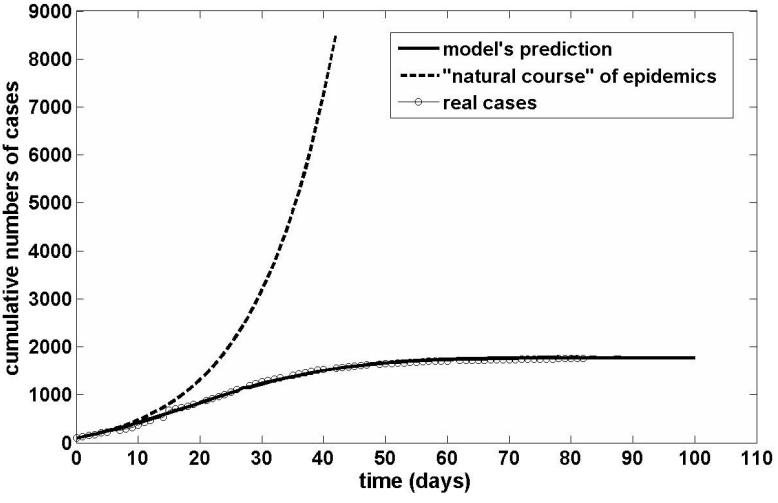

Fig. 6.2. The number of SARS real cases in Hong Kong, the model prediction and the natural course of the epidemics (Massad *et al.*, 2005b)

By that time, SARS has already become a global health hazard, and its high infectivity was alarming. Early recognition, prompt isolation, and appropriate precaution measures were considered to be key factors in combating this infection (Lee *et al.*, 2003). In figure 6.2 we show the simulation for the Hong Kong community.

The model mimics real data with good accuracy when considering adoption of control measures. The model's prediction demonstrated an epidemic that is, by far, milder than expected without control measures. The model projects that, in the absence of control, the final number of cases would be 320,000 in Hong Kong. In contrast, with control measures, which reduce the contact rate to about 25% of its initial value, the expected final number of cases is reduced to 1,778. In fact, the stability level predicted by the model was indeed attained in Hong Kong by the end of the outbreaks.

6.3.2 The Case of the Measles Epidemic in São Paulo

In June 21, 1997, the proposed vaccination strategy was implemented in the State of São Paulo. A total of 213,084 doses were applied to children between 9 months and 6 years of age. This figures represents a coverage of 6.5% of the entire population of the State of São Paulo in the target age interval. In the Metropolitan Region of São Paulo city, 7.5% of the entire population in the target age interval was vaccinated. In the interior of the State 5.1% of the population in the target age interval was vaccinated. There are no official data on the efficacy of the selection process, that is, it is not known whether the small proportion of children vaccinated were those previously unvaccinated or not.

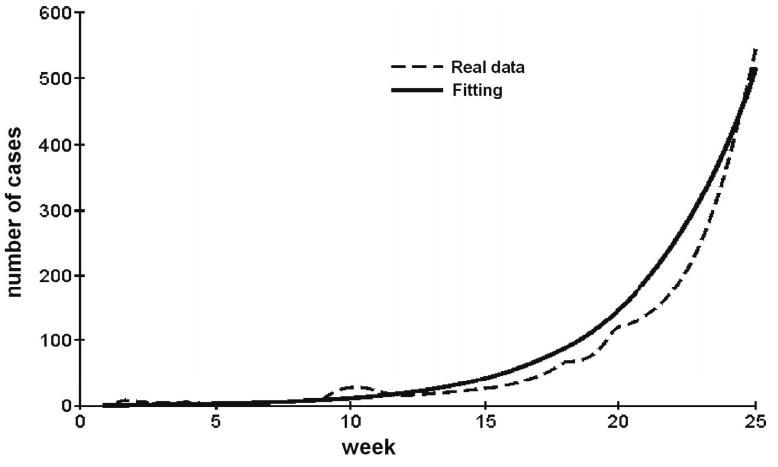

Fig. 6.3. Fitting of the continuous function to the initial phase of the actual epidemic until the last week before the first intervention (Massad *et al.*, 1999)

In order to estimate what would be the natural course of the epidemic we first fitted a continuous function to the initial phase of the actual epidemic until the last week before the first intervention. As expected, it resulted in an exponential curve, with a positive growing rate of 0.25/week. Figure 6.3 shows the result of this fitting.

Next, we calculated the effective contact rate, β, a composite rate describing the probability of contact between susceptible and infected individuals and the probability that such a contact will result in a new case. This was done by assuming that the number of new infections, $y(t)$, increase exponentially as seen in figure 6.3, according to:

$$y(t) = y(0)\exp\{[\beta\,\bar{x} - (\mu+\gamma)]t\} \tag{6.3}$$

where $\overline{x}$ is the expected proportion of susceptibles, assumed to be equal to 10%; μ is the natural mortality rate of the population, assumed to be equal to 0.0003/week and γ is the inverse of the infectiousness period of measles, assumed to be equal to 1 week. The term between square brackets resulted in a value of β equal to 12.5/week.

Those parameters then fed a dynamical system of the classical SIR type, in order to retrieve the natural course of the epidemic in the absence of vaccination. The model had the form:

$$\begin{aligned}
\frac{dx(t)}{dt} &= \mu[y(t)+z(t)] - \beta x(t)y(t) \\
\frac{dy(t)}{dt} &= \beta x(t)y(t) - (\mu+\gamma)y(t) \\
\frac{dz(t)}{dt} &= \gamma y(t) - \mu z(t)
\end{aligned} \tag{6.4}$$

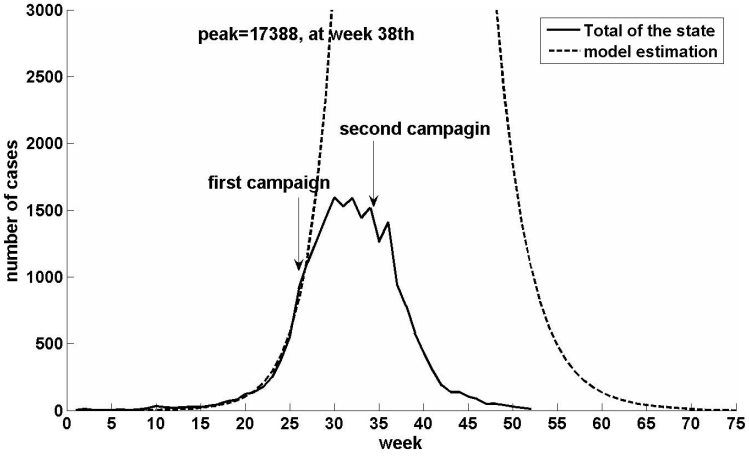

Fig. 6.4. The results of the model simulation and the actual epidemic underlying (Massad *et al.*, 1999)

where $z\,(t)$ represents the recovered (immune) individuals. The result of the simulation, with initial conditions $x(0) = 0.1$; $y(0) = 10^{-7}$ and $z(0) \cong 0.9$, with the actual epidemic underlying, can be seen in figure 6.4.

As can be noted from figure 6.4, the expected number of cases simulated by the model above would peak at around 17,500 cases at the 38^th week, totalizing almost 300,000 cases. This would represent an attack rate of around 8% of the susceptible population, a figure which is in the lower bound of others measles epidemic reported in the literature (Markowitz & Katz, 1994; Hutchins *et al.*, 1990; Weeks *et al.*, 1992). Also noteworthy in figure 6.4 is the striking concordance between the simulated curve and the actual epidemic until week 25. In this point, there is a significant deflection of the exponential trend of the epidemic curve, which occurred just after the first intervention.

By comparing the expected (simulated) number of cases with that seen in the actual epidemic we may conclude that the proposed vaccination strategy (carried out at week 25) had a significant impact on the epidemic in the city of São Paulo. However, as can be seen from figure 6.1, the number of cases in the interior of the State continued to raise after the first campaign, peaking around ten weeks after. Possible causes for this shall be discussed later on. Health authorities then decided to carry out a second campaign which differed from the first one by the virtual absence of costs constraints considerations. Strategy number eight, therefore, was the best choice available, because it has the highest adhesion, and it was implemented in August 16 (which corresponds to week 33). The total number of cases dropped significantly in all age strata and in the whole State soon after the second vaccination and the epidemic was then considered controlled.

In spite of a 95% efficient vaccine available for more than 25 years, measles still remains an important public health problem, killing every year more than one

million children in the developing regions (Murray & Lopez, 1996) and with a *Disability-Adjusted Life Years* (DALY) measure of $36.5x10^6$, which is even higher than malaria ($31.7x10^6$) for the same regions (Murray & Lopez, 1996). As a very transmissible infection with a Basic Reproduction Number (Anderson & May, 1991) usually above 15, it demands very high levels of vaccine coverage (above 93%) in order to be eliminated. However, these levels of coverage are rarely maintained in the routine schemes of immunization. Therefore, it is an usual control strategy, at least in developing countries, to carry out mass vaccination campaigns from time to time. In fact, this occurred in the State of São Paulo in 1987 and again in 1992, with a significant impact on measles incidence.

It is common to observe a severe dropping of cases shortly after a mass vaccination campaign. As time passes by, however, the residual fraction of non-responders to the vaccine and the immigration of susceptible individuals from other areas of the country, starts to accumulate in the population. This fact allied to the marked dropping in the coverage levels in the immunization routine observed in the last two years in the State of São Paulo, may explain the 1997 epidemic.

A subject of hot debate among public health authors, periodic mass vaccination has been considered an effective way to control measles epidemics (Nokes & Swinton, 1997). The design of such a vaccination strategy is based on the rate of replenishment of susceptibles into the population that follows the vaccination. In the case when the mass vaccination is intended to supplement an existing routine (the case of São Paulo State), the rationale is as follows (Nokes & Swinton, 1997): the replenishment of susceptibles equal the birth rate, $1/L$ (as in other works, L denotes the population life expectancy), reduced by a fraction $(1-p)$, where p is the proportion of newborn effectively vaccinated in the routine schedule. If we denote the proportion of children vaccinated in the campaign as p', then the interval, T_v, between two successive campaigns is given by:

$$T_v = \frac{p'A}{(1-p)}, \tag{6.5}$$

where A is the average age of the first infection.

In very populous countries like Brazil and, in particular, in regions like the State of São Paulo, where mass vaccination campaigns are aimed to cover millions of individuals, any reasonable estimate of the minimum number to be vaccinated could represent savings of millions of dollars to public money.

When the São Paulo epidemic was detected and the vaccination campaign decided, very few data was available to allow the application of dynamical modeling, a more structured approach, to the design of the optimal vaccination schedule (Massad *et al.*, 1994b). Moreover, the dynamics of a measles epidemic shortly after an intervention such as a mass vaccination campaign has been poorly documented in the literature. So, it would be very difficult to predict the impact of the intervention on the course of the epidemic. In addition, an important constraint was imposed - the total number of doses available was dangerously limited to 300,000. This scenario encouraged us to attempt, for the

first time (to the best of our knowledge), the use of fuzzy logic concepts to design the vaccination campaign.

The capacity of the fuzzy decision model in predicting the number of children that could be reached by the vaccination strategy can be evaluated by contrasting this number (177,763, which corresponds to 60% of the susceptibles in the targeted population) with the actual number of children who received the vaccine (213,084, which corresponds to 72% of the susceptibles in the targeted population). Therefore, the fuzzy model prediction of the number of children that should be vaccinated has an accuracy of more than 80%. As a result the efficacy of the strategy was significant, at least for the metropolitan region of São Paulo city (figure 6.4), notwithstanding the minor impact seen in the rest of the State. A possible explanation for this could be a lack of adequacy of the selectiveness criteria adopted (to vaccinate only previously unvaccinated children). As a matter of fact, another uncertainty, not forecasted by the initial model, was the decision of public health authorities to extend the measles campaign to a broader scope strategy that included other vaccines like diphtheria-pertussis-tetanus (DPT). However, shortly after midday of June, 21, the DPT vaccine run out of stock, which probably demobilized the population. The latter argument is intended only as an example of how unexpected facts can influence the final result of such a complex endeavor like a mass vaccination campaign. In conclusion, we think that the fuzzy logic approach for designing the control strategy against the measles epidemic in São Paulo was very useful in the sense that it allowed the combination of intuitive informations from public health experts and costs constraints into a coherent model. Moreover it proved to be very effective, in the sense that the strategy adopted resulted in a significant control of the epidemic. Our results, notwithstanding several intervenient factors out of our control during the implementation of the proposed strategy, are very encouraging in demonstrating the potential of new techniques for the designing of interventions in public health.

Maybe the great advantage of the making decision approach proposed by Bellman and Zadeh applied here is its simplicity, both from the practical and theoretical points of view (Bellman & Zadeh, 1970). This simplicity allowed that the fuzzy model for design a control strategy for vaccination against measles could be developed quickly. In fact, this model was elaborated, in a consensus form, in just two meetings. At the final of the second meeting the best strategy elected by the model was accepted by all experts and in few days it was implemented in whole São Paulo State. Clearly, from the sanitary surveillance point of view, the agility and the adhesion capacity are important characteristics desired in the mathematical models.

Stochastic Decision Trees is one of the most traditional approach to decision making that deals with uncertainty in health care applications (Mason *et al.*, 1995; Col *et al.*, 1997; Onho-Machado *et al.*, 2000). In order to compare the fuzzy decision making with other more traditional probabilistic methods, Onho-Machado and collaborators (2000) studied the same situation with the decision trees technique. The authors built a ranking of the strategies to control

the measles epidemic in 1997, in Brazil, considering the same structure proposed in the fuzzy decision making and compared them (Onho-Machado *et al.*, 2000). The models identify the same strategy as being the best one, but exhibit differences in the ranking starting from the fourth strategy. So, in terms of the health care decision making the fuzzy model and the stochastic decision trees were completely equivalent. Thus, the differences between the two approaches refer only to the mathematical structures and, in this case, the fuzzy decision approach presents the advantage of its mathematical simplicity, which resulted in a great adhesion power.

7 Fuzzy Rule-Based Models in Epidemiology

Mathematical models are, in essence, theoretical structures that describe the behavior of real systems through the quantification and manipulation of variables. These models have been widely applied in several areas, aiming to the elaboration of forecasts, analysis of information, treatment of data, control of systems, and evaluation of hypotheses and strategies. In this sense, mathematical modeling have importantly contributed to decision making process. Chapter 3 presents an overview of mathematical modeling and its aspects in epidemiology, introducing the basic concepts of some classic models, that is, models based on the classical logic axioms, and discuss the problems to modeling uncertainties when a more realistic scenario is considered.

Fuzzy rule-based models are systems whose variables are described by fuzzy sets rather than crisp numbers. They are based on the concept of fuzzy partitioning of the information and may be categorized in two general groups, depending on how the information is represented: 1) linguistic models (LM), whose most famous example is the Mamdani type model, and 2) the Takagi-Sugeno-Kang model (TSK). Both models are based on fuzzy rules and linguistic variables. However, while the linguistic models are essentially a qualitative description of the system behavior by using a natural language, the TSK models are a clever combination of fuzzy and non-fuzzy structures. In this chapter we present the basic structures of the fuzzy rule-based models and detail the LM and TSK models.

Due to its great success in the modeling of controllers systems, the most applied structure is the so-called *Fuzzy Linguistic Model* (FLM). FLM are based on Approximate Reasoning, which provides the framework for reasoning with uncertain information through adequate inference mechanisms (Yager & Filev, 1994). FLM could be defined as a particular expert system, since it is composed basically by a knowledge base and an inference engine, both of them allowing the influence of human expert knowledge. In the biomedicine and epidemiology, the majority of fuzzy applications is based on fuzzy linguistic systems. They have been widely used in the development of fuzzy controllers of medical machines, in risk evaluation and in medical diagnosis systems (Mahfouf *et al.*, 2001; Nascimento & Ortega, 2002; Castanho *et al.*, 2007; Duarte *et al.*, 2006). In

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 111–150, 2008.
springerlink.com

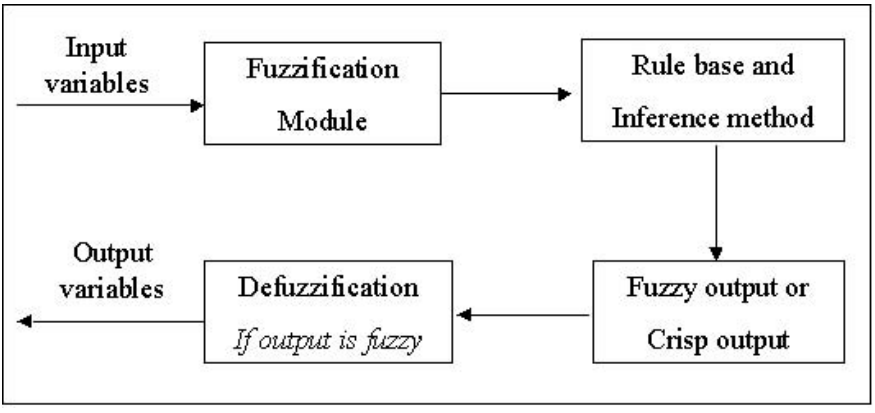

Fig. 7.1. A typical fuzzy linguistic model scheme

epidemiology there are few Fuzzy Linguistic Models published (Ortega *et al.*, 2000; Ortega *et al.*, 2003; Jafelice *et al.*, 2004a and 2004b; Ortega *et al.*, 2008b). In this chapter we shall present three proposals of fuzzy linguistic models applied to epidemic problems.

Fuzzy rule-based models have a simple structure and are composed by four main components: 1) a *fuzzification* module, which translates crisp inputs (classical measurements) into fuzzy values through *linguistic variables*; 2) a *If-Then* fuzzy rule base, which consists in a set of conditioned fuzzy propositions; 3) an *inference method*, which applies fuzzy reasoning mechanisms to obtain the outputs or, in other words, a way to compute with fuzzy rules; and 4) a *defuzzification* module, which translates fuzzy outputs back to crisp values, if necessary. Figure 7.1 shows a typical Fuzzy rule-based model scheme and the interconnection of its modules. Each rule-based system component will be detailed in detail.

7.1 Linguistic Variables

Linguistic terms are used to express concepts and knowledge in human communication, and in a several areas they are the most important form, when not the only way, to quantify/qualify the data and information. The use of linguistic terms is frequent in our daily life. We usually say "*hot day*", "*crowded bus*", "*the person is high, thin, healthy*", etc. All of these linguistic terms have a meaning and transmit an information that, in general, is context-dependent.

In the medical universe the use of linguistic terms permeates all the areas, including in laboratorial exams descriptions. The use of terms as *normal*, *slightly increased/decreased*, *weakened*, *good state*, and so on is very frequent (Sanchez,1998). Aiming to express the intensity of the observation, a numeric value is commonly associated with a linguistic term. It is useful to quantify the

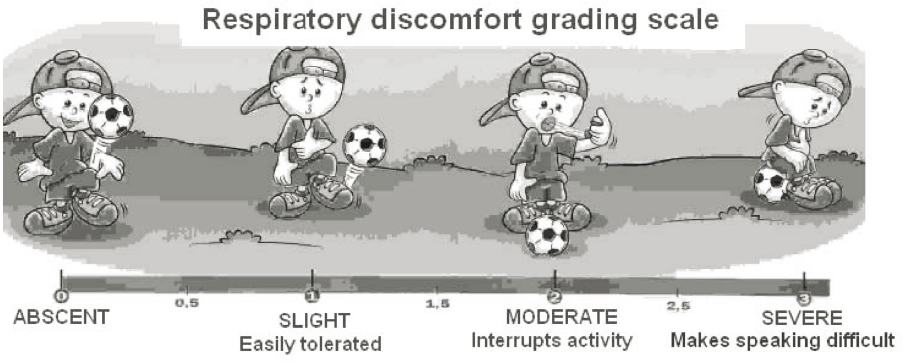

Fig. 7.2. Poster in use by epidemiologists in the São Paulo Hospital of Clinics to describe the breathing discomfort

clinical findings and to arrange the observations. A typical example of these quantities is the crosses measure, widely used by doctors to quantify the observation in the propedeutic examamination. As an example of the linguistic terms used in medicine, consider the illustration below (figure 7.2). This picture is a poster used by epidemiologists for evaluation of breathing discomfort in the Hospital of Clinics, of São Paulo University.

Notice that the **Degree of Breathing Discomfort** could be considered a linguistic variable that may receive the following linguistic terms: *absent*, *light*, *moderate* and *strong*. In addition, these linguistic terms are associated with the numeric values 0, 1, 2 and 3, respectively, and a color scale denotes the gradual and constant transition of one situation to another. This gradual transition of breathing discomfort degree, expressed in the poster by the color scale, reveals the inherent fuzziness of this classificatory process.

Thereby a fuzzy linguistic variable is, in an informal way, a variable whose value is qualitatively expressed by linguistic terms (that supplies a concept to the variable) and quantitatively expressed by membership functions, that is, by fuzzy sets. In this sense, the linguistic variable is composed by a symbolic and a numerical part. The symbolic part allows the description of the phenomenon using natural language and the numerical part allows to compute with them. We also point out that although numerical variables are widely applied in the exact sciences as engineering, physics and mathematics, the symbolic variables have been conquering larger importance due to the development of the areas of artificial intelligence and decision processes. The capacity to combine symbolic variables (linguistics) and numeric is one of the main reasons for the success of the applications of the fuzzy logic in intelligent systems, either in engineering or in other areas as biology, ecology, medicine and so on.

Formally, a linguistic variable is characterized by a quintuple denoted by (v, T, X, g, m) in which v is the *name* of the variable (for example *Temperature*, *Pressure*, *Fever*), T is the set of *linguistic terms* of v (as *high*, *small*, *medium*)

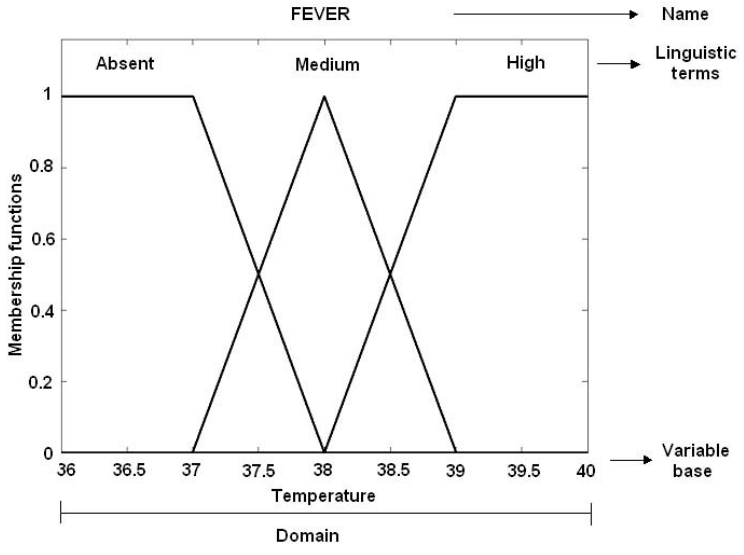

Fig. 7.3. An example of a linguistic variable

that refer to a base variable x, whose values range over a universal set X, g is a grammar for generating the linguistic terms, and m is a semantic rule for association each linguistic term $t \in T$ with its meaning $m(t)$, which is a fuzzy set on the universe X.

For example, consider a linguistic variable named *fever*, that is, $v = fever$, and showed in figure 7.3. This variable is defined in the universal set $X = [36, 41]$ and the base variable x is a measure of body temperature, $x \in X$. The set of terms associated with *fever* could be, in a given context, $T(fever) = absent, low, medium, high$, where each term in $T(fever)$ is a label of a linguistic value of *fever* generated by a syntactic rule (not explicitly shown in figure 7.3). Each linguistic term is assigned one of four fuzzy sets, whose membership functions have, in this example, trapezoidal and triangular shapes defined in the interval $[36, 41]$.

The concept of linguistic variable plays an important role in many applications of fuzzy sets and is essential in approximate reasoning. In fuzzy linguistic models they are used in the *fuzzification module*, where the relevant antecedents and consequents of the rule system, with their respective ranges, are identified and defined. In addition, the *fuzzification module* could be used to fuzzify each input variable to express the associated measurement uncertainty through a *fuzzification function*. However, in the majority of the applications the inputs are not fuzzified.

Essentially, fuzzy sets are characterized by functions of the form $X \rightarrow [0, 1]$, where X is a given universal set. Thus, the problem of constructing fuzzy sets

is reduced to the problem of translating the meanings of the relevant linguistic terms for the adequate membership functions, considering the context.

In biomedicine and epidemiology, fuzzy sets are commonly constructed through experts' judgment. This expert approach could involve one or more experts in the specific field of interest. The experts opinion could be extracted together, in consensus, or separately, but aggregated in an appropriate way. Besides, the membership functions could be found through *direct* or *indirect* methods, depending on the complexity of the meaning that should be described by the linguistic term. Clearly, in this context the experts assume a fundamental role in the fuzzy modeling development, particularly in the epidemiological and diagnose processes.

If a sample data is available, then it is possible to build the membership functions through the information contained in them. Again, there are many ways to find it and most of them could be classified by the *mathematical theory of curve fitting methods* (as *least-square error* and *Lagrange interpolation*) or *learning through artificial intelligence methods* (as *neural networks* and *genetic algorithms*). Other frequent situation is to combine the sample data information and experts'judgment in order to find more specific membership functions. The above methods and others can be found in the texts such as Klir and Yuan (1995), Kosko (1997) and, Pedrycz and Gomide (1998 and 2007).

7.2 Fuzzy Rules

Fuzzy rules are structures widely used in several approaches of fuzzy sets theory, since they provide a formal way to represent strategies and information from experiences and empirical associations. Thus, a fuzzy rule can be understood as a unit for capturing some specific knowledge that involves imprecision, vagueness and/or uncertainty.

Knowledge can be represented through fuzzy propositions, which can be unconditional or conditional. An unconditional fuzzy proposition is a simple statement as

Fever is high,

in which *fever* is a kind of an attribute of an object and *high* is its adjective (a qualitative value) usually described with a linguistic variable.

In contrast, a conditional fuzzy proposition is composed by two parts: an **if-part**, called *antecedent* part, which describes a condition (premises) that can be partially satisfied, and a **then-part**, called *consequent* part, which describes a conclusion or an action that can be found when the condition holds. Thus, a conditional fuzzy proposition has the following form:

IF fever is high, THEN illness is great.

Fuzzy rules are conditioned fuzzy propositions. So, the main difference between classical and fuzzy rules is that in the latter the rule's antecedent describes an elastic condition, while in the former it describes a rigid condition,

i.e., a classical rule does not work if its antecedents are not completely satisfied. This flexibility of the fuzzy rules allows us to model the vagueness and uncertainties of the statements, commonly used in the real world.

As in its classical counterpart, fuzzy rules may combine many simple conditions in its antecedent part using logic connectives (conjunction, AND, disjunction, OR, and negation, NOT) as:

IF the newborn has extremely low weight AND the gestational age is preterm, THEN the risk of neonatal death is high,

and can be chained or parallel. It is called chained when the consequent of one rule is the antecedent of the other, and parallel otherwise. An example of chained rules is showed below.

IF contact rate between infected and susceptible individuals is high, THEN the force of infection is high;
IF the force of infection is high, THEN the number of infected is great.

Another frequent fuzzy rule in biomedicine and epidemiology are the so-called gradual rules, which express the gradual relationships between concepts and usually reflects the commonsense reasoning, as:

THE MORE inclusive the vaccination campaign is, THE MORE expensive it is.

This rule expresses a continue and progressive variation of the campaign cost as it becomes more inclusive. Gradual rules are frequent in experts representation knowledge.

A rule-based system is a collection of fuzzy rules, which are parallel rules systems in most of the applications, that is, they have the following form:

$$\begin{gathered} \textit{If } X \textit{ is } A_1, \textit{ then } Y \textit{ is } B_1, \\ \textit{If } X \textit{ is } A_2, \textit{ then } Y \textit{ is } B_2, \\ \ldots \\ \textit{If } X \textit{ is } A_N, \textit{ then } Y \textit{ is } B_N. \end{gathered}$$

Clearly, it is desirable that this rule's collection presents neither inconsistencies nor conflicts. To avoid this, it is necessary that the rules contain neither mutually exclusive knowledge nor contradictory information. A very common situation that the rule's system is potentially inconsistent is when two or more rules have the same antecedents but different consequents or, when the progression of the changes is not preserved in the gradual rules systems. In these cases, the rules computation could lead to unsatisfactory and unexpected results. An example of potentially inconsistent rules is:

IF the newborn weight is low, THEN the neonatal risk of death is high;
IF the newborn weight is low, THEN the neonatal risk of death is medium.

From the modeling point of view, the antecedents of the fuzzy rule define a fuzzy area in the input variable space of the system. On the other hand, the

consequent part describes a fuzzy area in the output space. In this sense, the processes involved in the elaboration of the rule's antecedents consist rather in a classificatory work, while the elaboration of the rule's consequents demands a priori knowledge, though empiric, about the behavior of the system. So, it is expected that the elaboration of the consequents of the rules should be a more complex task than the antecedents ones, particularly when the model is expert knowledge based. However, since the rules of the model are defined it is possible to compute with them through any appropriated inference procedure. Clearly, the choice of the inference procedure depends, among other things, on the type of the fuzzy rules. In this text we will concentrate in the parallel rules computation, that is the most used in the applications. For more details about the inference process on other kinds of rules we recommend the book by Pedrycz and Gomide (1998 and 2007).

7.3 Inference Procedure in Fuzzy Rule-Based Models

Linguistic models could be understood as a mapping of a fuzzy input space into fuzzy output space. In this sense a set of rules associated to an inference procedure is analogous to a function able to describe both linear and non-linear systems. For this purpose, each fuzzy rule in the systems could be interpreted as a fuzzy relation (see chapter 2, section 2.5).

An If-Then rule is a conditioned proposition that can be written as:

$$p\text{: } \textit{If antecedents, Then consequents},$$

where antecedents and consequents are also fuzzy propositions, whose variables values are linguistic terms that establish a fuzzy constraint in some appropriated universe set. So, a fuzzy rule

$$\textit{If } x \textit{ is } A, \textit{ then } y \textit{ is } B$$

describes a fuzzy relation between the variables x and y, whose membership grade $R(x, y)$ represents the degree to which the pair $(x, y) \in X \times Y$ is compatible with the relation between those variables involved in the rule. Thus, the fuzzy rule "*If* x *is* A*, then* y *is* B" with $x \in X$ and $y \in Y$, may be understood as a fuzzy relation on the Cartesian product space $X \times Y$, as illustrated in figure 7.4.

Therefore, as in the fuzzy relations case, it is possible to elaborate new fuzzy propositions through conjunctions and/or disjunctions operators combining one or more simple propositions, like:

$$p\text{: } \textit{If } x_1 \textit{ is } A_1 \textit{ AND } x_2 \textit{ is } A_2 \textit{ AND } x_3 \textit{ is } A_3, \textit{ Then } y_1 \textit{ is } B_1 \textit{ AND } y_2 \textit{ is } B_2;$$
$$q\text{: } \textit{If } x_1 \textit{ is } A_1 \textit{ OR } x_2 \textit{ is } A_2 \textit{ OR } x_3 \textit{ is } A_3, \textit{ Then } y_1 \textit{ is } B_1,$$

where A_1, A_2 and A_3 are fuzzy sets of the variables $x_1 \in X_1$, $x_2 \in X_2$ and $x_3 \in X_3$ and similarly, B_1 and B_2 are fuzzy sets of the variables $y_1 \in Y_1$ and $y_2 \in Y_2$, respectively. In this case, the fuzzy rules p and q can be understood as inducing a fuzzy relation P and Q on the space $X_1 \times X_2 \times X_3 \times Y_1 \times Y_2$.

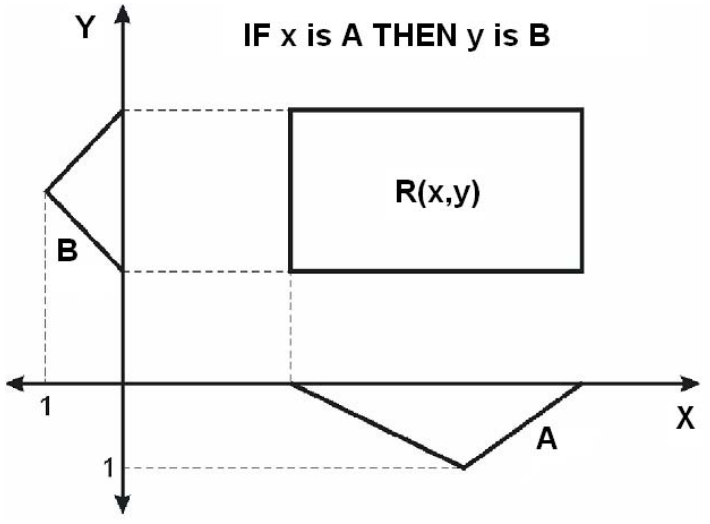

Fig. 7.4. Fuzzy relation induced by rule *"If x is A, then y is B"* (modified from Yager and Filev, 1994)

Clearly, there are many ways to proceed the inference in a collection of fuzzy rules. The choice for a certain method depends on the semantics of the rules, on the type of fuzzy model and on the characteristics of the phenomenon that it is being modeled. In general, linguistic fuzzy models can be classified into two groups, depending on the inference processes: constructive or destructive linguistic model. The former is the linguistic model in which the output is constructed by superimposing the individual outputs of each rule through a disjunctive approach, i.e., applying t-norms on the fuzzy sets. On the other hand, in destructive models the solution is formulated through the conjunctive operators, i.e., applying t-conorms on the fuzzy sets.

In a constructive linguistic model composed by a collection with N parallel rules such as:

$$\textit{If } x \textit{ is } A_i, \textit{ then } y \textit{ is } B_i,$$

each fuzzy rule can be seen as inducing a fuzzy relation R_i, and the set of rules as a global relation R found through the aggregation of the individuals' relation using disjunction operators. Therefore, the global relation is given by:

$$R = \cup_{i=1}^{N} R_i. \tag{7.1}$$

On the other hand, since A_i and B_i are fuzzy sets of X_i and Y_i, respectively, the binary fuzzy relation $R_i = A_i \times B_i$ can be interpreted as fuzzy subsets of $X \times Y$, whose membership function is

$$R_i(x, y) = A_i(x) \otimes B_i(y), \tag{7.2}$$

where $\otimes$ is any t-norm, which the most used is the *min* operator (see chapter 2).

Thereby, given the input fuzzy value D, the output set F inferred through the N parallel rules is such that,

$$\begin{aligned} F(y) &= \oplus_x \left[D(x) \otimes R(x,y) \right] \\ &= \oplus_x \left[\oplus_{i=1}^{N} \left(D(x) \otimes R_i(x,y) \right) \right] \\ &= \oplus_{i=1}^{N} \left[\oplus_x \left(D(x) \otimes A_i(x) \otimes B_i(y) \right) \right] \\ &= \oplus_{i=1}^{N} \left[\left(\oplus_x \left[D(x) \otimes A_i(x) \right] \right) \otimes B_i(y) \right] \\ &= \oplus_{i=1}^{N} \left[\tau_i \otimes B_i(y) \right], \end{aligned} \tag{7.3}$$

where $\oplus$ is any t-conorm, the most used of which is the *max* operator (see chapter 2), and τ_i is the possibility of the set B_i given the input set D, $Poss[B_i|D]$, called *degree of firing*, or *degree of match*, of the i^{th} such that (Yager & Filev, 1994):

$$\tau_i = \oplus_x \left[D(x) \otimes A_i(x) \right]. \tag{7.4}$$

A similar inference process can be obtained if each fuzzy rule in the collection of N parallel rules has several fuzzy sets in its antecedent part. These inference processes are detailed in the Mamdani linguistic model context (see section 7.5.1).

The inference procedure in destructive linguistic fuzzy models is quite different from its constructive counterpart. In this kind of models the output is obtained by the remotion of the possibilities that are not acceptable to the individual rule. So, in a destructive model composed by N parallel rules like "*If* x *is* A_i, *then* y *is* B_i", each rule is associated with a fuzzy relation R_i, defined in the Cartesian Product space $X \times Y$, that is found through the disjunction operators, and whose membership function is given by:

$$R_i(x,y) = \overline{A}_i(x) \oplus B_i(y), \tag{7.5}$$

where $\overline{A}_i(x) = 1 - A_i(x)$ is a fuzzy standard complement of the fuzzy set A_i, correspondent to the logical operation of negation.

Under this approach the individual rules are aggregated through the conjunction operator as:

$$R = \cap_{i=1}^{N} R_i, \tag{7.6}$$

whose membership function is given by:

$$R(x,y) = \otimes_{i=1}^{N} R_i(x,y) = \otimes_{i=1}^{N} \left[\ \overline{A}_i(x) \oplus B_i(y) \ \right]. \tag{7.7}$$

Thus, for a given input set D, the output fuzzy set F inferred by the rules is given by:

$$\begin{aligned} F(y) &= \oplus_x \left[D(x) \otimes R(x,y) \right] \\ &= \oplus_x \left[\otimes_{i=1}^{N} \left(D(x) \otimes R_i(x,y) \right) \right] \\ &= \oplus_x \left[\otimes_{i=1}^{N} D(x) \otimes \left(\overline{A}_i(x) \oplus B_i(y) \right) \right]. \end{aligned} \tag{7.8}$$

Expression (7.8) can be rewritten using the distributivity properties of fuzzy sets (see chapter 2) as:

$$\begin{aligned} F(y) &= \oplus_x \left[\otimes_{i=1}^{N} \left(D(x) \otimes \overline{A}_i(x) \right) \oplus \left(D(x) \otimes B_i(y) \right) \right] \\ &= \oplus_{i=1}^{N} \left[\otimes_x (\overline{A}_i(x) \otimes D(x)) \oplus (B_i(y) \otimes D(x)) \right]. \end{aligned} \tag{7.9}$$

Assuming now that $D(x)$ is a normal fuzzy set, i.e., there are $x' \in X$ that $D(x') = 1$, it is possible to find $F(y)$ as (Yager & Filev, 1994):

$$F(y) = \otimes_{i=1}^{N} \left[\overline{\tau}_i \oplus B_i(y) \right], \tag{7.10}$$

where $\overline{\tau}_i$ is the possibility of the set $\overline{B}_i$ given the input set D, $Poss[\overline{B}_i|D]$, that represents the *degree of firing* of the i^{th}, whose value is

$$\overline{\tau}_i = \oplus_x \left[\ \overline{A}_i(x) \otimes D(x) \ \right]. \tag{7.11}$$

In the case that the input value is a crisp number x^*, in other words, the input D is interpreted as a fuzzy singleton whose pertinence function is given by:

$$D(x) = \begin{cases} 0 \ if \ x \neq x^* \\ 1 \ if \ x = x^* \end{cases}, \tag{7.12}$$

and assuming the standard union and standard intersection for, respectively, disjunction and conjunction operators, it is possible to rewrite the degree of firing as:

$$\overline{\tau}_i = 1 - A_i(x^*) = 1 - \tau_i. \tag{7.13}$$

The destructive inference procedure of linguistic models, also called logical method, is illustrated in figure 7.5 and summarized in the following algorithm (Yager & Filev, 1994):

1. For each rule of the model do:

   - Calculate the degree of firing of the rule by:
     $\overline{\tau}_i = \oplus_x \left[\overline{A}_i(x) \otimes D(x) \right]$ if the input $D(x)$ is a fuzzy set; or

     $\overline{\tau}_i = 1 - B_i(x^*)$ if the input is a crisp number x^*.
   - Find the fuzzy set F_i inferred by each rule by:
     $F_i(y) = \overline{\tau}_i \oplus D_i(y)$

2. Find the fuzzy output set $F(y)$ through the aggregation of the individual F_i by *min* operation:
   $F(y) = \otimes_{i=1}^{N} F_i(y)$

An important feature of linguistic fuzzy models, under both constructive and destructive approaches, is that in spite of the input value, they provide a fuzzy output set. However, in most of the applications the fuzzy system require a crisp

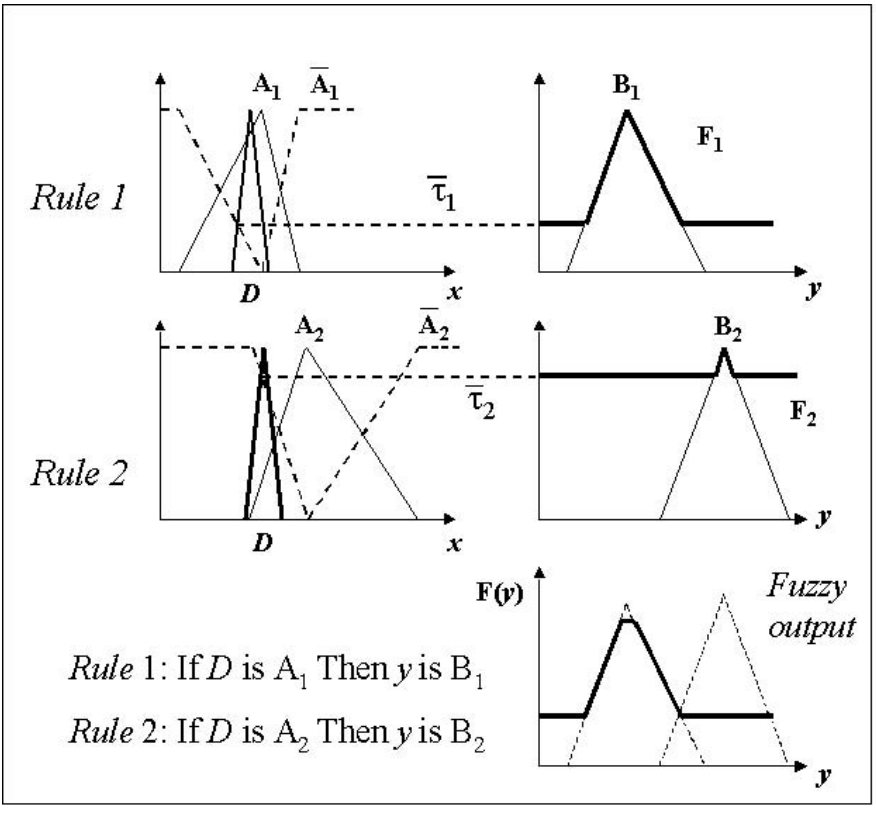

Fig. 7.5. A example of the inference procedure in the destructive linguistic model (modified from Yager and Filev, 1994)

number as the final output, since decisions and controls should be processed. In order to accomplish this, the fuzzy output $F(y)$ of the linguistic model must be defuzzified. There are many techniques to defuzzify a fuzzy set and the choice for one particular method is arbitrary. It is worth to point out that the defuzzification requirement of the linguistic models associated to its arbitrary choice consists one of the main disadvantages of these models when compared to the TSK models, which we will be present later on.

7.4 Defuzzification Methods

Defuzzification is a procedure that allows to interpret the possibility distribution of the output fuzzy sets in a quantitative way. In other words, it is a method that provides the most representative crisp number, in the variable domain, that captures the essential meaning of that possibility distribution. There are many

techniques to perform the defuzzification and the most common in practice is the *Center of Maximum*, the *Mean of Maximum* and the *Center of Area* methods.

7.4.1 Center of Maximum Method

Consider a fuzzy subset F in $\mathbb{R}$, whose membership function is $F(y)$. The Center of Maximum method provides the defuzzified value, $d_{CM}(F)$, defined as the average of the smallest and largest values of $y \in M$ given by:

$$d_{CM}(F) = \frac{inf(M) + sup(M)}{2}, \tag{7.14}$$

where M is a crisp set such that,

$$M = \{y \in Y | F(y) = hgt(F)\} \tag{7.15}$$

and $hgt(F)$ is the height of the fuzzy set F, defined in chapter 2 as the maximum value of the fuzzy set. For the discrete case, we have:

$$d_{CM}(F) = \frac{min\left[y_k \mid y_k \in M\right] + max\left[y_k \mid y_k \in M\right]}{2}, \tag{7.16}$$

where

$$M = \{y_k \mid F(y_k) = hgt(F)\}. \tag{7.17}$$

7.4.2 Mean of Maximum Method

The *Mean of Maximum* defuzzification is a method usually defined only for the discrete fuzzy sets, whose formulation is very similar to the *Center of Maximum*, particularly when the fuzzy output set is convex (see chapter 2). In fact, there are mathematical details involved in the difference among these two methodologies whose discussion is not in the scope of this book.

The defuzzified value for a fuzzy set F through the *Mean of Maximum* method, $d_{MM}(F)$, is the average of all values in the set M given by (7.17), such that:

$$d_{MM}(F) = \frac{\sum_{y_k \in M} y_k}{|M|}, \tag{7.18}$$

where $|M|$ is the cardinality of the crisp set M.

If M is a continuous set as described in (7.15), it is possible to define the $d_{MM}(F)$ as the arithmetic average of mean values of all intervals contained in M, including that intervals with zero length (Klir & Yuan, 1995). Figure 7.6 illustrates the *Mean of Maximum* (MM) defuzzification approach.

The principal limitation of the CM and MM defuzzification methods is that they do not consider the total shape of the fuzzy set F. In this sense, two different distributions of possibility that present the same M set will provide, after defuzzification procedure, the same crisp value, independently of the differences in the shape. Figure 7.7 illustrate this situation. Clearly, in this case the defuzzification approach disrespect a part of the information present in the F set. It is particularly important when the F set is the output of some linguistic model.

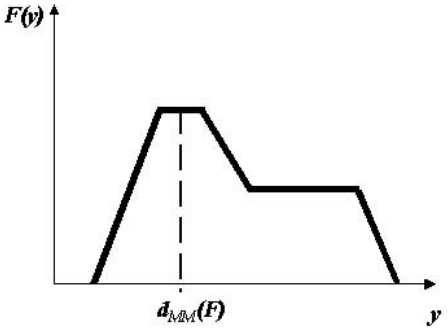

Fig. 7.6. Example of the Mean of Maximum defuzzification method for continuous and discrete cases

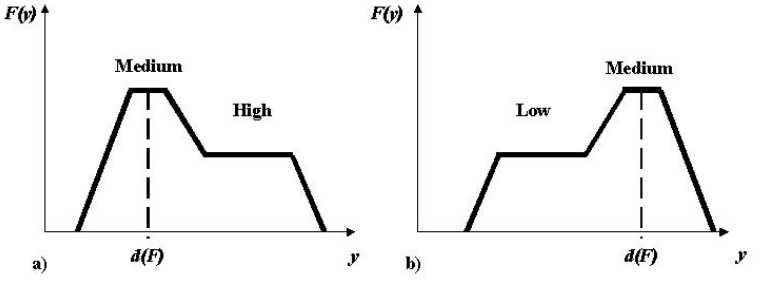

Fig. 7.7. Example of two different shape of F whose defuzzified value is the same through the CM and MM method. In the situation a) we have that the fuzzy output indicates that the fuzzy results is medium tending to high; in counterpart, in the situation b) the fuzzy results is medium tending to low. These defuzzification methodology are not able to distinct those two different fuzzy outputs.

7.4.3 Center of Area Method

Differently of the methods above, the *Center of Area* consider the entire possibility distribution of F to calculate the defuzzified value. This technique is similar to that applied to calculate a center of gravity in physical systems and, for this reason, it is also called the *Center of Gravity* or *Centroid* method. It consists in the division of the area under the F graph into two equal parts. Therefore, the defuzzified value $d_{CA}(F)$ of the fuzzy set F is found, in a discrete case, by:

$$d_{CA}(F) = \frac{\sum_{k=1}^{q} y_k F(y_k)}{\sum_{k=1}^{q} F(y_k)}, \tag{7.19}$$

where F is defined on a finite universal set Y whose cardinality is q.

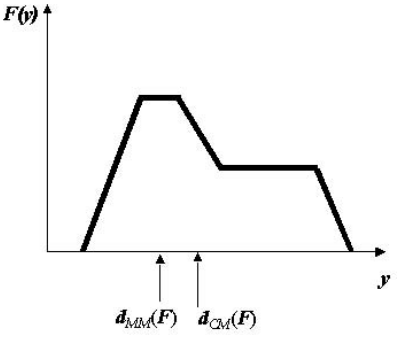

Fig. 7.8. Example of a fuzzy set and its defuzzified values trough the MM and CA methods. In this case d_{CM} and d_{MM} values are equal.

If F is defined in a continuous interval $[y_a, y_b]$ the expression for $d_{CA}(F)$ is given by:

$$d_{CA}(F) = \frac{\int_{y_a}^{y_b} yF(y)dy}{\int_{y_a}^{y_b} F(y)dy}. \tag{7.20}$$

The *Center of Area*(CA) method can also be understood as a weighted average where the value $F(y)$ can be understood as the weight of the value y. Although the CA defuzzification is the most applied method, it presents higher computational costs than the others two methods discussed, particularly for continuous domains. Figure 7.8 illustrate an example of a fuzzy set and its defuzzified values trough the methods above.

To end this section, we want to emphasize that, although defuzzification is widely applied in the linguistic modeling context, it is not applied exclusively for this approach. In fact, the defuzzification is, in a general view, all operations applied on fuzzy sets in order to elect one crisp value, under one domain, in which the decision is made. In this sense, the vaccination strategy chosen as the best strategy to vaccinate children against measles, presented in chapter 6, was found through the defuzzification procedure applied on the set of possible decisions D by the max operator.

7.5 Some Types of Fuzzy Rule-Based Models

As previously mentioned, there are basically two kinds of rule-based fuzzy systems: models whose output of the rule is a fuzzy set, and models whose output of the rule is a function of the input variables. When the output is fuzzy, the model is called Linguistic Model (LM), which from the inference point of view can be classified as constructive or destructive. This kind of model is commonly based on experts knowledge and be useful in the development of expert systems in areas where large databases are not available. The majority of the applications

using linguistic models is based on the Mamdani's inference or on the Standard Additive Model (SAM) model, both of them are constructive approaches. On the other hand, when the output is a function we have the so-called Takagi-Sugeno-Kang model (TSK), that is normally referred as just Sugeno model. The great advantage of the TSK model, when compared with the linguistic one, is its independence of experts in the formulation of the consequents of the rules. They are widely applied in engineering, where large data collections and functional information about the input variables relationships, as physical laws, are available.

7.5.1 The Mamdani Model

The Mamdani model was the first fuzzy rule-based system developed. It was proposed, in 1974, by E. H. Mamdani and S. Assilian in the control systems context (Mamdani & Assilian, 1975), but nowadays it is widely applied in several areas. Most of the fuzzy controllers elaborated in the 1980's are of the Mamdani type. However, due to the development of the TSK models in the 1990's, strengthened by the development of powerful techniques in the extraction of information from databases, the Mamdani control lost its prior importance for the controllers. Nevertheless, the crescent interests for fuzzy systems in the biomedicine and epidemiology have increased the Mamdani model relevance. Since in these fields the expert experience and knowledge are fundamental, the Mamdani model plays an important role in modeling of such systems.

The Mamdani model is composed by a collection of N parallel rules, such as:

$$\begin{gathered} \textit{If } x \textit{ is } A_1, \textit{ then } y \textit{ is } B_1, \\ \textit{If } x \textit{ is } A_2, \textit{ then } y \textit{ is } B_2, \\ \ldots \\ \textit{If } x \textit{ is } A_N, \textit{ then } y \textit{ is } B_N, \end{gathered}$$

whose inference is processed through the methodology described in section 7.3 for constructive linguistic models, using the *min* for the conjunction operator and *max* for the disjunction operator. Because of this $max - min$ inference, the Mamdani systems are also called "max-min" models.

So, in the Mamdani case, the fuzzy relation R_i, described by the i_th rule, has a membership function given by:

$$R_i(x, y) = \min\left(A_i(x), B_i(y)\right) \tag{7.21}$$

and the global relation R found through the aggregation of the rules is:

$$R = \max_{i=1}^{N}\left(R_i\right), \tag{7.22}$$

whose membership function is given by:

$$R(x, y) = \max_{i=1}^{N}\left[\min\left(A_i(x), B_i(y)\right)\right]. \tag{7.23}$$

Thus, for a given input fuzzy set D, the output fuzzy set F inferred by $max-min$ inference on the rules has the following membership function:

$$F(y) = \max_{i=1}^{N} [\min(\tau_i, B_i(y))], \tag{7.24}$$

where τ_i is the degree of firing of the i_th rule such that,

$$\tau_i = \max_x [\min(A_i(x), D(x))]. \tag{7.25}$$

In the case that the input value is a crisp number x_*, i.e., whose membership function is

$$D(x) = \begin{cases} 0 \; if \; x \neq x_* \\ 1 \; if \; x = x_* \end{cases}, \tag{7.26}$$

this degree of firing becomes

$$\tau_i = \max_x [\min(A_i(x), D(x))] = A_i(x_*). \tag{7.27}$$

Summarizing, the algorithm used to compute the fuzzy rules in the Mamdani system is as following (Yager & Filev, 1994):

1. For each rule of the model:

   - Calculate the degree of firing of the rule by:
     $\tau_i = \max_x [\min(A_i(x), D(x))]$ if the input $D(x)$ is a fuzzy set; or

     $\tau_i = A_i(x^*)$ if the input is a crisp number x^*.

   - Find the fuzzy set F_i inferred by each rule by:
     $F_i(y) = \min[\tau_i, B_i(y)]$

2. Find the fuzzy output set $F(y)$ through the aggregation of the individual F_i by min operation:
   $F(y) = \max_{i=1}^{N} F_i(y)$

Figure 7.9 illustrates an example of Mamdani inference procedure considering a fuzzy input set in the antecedent part of the *If-Then* rule, A_i, and a crisp input number, x^*.

Fuzzy systems composed by a collection of N rules as in Mamdani model above is normally called SISO models, since they have a Single Input and a Single Output variables. However, in most of the applications the Mamdani models are MISO, Multiple Input and Single Output, or MIMO, Multiple Input and Multiple Output, models. The inference procedure described above can be easily generalized for this MISO and MIMO models. Nevertheless, for simplicity, we will present here only inference processes to MISO Mamdani systems (Yager & Filev, 1994).

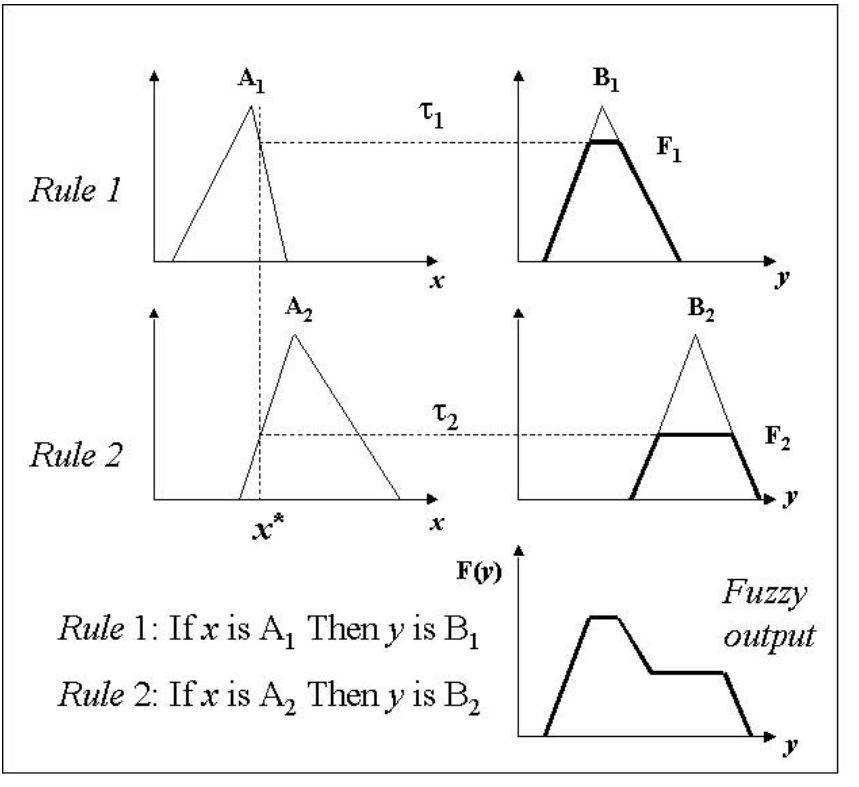

Fig. 7.9. A example of Mamdani inference procedure (modified from Yager and Filev, 1994)

Thus, consider the following rules:

$$\begin{gathered} \textit{If } x_1 \textit{ is } A_{11} \textit{ AND } x_2 \textit{ is } A_{12} \textit{ AND } x_M \textit{ is } A_{1M}, \\ \textit{Then } y \textit{ is } B_1, \\ \textit{If } x_1 \textit{ is } A_{21} \textit{ AND } x_2 \textit{ is } A_{22} \textit{ AND } x_M \textit{ is } A_{2M}, \\ \textit{Then } y \textit{ is } B_2, \\ \ldots \\ \textit{If } x_1 \textit{ is } A_{N1} \textit{ AND } x_2 \textit{ is } A_{N2} \textit{ AND } x_M \textit{ is } A_{NM}, \\ \textit{Then } y \textit{ is } B_N, \end{gathered}$$

where $x_1, x_2, ..., x_M$ are the linguistic input variables and y is the linguistic output variable, and A_{ij} and B_i ($i = 1, 2, ..., N$ and $j = 1, 2, ..., M$) are fuzzy subsets of the universal sets $X_1, X_2, ..., X_M; Y$ of the variables $x_1, x_2, ..., x_M$ and y, respectively. So, as in SISO models, each rule is associated to the individual fuzzy relation R_i defined in the Cartesian product space $X_1 \times X_2 \times ... \times X_M \times Y$, as:

$$R_i = A_{i1} \times A_{i2} \times ... \times A_{iM} \times B_i, \tag{7.28}$$

whose membership function is given by:

$$R_i(x_1, x_2, ..., x_M, y) = \min\left[A_{i1}(x_1), A_{i2}(x_2), ..., A_{iM}(x_M), B_i(y)\right]. \quad (7.29)$$

So, the membership function of the global fuzzy relation of that model is found through the *max* operator in such a way that

$$\begin{aligned} R(x_1, x_2, ..., x_M, y) &= \max_{i=1}^{N}\left[R_i(x_1, x_2, ..., x_M, y)\right] \quad (7.30)\\ &= \max_{i=1}^{N}\left[\min\left(A_{i1}(x_1), ..., A_{iM}(x_M), B_i(y)\right)\right]. \end{aligned}$$

In this case, a fuzzy output F for the input set of fuzzy variable $\mathbf{D} = D_1, D_2, ..., D_N$ is also found through the $max - min$ inference, such that

$$F(y) = \vee_{i=1}^{N}[\tau_i \wedge D_i(y)], \quad (7.31)$$

where $\vee$ is the *max* operator, $\wedge$ is the *min* operator and τ_i, with $(i = 1, ..., N)$, denotes the degree of firing of the $i - th$ rule, given by:

$$\tau_i = (\vee_{x_1}[A_{i1}(x_1) \wedge D_1(x_1)]) \wedge ... \wedge (\vee_{x_M}[A_{iM}(x_M) \wedge D_1(x_M)]). \quad (7.32)$$

In the particular case in which the inputs $D_i's$ are crisp numbers, τ becomes:

$$\tau_i = A_{i1}(x_1) \wedge A_{i2}(x_2) \wedge ... \wedge A_{iM}(x_M). \quad (7.33)$$

Clearly, all Mamdani models discussed above provide a fuzzy set as output. Therefore, all of them must be defuzzified if a crisp output is required. The advantage of these models is that they may be data-independent, or in other words, a phenomenon can be modeled only from experts knowledge, if a collection of data is not available. On the other hand, they may become expert-dependent. Furthermore, if the model has many input variables, resulting in a great number of rules, and if, in addition these input variables are strongly and no-linearly related, then the experts will show great difficulties to supply the output sets and the refinement of the model becomes a tough work. In this situation an alternative is to model the phenomenon through a Takagi-Sugeno-Kang system.

7.5.2 The Takagi-Sugeno-Kang Model

Nowadays the Takagi-Sugeno-Kang (TSK) model, normally referred simply Sugeno, is the most widely used model in engineering. It was proposed firstly by T. Takagi and M. Sugeno in 1983, and received later the attention of Kang, who works particularly in the identification of this kind of systems (Sugeno & Takagi, 1983; Sugeno & Kang, 1988). This model appeared as an alternative to the Mamdani model to deal with complex and high-dimensional systems, in the search of a reduction in the number of rules and in a more objective formulation of the consequent sets in the fuzzy rule.

The idea in the development of TSK's model is to apply explicitly the functional information about the relationship between the input variables in the fuzzy

rules, since this knowledge is available in many problems, as physical laws that govern the systems, for example. Thus, in this model a fuzzy rule is composed by fuzzy sets in its antecedent part, as in the Mamdani type, but by functions in its consequent part. So, in TSK model the fuzzy rules with multiple inputs can be as following:

$$\begin{gathered}
\textit{If } x_1 \textit{ is } A_{11} \textit{ AND } x_2 \textit{ is } A_{12} \textit{ AND } x_M \textit{ is } A_{1M}, \\
\textit{Then } y_1 = a_{10} + a_{11}x_1 + ... + a_{1M}x_M \\
\textit{If } x_1 \textit{ is } A_{21} \textit{ AND } x_2 \textit{ is } A_{22} \textit{ AND } x_M \textit{ is } A_{2M}, \\
\textit{Then } y_2 = a_{20} + a_{21}x_1 + ... + a_{2M}x_M, \\
... \\
\textit{If } x_1 \textit{ is } A_{N1} \textit{ AND } x_2 \textit{ is } A_{N2} \textit{ AND } x_M \textit{ is } A_{NM}, \\
\textit{Then } y_N = a_{N0} + a_{N1}x_1 + ... + a_{NM}x_M,
\end{gathered}$$

where A_{ij} $(i = 1, 2, ..., N$ and $j = 1, 2, ..., M)$ are fuzzy subsets of the universal set. Geometrically, the rules of the Sugeno model 7.5.2 can be seen as an approximation of the mapping $X_1 \times X_2 \times ... \times X_M \rightarrow Y$ by a piecewise linear function. In addition, there are no constraints on the functions $y_i's$. In fact, in a more general setting, these linear functions in the rules's consequents can be replaced by non-linear ones. In this sense, each rule in the TSK model above may be in the following format:

$$\begin{gathered}
\textit{If } x_1 \textit{ is } A_{i1} \textit{ AND } x_2 \textit{ is } A_{i2} \textit{ AND } x_M \textit{ is } A_{iM}, \\
\textit{Then } y_i = f_i\,(x_1, x_2, ..., x_N),
\end{gathered}$$

where the $y_i's$ are non-linear functions. Clearly, the simplest TSK model is that whose functions $y_i's$ are constant values, which is a special case of the system of fuzzy rules above, when the coefficients $a_{ij} = 0$ $(i = 1, 2, ..., N$ and $j = 1, 2, ..., M)$.

The inference procedure in the TSK approach is defined by the average of the crisp outputs y_i weighted by the degree of firing τ_i of each rule. Thus, the crisp output y inferred by this method is given by:

$$y = \frac{\sum_{i=1}^{N} \tau_i y_i}{\sum_{j=1}^{N} \tau_j}, \tag{7.34}$$

where the degree of firing, τ_i, of the i^{th} is, such that:

$$\tau_i = A_{i1}(x_1) \otimes A_{i2}(x_2) \otimes ... \otimes A_{iN}(x_M), \tag{7.35}$$

where the $\otimes$ is any t-norm operation (the most commonly used is the *min* operator) and the inputs to the model, $(x_1, x_2, ..., x_M)$, are necessarily crisp numbers.

The great advantage of the TSK models is its power to describe highly non-linear systems using a small number of rules. Moreover, due to the explicit functional form, the outputs can be found from a database using some learning algorithms, of which the most used is the neuro-fuzzy systems such as ANFIS (Jang, 1993). In health sciences, specially in epidemiology, the TSK approach

is not very applied due to, among other things, the absence of reliable data. Nevertheless, when the problem is not very complex it may be described by TSK models, whose outputs of rules are constant functions (Duarte *et al.*, 2006; Sousa *et al.*, 2006).

7.5.3 The Standard Additive Model

The *Standard Additive Model* (SAM) was introduced by B. Kosko in 1996 (Kosko, 1997) and consists in a fuzzy model composed by N parallel rules, whose antecedents and consequents are fuzzy sets. However, in spite of this fact, the SAM approach shows many differences in relation to the Mamdani inference procedure. The inference method of SAM is similar to that used in the TSK model, since both of them apply the weighted sum in order to aggregate the individual output of the set of rules into a final conclusion.

Both in the Mamdani and SAM models the inference methodology produces their fuzzy conclusions through transformations on the output sets y_i, considering the membership function of the i^{th} consequent set and the degree of firing of that rule. This transformation is called *clipping approach* in the Mamdani's case and *scaling approach* in the SAM's model. In the clipping method the output set has its membership function cut off in the top, whose $\alpha - cut$ value is equal to the degree of firing for that rule. On the other hand, in the scaling method the membership function is scaled down in the proportion of the degree of firing (Yen & Langari, 1999). Figure 7.10 illustrates the scaling method for fuzzy inference used in the SAM models, considering a *If-Then* rule with an antecedent fuzzy set, A_i, and a crisp input value, x^*.

In the SAM model the inputs are necessarily crisp numbers and the inference procedure produces an output fuzzy set that must be defuzzified by the *centroid* (Center of Area) method. Thus, considering a MISO SAM model composed by N rules, in which the $i-th$ rule is

$$If\ x_1\ is\ A_{i1}\ AND\ x_2\ is\ A_{i2}\ Then\ y\ is\ B_i;\ (i = 1, 2, .., N)$$

and a given crisp input numbers, $x_1 = x_1^0$ and $x_2 = x_2^0$, the output of the model is

$$y = centroid\left(\sum_{i=1}^{N}\left[A_{i1}(x_1^0)\ A_{i2}(x_2^0)\ B_i(y)\right]\right), \tag{7.36}$$

where the centroid is the function that performs the Center of Area defuzzification.

Equation (7.36) can be re-written in a more computable form:

$$y = \frac{\sum_{i=1}^{N}\left(A_{i1}(x_1^0)\ A_{i2}(x_2^0)\right)\ A_i\ c_i}{\sum_{i=1}^{N}\left(A_{i1}(x_1^0)\ A_{i2}(x_2^0)\right)\ A_i}\quad , \tag{7.37}$$

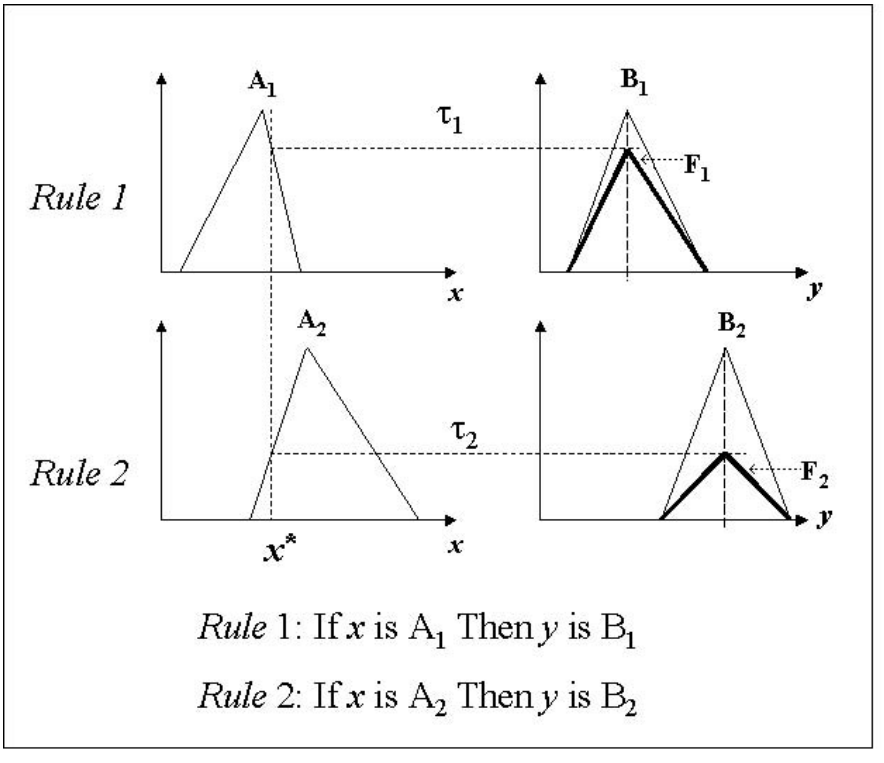

Fig. 7.10. The scaling method for fuzzy inference used in SAM models (modified from Yen and Langari, 1999)

where A_i is the area under the output B_i of the i^{th} rule and c_i is the centroid of the B_i, i.e.,

$$A_i = \int B_i(y)dy \tag{7.38}$$

and

$$c_i = \frac{\int yB_i(y)dy}{\int B_i(y)dy}. \tag{7.39}$$

Due to the aggregation form expressed in (7.37) this inference procedure proposed by Kosko is called *additive model*. It is based on the sup-product composition and the use of "addition" as a rule aggregation operator (Yen & Langari, 1999). In addition, note that the inference in the SAM model is easily computable, since A_i and c_i are constants once the rules are defined.

In this section we presented the three most important fuzzy rule-based systems: Mamdani, Takagi-Sugeno-Kang and SAM models. To illustrate the performance and the possibilities of the Mamdani inference in public health

problems, we will present three models in the next section: a *fuzzy linguistic model for HIV Natural History* (Massad *et al.*, 2003); a *fuzzy model to estimate the risk of neonatal death* (Nascimento & Ortega, 2002); and a *fuzzy model to quality of life evaluation* (Costa *et al.*, 2004). In chapter 8 we will present a TSK and a Mamdani models applied to study the canine rabies distribution in a dynamical context.

7.6 Modeling Health Decisions through Fuzzy Linguistic Systems

7.6.1 A Fuzzy Model for HIV Natural History

It is currently accepted that the human immunodeficiency virus (HIV) concentration in the circulating blood (viraemia) determines the clinical course of the infection. It has been demonstrated that the average time between infection and the development of the acquired immunodeficiency syndrome (AIDS) in untreated individuals is around ten years (Mellors *et al.*, 1996). However, a significant proportion of individuals progresses rapidly to AIDS within five or less years of infection (Mellors *et al.*, 1996; Bravo *et al.*, 1995). On the other hand, about 12% of infected individuals remain free of AIDS for at least 20 years. These different outcomes are related, among other factors, to the viral load attained soon after the infection (Mellors *et al.*, 1996).

Of central importance in therapy and control of the infection is the establishment of tools to predict the clinical course of individuals after the initial phase of the infection. However, the relationship between potential indicators and clinical course is still plagued by several uncertainties. In what follows we present a model, whose objective is to establish a relationship between the viral load, the $CD4+$ cells and the clinical progression to AIDS in HIV infected individuals.

HIV infected individuals can be classified *roughly* in 4 categories according to the length of the asymptomatic phase: a *relatively small* group of individuals progresses to AIDS within *approximately 5* years; the *majority* of HIV-infected individuals progresses to disease after *5 to 10* years; a *smaller* group progresses after *10 to 15* years and the *smallest* group remains asymptomatic for *longer than 15* years. Each class of progressors can be associated with a *general pattern* of the HIV viral load in peripheral blood during the course of infection. Rapid progressors have persistently *high levels* of viral load during the entire asymptomatic phase. In *most* individuals the RNA decreases after seroconversion, and the magnitude of this reduction, as well as the duration of the lowered RNA, are indicative of the rate of disease progression. Thus, the levels of viral RNA in peripheral blood *appear* to reflect the strength of the antiviral host response. The italicized words are intended to emphasize the uncertainties related to these processes.

Another important variable in the HIV natural history is the quantity of the $CD4+$ cells, a cell of the immunological system, in the infected individual blood. In this sense, its expected that if this quantity is high in the blood, then the

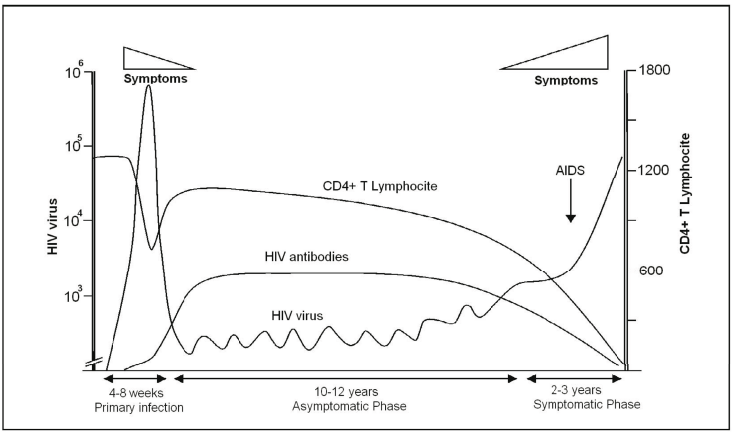

Fig. 7.11. Schematic view of the currently accepted natural history of HIV infection (Coutinho *et al.*, 2001; Perelson & Nelson, 1999)

progression to AIDS should not be fast. On the other hand, if this quantity is small, then the situation of the infected individual is not good. Figure 7.11 presents the typical behavior of the viral load and the $CD4+$ cells in the HIV infected individual.

It is possible to note in figure that in a first phase the concentration of viral load in the blood circulation grows quickly (one to two months after the infection) and it reaches a maximum, from which the viral load begins to decrease until it reaches a stationary level, where the individual stays for some years. After the incubation period, its concentration in the blood circulation starts to grow and the infected individual begins to show the first symptoms of AIDS.

This process is called *clinical progression*. It is known that the individual will progress more quickly to AIDS, the higher the level of the viral load reached in the stationary (non symptomatic) phase. For the therapeutic strategy point of view, it is important to estimate what is the HIV phase correspondent to a given infected individual. According to this, besides the exams of the viral load, it is also measured the concentration of $CD4+$, that is, the amount of lymphocyte T (antigen of $CD4+$ surface) in the blood. Thus, a high level of $CD4+$ is related with slow clinical progress, depending on the level of viral load.

In addition, the inference about what phase the patient remains requires a longitudinal follow up of the indicators mentioned, which would consume many months to analyze. In this section we present a linguistic fuzzy model that establishes relationships among the values of viral load and $CD4+$ with the clinical progression to AIDS in HIV infected individuals. The objective, therefore, it is to estimate the clinical progression of the individual, starting from a collection of laboratory tests in time.

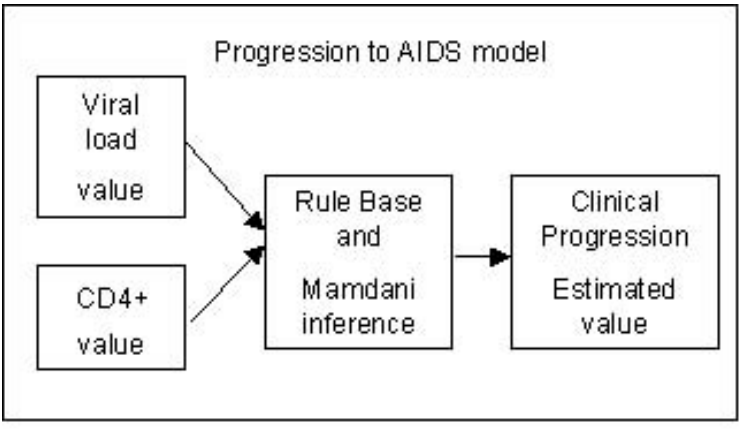

Fig. 7.12. Scheme of the Mamdani model to study the progression to AIDS based on two information: viral load and $CD4+$ counting

The model is composed by a collection of 20 parallel fuzzy rules, whose antecedent part assumes two linguistic variables: the viral load, expressed as the logarithm of the RNA-HIV counting in the plasma, and the $CD4+$ concentration, expressed as the counting of the lymphocyte $CD4+$; and the consequent part is the clinical progression, expressed as the time (in years) that the individual takes until presenting the first symptoms of AIDS (Massad *et al.*, 2003). Figure 7.12 illustrates the model scheme. One rule in this MISO Mamdani model has a following form,

If viral load is very low AND the $CD4+$ is very high, Then the progression to AIDS is very slow.

With an expert's aid we divided the domain of the input variables in the following way: the viral load into four categories, *Very Low*, *Low*, *Medium* and *High*; the $CD4+$ into five fuzzy sets, *Very High*, *High*, *Medium*, *Low* and *Very Low*; and the clinical progression to full-blown AIDS was defined in terms of the time taken from infection to the appearance of the first AIDS-defining clinical condition, as *Very Slow*, *Slow*, *Medium*, *Fast* and *Very Fast*. The membership functions of these variables were established as triangular and trapezoidal ones, whose functional forms are defined below and showed in figures 7.13, 7.14 and 7.15. The fuzzy rules of this model is presented in table 7.1.

VL: *The membership functions of viral load counting.*

$$VL \ \textit{Very Low}: \quad \mu_{VL_{VL}}(x) = \begin{cases} 1 & if \ \ x \leq 3 \\ 4 - x & if \ \ 3 < x \leq 4 \\ 0 & if \ \ x > 4 \end{cases}$$

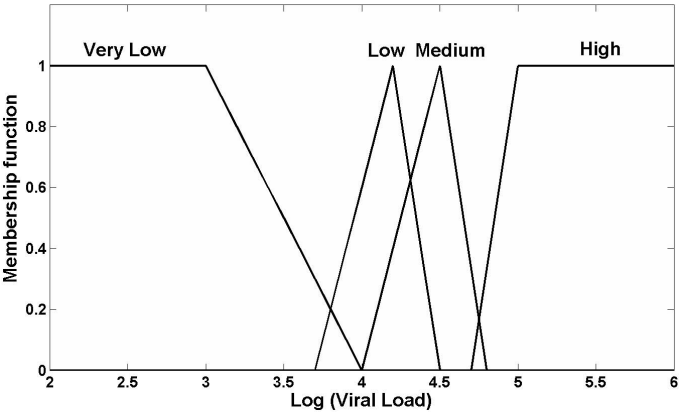

Fig. 7.13. Membership functions to Viral Load variable

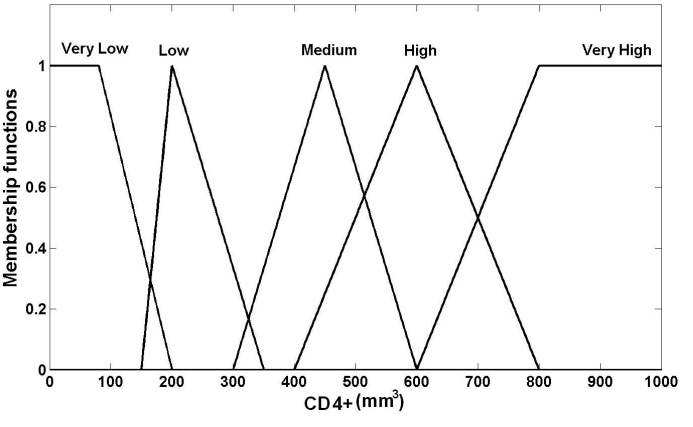

Fig. 7.14. Membership functions to $CD4+$ variable

$$VL\ Low: \quad \mu_{VL_L}(x) = \begin{cases} 0.0 & if \quad x < 3.7; x > 4.5 \\ 2.0(x - 3.7) & if \quad 3.7 \leq x < 4.2 \\ 3.\overline{3}(4.5 - x) & if \quad 4.2 \leq x \leq 4.5 \end{cases}$$

$$VL\ Medium: \quad \mu_{VL_M}(x) = \begin{cases} 0.0 & if \quad x < 4.0; x > 4.8 \\ 2.0(x - 4.0) & if \quad 4.0 \leq x < 4.5 \\ 3.\overline{3}(4.8 - x) & if \quad 4.5 \leq x \leq 4.8 \end{cases}$$

$$VL\ High: \quad \mu_{VL_H}(x) = \begin{cases} 0.0 & if \quad x < 4.7 \\ 3.\overline{3}(x - 4.7) & if \quad 4.7 \leq x \leq 5.0 \\ 1.0 & if \quad x > 5.0 \end{cases}$$

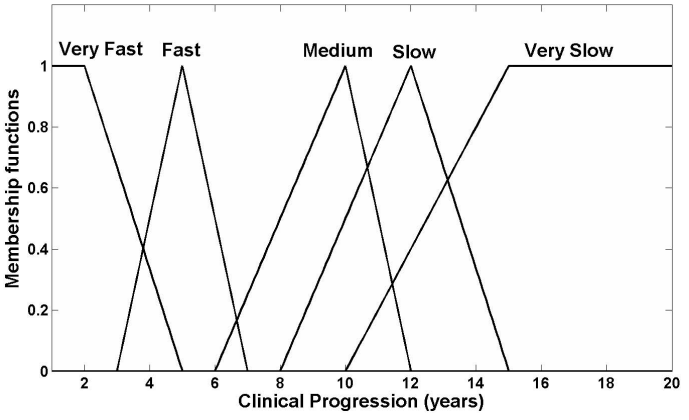

Fig. 7.15. Membership functions to Clinical Progression variable

CD4+: The membership functions of CD4+ cells.

$$CD4+ \ \ Very\ Low: \ \ \mu_{CD4+_{VL}}(x) = \begin{cases} 1.0 & if \ \ x < 80 \\ 0.008\overline{3}(200 - x) & if \ \ 80 \leq x \leq 200 \\ 0.0 & if \ \ x > 200 \end{cases}$$

$$CD4+ \ \ Low: \ \ \mu_{CD4+_{L}}(x) = \begin{cases} 0.0 & if \ \ x < 150; x > 350 \\ 0.02(x - 150) & if \ \ 150 \leq x \leq 200 \\ 0.00\overline{6}(350 - x) & if \ \ 200 \leq x \leq 350 \end{cases}$$

$$CD4+ \ \ Medium: \ \ \mu_{CD4+_{M}}(x) = \begin{cases} 0.0 & if \ \ x < 300; x > 600 \\ 0.00\overline{6}(x - 300) & if \ \ 300 \leq x \leq 450 \\ 0.00\overline{6}(600 - x) & if \ \ 450 \leq x \leq 600 \end{cases}$$

$$CD4+ \ \ High: \ \ \mu_{CD4+_{H}}(x) = \begin{cases} 0.0 & if \ \ x < 400; x > 800 \\ 0.005(x - 400) & if \ \ 400 \leq x \leq 600 \\ 0.005(800 - x) & if \ \ 600 \leq x \leq 800 \end{cases}$$

$$CD4+ \ \ Very\ High: \ \ \mu_{CD4+_{VH}}(x) = \begin{cases} 0.0 & if \ \ x < 600 \\ 0.005(x - 600) & if \ \ 600 \leq x \leq 800 \\ 1.0 & if \ \ x > 800 \end{cases}$$

Table 7.1. Fuzzy rules of the clinical progression to AIDS model

Rule	IF CV is	AND CD4+ is	THEN CP is
1	*very low*	*very high*	*very slow*
2	*very low*	*high*	*very slow*
3	*very low*	*medium*	*slow*
4	*very low*	*low*	*slow*
5	*very low*	*very low*	*medium*
6	*low*	*very high*	*very slow*
7	*low*	*high*	*slow*
8	*low*	*medium*	*medium*
9	*low*	*low*	*medium*
10	*low*	*very low*	*fast*
11	*medium*	*very high*	*medium*
12	*medium*	*high*	*medium*
13	*medium*	*medium*	*medium*
14	*medium*	*low*	*fast*
15	*medium*	*very low*	*very fast*
16	*high*	*very high*	*fast*
17	*high*	*high*	*fast*
18	*high*	*medium*	*fast*
19	*high*	*low*	*very fast*
20	*high*	*very low*	*very fast*

CP: *The membership functions of clinical progression to AIDS.*

$$CP\ Very\ Slow: \quad \mu_{CP_{VS}}(x) = \begin{cases} 0.0 & if \quad x < 10 \\ 0.20(x-10) & if \quad 10 \leq x \leq 15 \\ 1.0 & if \quad x > 15 \end{cases}$$

$$CP\ Slow: \quad \mu_{CP_{S}}(x) = \begin{cases} 0.0 & if \quad x < 8; x > 15 \\ 0.25(x-8) & if \quad 8 \leq x \leq 12 \\ 0.\overline{3}(15-x) & if \quad 12 \leq x \leq 15 \end{cases}$$

$$CP\ Medium: \quad \mu_{CP_{M}}(x) = \begin{cases} 0.0 & if \quad x < 6; x > 12 \\ 0.25(x-6) & if \quad 6 \leq x \leq 10 \\ 0.5(12-x) & if \quad 10 \leq x \leq 12 \end{cases}$$

$$CP\ Fast: \quad \mu_{CP_{F}}(x) = \begin{cases} 0.0 & if \quad x < 3; x > 7 \\ 0.5(x-3) & if \quad 3 \leq x \leq 5 \\ 0.5(7-x) & if \quad 5 \leq x \leq 7 \end{cases}$$

$$CP\ Very\ Fast: \quad \mu_{CP_{VF}}(x) = \begin{cases} 1.0 & if \quad x < 2 \\ 0.\overline{3}(5-x) & if \quad 2 \leq x \leq 5 \\ 0.0 & if \quad x > 5 \end{cases}$$

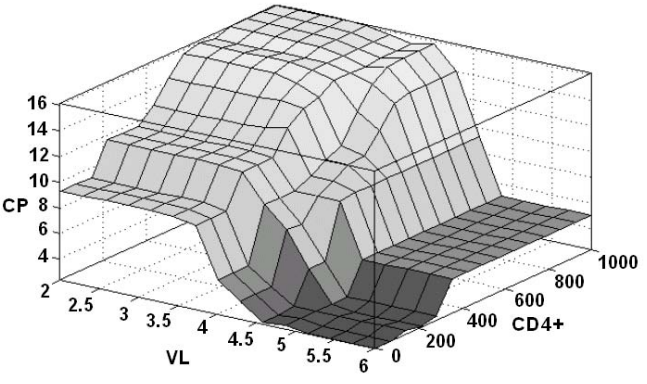

Fig. 7.16. Surface found by the HIV model mapping

The procedure of the fuzzy model consists in, given the Viral load, VL_i, and $CD4+$, $CD4+_i$ values of the individual i, the system compute the estimated value of their clinical progression to AIDS in years, through the rules base and the Centroid defuzzification method. Figure 7.16 presents the surface found by the model mapping.

Thus, as an example, suppose an individual that present a crisp value for the logarithm of HIV-RNA counts like 4.3 and the value of $CD4+$ equal to 650 counts for mm^3 of blood, then by applying the model described above the estimated output value to clinical progression is 12.3 years, considering the HIV natural history. Table 7.2 shows the estimated values to clinical progression for some pairs of viral load and $CD4+$.

Table 7.2. Some values estimated to clinical progression through the AIDS model

log(VL)	**CD4 + (count/mm^3)**	**CP(years)**
4.3	650	12.3
4.0	100	5.0
6.0	200	1.8
2.0	350	11.6
4.5	400	9.28

Its important to point out that interventions in the natural course of the disease were not consider in this model, as drug treatment. In this sense, the estimated value to the clinical progression would be useful only to guide some initials clinical interventions, helping the clinician to position their patient, at the beginning, in the curve shown in figure 7.11. Nevertheless, once the treatment has started the model no longer can provide information about the patient's clinical progression.

Drugs for treatment of HIV seropositives have always been in improvement and the reality of the AIDS disease today are quite different than that when this model was developed, in April of 1997. In fact, from the individual point of view, AIDS could be consider reasonably under control due to the development of the drugs cocktails, that have been prolonging to the maximum (and with life quality) the survival of the HIV seropositives. This is particularly true in Brazil where the health program of attention to HIV seropositives is one of the best of the world.

Although the model presented was not validated due to the lack of drugs-free data, the method proposed has demonstrated a good predictive capacity of the clinical course of HIV infection to AIDS basing on a qualitative description of the current knowledge of AIDS epidemiology. Its major advantage is the possibility of including the subjective opinion of HIV/AIDS experts on an analytical model and its simplicity. This is one of the first attempts to apply fuzzy logic and approximate reasoning to tackle a specific HIV/AIDS epidemiology problem, to the best of our knowledge. In the next section we present another example of Mamdani model applied to risk evaluation of neonatal death.

7.6.2 A Fuzzy Model to Estimate the Risk of Neonatal Death

Neonatal mortality is defined as the death that occurs up to 28 days of life and is a very important indicator of a population's heath, since it informs about social welfare, ethical and political aspects of a determined population under certain conditions. Among the main causes of neonatal mortality, low birthweight (LBW) and preterm newborn (PT) are the most important. There is a crisp classification which is used to classify children in preterm and low birthweight that is commonly used in neonatology. In this scale, those who are born with weight under 2500 g are considered to be of low birthweight and among these, those that are born with a weight lower than 1500 g are considered to be of very low birthweight. In the same way, children who are born before having completed 37 weeks of gestation are considered preterm, and extreme pre-term those who are born before having completed 32 weeks of gestation (Abrams & Newman, 1991).

The incidence of LBW and the incidence of PT in Brazil are around 10% each (Costa & Gotlieb, 1998; Bettiol *et al.*, 2000). So, estimating the risk of neonatal death can supply important information to pediatrics and especially to the neonatal intensive care physician, with respect to the attention that should be dedicated to the newborn. In São Paulo State, the most developed state of Brazil, the neonatal mortality in 2000 was 11,45/1000 livebirths. Nevertheless, a possible generator of confusion, in an automatic decision making perspective, is the Boolean classification for PT and LBW described above, because a child born with 2600 g, for instance, could not receive the necessary attention by not being considered LBW. The same could happen with a child born with 38 weeks of gestation. The main advantage of the fuzzy theory is to consider a smooth and more realistic classification of the children with respect to that two variables. Low

birthweight, extreme low birthweight, preterm and extreme pre-term newborns are the main risk factors to neonatal mortality.

Clearly the care with the newborn child could vary depending on the hospital and on its location (areas more or less developed, more populous, rural or urban zone, etc.) Not rarely, in modest hospitals the pediatrician is not there at the moment of birth, and other professionals carry out the evaluation of the newborn. This may happen even in developed countries and is the reality in most countries in development. Considering this scenario, the elaboration of a simple model that is able to evaluate more appropriately the risk of neonatal death may become an important tool, particularly if it requires low computational investment in its implementation. This is the case of the fuzzy model presented in this section.

This fuzzy linguistic model to evaluate the risk of neonatal death was also based in experts knowledge and has two antecedents: *birthweight*, fuzzily classified in *Very low birthweight* (VLBW), *Low birthweight* (LBW), *Insufficient birthweight* (IBW) and *Normal birthweight* (NBW); and *gestational age*, fuzzily classified in *Very preterm* (VPT), *Preterm* (PT) and *Term* (T). These fuzzy sets were built with fuzzification of the classical pediatrics classification. The fuzzy consequent of the model is the risk of death until 28 days, which was considered as *Very low* (VLR), *Low* (LR), *Slightly high* (SHR) and *high* (HR). Again, by simplicity, the membership functions have triangular and trapezoidal's shapes, as showed in figures 7.17, 7.18 and 7.19. In this case, the base rule was composed by 10 rules, as showed in table 7.3.

Note that, combing all possible inputs we are able to build 12 rules, but we considered as relevant only 10 rules, since there are situations that do not occur in the reality. For instance, it is impossible to find a very pre-term newborn with a normal birthweight or with insufficient birthweight. Normally a baby in this situation presents low or very low birthweight. So, although this case is mathematically possible it was not considered in the rule bases, reducing the

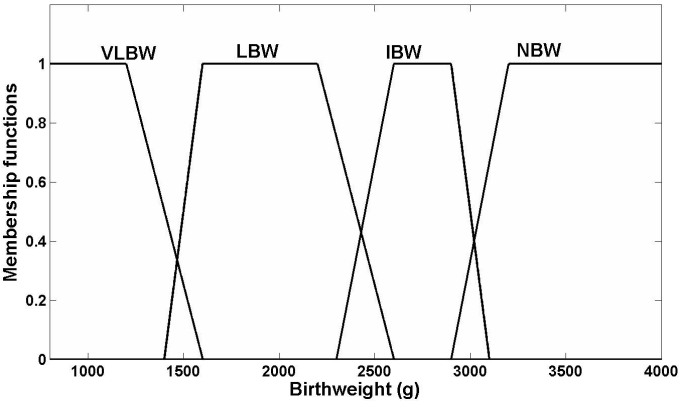

Fig. 7.17. Membership functions for the variable *birthweight* in the risk of neonatal death model (Nascimento & Ortega, 2002)

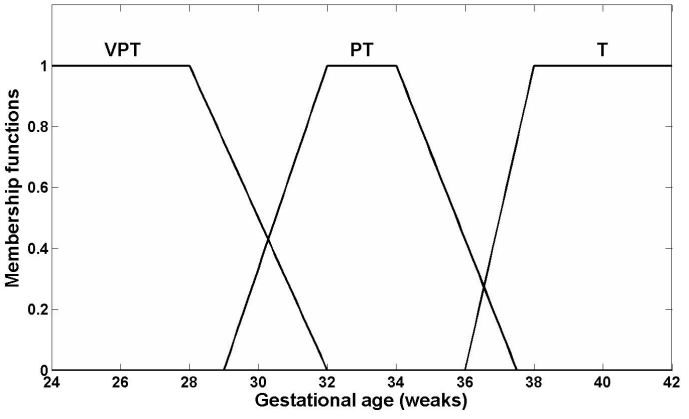

Fig. 7.18. Membership functions for the variable *gestational age* in the risk of neonatal death model (Nascimento & Ortega, 2002)

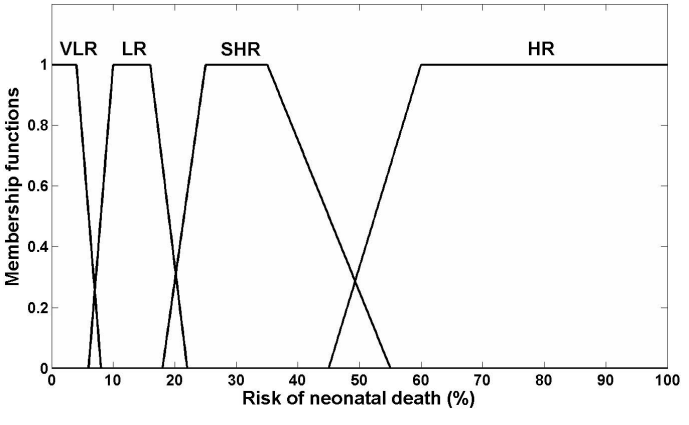

Fig. 7.19. Membership functions for the variable *risk of neonatal death* (Nascimento & Ortega, 2002)

Table 7.3. Illustration of the rules in the Mamdani model for estimation of the risk of neonatal death

Gestational age	**Birthweight**			
	Very Low	**Low**	**Insufficient**	**Normal**
Very Preterm	High	High	–	–
Preterm	High	Slightly high	Low	Low
term	Slightly high	Low	Very low	Very low

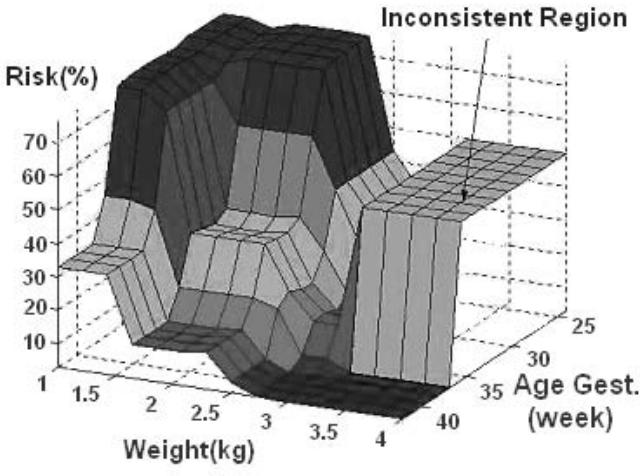

Fig. 7.20. Surface found by the mapping of the model to estimate the risk of neonatal death (Nascimento & Ortega, 2002)

number of the rules. Table 7.3 illustrates also that the rules of the model can be expressed in a matrix way, if the system has at the most two input variables.

The procedure of this fuzzy linguistic model consists in, given two of the above input for any child, calculating the membership degree of these values in all fuzzy sets of birthweight and gestational age. After that, the risk of neonatal death is determined by inference of the fuzzy rule set, using Mamdani inference, and with defuzzification of the fuzzy output. Figure 7.20 presents the results of the mapping through this fuzzy model.

We can see in this figure that the risk of neonatal death decreases monotonically when the birthweight or the gestational age increases, which is expected. The inconsistent region in this figure corresponds to the excluded rules discussed above. It means that it is impossible to occur, for instance, a newborn with a birthweight of 3200g with a gestational age of 30 weeks or a newborn with a birthweight of 4000g with a gestational age of 34 weeks.

In order to evaluate the performance of the model, 15 cases were analyzed by four other experts, whose results were compared with the model's ones. The Spearman correlation coefficient between model results and the experts opinion varies of 0.91 to 0.97 ($p < 0.001$). Considering the average of experts opinion and model results we found the Spearman correlation coefficient equal to 0.96. In addition, it was calculated the agreement test between all experts and the model as much as among experts each other. All Kappa tests were statistically significant considering the significancy level of 5% (Kaplan *et al.*, 1976). The larger p value was 0.024 for the agreement between the model versus one

expert, however, in all the remanining tests the p values were smaller that 0.01. The Kappa values were slightly better in the agreement among the experts than between the model and the experts. Nevertheless, all values were statistically equivalent ones when we considered the 95% confidence intervals.

Therefore, this fuzzy model, based only in two input variable, was sufficiently robust to estimate the risk of neonatal death when compared with the experts opinion. It is important to point out that the variables are simple measures, both of them easily available in the delivery room, even in very modest circumstances.

As expected, the agreement between the model and the experts is better in extreme situations, since this cases present less uncertainties involved. For instance, when the birthweight and the gestational age are ideal and when the birthweight and the gestational age are very critical there are few doubts about the expected outcome. On the other hand, when the birthweight and the gestational age are in intermediate situations (doubtful ones), the experts providing several and different opinions which result from their feelings and personal experiences.

Clearly, this model could be improved through the introduction of new variables, as the Apgar score. However, it is important to consider that the number of fuzzy rules grows in exponential aspects and it can compromise the model performance. Besides, the inclusion of new variables does not ensure the improvement and robustness of the model.

The application of fuzzy sets theory in pediatrics is a recent area of research (Reis *et al.*, 2004). Nevertheless, this approach has provided promising results in several medical applications, proposing a paradigmatic shift in medicine (Sadegh-Zadeh, 1999 and 2000). In this sense, the fuzzy model proposed here represents a modest contribution to this changing scenario, since the results show that fuzzy sets theory can be a powerful tool to estimate neonatal mortality and other important health indicators (see chapter 5). In fact, the fuzzy rule-based models can play a wide role in the development of systems to health decision support.

The measures of human development and of a nation's health depend on the evaluation of the life quality and of its citizens' health. For a long time the indicators of health were based on the measures of neonatal mortality, infant mortality and life expectation. However, the scientific progress in the medical fields, the control and prevention of diseases, the largest attention to childhood, the fast attendance in the case of accidents, among other actions, have been altering the profile of the societies with relationship to these traditional indicators of health. In this new scenario it becomes more and more necessary the development of indicators that take into account the citizens' health in a wider viewpoint. So, modeling emotions, feelings and the quality of life can contribute in an important way to the understanding and the evaluation of the health status in the individual and in the population as a whole. Clearly, this kind of modeling involves several identification uncertainties and fuzzy logic consists in one of the most appropriate tools to treat them.

Although the modeling of human emotions and feelings can provide interesting analysis in the behavior and psychological studies, they have been basically

applied in robotics (Shirahama *et. al*, 1999). This kind of research has as main goal the development of robots that are able to manifest human reactions like happiness, sadness, astonishment, and so on. Modeling the quality of life, however, can provide results about the individuals' health, acting directly in public health decisions (Costa *et al.*, 2004). In the next section we present an example of a model able to evaluate the disability degree, where the perception of disability is treated in a public health perspective.

7.6.3 A Fuzzy Model to Quality of Life Evaluation

Measures of health have played an essential role in the analysis of health status and quality of life, at individual and population levels. So far, they have been used, mainly, in the setting of health policy priorities and goals, and in the monitoring of medical and health care effectiveness.

As societies evolve, health problems change and new health measures are needed to adequately reflect such changes. Usually, death rates are suitable indicators of health where high levels of mortality predominate. But the increase of life expectancy, as well as of general morbidity prevalence levels, have changed the emphasis on death rates, alone, as suitable indicators of health (McDowell & Newell, 1987). Another reason for this is that the aims of the health care system have expanded in order to incorporate the achievement of physical, mental and social well being (Fanshel & Bush, 1970).

The need for the development of new health indicators based on both fatal and non-fatal health outcomes has been stressed since the late 1960's (Bergner, 1985), when functional disability began to emerge as a major public health problem worldwide due to its hazardous consequences upon economic production, social welfare and population well-being. Since then, efforts have been aimed to develop composite measures of morbidity and mortality, expressed as units of time lived, adjusted for different functional levels and ranging on a continuum from perfect health to death. Such measures are intended to provide a rationale for the allocation of health care and research resources - encompassing the prevention, treatment and rehabilitation of functional disability due to both fatal and non fatal diseases and injuries -, as well as social assistance expenditures aimed at the disabled population. Considering the limited availability of resources, policy makers need to reduce uncertainty to an utmost degree when establishing priorities and goals based on the assessment of the health status of populations (Murray,1996).

So far, a variety of methods has been developed in order to adjust the time lived in different functional levels. In general, these methods are intended to provide quantitative estimates of subjective phenomena, such as value judgements or preferences for different health states or functional levels. Opinions on which method is best vary widely, although a consensus exists around the need of a thorough understanding of the underlying mechanisms involved in functional disability measurement within quality of life research (Murray,1996).

In addition, the classic, current view of disease is that health and disease are opposites and that they are dual and contradictory attributes. It is said

that health is the absence of disease and vice versa. The fuzzy logic approach, otherwise, considers health and disease as, at least partially, complementary states, and better fits the concepts of health states and health related quality of life, as used in public health and quality of life research (Patrick & Erickson, 1993). Nevertheless, none of the methods for the assessment of quality of life developed so far has been based on a fuzzy framework. In this section we present a fuzzy linguistic model developed for measuring the degree of functional disability, and illustrate its potential use in public health.

A fuzzy linguistic model based on experts opinions was developed for measuring the degree of functional disability. Three fuzzy input variables were considered, according to social activity, mobility and physical activity.

Social activity (S) refers to the performance of activities usual for a person's age and social role, according to: play for pre-schoolers below 6 years old, study for the 6 to 17 years old age group, work and/or house keeping for the 18 to 64 years old age group, and house keeping and leisure from 65 years on. *Mobility* (M) is related to the range and to the freedom to travel from one place to another. *Physical activity* (P) is concerned mainly with walking, but includes other physical movements of the trunk and extremities, such as standing and stooping. These three dimensions of health are as those defined in the Quality of Well-being Scale, QWB (Patrick *et al.*, 1973).

A set with 100 fuzzy rules was derived, and considered as consequent for each rule the degree of functional disability (D), which describes the overall functional level of an individual, based on the previous three dimensions. The set of 100 fuzzy rules was derived by relating the fuzzy sets representing the functional levels of each input variable, namely, social activity (5 levels), mobility (5 levels) and physical activity (4 levels). A knowledge base was then developed according to the definitions presented in the original QWB scale approach (Patrick *et al.*, 1973).

The fuzzy sets related to each linguistic variables were also derived from the original QWB scale framework - not as a direct translation of its definitions -, taking into account an underlying ordinal measurement scale. Expert knowledge was acquired through a standard questionnaire. Following the presentation of essential concepts regarding fuzzy sets theory, the experts were asked to: (i) define the membership functions related to the fuzzy sets representing the function levels of each input variable as well as the output variable (functional disability); (ii) define the consequent part of each of the 100 fuzzy rules; (iii) assign a value within the [0,10] interval - the fuzzy rating scale - to each of the three fuzzy input variables; and (iv) assign a value within the [0,10] interval to each of the 100 combinations of the functional input variables, as in the QWB scale approach.

Each fuzzy rule has the form:

$$IF\ S\ is\ S_i\ \ AND\ M\ is\ M_j\ \ AND\ P\ is\ P_k\ \ THEN\ D\ is\ D_l,$$

where S, M, P and D are the fuzzy representation of social activity, mobility, physical activity and the degree of functional disability, and S_i, M_j, P_k and D_l are the fuzzy sets concerning the magnitude of S, M, P and D, respectively (see table 7.4). The linguistic terms represent different levels of functional disability,

Table 7.4. Linguistic variables and respective fuzzy sets related to functional levels (Costa et al., 2004)

Linguistic variables	**Fuzzy sets of functional levels**	**Symbol**
Social activity (S)	Extremely limited	S_1
	Very much limited	S_2
	Limited	S_3
	Somewhat limited	S_4
	Plenty	S_5
Mobility (M)	Extremely limited	M_1
	Very much limited	M_2
	Limited	M_3
	Somewhat limited	M_4
	Plenty	M_5
Physical activity (P)	Extremely limited	P_1
	Very much limited	P_2
	Somewhat limited	P_3
	Plenty	P_4
Degree of functional disability (D)	Very high	D_1
	High	D_2
	Moderate	D_3
	Low	D_4
	None	D_5

and reflect the uncertainty and imprecision that underlie the measurement of such subjective concepts.

The rules were defined by a neurology expert and only the combinations of fuzzy sets considered to be clinically meaningful and plausible were considered.

Each of the fuzzy linguistic variables are numerically represented by the set of real numbers, included in the interval [0,10]. Disposed as a rating scale, such numbers express the functional level for each dimension of health, along a continuum that ranges from optimal function to death, represented by the extreme values 10 and 0, respectively. These membership functions of input and output variables are shown in figure 7.21.

For each functional level presented in table 7.5, a crisp value was assigned on the proper fuzzy rating scale. Multidimensional functional levels, derived from the combination of single dimensional functional levels, as in the Quality of Well-being Scale (Patrick *et al.*, 1973), were evaluated through the fuzzy model, as already explained.

As an example, it is presented the evaluation of one multidimensional functional level, derived from the combination of the B categories of each single dimensional levels (see table 7.5). The crisp values (x, y, z) assigned to the social activity, mobility and physical activity fuzzy rating scales were 8, 7 and 6, respectively. In words, social activity was judged to be between *somewhat limited* and *limited*, and mobility and physical activity were evaluated as being

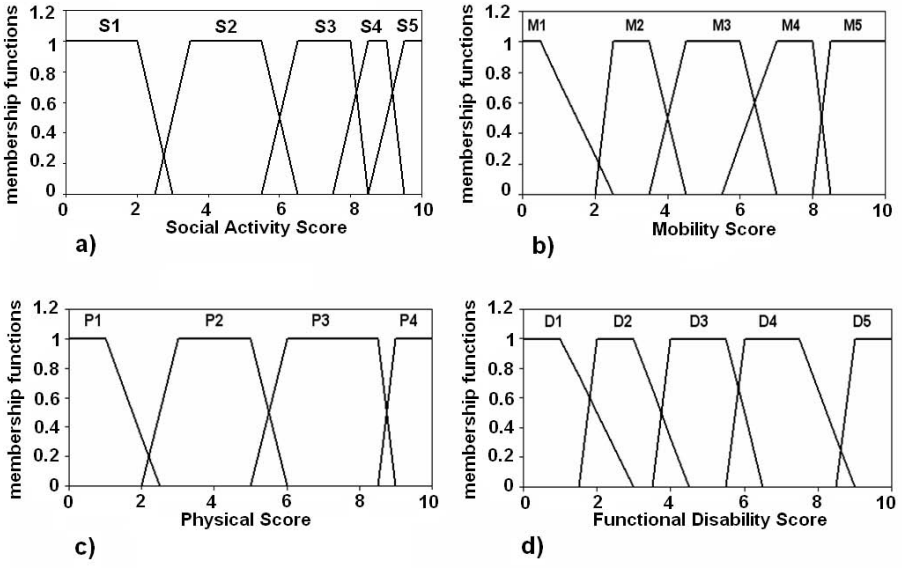

Fig. 7.21. Membership functions of *Social activity*, *Mobility*, *Physical activity* and *Degree of functional disability* (modified from Costa *et al.*, 2004)

Table 7.5. Function level of fuzzy input variables (Costa et al., 2004)

Function level	Social activity	Mobility	Physical activity
A	Did work, school, or housework, and other activities	Drove car or used public transport without help	Walked without physical limitations
B	Did work, school, or housework, but other activities were limited	Did not drive, or had help to use public transport	Walked with physical limitations
C	Limited in amount or kind of work, school or housework	In house	Moved own wheelchair without help
D	Performed self-care, but not work, school or housework	In hospital	In bed or chair
E	Had help with self-care	In special care unit	–

somewhat limited. From the membership functions presented in figure 7.21, the fuzzy output set D_l was derived. A crisp value d, expressing the degree of disability associated to the multidimensional functional level (B, B, B), was determined through the center of area defuzzification method.

In this example, d was equal to 5.75. When normalized with reference to a scale bounded by the limits 0 and 1, it simply becomes equal to 0.575. From the original Quality of Well-being Scale, d was estimated as 0.5402, a value which can be considered fairly close to the one obtained with the use of the fuzzy model.

With reference to a well known composite health indicator, the health status index (Patrick *et al.*, 1973), this estimate means that one year lived in such functional state is worth 0.575 years in perfect health, for which $d = 1$. Inversely, the value given by $1 - d$ expresses the amount of time lost in one year, due to a determined functional level (0.325 years, in the present example). Consequently, ten years lived in functional state (B, B, B), irrespective of its underlying cause, implies a loss of 3.25 years, with reference to perfect health. In case of death, d is estimated as 0.

The estimates of d varied according to the different functional levels evaluated. For less severe functional disabilities, characterized by functional states (A, B, A) and (A, A, B) (see table 7.5), d was estimated as 7.00; for more severe functional disabilities, such as states (B, C, B) and (B, B, C), lower estimates of d were obtained (4.90 and 2.80, respectively). Considering the whole range of conditions evaluated, the estimates of d tended to decrease as the functional states moved away from the absence of disabilities (*perfect health*) towards the death extremes, establishing, thus - as expected from any consistent method -, a grading system for quality of life evaluation.

In order to assess the model, we submitted all functional levels to the evaluation of two other experts. An agreement analysis was carried out, concerning the estimates of the degree of functional disability for different (three dimensional) conditions obtained through the fuzzy model and through a direct assignment on the fuzzy rating scale; this latter approach - named *direct model* - resembles the method used in the original QWB scale.

Agreement was assessed by means of the intraclass correlation coefficient (ICC), in two different ways, as follows: (i) considering the results obtained from each of the two approaches - i.e. the fuzzy model and the direct approach - separately, as defined by each of the three experts - namely, between observers agreement; and (ii) considering the results obtained from each of the two approaches for each neurology expert separately - namely, between methods agreement.

Between experts agreement was higher for the results obtained from the fuzzy model (ICC=0.666; 95%CI: 0.568 - 0.751) when compared to the estimates originated from the direct model (0.496; 95% CI: 0.220 - 0.681), considering the estimates provided by the three neurologists altogether. In this sense, the fuzzy model constitutes a consistent alternative for disability measurement, since it provides more stable results when compared with the original crisp approach.

Comparing the performance of the fuzzy model with the direct (QWB) method at the level of the single expert neurologist, results varied widely. While there was poor agreement between the two methods for expert 1 (ICC = 0.309; 95%CI: -0.065 - 0.666), the estimates provided by expert 3 were found to be reasonably reliable (ICC = 0.791; 95%CI: 0.218 - 0.919) (Landis & Koch, 1977).

Considering the average of experts opinions and model results we found the intraclass correlation coefficient equal to 0.742 (95%CI: 0.112 - 0.899). The high ICC estimates suggest that the fuzzy model is as good as the original QWB approach, in what concerns functional disability measurement. Such findings support the hypothesis about the validity of the results obtained from the fuzzy model, since the QWB scale is widely recognized as a consistent and valid approach in quality of life research (Patrick & Erickson, 1993). The estimates of d for a selected sample of conditions evaluated by the experts, as derived from the fuzzy and the *direct models*, are shown in table 7.6.

Table 7.6. Estimates of the degree of functional disability (d) obtained from the fuzzy and *direct models* (Costa et al., 2004)

Functional level	**Fuzzy model** d	**Direct model** d
AAA	0.94	1.00
BAB	0.73	0.75
CBA	0.73	0.76
CCB	0.51	0.53
ECA	0.51	0.52
ECC	0.30	0.28
EDC	0.30	0.30
EDD	0.10	0.12

An additional advantage of the fuzzy model over the crisp approach is that it allows a better understanding of how different scales - qualitative, linguistic, ordinal scales and quantitative, numerical, continuous scales - are related. The fuzzy rating scales, together with the whole set of fuzzy rules, showed how linguistic expressions, inherently vague, may be represented by numbers. By closely depicting the complex rationale underlying functional disability measurement, the fuzzy model may serve as an important tool to achieve a thorough comprehension of the mechanisms related to functional disability measurement - a central and much controversial question within quality of life research (Patrick & Erickson, 1993).

The tests performed suggest that the results obtained from the fuzzy model match those from the Quality of Well-being Scale, from which its conceptual and structural framework was derived. Such findings are encouraging, since the Quality of Well-being Scale is considered a consistent and valid approach for disability assessment and quality of life evaluation (Kaplan *et al.*, 1976). In addition, the fuzzy model provided comparable estimates of disability degree with the neurologists opinions.

The results obtained from the agreement assessment are encouraging, although they must be carefully interpreted. The fuzzy model showed a better performance in terms of between observers agreement when compared to the direct model. Since the latter approach is much similar to the original QWB

scale method, we interpreted this finding as an evidence which supports the assumption about the appropriateness of the fuzzy model to deal with an essentially subjective measurement process. However, functional disability estimates derived from experts 2 and 3 fuzzy models showed a little variation in the whole range of functional conditions evaluated; this might partially explain the higher ICC estimates observed. In other words, the results derived from experts 2 and 3 fuzzy models showed limited ability to discriminate among rather different functional levels, defined according to the three dimensions already mentioned. One possible explanation for these findings is the close resemblance among the fuzzy variables sets and the functional levels defined in the original QWB scale, which, in fact, provided the basis for the fuzzy model's structure. Such resemblance might have induced experts 2 and 3 to superpose the original QWB scale functional levels with the fuzzy variables sets, thus restricting the expert's capacity to differentiate among the varying functional states.

Another word of caution is needed as related to the intrinsic limitation of our inference model. As any inference set of rules, ours is context dependent and the variation found between experts could be partially dependent on this fact. However, the fuzzy model is very robust and allows a great deal of generalization about subjective information from patients.

Finishing this chapter we want to highlight that all fuzzy linguistic models were presented here to illustrate the power and usefulness of the fuzzy rule-based systems in epidemiology are static models, that is, they have no dynamics involved in their structures. However, the majority of models in modern epidemiology are based on dynamical systems. In fact, fuzzy dynamical systems consist in a hard area and to develop this models require the knowledge of more sophisticated mathematics, as will be seen in chapter 9. However, there is a simple way to elaborate fuzzy dynamical systems that it is based on linguistic models, as will be showed in chapter 8.

8 Fuzzy Rule-Based Dynamical Models

Epidemic dynamical systems theorists have been facing several hurdles in trying to validate their models, in particular due to several uncertainties related to variables, initial states and parameters values. These should ideally be taken from experimental work which are, quite to the contrary, demonstrating the extreme vagueness in the definition of such concepts like the force of infection, contact patterns or infected status. Therefore, a possible alternative approach could be the combination of fuzzy logic techniques with non-linear dynamical systems in order to provide a comprehensive analysis and the development of predictive tools in the epidemiology of infectious diseases.

Fuzzy dynamical systems comprise a relatively new area of research, the fundamental idea being to take a standard dynamical system modeled by a difference, or a differential equation, and then to extend this into a fuzzy set theoretical framework. The methods allow one to take into account the uncertainties related to the variables, parameters and initial states and to model their evolution whilst respecting the underlying dynamics of the system.

In linear fuzzy models the system could be modeled by a differential equation structure as proposed by Pearson (Pearson, 1997; Pearson *et al.*, 1997), based on $\alpha - levels$ and Seikkala's work (1987). However it is difficult to apply it in epidemiology because epidemic systems, in particular those dealing with infectious diseases, have strong non-linearities and should be treated in a different way. These non-linearities are due to the fact that the course of epidemic of an infectious agent, in contrast with chronic diseases, depends, among other things, on the fraction of susceptible individuals and the fraction of infectious individuals. Both susceptibility and infectiousness are intrinsically fuzzy concepts and are, therefore, ideal subjects for fuzzy logic analysis.

An approach to deal with non-linear fuzzy dynamical systems was proposed by Barros and collaborators (Bassanezi & Barros, 1995; Barros *et al.*, 2001 and 2003). They have treated ecological and epidemic systems applying fuzzy parameters in differential equations. In this case the solution of the set of equations is found from the fuzzy expected value technique (FEV)(see chapter 4). However, applying such an approach is not easy because several details should be treated

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 151–179, 2008.
springerlink.com

carefully and the calculus is very complex. So, although the FEV methodology consists in a possible way to model more realistically an epidemic system of an infectious disease, we are convinced that another approach, capable to deal with the non-linearities, should be developed.

The most common fuzzy structure applied to real problems is the rule-based model. As discussed in chapter 7, linguistic models have played an important role in fuzzy modeling, particularly in systems control. This is due to the simplicity of their inference methods and their powerful interpretation by experts. Thus, it is very convenient to develop dynamical systems using rule-based approaches. In this chapter we present a structure of dynamic linguistic fuzzy models and show some examples applied to epidemiology.

8.1 The Fuzzy Rule Dynamic Structure

Fuzzy linguistic dynamical systems are usually discrete and deterministic models. In general, a discrete and deterministic dynamical system can be represented by a set of state equations like:

$$\omega(k+1) = f\left(\omega(k), u(k)\right) \tag{8.1}$$

and

$$y(k) = g\left(\omega(k), u(k)\right), \tag{8.2}$$

where $u(k)$ and $y(k)$ are the input and output variables of the system, and $\omega(k) = [\omega_1(k), \omega_2(k), ..., \omega_n(k)]$ is the vector of state variables in the instant k (Yager & Filev, 1994). In this sense, the state value $\omega(k+1)$ and the output value $y(k)$ are completely determined by the values of $\omega(k)$ and $u(k)$, if the functions f and g are known. Usually, in engineering applications, the state variables have physical meaning as, for instance, speed, temperature, volume, etc. From the epidemic point of view, these state variables can represent the parameters of the epidemic spreading as the force of infection, vaccinating rate, recovering rate etc.

The mapping f and g in equations (8.1) and (8.2) describe the analytical relations between the input, the output and the states variables, based on the specific knowledge as the physical laws, chemistry reactions, economy theories and so on. However, there are many situations in which this kind of mapping is not available, due to the lack of knowledge about the phenomenon or the uncertainty identification inherent to the process. Besides, there are variables that are not mensurable or difficult to estimate, particularly in biomedicine and epidemiology. In these cases, an alternative is to incorporate variables as linguistic terms which can denote their values, using fuzzy sets. In this way, its is possible to transform the mathematical mapping f and g into logical rules that manipulate these linguistic terms, constructing, therefore, a discrete dynamical system where the non-linearities are described through fuzzy values of the states variables, $\omega(k+1)$, and the output, $y(k)$. Then, it is possible to use the structure of the linguistic models to formulate linguistic alternative for equations (8.1) and (8.2).

The main idea behind this kind of approach is that the system dynamics can be described by a set of rules applied iteratively, as described in Yager and Filev

(1994) and, Klir and Yuan (1995). Each rule assumes an input and an output as fuzzy sets. From the empirical experience of experts we can generate a fuzzy membership function for each variable and/or parameter, as well as the linguistic rules. Therefore, the fuzzy model consists of a set of rules and an appropriate inference machine. This linguistic model has the form:

IF U is B_1 AND W_1 is A_{11} AND ... AND W_n is A_{1n}
THEN $\overline{W}_1$ is $\widehat{A}_{11}$ AND...AND $\overline{W}_n$ is $\widehat{A}_{1n}$ AND V is D_1

IF U is B_2 AND W_1 is A_{21} AND ... AND W_n is A_{2n}
THEN $\overline{W}_1$ is $\widehat{A}_{21}$ AND ... AND $\overline{W}_n$ is $\widehat{A}_{2n}$ AND V is D_2

...

IF U is B_m AND W_1 is A_{m1} AND ... AND W_n is A_{mn}
THEN $\overline{W}_1$ is $\widehat{A}_{m1}$ AND ... AND $\overline{W}_n$ is $\widehat{A}_{mn}$ AND V is D_m,

were U is the input and W_i are the state-variables of the system; V and $\overline{W}_i$ are the output and the state variables after each iteration, respectively; B_i and A_{ij} are the input fuzzy sets and D_i and $\widehat{A}_{ij}$ are the output fuzzy sets. Therefore, by choosing an appropriate inference and a defuzzification method, after running the model each step, we get the value of each state variable that will be the input variable of the system in the following step, and so on, iteratively. It follows that:

$$U(k+1) = V(k) \tag{8.3}$$

and

$$W_i(k+1) = \overline{W}_i(k), \tag{8.4}$$

where $k+1$ is the next step after k. This model is a kind of a Markovian process and is commonly applied in the Mamdani approach. However, the same structure can be applied in TSK models, if there are enough information about the dynamics of the system. In this case part of the system behavior is known and the rules take the form:

IF $U(k)$ is B_m AND $W_1(k)$ is A_{m1} AND ... AND $W_n(k)$ is A_{mn}
THEN $y(k+1) = f(U(k), W_1(k), ..., W_n(k))$;

where $y(k+1)$ is some *a priori* function known from the system dynamics (Yager & Filev, 1994).

In order to illustrate the epidemic situations that could be treated through these dynamical structures we present two examples: a SIS (Susceptible-Infected-Susceptible) model, applied to a study about canine rabies seroprevalence in São Paulo City and, a SIR (Susceptible-Infected-Recovered) model, applied to Measles (Ortega *et al.*, 2000). We then return to the canine rabies model to discuss the problems involved in the task of elaborating the fuzzy sets in the consequent part of the rules, presenting two techniques based on the Principle of Extension (Ortega *et al.*, 2003).

8.2 A Model for Canine Rabies Seroprevalence

In this section we present an attempt to model the dynamics of rabies among a population of dogs. This study demonstrates how a dynamical system can be modeled by fuzzy linguistic rules as compared to the classical differential equations approach.

A sample of 600 street dogs from São Paulo municipal service of zoonosis control was analyzed for the presence of antibodies against the rabies virus as compared to a control sample of 50 dogs from the kennel of the São Paulo police, whose age and vaccine records are very reliable. Seroprevalence data from both samples were stratified into 4 age intervals and the age from the street dogs sample estimated by general aspects and dental observation according to the technique described in (Sallum *et al.*, 2000). The model assumes no subclinical infection since rabies is a 100% lethal infection. Therefore, animals seropositives were assumed as vaccinated and will be denoted $S^+(a)$ hereafter (a stands for *age*). Seronegative animals were assumed as susceptible to the infection and will be denoted $S^-(a)$ hereafter. The force of vaccination rate, $\nu(a)$, was also considered to be age-dependent and an additional rate, τ, meaning the loss of antibodies in the absence of new vaccinations, was also considered in this model. The system was assumed as isolated, without demographic structure, and the total population considered as a constant. Therefore, working with proportion we have that

$$S^+(a) + S^-(a) = 1. \tag{8.5}$$

The system is described, in the classical approach, by the following system of ordinary differential equations:

$$\frac{dS^-(a)}{da} = -\nu(a)S^-(a) + \tau S^+(a) \tag{8.6}$$

and

$$\frac{dS^+(a)}{da} = \nu(a)S^-(a) - \tau S^+(a), \tag{8.7}$$

and assuming (8.5) we can solve equation (8.6) and (8.7) for $\nu(a)$:

$$\nu(a) = \frac{\frac{dS^-(a)}{da} + \tau S^+(a)}{1 - S^-(a)}. \tag{8.8}$$

Table 8.1. Rabies seroprevalence of street dogs (Ortega *et al.*, 2000)

Age (years)	Proportion of seropositives
0.5	0.09
1.5	0.13
3.5	0.22
6.0	0.32

The results of the epidemiological data from the São Paulo street dogs sample are presented in table 8.1 and can be fitted to the linear equation:

$$S^{+}(a) = 0.04 + 0.0486a. \tag{8.9}$$

Therefore, the force of vaccination $\nu(a)$ can be calculated, resulting in the expression bellow:

$$\nu(a) = \frac{(0.04 + 0.0486a)\,\tau + 0.0486}{1 - (0.04 + 0.0486a)}. \tag{8.10}$$

On the other hand, the results of the epidemiological data from the São Paulo police kennel sample are presented in table 8.2 and can be fitted by the following non-linear equation:

$$S^{+}(a) = 0.9\,(1 - \exp(-0.95a))\,. \tag{8.11}$$

Table 8.2. Rabies seroprevalence of police kennel dogs (Ortega *et al.*, 2000)

Age (years)	Proportion of seropositives
0.5	0.57
1.5	0.80
3.5	0.80
6.0	0.90

The force of vaccination $\nu(a)$ can also be calculated, resulting in:

$$\nu(a) = \frac{0.9\,(1 - \exp(-0.95a))\,\tau + 0.86\exp(-0.95a)}{1 - (0.9\,(1 - \exp(-0.95a)))}. \tag{8.12}$$

These parameters were applied to the equations (8.6) and (8.7), in order to test the retrieving capacity of the classical model. Figures 8.1 and 8.2 show the real data, the fitting and the recovered data for the street dogs and the police kennel dogs, respectively.

As discussed above, the fuzzy model is comprised by a set of rules that attempts to reconstruct the system dynamics, taking into account the vagueness involved in the system variables. The rules and the fuzzy sets associated with

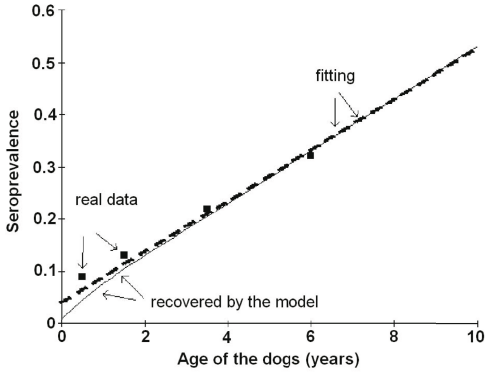

Fig. 8.1. Real data of seroprevalence of canine rabies antibodies, the fitting and the recovered data, for the street dogs sample

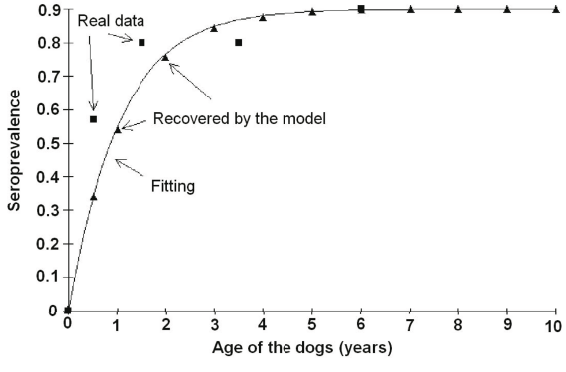

Fig. 8.2. Real data of seroprevalence of canine rabies antibodies, the fitting and the recovered data, for the police kennel dogs sample

the system were constructed according to the empirical experience of a rabies expert. In this work we developed two fuzzy dynamic systems: one model based on the Mamdani approach and, another one based on the TSK approach.

8.2.1 A TSK Model for Canine Rabies

In this model we assumed three membership functions for each of the rabies system variables and parameters: $S^+(a)$, the age-dependent proportion of seropositive dogs, $\nu(a)$, the age-dependent force of vaccination, and τ, the age-independent rate of loss of antibodies anti-rabies (as in the classical dynamical system τ^{-1} is the average period of time animals remain in the protected state, $S^+(a)$). The membership functions used for the fuzzy sets in the antecedent part of the rules, in both TSK and Mamdani approaches, were the following triangular functions:

$S^+(a)$: *The age-dependent proportion of seropositive dogs*

$$S\ Low: \quad \mu_{S_L}(x) = \begin{cases} 5x & if \ \ 0 \le x \le 0.2 \\ -5(x-0.4) & if \ \ 0.2 < x \le 0.4 \\ 0 & if \ \ x > 0.4 \end{cases}$$

$$S\ Medium: \quad \mu_{S_M}(x) = \begin{cases} 5(x-0.3) & if \ \ 0.3 \le x \le 0.5 \\ -5(x-0.7) & if \ \ 0.5 < x \le 0.7 \\ 0 & if \ \ x < 0.3; x > 0.7 \end{cases}$$

$$S\ High: \quad \mu_{S_H}(x) = \begin{cases} 5(x-0.6) & if \ \ 0.6 \le x \le 0.8 \\ 5(x-1.0) & if \ \ 0.8 < x \le 1.0 \\ 0 & if \ \ x < 0.6; x > 1.0 \end{cases}$$

$\nu(a)$: *The age-dependent force of vaccination*

$$\nu\ Weak: \quad \mu_{\nu_W}(x) = \begin{cases} 2.22(x-0.05) & if \ \ 0.05 \le x \le 0.5 \\ -3.33(x-0.8) & if \ \ 0.5 < x \le 0.8 \\ 0 & if \ \ x > 0.5 \end{cases}$$

$$\nu\ Medium: \quad \mu_{\nu_M}(x) = \begin{cases} 1.66(x-0.6) & if \ \ 0.6 \le x \le 1.2 \\ -1.25(x-2.0) & if \ \ 1.2 < x \le 2.0 \\ 0 & if \ \ x < 0.6; x > 2.0 \end{cases}$$

$$\nu\ Strong: \quad \mu_{\nu_S}(x) = \begin{cases} x-1.0 & if \ \ 1.0 \le x \le 2.0 \\ -1(x-3.0) & if \ \ 2.0 < x \le 3.0 \\ 0 & if \ \ x < 1.0; x > 3.0 \end{cases}$$

τ: *The age-independent rate of loss of antibodies*

$$\tau\ Low: \quad \mu_{\tau_L}(x) = \begin{cases} 10(x-0.3) & if \ \ 0.3 \le x \le 0.4 \\ -10(x-0.5) & if \ \ 0.4 < x \le 0.5 \\ 0 & if \ \ x < 0.3x > 0.5 \end{cases}$$

$$\tau\ Medium: \quad \mu_{\tau_M}(x) = \begin{cases} 3.33(x-0.4) & if \ \ 0.4 \le x \le 0.7 \\ -3.33(x-1.0) & if \ \ 0.7 < x \le 1.0 \\ 0 & if \ \ x < 0.4; x > 1.0 \end{cases}$$

$$\tau\ High: \quad \mu_{\tau_H}(x) = \begin{cases} 1.43(x-0.5) & if \ \ 0.5 \le x \le 1.2 \\ -1.25(x-2.0) & if \ \ 1.2 < x \le 2.0 \\ 0 & if \ \ x < 0.5; x > 2.0 \end{cases}$$

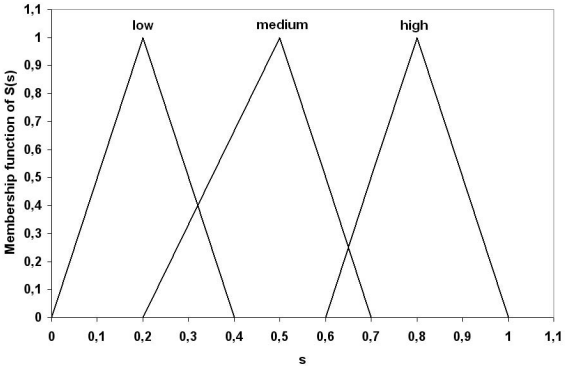

Fig. 8.3. Membership functions of fuzzy sets to proportion of protected dogs, $S(a)$

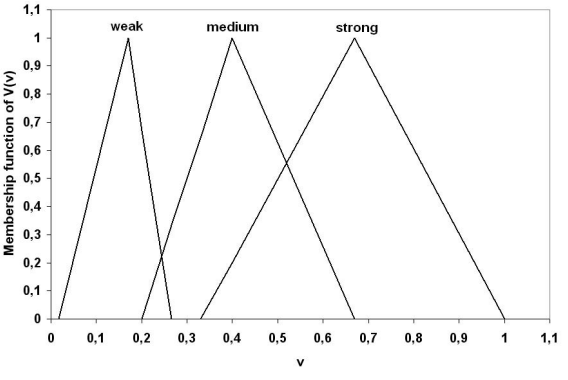

Fig. 8.4. Membership functions of fuzzy sets to force of vaccination, ν

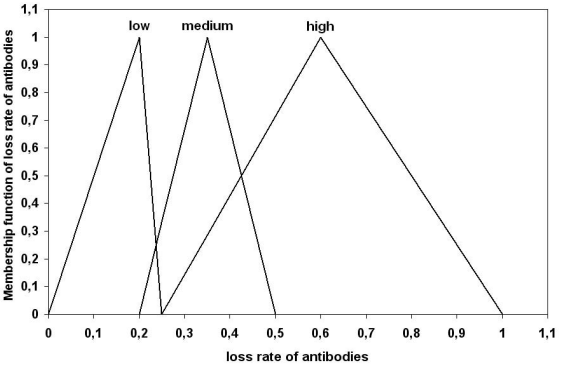

Fig. 8.5. Membership functions of fuzzy sets of rate of loss of antibodies, τ

The fuzzy sets described by the membership functions for $S^+(a)$, ν and τ are illustrated in figures 8.3, 8.4 and 8.5, respectively.

By combing all the above membership functions we elaborated 19 linguistic rules whose antecedent parts are presented in table 8.3.

Table 8.3. Antecedent part of the rules to the TSK model for canine rabies system (Ortega *et al.*, 2000)

Rule	**Antecedent part**
1	IF $S^+(a)$ *high* AND τ *low* AND $\nu(a)$ *medium*
2	IF $S^+(a)$ *high* AND τ *medium* AND $\nu(a)$ *medium*
3	IF $S^+(a)$ *high* AND τ*high* AND $\nu(a)$ *strong*
4	IF $S^+(a)$ *high* AND τ *low* AND $\nu(a)$ *strong*
5	IF $S^+(a)$ *medium* AND τ *low* AND $\nu(a)$ *medium*
6	IF $S^+(a)$ *medium* AND τ *low* AND $\nu(a)$ *weak*
7	IF $S^+(a)$ *medium* AND τ *medium* AND $\nu(a)$ *medium*
8	IF $S^+(a)$ *medium* AND τ *high* AND $\nu(a)$ *medium*
9	IF $S^+(a)$ *medium* AND τ *high* AND $\nu(a)$ *strong*
10	IF $S^+(a)$ *medium* AND τ *high* AND $\nu(a)$ *weak*
11	IF $S^+(a)$ *medium* AND τ *low* AND $\nu(a)$ *strong*
12	IF $S^+(a)$ *low* AND τ *high* AND $\nu(a)$ *weak*
13	IF $S^+(a)$ *low* AND τ *low* AND $\nu(a)$ *weak*
14	IF $S^+(a)$ *low* AND τ *low* AND $\nu(a)$ *strong*
15	IF $S^+(a)$ *low* AND τ *medium* AND $\nu(a)$ *weak*
16	IF $S^+(a)$ *low* AND τ *medium* AND $\nu(a)$ *medium*
17	IF $S^+(a)$ *low* AND τ *high* AND $\nu(a)$ *medium*
18	IF $S^+(a)$ *low* AND τ *high* AND $\nu(a)$ *strong*
19	IF $S^+(a)$ *low* AND τ *low* AND $\nu(a)$ *medium*

Since TSK models require a functional mapping between antecedent variables it is necessary to find a function that express these mathematical relations. This function is usually found through linear regression techniques, or other tools available in artificial intelligence fields, as neural networks or genetic algorithms. However, these techniques require, in general, a large data collection. In this example, where we are simply aiming at to compare the dynamics of fuzzy and classical models, its is possible to guess the output functions in the rule consequent part observing the behavior of the classical equations solutions.

In the classical approach, the iteration updating given by discrete version of the differential equation (8.7) is:

$$\Delta S^+ = \nu(a)(1 - S^+(a)) + \tau S^+(a) \tag{8.13}$$

and the consequent output inspired in this classical solution can be:

$$\Delta S^+ = \mu_{\nu_i}(\nu(a)) \left[1 - \mu_{S_i}(S^+(a))S^+(a)\right] - \tau \mu_{\tau_i}(\tau)\mu_{S_i}(S^+(a))S^+(a). \tag{8.14}$$

where $\mu_{\nu_i}(\nu(a))$ is the pertinence degree of the ν value in the fuzzy set of force of vaccination in the $i-th$ rule, $\mu_{S_i}(S)$ is the pertinence degree of the S^+ value

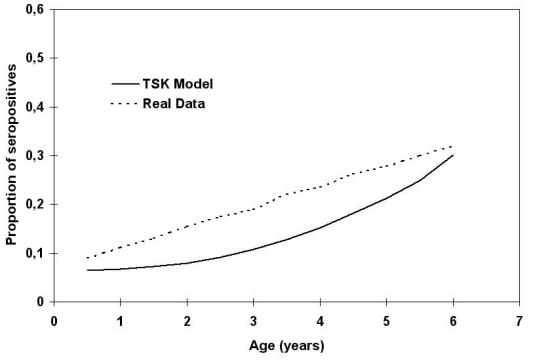

Fig. 8.6. TSK model result for the street dogs sample

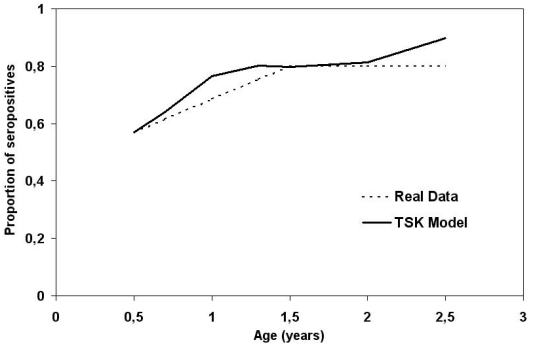

Fig. 8.7. TSK model result for the police kennel dogs sample

in the fuzzy set of the proportion of seropositive dogs, and μ_τ is the pertinence degree of the τ value in the fuzzy set of the rate of loss of antibodies.

So, all rules presented above have the structure exemplified bellow:

$$\text{IF } S^+(a) \text{ is } \textit{low} \text{ AND } \nu(a) \text{ is } \textit{weak} \text{ AND } \tau \text{ is } \textit{low}, \text{ THEN}$$
$$\Delta S_i^+ = \mu_{\nu_W}(\nu(a))\left[1 - \mu_{S_L}(S^+(a))S^+(a)\right] - \tau\mu_{\tau_L}(\tau)\mu_{S_L}(S^+(a))S^+(a)$$

where ΔS_i^+ is the increment of $S^+(a)$ according to $i-th$ rule.

After the system run over all the rules the increment is calculated by the *degree of firing* (see chapter 7), dof_i, which is then a measure of the relative importance of the $i-th$ rule, according to the t-norm operator minimum (Yager & Filev, 1994):

$$\Delta S^+ = \frac{\sum_{i=1}^{19} dof_i \Delta S_i}{\sum_{i=1}^{19} dof_i} \tag{8.15}$$

and finally we have

$$S^+(a+1) = S^+(a)\Delta S^+. \tag{8.16}$$

The updating procedures for the force of vaccination, $\nu(a)$, were, first, the computation of $\nu(a)$ from equation (8.8) and, second, the estimation of $\nu(a)$ from the numerical simulation of system (8.5). The results of the TSK model for the street dogs sample and for the police kennel sample (with $\tau = 0.33$) can be seen in figures 8.6 and 8.7, respectively.

As can be noted from figures 8.6 and 8.7, the TSK model worked reasonably well for the street dogs sample, but the results were not so good when the model was applied to the police kennel sample. We may conclude that the equation (8.14) was not capable to describe the dynamics appropriately in this case. In fact, the difficult to develop TSK models in epidemiology, and other health sciences, is related to building the consequent functions appropriately. For these reasons, in general, the linguistic dynamical models applied in these fields are based on the experts knowledge using Mamdani approach. To demonstrate the differences among those two types of models, we present in the next session a Mamdani model applied to the same study.

8.2.2 A Mamdani Model for Canine Rabies

In this model the updating increment of the proportion of seropositive animals is treated in a different way. Instead of using a mathematical function, the consequent of the rules are also fuzzy sets built by experts in this issue. So, the antecedent part in both models are the same presented in table 8.3, and the difference is only in the consequent part.

Let us take, as an example, the antecedent part of the rule number 1:

$$\text{IF } S^+(a) \text{ is } \textit{high} \text{ AND } \tau \text{ is } \textit{low} \text{ AND } \nu(a) \text{ is } \textit{medium}.$$

In this case, we expect from what is known about the actual epidemiological situation, that the variation in the proportion of seropositive animals is very low. This means that the epidemiological evidences from a scenario like the state by the rule above, points to a quasi-steady-state of the proportion of seropositives. So, basing on this kind of approach, we constructed a set of membership functions that express the epidemiological evidences for the increment in $S^+(a)$. We consider five fuzzy sets for the increment of $S^+(a)$, whose triangular membership functions are defined bellow and are illustrated in figure 8.8.

$$\Delta S^+ \ \textit{Very Small} \ : \ \mu_{\Delta S^+_{VS}}(x) = \begin{cases} 20x & if \ \ 0 \leq x \leq 0.05 \\ -20(x-0.1) & if \ \ 0.05 < x \leq 0.1 \\ 0 & if \ \ x > 0.1 \end{cases}$$

$$\Delta S^+ \ \textit{Small} \ : \ \mu_{\Delta S^+_{S}}(x) = \begin{cases} 20(x-0.05) & if \ \ 0.05 \leq x \leq 0.1 \\ -20(x-0.15) & if \ \ 0.1 < x \leq 0.15 \\ 0 & if \ \ x < 0.05; x > 0.15 \end{cases}$$

$$\Delta S^+ \ \textit{Null} \ : \ \mu_{\Delta S^+_{N}}(x) = \begin{cases} -20x + 1 & if \ \ 0.0 \leq x \leq 0.05 \\ 0 & if \ \ x > 0.05 \end{cases}$$

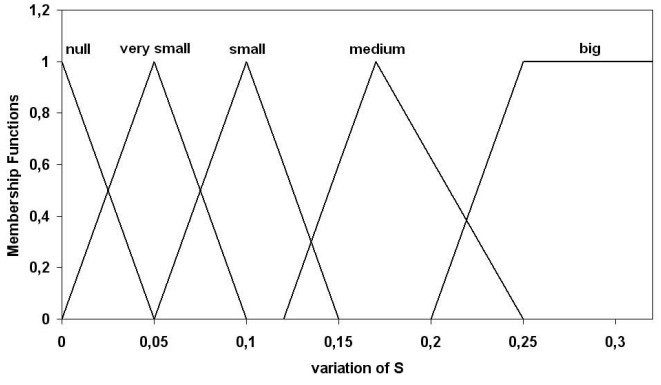

Fig. 8.8. Membership functions of fuzzy sets for the increment of $S^+(a)$

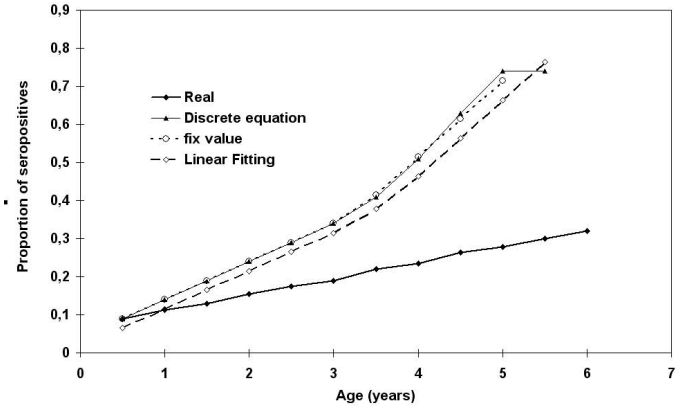

Fig. 8.9. Results found by Mamdani model with the fixed value $\tau = 0.33$ and updating the $\nu(a)$ by several methods, for street dogs

$$\Delta S^+ \ Medium \ : \ \mu_{\Delta S_M^+}(x) = \begin{cases} 20(x-0.12) & if \ \ 0.12 \leq x \leq 0.17 \\ -12.5(x-0.25) & if \ \ 0.17 < x \leq 0.25 \\ 0 & if \ \ x < 0.12; x > 0.25 \end{cases}$$

$$\Delta S^+ \ Big \ : \ \mu_{\Delta S_B^+}(x) = \begin{cases} 14.29(x-0.2) & if \ \ 0.2 \leq x \leq 0.27 \\ 1 & if \ \ 0.27 < x \leq 0.32 \\ 0 & if \ \ x < 0.2; x > 0.32 \end{cases}$$

The inference method applied was the maximum-minimum operators and the crisp output was obtained by the center of area defuzzification method (see chapter 7). The set of rules applied in this approach is presented in table 8.4.

Table 8.4. Set of rules of Mamdani approach (Ortega *et al.*, 2000)

Rule	IF $S^+(a)$ is	AND τ is	AND $\nu(a)$ is	THEN ΔS^+
1	*high*	*low*	*medium*	*very small*
2	*high*	*medium*	*medium*	*very small*
3	*high*	*high*	*strong*	*null*
4	*high*	*low*	*strong*	*null*
5	*medium*	*low*	*medium*	*small*
6	*medium*	*low*	*weak*	*small*
7	*medium*	*medium*	*medium*	*small*
8	*medium*	*high*	*medium*	*very small*
9	*medium*	*high*	*strong*	*null*
10	*medium*	*high*	*weak*	*very small*
11	*medium*	*low*	*strong*	*medium*
12	*low*	*high*	*weak*	*null*
13	*low*	*low*	*weak*	*very small*
14	*low*	*low*	*strong*	*big*
15	*low*	*medium*	*weak*	*null*
16	*low*	*medium*	*medium*	*small*
17	*low*	*high*	*medium*	*very small*
18	*low*	*high*	*strong*	*medium*
19	*low*	*low*	*medium*	*small*

As in the TSK model the force of vaccination was updating with several techniques, keeping the rate of loss of antibodies constant, $\tau = 0,33$. Figure 8.9 shows the results for the street dogs for the following situations: a) $\nu(a)$ was updated from the equation (8.13); b) $\nu(a)$ was kept constant; and c) $\nu(a)$ was updated through a linear fit given by equation (8.10). It is possible to note that this model was not sensitive to small alterations in the values of $\nu(a)$, which is due to the structure of the model. Since the different methodologies applied to treat the $\nu(a)$ values provide very similar behavior, it is reasonable to choose the simplest method and consider the variable ν as age-independent.

In the Brazilian reality, although there are annual immunization campaign against canine rabies, the street dogs are immunized just if somebody collect the dog and guide it for the vaccination (the so-called temporary adoption of street dogs). Thus, it is not expected that the force of vaccination is strong in this case. In this sense, we fixed $\nu = 0.12$, which correspond to the weak force of vaccination in the fuzzy sets. With fixed ν value we can vary the value of the rate of loss of antibodies, searching for the best parameter value for τ. Figure 8.10 illustrates the results, showing that the best value is $\tau = 0.49$. This τ value corresponds to 2.04 years of protection, what is considered a reasonable value, according to the available knowledge about the vaccines used in the Brazilian campaigns.

Assuming the value $\tau = 0.49$ it is possible to choose the best value for the variable ν. Figure 8.11 shows that, in fact, the best results are the combination of $\tau = 0.49$ and $\nu = 0.12$.

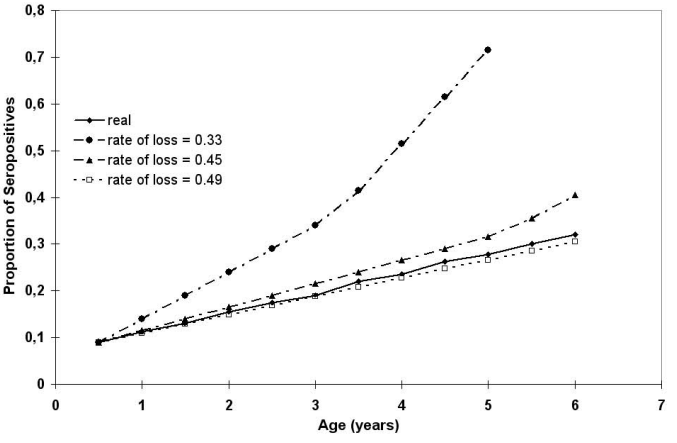

Fig. 8.10. Results found by Mamdani model with the fixed value $\nu = 0.12$ and varying the values for τ, for the street dogs

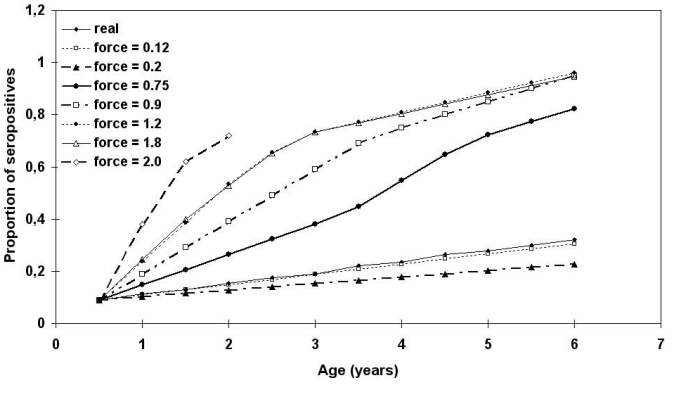

Fig. 8.11. Results found by Mamdani model with the fixed value $\tau = 0.49$ and varying the values for ν, for the street dogs

Figure 8.12 presents the correlation between the real data and the fuzzy model for the street dogs sample ($r = 0.9997$). Clearly, the Mamdani model was able to describe the dynamic of the system for the street dogs case.

The conditions of the police kennel dogs are completely different from that experienced by the dogs that live in the streets. The police kennel dogs are registered, fed correctly, taken care by specialists and vaccinated yearly. Therefore, it is expected that these dogs stay protected by a longer period and the vaccination force should be strong. Repeating the process applied in the Mamdani study, we found that for this case the best values for the vaccination force and the rate of loss of antibodies were $\tau = 0.33$ and $\nu = 1.2$, respectively. This result is showed in figure 8.13.

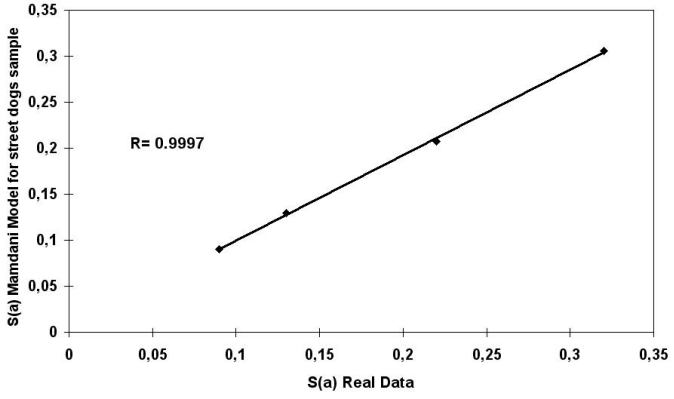

Fig. 8.12. Correlation between the real data and the results found by Mamdani model for the street dogs, $\tau = 0.49$ and $\nu = 0.12$ (Ortega *et al.*, 2000)

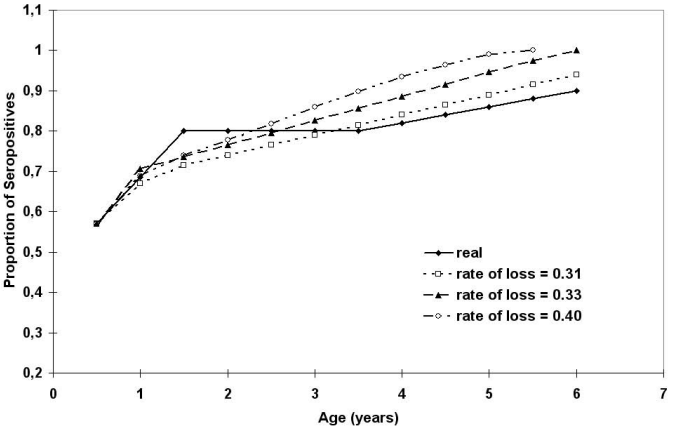

Fig. 8.13. Results found by Mamdani model with the fixed value $\nu = 1.2$ and varying the values for τ, for the police kennel dogs (Ortega *et al.*, 2000)

Therefore, the model provides results that point to a large period of protection (3.23 years) and a strong force of vaccination, which is completely in accord to the police kennel dogs reality. Figure 8.14 presents the correlation between the real data and the fuzzy model for the police kennel dogs sample ($r = 0.939$).

As happened in the TSK approach, the Mamdani model provided best results, comparing with real data, for street dogs analysis than for the police kennel dogs. This is due, in part, to the linear behavior observed in the street dogs sample. When we compared both approaches, TSK and Mamdani models, we notice that the Mamdani model has a better performance. This is a direct consequence of the fact that the chosen function to compose the consequent part of the rule

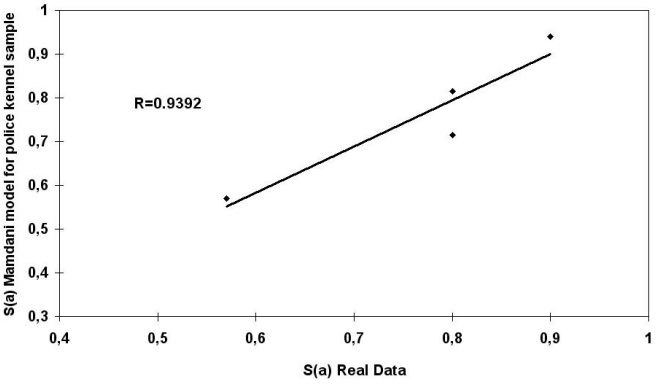

Fig. 8.14. Correlation between the real data and the results found by Mamdani model for the police kennel dogs, $\tau = 0.31$ and $\nu = 1.2$ (Ortega *et al.*, 2000)

in the TSK model is not very adapted. Clearly, another function could provide better results. So, it is important to point out that the fact that the Mamdani model to have provide better results than the TSK model does not mean that it is better than the other. On the other hand, from the Mamdani's point of view, we can conclude that the experts were able to describe the relations between the variables. Figures 8.15 and 8.16 show the experimental data compared with the results provided by the classical and Mamdani approaches for the street dogs and police kennel dogs sample, respectively.

It is possible to note that, although the classical differential equations approach presents good results, the Mamdani system is much better for recovering

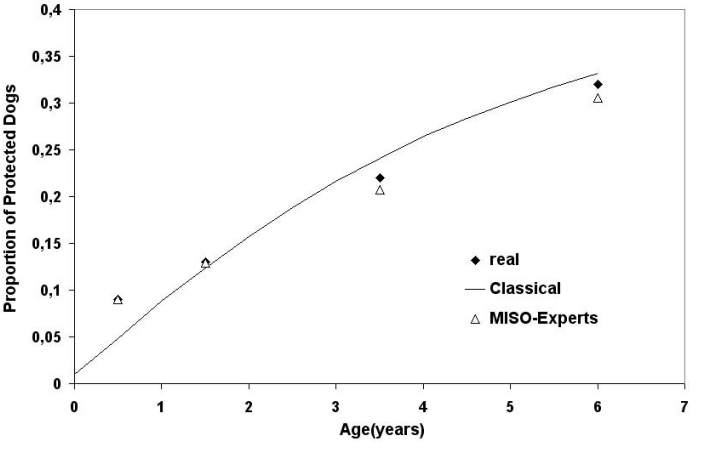

Fig. 8.15. Experimental data compared with the results provided by Mamdani and Classical approaches, for the street dogs

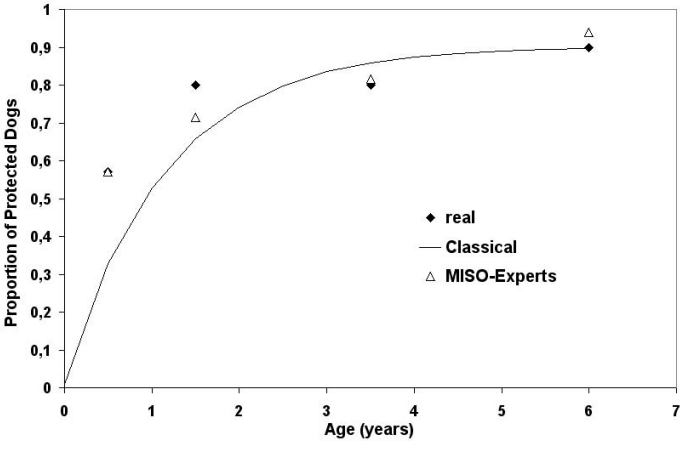

Fig. 8.16. Experimental data compared with the results provided by Mamdani and Classical approaches, for the police kennel dogs

the real data. This result demonstrates that the fuzzy linguistic dynamic systems, based on the iterative process of the fuzzy rules, could be an appropriate structure to describe the dynamical behaviors, even in non-linear systems. In order to analyze this ability in a more complex context we present in the next section an application of this structure to a SIR epidemic model.

8.3 A Fuzzy Dynamical Model for a SIR Epidemic

As discussed in chapter 3, a SIR (Susceptible-Infected-Recovered) model is, in the binary logic context, a compartmental system commonly used to describe the spreading of a microparasitic infection. This system is composed by a set of differential equations whose variables represent three possible status of the individuals, like susceptible, infected, recovered, immunized and so on (Anderson & May, 1991; Massad *et al.*, 1995). A typical example of this kind of epidemic system is for measles, whose structure allows to consider the role that the vaccination campaign plays. A measles epidemic course may be expressed by the following set of non-linear differential equations:

$$\frac{ds}{dt} = \alpha n - \alpha s - \beta si - vs \tag{8.17}$$

$$\frac{di}{dt} = -\alpha i + \beta si - \gamma i \tag{8.18}$$

$$\frac{dr}{dt} = \gamma i - \alpha r + vs \quad , \tag{8.19}$$

where s is the proportion of susceptible individuals, i is the proportion of infected individuals and r is a proportion of protected individuals, that became immunized either through vaccination or through natural infection.

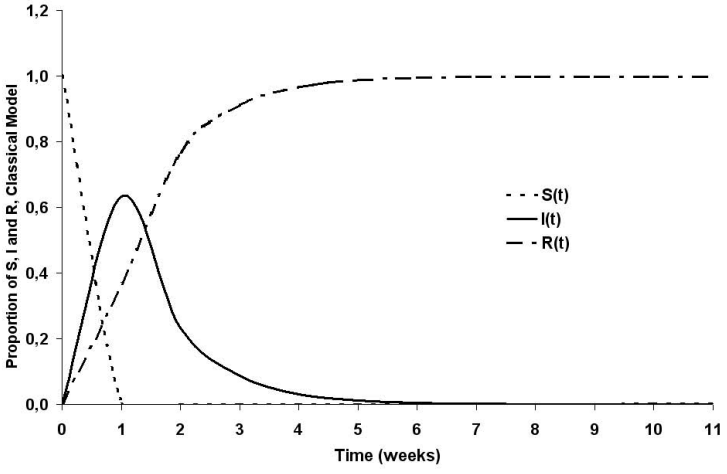

Fig. 8.17. Typical dynamic behavior of the proportions s, i and r provided by the classical model, based on the set of differential equations for the measles epidemic spreading (Ortega *et al.*, 2000)

This model has four parameters: 1) α, the rate of natural mortality in the population; 2) β, the rate of infective contact between infected and susceptible individuals; 3) ν, the force of vaccination; and 4) γ, the rate with which sick individuals recover. The number of total individuals is assumed to be fixed, that is, $n = s + i + r$, which allows to reduce a set of equations above to only two equations. Figure 8.17 shows a typical dynamical behavior of the proportions of the susceptible, infected and protected individuals.

In order to study the dynamical behavior of this model we chose the values for parameters based in a real situation of measles in São Paulo City and simulated numerically the set of differential equations above for several values of the parameters. The analysis of this simulations supplied the base to build the fuzzy sets in the linguistic dynamical model proposed.

The great advantage of a fuzzy SIR model is that it allows to consider in the dynamics the vagueness and the imprecision inherent to the individuals statuses. In this case the Mamdani model has four input variables (in the antecedent part of the rules) and two output variables (in the consequent part). The input variables are: the proportion of susceptible individuals, s; the proportion of infected individuals, i; the rate of infective contact between infected and susceptible individuals, β; and the force of vaccination against measles, ν. The fuzzy rules consequents are: the variation of s, Δs and the variation of i, Δi. We consider the following triangular membership functions for these linguistic variables:

β: *The rate of infective contact.*

$$\beta \; small: \quad \mu_{\beta_{sm}}(x) = \begin{cases} 1.0 & if \;\; 1.0 \leq x \leq 2.0 \\ -1.0(x - 3.0) & if \;\; 2.0 < x \leq 3.0 \\ 0.0 & if \;\; x > 1.0; x < 3.0 \end{cases}$$

$$\beta \ medium: \quad \mu_{\beta_m}(x) = \begin{cases} 0.2(x-2.0) & if \ \ 2.0 \le x \le 7.0 \\ -0.3(x-10.0) & if \ \ 7.0 < x \le 10.0 \\ 0.0 & if \ \ x < 2.0; x > 10.0 \end{cases}$$

$$\beta \ big: \quad \mu_{\beta_b}(x) = \begin{cases} 0.08\overline{3}(x-8.0) & if \ \ 8.0 \le x \le 20.0 \\ -0.1(x-30.0) & if \ \ 20.0 < x \le 30.0 \\ 0.0 & if \ \ x < 8.0; x > 30.0 \end{cases}$$

$$\beta \ very \ big: \quad \mu_{\beta_{vb}}(x) = \begin{cases} 0.04(x-25.0) & if \ \ 25.0 \le x \le 50.0 \\ 1.0 & if \ \ 50.0 < x \le 100.0 \\ 0.0 & if \ \ x < 25.0; x > 100.0 \end{cases}$$

ν: *The force of vaccination.*

$$\nu \ weak: \quad \mu_{\nu_w}(x) = \begin{cases} 1.0 & if \ \ 0.0 \le x \le 0.01 \\ -50.0(x-0.03) & if \ \ 0.01 < x \le 0.03 \\ 0.0 & if \ \ x > 0.03 \end{cases}$$

$$\nu \ medium: \quad \mu_{\nu_m}(x) = \begin{cases} 20.0(x-0.02) & if \ \ 0.02 \le x \le 0.07 \\ -33.\overline{3}(x-0.10) & if \ \ 0.07 < x \le 0.10 \\ 0.0 & if \ \ x < 0.02; x > 0.10 \end{cases}$$

$$\nu \ strong: \quad \mu_{\nu_{st}}(x) = \begin{cases} 4.0(x-0.05) & if \ \ 0.05 \le x \le 0.30 \\ 1.0 & if \ \ 0.30 < x \le 1.0 \\ 0.0 & if \ \ x < 0.05; x > 1.0 \end{cases}$$

s: *The proportion of susceptible individuals.*

$$s \ very \ small: \quad \mu_{s_{vs}}(x) = \begin{cases} 1.0 & if \ \ 0.0 \le x \le 0.10 \\ -10.0(x-0.20) & if \ \ 0.10 < x \le 0.20 \\ 0.0 & if \ \ x > 0.20 \end{cases}$$

$$s \ small: \quad \mu_{s_{sm}}(x) = \begin{cases} 7.7(x-0.12) & if \ \ 0.12 \le x \le 0.25 \\ -6.\overline{6}(x-0.40) & if \ \ 0.25 < x \le 0.40 \\ 0.0 & if \ \ x < 0.12; x > 0.40 \end{cases}$$

$$s \ medium: \quad \mu_{s_m}(x) = \begin{cases} 5.0(x-0.30) & if \ \ 0.30 \le x \le 0.50 \\ -10.0(x-0.60) & if \ \ 0.50 < x \le 0.60 \\ 0.0 & if \ \ x < 0.30; x > 0.60 \end{cases}$$

$$s\ big: \quad \mu_{s_b}(x) = \begin{cases} 5.0(x-0.50) & if \ \ 0.50 \le x \le 0.70 \\ -10(x-0.80) & if \ \ 0.70 < x \le 0.80 \\ 0.0 & if \ \ x < 0.50; x > 0.80 \end{cases}$$

$$s\ very\ big: \quad \mu_{s_{vb}}(x) = \begin{cases} 5.0(x-0.70) & if \ \ 0.70 \le x \le 0.90 \\ 1.0 & if \ \ 0.90 < x \le 1.0 \\ 0.0 & if \ \ x < 0.70 \end{cases}$$

i: *The proportion of infected individuals.*

$$i\ very\ small: \quad \mu_{i_{vs}}(x) = \begin{cases} 1.0 & if \ \ 0.0 \le x \le 0.03 \\ -50.0(x-0.05) & if \ \ 0.03 < x \le 0.05 \\ 0.0 & if \ \ x > 0.05 \end{cases}$$

$$i\ small: \quad \mu_{i_{sm}}(x) = \begin{cases} 25.0(x-0.03) & if \ \ 0.03 \le x \le 0.07 \\ -12.5(x-0.15) & if \ \ 0.07 < x \le 0.15 \\ 0.0 & if \ \ x < 0.03; x > 0.15 \end{cases}$$

$$i\ medium: \quad \mu_{i_m}(x) = \begin{cases} 5.0(x-0.10) & if \ \ 0.10 \le x \le 0.30 \\ -10.0(x-0.40) & if \ \ 0.30 < x \le 0.40 \\ 0.0 & if \ \ x < 0.10; x > 0.40 \end{cases}$$

$$i\ big: \quad \mu_{i_b}(x) = \begin{cases} 10.0(x-0.30) & if \ \ 0.30 \le x \le 0.40 \\ -5.0(x-0.60) & if \ \ 0.40 < x \le 0.60 \\ 0.0 & if \ \ x < 0.30; x > 0.60 \end{cases}$$

$$i\ very\ big: \quad \mu_{i_{vb}}(x) = \begin{cases} 4.0(x-0.50) & if \ \ 0.50 \le x \le 0.75 \\ 1.0 & if \ \ 0.75 < x \le 1.0 \\ 0.0 & if \ \ x < 0.50 \end{cases}$$

Δs: *Variation of the proportion of susceptible individuals.*

$$\Delta s\ very\ small: \quad \mu_{\Delta s_{vs}}(x) = \begin{cases} 1.0 & if \ \ 0 \le x \le 0.03 \\ -33.\overline{3}(x-0.06) & if \ \ 0.03 < x \le 0.06 \\ 0.0 & if \ \ x > 0.06 \end{cases}$$

$$\Delta s\ small: \quad \mu_{\Delta s_{sm}}(x) = \begin{cases} 20.0(x-0.05) & if \ \ 0.05 \le x \le 0.10 \\ -10.0(x-0.20) & if \ \ 0.10 < x \le 0.20 \\ 0.0 & if \ \ x < 0.05; x > 0.20 \end{cases}$$

$$\Delta s \ null: \quad \mu_{\Delta s_{null}}(x) = \begin{cases} 1.0 & if \quad x = 0.0 \\ 0.0 & if \quad x \neq 0.0 \end{cases}$$

$$\Delta s \ medium: \quad \mu_{\Delta s_m}(x) = \begin{cases} 20.0(x - 0.15) & if \quad 0.15 \leq x \leq 0.20 \\ -10.0(x - 0.30) & if \quad 0.20 < x \leq 0.30 \\ 0.0 & if \quad x < 0.15; x > 0.30 \end{cases}$$

$$\Delta s \ big: \quad \mu_{\Delta s_b}(x) = \begin{cases} 6.\overline{6}(x - 0.25) & if \quad 0.25 \leq x \leq 0.40 \\ 1.0 & if \quad 0.40 < x \leq 1.0 \\ 0.0 & if \quad x < 0.25 \end{cases}$$

and, finally:

Δi: *Variation of the proportion of infected individuals.*

$$\Delta i \ very \ small \ < 0: \quad \mu_{\Delta i_{vs<0}}(x) = \begin{cases} 1.0 & if \quad -0.02 \leq x \leq 0.00 \\ 25.0(x + 0.06) & if \quad -0.06 < x \leq -0.02 \\ 0.0 & if \quad x < -0.06; x > 0.00 \end{cases}$$

$$\Delta i \ small \ < 0: \quad \mu_{\Delta i_{sm<0}}(x) = \begin{cases} -25.0(x + 0.04) & if \quad -0.08 \leq x \leq -0.04 \\ 25.0(x + 0.12) & if \quad -0.12 < x \leq -0.08 \\ 0.0 & if \quad x < -0.12; x > -0.04 \end{cases}$$

$$\Delta i \ medium \ < 0: \quad \mu_{\Delta i_{m<0}}(x) = \begin{cases} -25.0(x + 0.08) & if \quad -0.12 \leq x \leq -0.08 \\ 25.0(x + 0.16) & if \quad -0.16 < x \leq -0.12 \\ 0.0 & if \quad x < -0.16; x > -0.08 \end{cases}$$

$$\Delta i \ big \ < 0: \quad \mu_{\Delta i_{b<0}}(x) = \begin{cases} -6.\overline{6}(x + 0.15) & if \quad -0.30 \leq x \leq -0.15 \\ 1.0 & if \quad -1.0 < x \leq -0.30 \\ 0.0 & if \quad x < -1.0; x > -0.15 \end{cases}$$

$$\Delta i \ null: \quad \mu_{\Delta i_{null}}(x) = \begin{cases} 1.0 & if \quad x = 0.0 \\ 0.0 & if \quad x \neq 0.0 \end{cases}$$

$$\Delta i \ very \ small \ > 0: \quad \mu_{\Delta i_{vs>0}}(x) = \begin{cases} 1.0 & if \quad 0.00 \leq x \leq 0.02 \\ -25.0(x - 0.06) & if \quad 0.02 < x \leq 0.06 \\ 0.0 & if \quad x < 0.00; x > 0.06 \end{cases}$$

$$\Delta i \ small \ > 0: \quad \mu_{\Delta i_{sm>0}}(x) = \begin{cases} 25.0(x - 0.04) & if \quad 0.04 \leq x \leq 0.08 \\ -25.0(x - 0.12) & if \quad 0.08 < x \leq 0.12 \\ 0.0 & if \quad x < 0.04; x > 0.12 \end{cases}$$

$$\Delta i \ medium \ > 0: \quad \mu_{\Delta i_{m>0}}(x) = \begin{cases} 25.0(x - 0.08) & if \ \ 0.08 \leq x \leq 0.12 \\ -25.0(x - 0.16) & if \ \ 0.12 < x \leq 0.16 \\ 0.0 & if \ \ x < 0.08; x > 0.16 \end{cases}$$

$$\Delta i \ big \ > 0: \quad \mu_{\Delta i_{b>0}}(x) = \begin{cases} 6.\overline{6}(x - 0.15) \ if & 0.15 \leq x \leq 0.30 \\ 1.0 & if \ \ 0.30 < x \leq 1.00 \\ 0.0 & if \ \ x < 0.15; x > 1.00 \end{cases}$$

The combination of all fuzzy sets results in 300 possible rules. Each situation, described by each rule, was carefully analyzed and 157 rules were considered relevant in the model. Each selected rule has a form like:

IF s is *big* AND β is *medium* AND ν is *weak* AND i is *small*,
THEN Δs is *medium* AND Δi is *medium positive*.

The model was run with maximum-minimum inference and the defuzzification method applied was the center of area. Figure 8.18 presents a typical dynamical behavior of the proportions s, i and r for the fuzzy model.

As can be noted from figures 8.17 and 8.18, the Mamdani model provides, in a qualitative way, a dynamical behavior very similar to that found from the classical differential equations. This model was simulated for several parameters values and the results were compared with those from classical model. The fuzzy model showed, in all simulations, a dynamical behavior compatible with those presented by the classical SIR model.

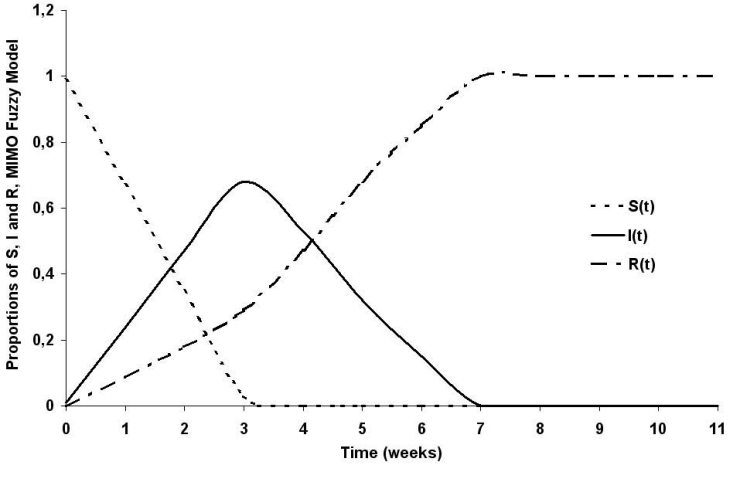

Fig. 8.18. Typical dynamic behavior of the proportions s, i and r provided by fuzzy model for the measles epidemic spreading (Ortega *et al.*, 2000)

The results provided by the fuzzy SIR model suggest that the linguistic fuzzy dynamic models can be an interesting approach to treat non-linear systems, when uncertainties and imprecisions are involved. Its structure allows to aggregate into the model experts knowledge and to build mathematical mappings, even when the functional informations are not available. However, this example showed also that this kind of modeling can became a tough task due to the large number of the possible rules. In fact, if the number of input variables considered in the system is very large, the development of a fuzzy linguistic dynamical model could be not viable, particularly because of the expert dependence that this kind of models have. However, it is important to point out that this limitation is due not only to the dynamical aspects. As discussed in chapter 7, the most important limitation of the fuzzy rule-based models, particularly Mamdani ones, is the explosion of the number of the possible rules caused by the combination of the input variables.

Nevertheless, since linguistic models assume great importance in the biomedical fuzzy modeling, it is worthwhile to study alternatives and develop techniques that allow to build the fuzzy rules in a way less dependent on the information supplied by experts (Ortega *et al.*, 2003). In the next section we present a deeper discussion about the role of experts in the elaboration of fuzzy linguistic models and present an alternative to treat these problems using the concept of the *Extension Principle*.

8.4 A Fuzzy Linguistic Dynamical Model Based on the Extension Principle

As pointed out before, there are some situations where we may have functional informations about the behavior of the system. In this case the model becomes less dependent on the experts opinion and this may be an advantage. Sometimes one has a large data collection about the system behavior besides the opinion of experts. As discussed by Wang and Mendel (1992), in this situation an interesting way to design a model is to mix the experience of the human controller and a sample of input-output pairs of the system. This situation could be treated with hybrid models like fuzzy genetic algorithms and fuzzy neural network (Wang & Mendel, 1992; Jang, 1993; Bastian, 2000). This is the case of several engineering applications but, unfortunately, this is not the case of the majority of epidemiological problems.

Our own experience in dealing with biomedical problems and fuzzy modeling have demonstrated that in the case of biological systems, and particularly in epidemiology, it is often hard to find functional information about the dynamics of the systems. In this context, the academic and heuristic knowledge of experts, as well as their experience, assume a fundamental role in this kind of modeling. So, the majority of linguistic epidemiological models are built based on the empirical experience of a panel of experts. With this we are able to generate a membership function, which defines the fuzzy sets for each variable and/or parameters, as well as the linguistic rules that govern the system

dynamics. These membership functions describe all possible values that the variable could assume (a physically realistic domain). So, in this case, the experts elaborate the fuzzy sets input and output, besides the linguistic rules, which clearly cause an important dependence on the experts ability to describe the phenomenon.

8.4.1 Dealing with the Opinions of Experts

Elaborating fuzzy models with physicians and epidemiologists requires, in general, interdisciplinary relationships. In order to extract their knowledge it is very important to leave them quite free to build the fuzzy sets. Commonly, the fuzzy sets in epidemiological problems elaborated by experts are not "well-behaviored" sets. They tend to be asymmetric and irregular, different from that found in engineering applications. In addition, experts may have serious problems to insighting both the antecedents and the consequents of the rules when the model is too complex. Also, creating the consequents is a much more difficult task than the antecedents because in the former the expert needs to consider the dynamics of the system, weighting all influences that could concur, generating one specific output and its corresponding membership function. In contrast, in order to create the antecedents, the expert needs only to classify the variables in groups, elaborating their membership functions. Therefore, in general, the expert has more facility to elaborate the antecedents than the consequents. In this sense, a method that allows the elaboration of the consequents of the linguistic rules could represent an important progress in the modeling of systems, which have a high level of uncertainties, impreciseness and/or vagueness in the variables, parameters or both.

In the great majority of epidemic studies the most important result is the prediction capacity of the model. In general, we are interested in predicting the future space-time conditions based on analysis of the model results or of experimental data, to decide which strategies to apply or which public heath decisions to take. In this sense, it is important to consider that the mathematical model should be the most comprehensible possible, because the adhesion to the model's results by the decision makers will depend on this understanding. This goal is usually reach by models based on fuzzy sets theory in which linguistic variables are applied.

Considering both the difficulties of experts to build the consequents of the fuzzy rules and the specific conditions of real epidemic case, we studied the application of the Extension Principle to elaborate the consequent fuzzy sets in a rule-based model. This methodology was firstly proposed by Dubois and collaborators (Dubois *et al.*, 1995) and reformulated by Ortega and collaborators (Ortega *et al.*, 2003). To illustrate how this technique works we will return to the canine rabies dynamic study, for its simplicity. In the next section we briefly discuss the extension principle and the Dubois *et al.* proposal.

8.4.2 The Extension Principle Methodology Applied in the Canine Rabies Study

As discussed in chapter 2, the Extension Principle, which was first proposed by Zadeh in 1975, is used to produce a functional that maps fuzzy sets to fuzzy sets from a crisp function (Zadeh, 1975a, 1975b and 1975c). This tool emerges from the necessity to apply crisp functions f to imprecise arguments. In this sense, if we have a crisp function f and we need to apply this function to fuzzy arguments, we can use the extension principle. This fuzzy argument could be described as a possibility distribution of the argument of the function f. So, for each possible value the function's variable may assume, the functional produces its possible image providing the distribution of the possibility of this image. Depending on the function it could happen that different input values are mapped to the same output value, for instance, to a non-injective function. In this case we need to determine the possibility of such an output value, by combining the possibility degree of all inputs that map to the same output value, which could be done by using a disjunction operator.

Furthermore, we may end up with different possibility distributions for each argument $(x_1, x_2, ..., x_n)$ of the function f. In this case, we only need to apply a fuzzy conjunction operator to decide what would be the possibility degree of its image.

Thus, the extension principle is a very useful tool in the investigation of the action of a function over fuzzy sets, that is, to find the images of the fuzzy sets from a crisp function. This concept was defined in chapter 2 for a univariate function, but it could be generalized for a multivariate one.

Consider a crisp function

$$f : X_1 \times X_2 \times ... \times X_n \to Y$$

and $A_1, A_2, ..., A_n$, fuzzy subsets of $X_1, X_2, ..., X_n$, respectively. The extension principle produces a function $\widehat{f}$, that is the fuzzification of f, whose image $\widehat{f}(A_1, A_2, ..., A_n)$ is the fuzzy subset of Y, whose membership function is given by:

$$\mu_{\widehat{f}(A_1,...,A_n)}(y) = \begin{cases} \sup_{\bar{x} \in f^{-1}(y)} [\min(\mu_{A_1}(x_1), ..., \mu_{A_n}(x_n))] \; if \; f^{-1}(y) \neq \emptyset \\ \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad if \; f^{-1}(y) = \emptyset \end{cases} \tag{8.20}$$

where $\bar{x} = (x_1, x_2, ..., x_n)$ and $f^{-1}(y) = (x_1, x_2, ..., x_n) \in X_1 \times X_2 \times ... \times X_n$ such that $f(x_1, x_2, ..., x_n) = y$.

The fuzzy function $\widehat{f}$ estimated by this method presents many properties (Pedrycz & Gomide, 1998), one of the most important of which is that $\widehat{f}$, in fact, recovers the classical values in the sense that $\widehat{f}(\widehat{x}_1, \widehat{x}_2, ..., \widehat{x}_n) = f(x_1, x_2, ..., x_n)$ to all $(x_1, x_2, ..., x_n) \in X_1 \times X_2 \times ... \times X_n$ where

$$\widehat{x}_i(t) = \begin{cases} 1 \; if \; t = x_i \\ 0 \; if \; t \neq x_i \end{cases}.$$

Therefore, the extension principle allows us to find a possibility distribution of the function's image whose arguments are assumed to be fuzzy sets. To simplify our notation we will represent the membership function $\mu_{A_i}(x_i)$, of fuzzy set A_i, by $A_i(x_i)$.

The aim of this approach is to build fuzzy sets of the consequents of the rules of a linguistic model by applying the extension principle. Dubois *et al.* (1995) showed that computing $f(A)$, in the sense of the extension principle, where A is a fuzzy number, is equivalent to the statement *if* $X \in A$ *then* $Z \in f(A)$ in the sense of a gradual rule. Gradual rules correspond to the case when the rule means *the more* X *belongs to* A*, the more* Z *belongs to* C, which is a reasoning commonly applied in epidemic rule-based fuzzy models. In their work they proposed that their methodology could be interesting for applications in nonlinear systems.

The process consists of, to each rule, fixing a value r and varying the values of x, y and z such that $f(x, y, z) = r$. Then, to each triplet (x, y, z) we evaluate the membership degree of the x, y and z in their respective fuzzy sets and choose the minimum value of membership degree. After finding this minimum value, we choose the maximum value of membership degrees, over all triplets such that $f(x, y, z) = r$, applying $\max[\min(X_i(x), Y_j(y), Z_k(z))]$. The result will be the membership function of r in the consequent of the $i - th$ rule. Varying the value of r we are able to build the fuzzy set that is the consequent of this rule. Finally, executing the method to each rule we elaborate all consequents, finding the complete *if-then* rule set of the model.

To illustrate how this technique works, we will return to the fuzzy model to describe the canine rabies seropositive prevalence (see section 8.2). Remember that in this problem both TSK and Mamdani fuzzy modeling have three input variables: the proportion of seropositive dogs in the age a, $S(a)$, the force of vaccination, ν, and the rate of loss of antibodies, τ. Experts built three triangular membership functions for each input variable that are shown in figures 8.3, 8.4 and 8.5. Considering all combinations of these fuzzy sets we elaborated 27 rules, each one describing the different situations that could occur. Differently from experts-based models, in this case we did not suppress any rule, leaving under the responsibility of the methodology the treatment and decision about the rules through the construction of its consequent.

As discussed early, processing the *Extension Principle* requires a classical mathematical function. So, to build the rule's consequents we applied again the solution of the classical differential equations:

$$\Delta s(a+1) = \nu(1 - s(a)) - \tau s(a), \tag{8.21}$$

where $(1-s(a))$ is the proportion of seronegative dogs. Thus, using the procedure described above were generated all consequents of the rules. In this case all the consequent are different. In other words, each rule possesses a specific fuzzy set as output variable. All fuzzy outputs have triangular shapes, as we can see in figure 8.19. This result is not expected since the classical function considered in the process is not linear.

After the consequents are built, the next step consisted in choosing the appropriated fuzzy inference method. First, we applied the inference method of

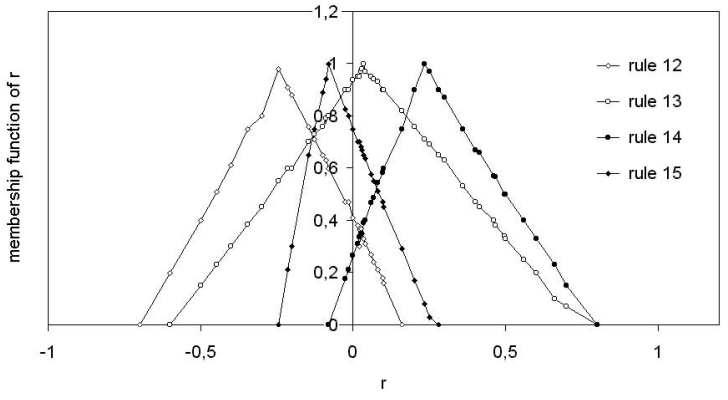

Fig. 8.19. Typical fuzzy sets of rule's consequents through Extension Principle methodology (Ortega *et al.*, 2003)

Mamdani and the output was defuzzified with the Center of Area technique. This model was named $MISO-ExtMamdani$. It is important to emphasize that one of the aims of this approach was to compare the capacity of the extension approach to recover the experts opinion. In this sense, it is interesting to maintain the same antecedents, inference and defuzzification methods applied previously in the Mamdani expert-based model, called here by $MISO - Expert$ model, changing only the construction of the rule's consequents (see section 8.2.2).

Second, we applied the Dubois *et al.* inference (1995), which applies a concept of *gradual fuzzy rule* set. In this case, called here by $MISO - ExtDubois$, the output of each rule is a crisp set, U_i, and not a fuzzy set as in Mamdani's case. In fact, in Dubois case the rule output U_i is a numerical interval. After processing all the rules, the final output is found by the intersection of the outputs of each rule, $U = \cap U_i$. So, in this inference method there is no defuzzification but, in counterpart, it is necessary to decide what crisp number, in the output region, is the best representative output. In this work we choose the mid point in the output interval. In all models the dynamical process was implemented through an iterative procedure described previously. Figures 8.20 and 8.21 show the $MISO - Experts$, $MISO - ExtMamdani$, $MISO - ExtDubois$ and real data results for the street and Police kennel dogs samples.

As can be noted from figures, the $MISO - Expert$ and the $MISO - ExtMamdani$ models are more associated with real data than the $MISO - ExtDubois$. In both $MISO - ExtMamdani$ and $MISO - ExtDubois$ models the antecedents of fuzzy sets (showed in figures 8.3, 8.4 and 8.5) were found via normalization of the sets created by the experts in the $MISO - Experts$ model. This was necessary because we worked with the same fuzzy input sets in all MISO models. However, in the extension application we had to guarantee that $-1 < \Delta s < 1$, such that the ΔS really meant a proportion of increment of protected dogs. The best fitting of the real data led to almost the same values for ν and τ parameters in the case of $MISO - Experts$ and $MISO - ExtMamdani$,

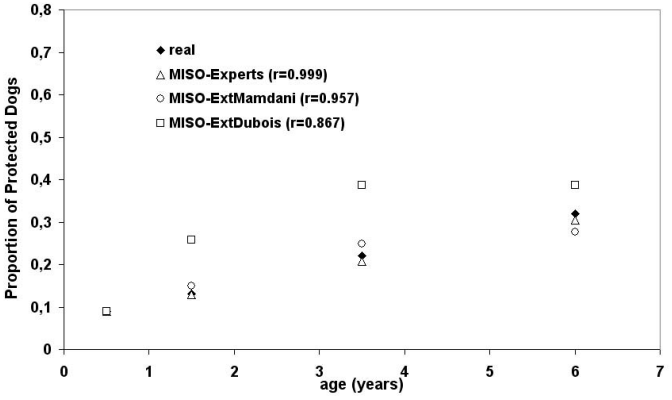

Fig. 8.20. Comparison between real data and the results provided by $MISO - Experts$, $MISO - ExtMamdani$ and $MISO - ExtDubois$ models for street dogs sample (Ortega *et al.*, 2003)

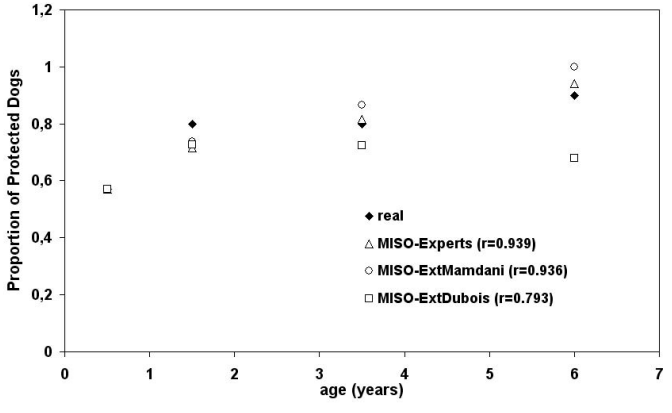

Fig. 8.21. Comparison between real data and the results provided by $MISO - Experts$, $MISO - ExtMamdani$ and $MISO - ExtDubois$ models for police kennel dogs sample (Ortega *et al.*, 2003)

considering this normalization, in both street and police kennel dogs samples. However, applying those same parameters values to $MISO - ExtDubois$ we have that the fitting is worse than the others methods. In fact, in the case of this example, the inference methods have strikingly different trends. The results obtained with Dubois *et al.* approach point to a saturation in the seroprevalence curve of street dogs. This contrasts with the Mamdani's method, which presents a more realistic trend as seroprevalence always increase with age. According to the experts opinion, this latter trend is more in accord with their expectations since the older the dog, the higher its chance of being vaccinated.

It is not the focus here to discuss the theoretical and philosophical aspects involved in the inference techniques when the Extension Principle is applied. The comparison between the Dubois *et al.* inference proposal and that based on Mamdani presented here requires more technical and deeper discussion. For interested readers we recommend the reading of Ortega and collaborators work (Ortega *et al.*, 2003).

We want to conclude this chapter by highlighting that linguistic models are probably the most important contribution of fuzzy sets theory to epidemiology and medical fields, and dynamical linguistic models can assume a prominent role in those areas. They mimic with astonishing accuracy medical reasoning and, therefore, are promptly accepted by public health professionals. In contrast with other mathematical techniques, like differential equations models, fuzzy sets are easily understood by decision makers in public health. Since those who make decisions are not, in general, those who elaborate the models, it is fundamental that experts and public health authorities understand the contents of the model, even if in a superficial way, in order to accept them and implement them. As the language of the linguistic model is quite similar to the natural language of the experts, this type of model is getting generally accepted and applied as a useful tool for treating medical and epidemic problems. An example of this acceptance was the fuzzy decision model to determine the best strategy of vaccination against measles presented in chapter 6. The strategy proposed by that fuzzy model was applied to a real epidemic of measles in the state of São Paulo, Brazil, providing good results.

Nevertheless, considering the limitations of dynamical linguistic models exposed here, we present in chapters 9 and 10 others alternatives to treat dynamical systems where uncertainties are considered in the epidemic process. However these methodologies are based on more complex mathematical structures. In chapter 11 we recover the fuzzy dynamical systems development applying the composition of fuzzy relations concepts in the simulation approach, which consists in a generalization of the simplest epidemic model proposed by Reed and Frost (Abbey, 1952). In chapter 12 we return to fuzzy rule-based models in a dynamical viewpoint, presenting a hybrid methodology in which this kind of model is important for the understanding of another epidemic spreading problem.

9 Fuzzy Dynamical Systems in Epidemic Modeling

As mentioned in the previous chapters, mathematical models are always subject to inaccuracies related to the nature of the state variables involved, parameters and/or initial conditions. In these models, the estimation of the parameters is usually based on statistical methods, starting from data obtained experimentally to the choice of the method adopted to their identification.

In biomathematical stochasticity can arise as an imposition of the state variable or due to some of the parameters, initially deterministic, but subjected to random fluctuations. These two cases are called demographic and environmental stochasticity, respectively (May, 1974; Turelli, 1986).

We propose in this chapter, to some extent, an adaptation of the concepts of demographic and environmental stochasticity, using fuzzy sets theory, an efficient tool to deal with the subjectivity that comes from the "fuzziness" of the biological phenomenon (Barros *et al.*, 2000). Both concepts of demographic and environmental fuzziness can be analyzed by *fuzzy differential equations* or by *fuzzy differential inclusion*, since they are in essence dynamical systems. In chapter 8 the fuzzy rule-based dynamical systems were widely discussed with examples of SIS and SIR epidemic models. Here we present other ways to treat these epidemic models using differential equations and differential inclusions approaches.

9.1 Demographic Fuzziness

Individuals of a given species usually show variations in their characteristic behavior. These differences remain bounded in a relatively small set that represents the characteristic behavior of the group as a whole. For example, a predator with a certain predatory level can possibly become a prey, depending on environmental circumstances. In such cases we should take into account the predatory degree of each individual of the species. In general, if we want to quantify the subjective quality that is being studied, we should attribute values or degrees to represent this quality satisfactorily. It is something that cannot be always obtained through objective mensuration or statistics. Therefore, when the state variables of a given

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 181–206, 2008.
springerlink.com

demographical system are uncertain we should include fuzziness in them, which are then represented by what we call fuzzy variables.

The modeling tools presented here are the fuzzy differential equations and the fuzzy differential inclusions. The fuzzy structures are introduced because of the eventual subjectiveness of individual characteristics in the initial population.

First we introduce the concept of fuzzy differential equations.

Definition 9.1. *Consider the space of fuzzy numbers denoted by* $\Im(I\!R)$*. Let* $F : I \subset I\!R \to \Im(I\!R)$ *be a fuzzy function with* $\alpha - levels$

$$[F(t)]^{\alpha} = [F_1^{\alpha}(t), F_2^{\alpha}(t)], \tag{9.1}$$

where $F_1^{\alpha}, F_2^{\alpha} : I \to I\!R$ *are ordinaries functions and* $\Im(I\!R)$ *represents the class of fuzzy numbers.*

The derivative of F, with respect to t, is given from its $\alpha - levels$ by:

$$[F'(t)]^{\alpha} = [(F_1^{\alpha})'(t), (F_2^{\alpha})'(t)], \tag{9.2}$$

where $(F_i^{\alpha})'(t)$ means the derivative of functions F_i^{α} (Puri & Ralescu, 1983). The concept of fuzzy differential equation can be found in Kaleva (1987) and Seikkala (1987).

In what follows we illustrate the concept of demographic fuzziness, where the initial condition is a fuzzy set in $\Im(I\!R)$.

Example 9.2. - Fuzzy Malthus Continuous Model

Let us suppose that a population grows according to the Malthusian model:

$$\begin{aligned} n^{'}(t) &= \lambda n(t) \\ n(0) &= n_0 \in \Im(I\!R) \end{aligned} \tag{9.3}$$

with $\lambda \in I\!R$.

Since n_0 is a fuzzy set, the field $F(n) = \lambda n$ associates fuzzy sets to fuzzy sets. In this sense, equation (9.3) results in a fuzzy differential equation.

According to the representation theorem (see chapter 2), to find the solution of (9.3) we make $[n(t)]^{\alpha} = [n_1^{\alpha}(t), n_2^{\alpha}(t)]$, where $n_1^{\alpha}(t)$ and $n_2^{\alpha}(t)$ are extremes values of the classical set $n^{\alpha}(t)$. Therefore, to solve (9.3) we must solve the deterministic system below (Seikkala, 1987):

$$\begin{cases} (n_1^{\alpha})'(t) = \lambda n_1^{\alpha}(t), \quad n_1^{\alpha}(0) = n_{01}^{\alpha} \\ \\ (n_2^{\alpha})'(t) = \lambda n_2^{\alpha}(t), \quad n_2^{\alpha}(0) = n_{02}^{\alpha} \end{cases} \qquad \lambda \geq 0 \tag{9.4}$$

or

$$\begin{cases} (n_1^\alpha)'(t) = \lambda n_2^\alpha(t), \quad n_1^\alpha(0) = n_{01}^\alpha \\ \\ (n_2^\alpha)'(t) = \lambda n_1^\alpha(t), \quad n_2^\alpha(0) = n_{01}^\alpha \end{cases} \qquad \lambda < 0 \tag{9.5}$$

for each $\alpha \in [0, 1]$.

The solutions of (9.4) and (9.5) are given, respectively, by:

$$\begin{cases} n_1^\alpha(t) = n_{01}^\alpha e^{\lambda t} \\ \\ n_2^\alpha(t) = n_{02}^\alpha e^{\lambda t} \end{cases} \qquad \lambda \geq 0 \tag{9.6}$$

and

$$\begin{cases} n_1^\alpha(t) = \dfrac{(n_{01}^\alpha - n_{02}^\alpha)}{2} e^{-\lambda t} + \dfrac{(n_{01}^\alpha + n_{02}^\alpha)}{2} e^{\lambda t} \\ \\ n_2^\alpha(t) = \dfrac{(n_{02}^\alpha - n_{01}^\alpha)}{2} e^{-\lambda t} + \dfrac{(n_{01}^\alpha + n_{02}^\alpha)}{2} e^{\lambda t} \end{cases} \qquad \lambda < 0 \ . \tag{9.7}$$

So, the solution $n(t)$ has $\alpha - levels$ given by (9.6) if $\lambda \geq 0$ or (9.7) if $\lambda < 0$.

The solution of a fuzzy differential equation is a fuzzy process in the sense that for each instant of time $t > 0$, $n(t)$ is a fuzzy set. In the deterministic case, at each t, the solution is a real (or a crisp) number.

9.2 Environmental Fuzziness

To model environmental fuzziness, we initially use differential equations, which are formally deterministic, but with some of their coefficients modeled by fuzzy sets. In this case the equations can be treated as standard differential equations with uncertainty in the parameters (Boxler, 1988) or through the calculus developed by Kaleva (1987), depending on the modeling approach adopted. So, derivatives can be classical or that introduced by Puri and Ralescu (1983) to deal with fuzzy functions, depending on the problem.

Like the case of demographic fuzziness, we illustrate the concept of environmental fuzziness through examples. In the Example below we consider poverty as a factor that supposedly influences life expectancy of a population (Bassanezi & Barros, 1995). A similar study is due to Kandel (1986) who analyzed the case of smoking.

Example 9.3. - Life Expectancy

Let A be a population with $n(t)$ individuals at instant t. Supposing that there is no birth in this group and that the dynamics of the number of individuals is modeled by the following differential equation:

$$\begin{cases} n'(t) = \lambda n(t) \\ n(0) = n_0 \in I\!R \end{cases}, \tag{9.8}$$

where n_0 is the initial number of individuals and λ is the mortality rate.

In which way does the environment or even the individuals way of living influence the population's life expectancy? A possible answer will be achieved by supposing that the environment "acts" in the population as a whole. That is, we will not take into account individual characteristics such as gender, skin color, resistance to diseases, etc. This is the main characteristic of fuzziness only in the parameters, originally deterministic, but subject to environmental stochasticity (May, 1974; Turelli, 1986).

To incorporate fuzziness in parameter λ we suppose that $\lambda = \lambda_1 + A_k(r)\lambda_2$, where λ_1 is the natural mortality rate (taken from a population with satisfactory conditions of life) and $A_k(r)\lambda_2$ is the coefficient that represents the influence of poverty in the mortality rate, λ, of the group. The mortality rate is maximal and equal to $\lambda_1 + \lambda_2$ when $A_k(r) = 1$. For model (9.8) we have chosen as a "poor set" the fuzzy set given by:

$$A_k(r) = \begin{cases} \left[1 - \left(\frac{r}{r_0}\right)^2\right]^k & if\ 0 \le r < r_0 \\ 0 & if\ r \ge r_0 \end{cases}. \tag{9.9}$$

As introduced in chapter 2, in example 2.5, the parameter r is the income, r_0 is the income threshold above which the life expectancy of the group is not affected, and k is an environmental parameter that characterize the group (for instance, rural, urban, Indians, etc.).

Assume that r is proportional to the salaries of the individuals in the studied population: $r = cs^m$, with c and m being two constants. So we have the following fuzzy set:

$$B_k(s) = A_k(cs^m) = \begin{cases} \left[1 - \left(\frac{s}{s_0}\right)^{2m}\right]^k & if\ 0 < s < s_0 \\ 0 & if\ s \ge s_0 \end{cases},$$

where $s_0 = \left(\frac{r_0}{c}\right)^{\frac{1}{m}}$.

To obtain the values of λ_1, λ_2 and s_0 we used the life expectancy table based on distinct salary levels (table 9.1). The values found are:

$$\lambda_1 = \frac{1}{54.4}, \quad \lambda_2 = 6.618 \times 10^{-3} \quad and \quad s_0 = 3.2.$$

According to the salary distribution of a group of workers from the same region, for which we got λ_1, λ_2 and s_0, assuming $c = 1$, we have $k = 1.51$ and $m = 0.4435$.

Table 9.1. Life expectancy in the central northeast (Brazil) according to income per capita and family class (urban zone) in 1970 (Fava, 1983)

Class of income	**Income** ($)	**Salary**	**Life expectancy (years)**
1	1-150	Less than 0.94	40.0
2	151-300	0.95-1.88	45.9
3	301-500	1.89-3.26	50.8
4	Over 500	Over 3.26	54.4

Thus, problem (9.8) can be solved by using the classical ordinary differential equations, whose solution is given by

$$n(t) = n_0 e^{-[\lambda_1 + B_k(s)\lambda_2]t}$$

for each value of s. In this way, we obtain a family of solutions for problem (9.8). The analysis of these solutions, as well as the average between a classical and a fuzzy setting (accordingly chapter 4) represents the life expectancy of this population (Bassanezi & Barros, 1995).

Let us now consider equation (9.8) as a fuzzy Cauchy problem:

$$\begin{cases} n'(t) = -(\lambda_1 + B_k(s)\lambda_2)n(t) \\ \\ n(0) = n_0 \in I\!R_+ \end{cases}, \tag{9.10}$$

whose solution for each instant of time is a fuzzy set.

The $\alpha - levels$ of B_k and $n(t)$ are (see chapter 2):

$$[B_k]^\alpha = [0, s_0(1-\alpha^{\frac{1}{k}})^{\frac{1}{2m}}] \quad and \quad [n]^\alpha = [n_1^\alpha, n_2^\alpha]$$

for each $\alpha \in [0,1]$.

So, as a result of the addition and multiplication operations, we have

$$[-(\lambda_1 + B_k(s)\ \lambda_2)\ n]^\alpha = [-(\lambda_1 + \lambda_2 s_0(1-\alpha^{\frac{1}{k}})^{\frac{1}{2m}})\ n_2^\alpha,\ -\lambda_1 n_1^\alpha]$$

and the solution of (9.10) is obtained from the bi-dimensional deterministic system:

$$\begin{cases} \dot{n}_1^\alpha = -(\lambda_1 + \lambda_2 s_0(1-\alpha^{\frac{1}{k}})^{\frac{1}{2m}} n_2^\alpha = -b n_2^\alpha\ , n_0 \\ \\ \dot{n}_2^\alpha = -\lambda_1 n_1^\alpha \qquad\qquad\qquad\qquad , n_0 \end{cases} \tag{9.11}$$

for all $\alpha \in [0,1]$.

From (9.11) we have:

$$\ddot{n}_1^\alpha = -b\dot{n}_1^\alpha \quad or \quad \ddot{n}_1^\alpha = \lambda_1 b n_1^\alpha$$

obtaining the solutions

$$\begin{cases} n_1^\alpha = n_0 \left[\dfrac{\left(1+\sqrt{\frac{b}{\lambda_1}}\right)}{2} \exp(-\sqrt{\lambda_1 b}t) - \dfrac{\left(\sqrt{\frac{b}{\lambda_1}}-1\right)}{2} \exp(\sqrt{\lambda_1 b}t) \right] \\ n_2^\alpha = n_0 \left[\dfrac{\left(\sqrt{\frac{\lambda_1}{b}}\right)+1}{2} \exp(-\sqrt{\lambda_1 b}t) + \dfrac{\left(1-\sqrt{\frac{\lambda_1}{b}}\right)}{2} \exp(\sqrt{\lambda_1 b}t) \right] \end{cases}. \tag{9.12}$$

We can note that problem (9.10) has a unique solution with $\alpha - levels$ given by (9.12), and that the diameter of each $\alpha - levels$ of this solution is given by:

$$d(\alpha, t) = n_2^\alpha - n_1^\alpha = n_0 \left(\sqrt{\frac{b}{\lambda_1}} - \sqrt{\frac{\lambda_1}{b}} \right) \ sinh(\sqrt{\lambda_1 b}t).$$

Therefore, the fuzziness of solution of (9.8) increases with the time t.

In the next section we present an epidemic model for a directly transmitted infection considering heterogeneities in some of the characteristics of the population.

9.3 Epidemiology with Heterogeneity

As mentioned in chapter 3, classical epidemiological models are "first approximations" since we assume the law of mass action, in addition to the hypothesis that all infective individuals have the same capacity of transmission of the infection. That means that those models assume homogeneity in both susceptibility and infectiousness.

In this section we consider that the contact between infected and susceptible individuals occurs with the same chance. However, we consider different degrees in the chance that such contact will result in new cases. In other words, a new case will occur with more or less likelihood, depending on some characteristics of those individuals who meet each other in a potentially infectious contact.

These heterogeneities turn our case a typical case of demographic fuzziness. However, in order to avoid the increasing diameter of fuzzy derivatives, we will consider classical dynamical equations. This is justified because we are interested only in the number of new cases of infection. We also assume that individuals with

great "power" of infectiousness contribute more than those with low "power" of infectiousness.

In order to describe the dynamics of the disease transmission in a more realistic way, several models that incorporate heterogeneities have been proposed. According to Sattenspiel and Simon (1988) there are five main sources of heterogeneity in the mathematical models: variable susceptibility, variable infectivity, age structure, variable number of contacts and different patterns of contact distribution among the subgroups. Coutinho *et al.* (1999) analyzed a very interesting model taking into account several heterogeneities in a frailty setting. Greenhalgh (1990) studied models with age structure to model child diseases. Hethcote and VanArk (1987) presented a model where the population is divided into subpopulations according to the number of contacts between the individuals. Diekmann *et al.* (1990) considered the heterogeneity of the contact distribution to study disease invasion in a population. Boylan (1991) studied disease propagation considering the variable susceptibility in the population arguing that this difference can be due to biological, behavioral or environmental effects. Generally, each of these models consider a discrete population utilizing compartments to characterize the diversities. However, the classification of individuals, and the consequent inclusion in a determined stage and the transition from one stage to another, is not a simple task and can result in a mathematically very complex global model as the number of compartments increases (see Leite *et al.*, 2000).

Barros and collaborators studied an epidemic model considering the different degrees of infectivity that each individual of a population can present without subdividing the infected class into compartments (Barros *et al.*, 2003). In the next section we will present the way to incorporate fuzziness in epidemic dynamical models, analyzing the case of SI and SIS models of transmission (Barros *et al.*, 2001, 2002 and 2003).

9.3.1 The *SI* Model

The simplest classical model to describe the dynamics of directly transmitted diseases with interaction among susceptible and infected individuals is the *SI* model without neither vital dynamics (i.e, the rates of birth and mortality are not considered), nor immunity, nor additional disease fatality rate. The model can be represented by the diagram showed in figure 9.1.

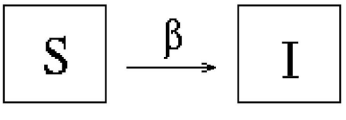

Fig. 9.1. SI model diagram where the flow between the susceptible and infective compartments are explicit

The classical normalized differential equations which describe such dynamics are given by:

$$\begin{cases} \dfrac{dS}{dt} = -\beta SI \\ \dfrac{dI}{dt} = \beta SI \end{cases}, \tag{9.13}$$

where $S+I=1$, S is the proportion of susceptible individuals, I is the proportion of infected individuals at each instant and β is the transmission coefficient of the disease.

A fundamental assumption in this formulation is that the population is homogeneous. That is, each infected individual transmits the disease with the same chance, given by the real number β. So, from (9.13) the number of infected individuals at any instant t is given by:

$$I = \frac{I_0 e^{\beta t}}{S_0 + I_0 e^{\beta t}}. \tag{9.14}$$

where S_0 and I_0 are the initial conditions.

Both concepts of susceptible and infectious are uncertain in the sense that there are different degrees in susceptibility and infectivity among the individuals of the population. Such differences can arise, for example, when we consider the population's distinct habits and customs, different degrees of resistance, etc. In this way, we could consider more realistic models, which consider different degrees of susceptibility and/or infectivity of the individuals. We consider the parameter β (which represents the chance that in one contact between a susceptible and an infected individual the transmission of the disease occurs) as a fuzzy number.

The SI fuzzy model

We assume that the population heterogeneity is given by the parasite load of infected individuals (see the fuzzy notion of disease described in Sadegh-Zadeh, 1999). Thus, the higher the parasite load, the higher will be the chance of disease transmission. In other words, we assume that $\beta = \beta(\nu)$ measures the chance of a transmission to occur in a meeting between a susceptible and an infected individual with an amount of pathogens ν. In this way, some values of β are more possible than others and that turns β into a membership function of a fuzzy number (see chapter 4).

To obtain the membership function β we assume that when the amount of pathogens in an individual is relatively low, the chance of transmission is negligible, and that there is a minimum amount of pathogens $\nu_{\min}$ needed to cause disease transmission. Furthermore, for a certain amount of pathogens ν_M, the chance of disease transmission is maximum and equal to 1. Finally, we suppose that the individual's amount of pathogens is always limited by $\nu_{\max}$ for each disease.

We have chosen for the fuzzy subset, the following membership function:

$$\beta(\nu) = \begin{cases} 0 & if \quad \nu < \nu_{\min} \\ \dfrac{\nu - \nu_{\min}}{\nu_M - \nu_{\min}} & if \quad \nu_{\min} \leq \nu \leq \nu_M \\ 1 & if \quad \nu_M < \nu < \nu_{\max} \end{cases}, \tag{9.15}$$

where $\nu_{\min}$ represents the minimum amount of pathogens needed for disease transmission to occur. This value can be understood as the one which gives the susceptibility of a particular population. In fact, the higher the $\nu_{\min}$ value, the higher the amount of pathogens needed for transmission to occur and it means that the populations has a low susceptibility to the disease. The graphic of $\beta(\nu)$ is presented in figure 9.2.

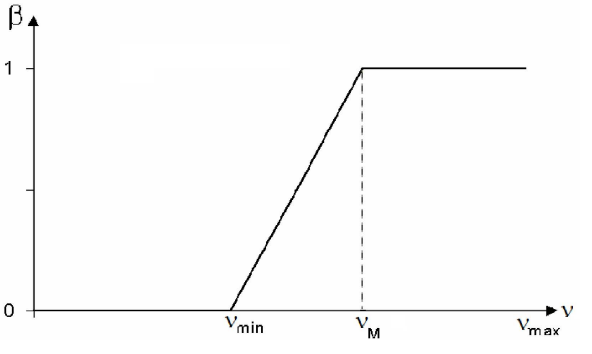

Fig. 9.2. Fuzzy coefficient of transmission - $\beta = \beta(\nu)$ (Barros *et al.*, 2003)

We also consider that the amount of pathogens can be different for different individuals. In this sense, ν can be seen as a *fuzzy number* with a triangular shape, according to the following membership function:

$$\rho(\nu) = \begin{cases} 1 - \dfrac{|\,\nu - \bar{\nu}|}{\delta} & if \quad \nu \in [\bar{\nu} - \delta, \bar{\nu} + \delta] \\ 0 & if \quad \nu \notin [\bar{\nu} - \delta, \bar{\nu} + \delta] \end{cases}. \tag{9.16}$$

The parameter $\bar{\nu}$ is a central value and δ gives the dispersion of each one of the fuzzy sets assumed by ν. Figure 9.3 shows the graphic of $\rho(\nu)$. For a fixed $\bar{\nu}$, $\rho(\nu)$ can has a linguistic meaning, given by a expert, such as *low*, *medium* and so on.

In what follows we study the evolution of the disease with $\beta = \beta(\nu)$, as discussed above.

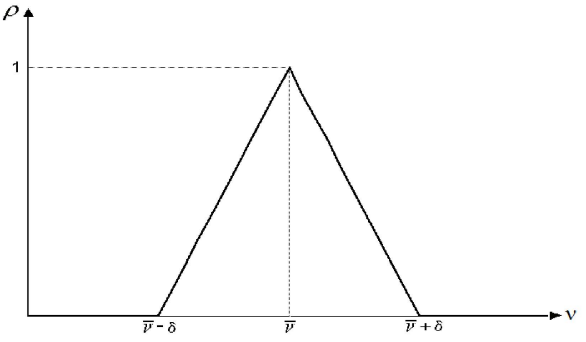

Fig. 9.3. Membership function of the variable ν, amount of pathogens - ρ (Barros *et al.*, 2003)

• *Analysis and interpretation of the SI fuzzy model*

In the deterministic model (9.13) if we consider $\beta = \beta(\nu)$, then the solution is given by:

$$I(\nu, t) = \frac{I_0 e^{\beta(\nu)t}}{S_0 + I_0 e^{\beta(\nu)t}}. \tag{9.17}$$

So, $I(\nu, t)$ can be considered as a family of solutions of (9.13) for each fixed ν. It represents the number of infected individuals at any instant t produced by the contact between susceptible and infected individuals with an amount of pathogens ν. On the other hand, for each fixed t, $I(\nu, t)$ is a membership function of a fuzzy number, because $0 \leq I(\nu, t) \leq 1$. We will represent both the fuzzy sets and its membership function by $I(\nu, t)$. In this way we can estimate a mean value of the number of infected individuals at each instant using some defuzzification procedure on the $I(\nu, t)$.

The mean number will be expressed through the fuzzy expected value of $I(\nu, t)$ (see chapter 4).

• *Fuzzy expectancy of the number of infected individuals*

The fuzzy expected value (FEV) of the number of infected individuals, $I(\nu, t)$, is given by:

$$FEV[I(\nu, t)] = \sup_{0 \leq \alpha \leq 1} \inf[\alpha, \mu\{I(\nu, t) \geq \alpha\}].$$

Let $H(\alpha) = \mu\{I(\nu, t) \geq \alpha\}$, for each $t > 0$. Remember that $FEV[I(\nu, t)]$ is the fixed point of $H(\alpha)$. So, it is easy to see that for $\alpha = 0$ and $\alpha = 1$, then $H(0) = 1$ and $H(1) = 0$.

For $0 < \alpha < 1$, and calling $k = S_0/I_0$,

$$H(\alpha) = \mu\{I(\nu,t) \geq \alpha\} = \mu\left\{\nu : \beta(\nu) \geq \ln\left(\frac{\alpha k}{1-\alpha}\right)^{1/t}\right\}$$

$$= \begin{cases} 1 & if \quad \ln\left(\frac{\alpha k}{1-\alpha}\right)^{\frac{1}{t}} \leq 0 \\ \mu[a, \nu_{\max}] & if \quad 0 < \ln\left(\frac{\alpha k}{1-\alpha}\right)^{\frac{1}{t}} \leq 1 \\ 0 & if \quad \ln\left(\frac{\alpha k}{1-\alpha}\right)^{\frac{1}{t}} > 1 \end{cases}$$

$$= \begin{cases} 1 & if \quad 0 \leq \alpha \leq I_0 \\ \mu[a, \nu_{\max}] & if \quad I_0 < \alpha \leq \frac{I_0 e^t}{S_0 + I_0 e^t} \\ 0 & if \quad \frac{I_0 e^t}{S_0 + I_0 e^t} < \alpha \leq 1 \end{cases}, \tag{9.18}$$

where $a = \nu_{\min} + (\nu_M - \nu_{\min}) \ln\left(\frac{\alpha k}{1-\alpha}\right)^{\frac{1}{t}}$. Note that $\nu_{\min} < a \leq \nu_M$.

Next, we assume that for any subset A of real numbers, its fuzzy measure (chapter 4) is given by

$$\mu(A) = \frac{1}{\delta}\int_A \rho(\nu)d\nu = \int_A \frac{\rho(\nu)}{\delta}d\nu.$$

Note that $\mu(A)$ is exactly the probability of A, since $\frac{\rho(\nu)}{\delta}$ can be seem as a density function of probability.

To study the $FEV[I(\nu,t)]$ we consider three different cases, according to three different linguistic meanings of ν, whose were classified as *low*, *medium* and *high*. Note that each of this classification is a fuzzy number based on the values $\nu_{\min}$, ν_M and $\nu_{\max}$ which appear in the definition of β (see figure 9.4). So, for each fuzzy triangular set with parameters $\nu_{\min}$, ν_M and $\nu_{\max}$ we can study the FEV $[I(\nu,t)]$.

Case a) *Low amount of pathogens*: In this case, we take $\bar{\nu}_{\min} > \bar{\nu} + \delta$.
As $a > \nu_{\min}$, we have $\mu[a, \nu_{\max}] = 0$ and thus,

$$H(\alpha) = \begin{cases} 1 & if \quad 0 \leq \alpha \leq I_0 \\ 0 & if \quad I_0 < \alpha \leq 1 \end{cases}.$$

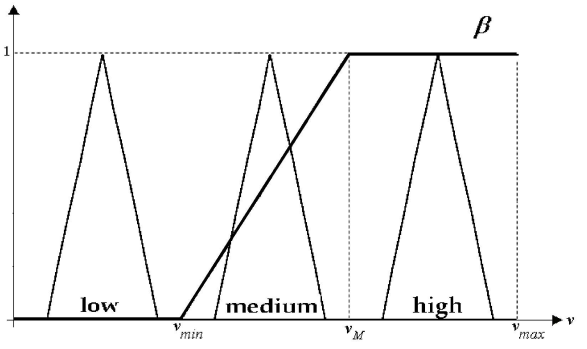

Fig. 9.4. Three linguistic meaning for the amount of pathogens, compared with the β values: *low*, *medium* and *high* (Barros *et al.*, 2003)

Therefore,

$$FEV[I(\nu,t)] = I_0.$$

Thus, as all infected individuals present an amount of pathogens less than $\nu_{\min}$, the disease propagation does not occur. We could interpret this situation as a highly resistant group ($\nu_{\min}$ is high), which makes the susceptibility very low. In this case the initial quantity I_0 of infected individuals remains unchanged.

Case b) *High amount of pathogens*: In this case, we have $\nu_M \leq \bar{\nu} - \delta$ and $\bar{\nu} + \delta \leq \nu_{\max}$.

For this situation, as $a \leq \nu_M$, we obtained $\mu[a, \nu_{máx}] = 1$, therefore

$$H(\alpha) = \begin{cases} 1 & if \quad 0 \leq \alpha \leq \dfrac{I_0 e^t}{S_0 + I_0 e^t} \\ \\ 0 & if \quad \dfrac{I_0 e^t}{S_0 + I_0 e^t} < \alpha \leq 1 \end{cases}$$

and, so

$$FEV[I(\nu,t)] = \frac{I_0 e^t}{S_0 + I_0 e^t}.$$

Note that the expression above coincides with the classical solution (9.14) when we consider the transmission coefficient β constant and equal to 1.

Case c) *Medium amount of pathogens*: In this case we have taken $\bar{\nu} - \delta > \nu_{\min}$ and $\bar{\nu} + \delta < \nu_M$

From (9.18), a direct calculation gives

$$H(\alpha) = \begin{cases} 1 \quad if \quad 0 \leq \alpha \leq I(\bar{\nu} - \delta, t) \\ 1 - \frac{1}{2}\left(\frac{a - \bar{\nu}}{\delta} + 1\right)^2 \quad if \quad I(\bar{\nu} - \delta, t) < \alpha \leq I(\bar{\nu}, t) \\ \frac{1}{2}\left(\frac{\bar{\nu} - a}{\delta} + 1\right)^2 \quad if \quad I(\bar{\nu}, t) < \alpha \leq I(\bar{\nu} + \delta, t) \\ 0 \quad if \quad I(\bar{\nu} + \delta, t) < \alpha \leq 1 \end{cases}.$$

According to the expression above we conclude that $H(\alpha)$ is a continuous and decreasing function with $H(0) = 1$ and $H(1) = 0$. Therefore, H has a unique fixed point which coincides with $FEV[I(\nu, t)]$. Figure 9.5 shows this fact:

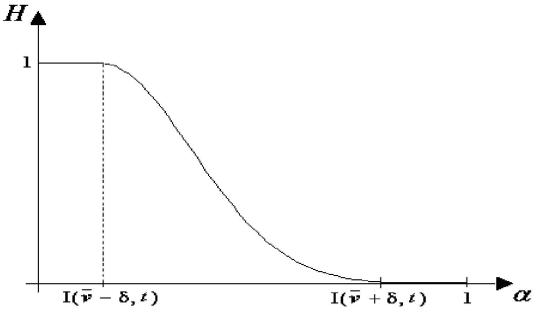

Fig. 9.5. Graphic of Function $H(\alpha)$ (Barros *et al.*, 2003)

Knowing the parameters δ, $\bar{\nu}$, $\nu_{\min}$, ν_M and $\nu_{\max}$, characteristic of each disease, the fixed point of $H(\alpha)$ can be obtained. This value provides an estimate of the number of infected individuals at instant t, as shown in chapter 4.

We are now going to compare the $FEV[I(\nu, t)]$ with the trajectory $I(\bar{\nu}, t)$.

From the expression of H we can conclude that $H(I(\bar{\nu}, t)) = \frac{1}{2}$ for all t. Thus, $FEV[I(\nu, t)] = I(\bar{\nu}, t)$ when $I(\bar{\nu}, t) = \frac{1}{2}$.

Since $FEV[I(\nu, t)]$ is the fixed point of H we have:

$$\begin{cases} FEV[I(\nu, t)] > I(\bar{\nu}, t) \quad if \quad I(\bar{\nu}, t) < \frac{1}{2} \\ FEV[I(\nu, t)] < I(\bar{\nu}, t) \quad if \quad I(\bar{\nu}, t) > \frac{1}{2} \end{cases}.$$

In this way, $I(\bar{\nu}, t)$ does not provide the average number of infected individuals (given by $FEV[I(\nu, t)]$). Therefore, it is not correct to use the average parasite

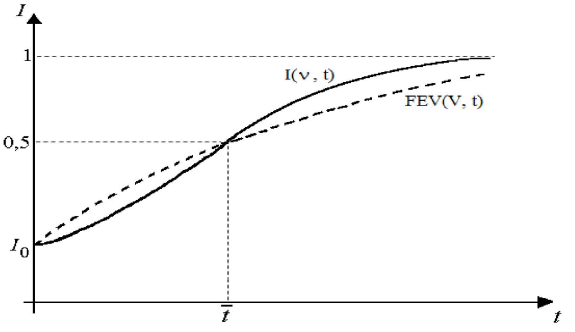

Fig. 9.6. Deterministic solution $I(\bar{\nu}, t)$ and fuzzy expectancy $FEV[I(V,t)]$ (Barros *et al.*, 2003)

load $\bar{\nu}$ (value used to obtain the parameter β in the deterministic model) to analyze the evolution of disease in the whole population, since $I(\bar{\nu}, t) = FEV[I(\nu, t)]$, only at the instant $t = \bar{t} = \dfrac{\nu_M - \nu_{\min}}{\overline{\nu} - \nu_{\min}} \ln\left(\dfrac{S_0}{I_0}\right)$ with $S_0 \geq I_0$, for which the increment of the increase rate of the trajectory $I(\bar{\nu}, t)$ is the largest and $I(\bar{\nu}, \bar{t}) = \dfrac{1}{2}$ (see figure 9.6).

Note that, instead of $FEV[I(\nu, t)]$ we could have applied the classical expectancy $E[I(\nu, t)]$. In this case, the results are exactly equal to those obtained before (Barros *et al.*, 2003).

Finally, we note that $FEV[I(\nu, t)]$ can be obtained for an arbitrary fuzzy measure. For example, we could have adopted μ as a measure of the possibility (see chapter 4).

$$\mu(A) = \sup_{\nu \in A} \rho(\nu), \qquad A \subset R.$$

In this case, a straightforward calculation provides

$$H(\alpha) = \begin{cases} 1 & if \quad 0 \leq \alpha \leq I_o \\ \sup\limits_{\nu \in [a, \nu_{\max}]} \rho(\nu) & if \quad I_o < \alpha \leq \dfrac{I_o e^t}{S_o + I_o e^t} \\ 0 & if \quad \dfrac{I_o e^t}{S_o + I_o e^t} < \alpha \leq 1 \end{cases} \tag{9.19}$$

and the $FEV[I(\nu, t)]$ is the fixed point of the function $H(\alpha)$.

The basic reproduction number (R_0)

An essential parameter concerned with disease evolution is *the basic reproduction number*, R_0, which gives the number of secondary cases caused by an infected

individual introduced into an entirely susceptible population, along his/her infectiousness period. This parameter indicates under which conditions the disease propagates in the population. In the deterministic framework, if an infected individual generates more than one secondary case (that is $R_0 > 1$), then the disease is propagated. On the other hand, when $R_0 < 1$, the disease is not able to establishes itself. For simple epidemiological models, the expression for this parameter can be obtained from the condition $dI/dt > 0$, which is the condition for an increase in the number of infected individuals. In this way, for the classical SI model, we would have:

$$\frac{dI}{dt} > 0 \Longleftrightarrow \beta SI = \beta(1-I)I > 0, \tag{9.20}$$

which is always satisfied because there are susceptible in the population, provided that $\beta > 0$. In other words, we will have $R_0 > 1$ whenever $\beta > 0$ and $I < 1$.

However, when we use a fuzzy set to describe the β parameter, other situations can occur. In our case, according to the analysis done in the previous section, to maintain the mean number of infected individuals constant and equal to I_0, we should have $\mu[a, \nu_{\max}] = 0$. Therefore, any of the expressions (9.18) or (9.19), implies $\rho(\nu) = 0$ for all $\nu \in [a, \nu_{\max}]$, namely, $\bar{\nu} + \delta < a$ for all $\alpha \in [0, 1]$. Thus, we conclude that the disease transmission will not occur if $\bar{\nu} + \delta < \nu_{\min}$, which means that no infected individual has the minimum amount of pathogens necessary to transmit the disease.

Based on the above conclusions we can define the *fuzzy basic reproduction number* for our model as

$$R_0^f = \frac{\bar{\nu} + \delta}{\nu_{\min}}.$$

The aim of control measures to prevent an epidemic is to obtain $R_0 < 1$. In the classical SI model there are no such possibility because the variation of infected individuals is always positive provided that $\beta > 0$ and $S(0) \neq 0$. This is mainly due to the fact that the parameter β is considered as a real number, as we have seen above (9.20). On the other hand, if β is considered as a fuzzy set, even this simple model gives additional information on the disease dynamics. For example, it is possible to interfere in the disease evolution by reducing the parameter value R_0^f. This can be done in two ways:

1) Increasing the value of $\nu_{\min}$. This is a consequence of an increase in the resistance of susceptible individuals (decreasing their susceptibility) which could be done, for example, through vaccination, sanitation, etc. Since $\nu_{\min}$ parameter is related to susceptible individuals, this way of reducing R_0^f is related to control policies to prevent the disease.

2) The other option to reduce R_0^f is through the reduction of $(\bar{\nu}+\delta)$. Reducing δ could be done through control policies with respect to the infected population, for example, quarantine. Reducing $\bar{\nu}$ is related to treatment as, for example, by using drugs.

The above situations show two possible strategies of public health. While (1) indicates an action in the whole population, (2) acts directly on infected

individuals. It is obvious that a combination of both has a better efficiency in the prevention and control of the disease.

It should be noted that the basic reproduction number can be obtained from $\frac{d(FEV[I(\nu,t)])}{dt} > 0$, instead of $\frac{dI}{dt} > 0$. The reader is invited to do this calculation.

In the conclusion section we present other methods to obtain R_0^f for the SI fuzzy model.

We now present the SIS fuzzy model and analyze the main consequences of considering fuzziness in the parameters β and γ (the recovery rate), as we did in the SI fuzzy model. For a more complete study, see Barros *et al.* (2001, 2002 and 2003).

9.3.2 The *SIS* Model

The simplest model to describe a disease in which the individual recovers but does not develop any kind of immunity, that is, he/she becomes susceptible again, can be seen in the figure 9.7.

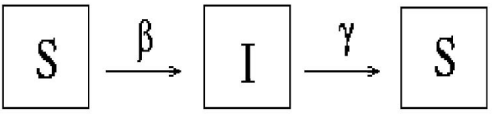

Fig. 9.7. SIS model diagram showing the flow between susceptible and infective compartments

In chapter 3 it was considered that the flow of an individuals from S class to I class occurs at a rate β depending only on the contact of a susceptible with an infected individual and that the individual recovers at a rate γ, returning to the susceptible condition. The dynamical system is described by the following system of differential equations:

$$\begin{cases} \frac{dS}{dt} = -\beta SI + \gamma I \\ \frac{dI}{dt} = \beta SI - \gamma I \end{cases}, \tag{9.21}$$

where $S + I = 1$, S is the proportion of susceptible individuals, I is the proportion of infected individuals, β is the contact rate and γ the recovering rate. Consequently, γ^{-1} is the average period of infectiousness.

In the following we propose an extension of the SIS model (9.21) incorporating heterogeneities, considering that individuals with different amount of pathogens contribute differently to the disease propagation.

The SIS fuzzy model

As in the SI fuzzy model, we will assume that $\beta = \beta(\nu)$, and that the individuals recovery rate (γ) is also a function of the parasite load. The higher the parasite load, the longer it will take to recover from infection. Consequently, γ should be a decreasing function of ν:

$$\gamma(\nu) = \frac{(\gamma_0 - 1)}{\nu_{máx}}\nu + 1, \tag{9.22}$$

where $\gamma_0 > 0$ is the lowest recovery rate.

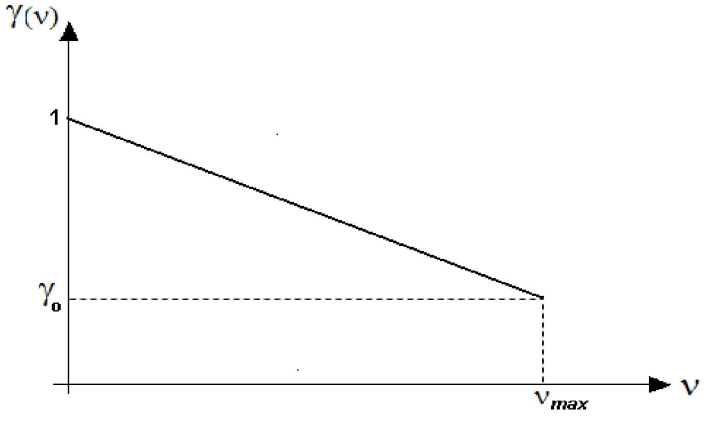

Fig. 9.8. Recovering fuzzy rate $\gamma = \gamma(\nu)$ (Barros *et al.*, 2002)

• *Solution and Equilibrium Points*

From system (9.21), we have:

$$\frac{dI}{dt} = \beta I\left[\left(1 - \frac{\beta}{\gamma}\right) - I\right]. \tag{9.23}$$

Thus we obtain the equilibrium solution, I^*,

$$I^* = \frac{\beta - \gamma}{\beta + \left[\frac{\beta-\gamma}{I_0} - \beta\right] e^{-(\beta-\gamma)t}} \tag{9.24}$$

so that, to make biological sense, this occurs when $\beta \geq \gamma$.

From the hypothesis of our model, β and γ depend on the parasite load ν. In this way, the number of infected people, at each instant of time, is given by

$$I(\nu, t) = \frac{\beta(\nu) - \gamma(\nu)}{\beta(\nu) + \left[\frac{\beta(\nu)-\gamma(\nu)}{I_0} - \beta(\nu)\right] e^{-[\beta(\nu)-\gamma(\nu)]t}}. \tag{9.25}$$

So, as in the SI fuzzy model, we can get the average number of infected individuals from $FEV[I(\nu,t)]$ or $E[I(\nu,t)]$. However, our aim here is to analyze the stability of the disease.

To study the temporal evolution of the number of infected people, that is, if the number of infected increases indefinitely or not, we should study the stability of the equilibrium points.

Making

$$\frac{dS}{dt} = 0 \qquad and \qquad \frac{dI}{dt} = 0$$

we obtain the equilibrium points $P_1 = (1,0)$ and $P_2 = \left(\frac{\gamma}{\beta}, 1 - \frac{\gamma}{\beta}\right)$ for system (9.21).

The analysis of stability of the classical model (9.21) shows that P_1 is unstable while P_2 (with $\beta \geq \gamma$) is asymptotically stable, indicating in this way that even if the number of infected increases (supposing initially I_0 small) this number will stabilize in $1 - \frac{\gamma}{\beta}$. Moreover, the fraction $\frac{\gamma}{\beta}$ of the population will not be affected.

Now, taking into account the parasite load we have:

$$P_2 = \left(\frac{\gamma(\nu)}{\beta(\nu)}, 1 - \frac{\gamma(\nu)}{\beta(\nu)}\right).$$

As mentioned above, while $\frac{\gamma(\nu)}{\beta(\nu)} < 1$, we have P_2 asymptotically stable. Therefore a value of bifurcation for ν is ν^*, the solution of the equation $\beta(\nu) = \gamma(\nu)$.

A direct calculation shows that:

$$\nu^* = \frac{\nu_M \nu_{\max}}{\nu_{\max} + (1 - \gamma_0)(\nu_M - \nu_{\min})} \tag{9.26}$$

and that $\nu_{\min} \leq \nu^* \leq \nu_M$.

The structural stability for this model is shown by the bifurcation diagram in figure 9.9, where P are the equilibrium points.

The parasite load ν^* is the value of the bifurcation of the model since for the values $\nu < \nu^*$, the model has only one unstable equilibrium point P_1 and, if $\nu > \nu^*$ the model also allows the asymptotically stable equilibrium point P_2. In this way we can think of ν^* as a parameter related to the disease control in the sense that, if a disease is installed in a population, it should be guaranteed that the parasite load ν is not higher than ν^*.

The basic reproduction number (R_0)

In general, the basic reproduction number is obtained through the local analysis of the stability of the trivial equilibrium point.

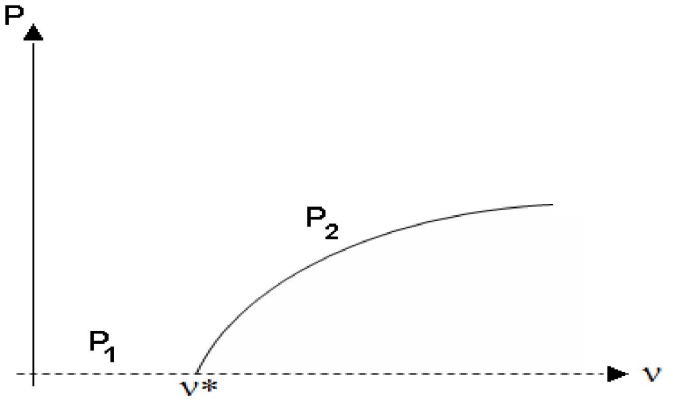

Fig. 9.9. Bifurcation diagram (Barros *et al.*, 2002)

If we consider the SI_nS model with n different infectious stages and without vital dynamics, it is not possible to obtain R_0 explicitly through the analysis of stability or through the condition $dI_n/dt > 0$ (Leite *et al.*, 2000).

For the classical SIS model we have $R_0 = \dfrac{\beta}{\gamma}$. Therefore, the disease will not establish itself if $\dfrac{\beta}{\gamma} < 1$ and it will invade the population if $\dfrac{\beta}{\gamma} > 1$.

Like the SI model above, we consider different degree of infectiousness.

As in this case we have $\beta = \beta(\nu)$ and $\gamma = \gamma(\nu)$, we could write $R_0(\nu) = \dfrac{\beta(\nu)}{\gamma(\nu)}$. However, $R_0(\nu)$ presupposes the knowledge of the parasite load to the whole population. So, in order to control the disease we can impose $\max R_0(\nu) < 1$. But this can be an extreme attitude. Perhaps it is better to adopt an "average" value of $R_0(\nu)$. For this, we consider the distribution of the parasite load as given by a triangular fuzzy number $\rho(\nu)$, as in we did for the SI fuzzy model.

We define the fuzzy basic reproduction number by

$$R_0^f = \frac{1}{\gamma_0} FEV[\gamma_0 R_0(\nu)].$$

Note that $R_0(\nu)$ can be greater than 1, but $\gamma_0 R_0(\nu) \leq 1$, so that the value R_0^f is well defined. This parameter can be thought as the average number of secondary cases caused by just one infected individual introduced into an entirely susceptible population. As mentioned in chapter 4, to get the $FEV[\gamma_0 R_0(\nu)]$ we need to define a fuzzy measure μ.

Here we will use the *possibility measure*:

$$\mu(A) = \sup_{\nu \in A} \rho(\nu), \qquad A \subset R.$$

This is a caution measure in the sense that the infectivity of a group is the one presented by the individual belonging to the group with the maximal infectivity.

In what follows we estimate R_0^f, assuming again that the amount of pathogen ν in the population has a linguistic meaning classified as *low*, *median* and *high*. The fuzzy sets given by the membership function $\rho(\nu)$ for the different cases are:

a) *low* if $\bar{\nu} + \delta < \nu_{\min}$;
b) *median* if $\bar{\nu} - \delta > \nu_{\min}$ and $\bar{\nu} + \delta \leq \nu_M$; and
c) *high* if $\bar{\nu} - \delta > \nu_M$.

Case (a): It is easy to see that $R_0^f < 1$ if ν is *low*.

Now, to obtain R_0^f for cases (b) and (c), we recall that $R_0(\nu) = \dfrac{\beta(\nu)}{\gamma(\nu)}$ is an increasing function of ν, then $H(\alpha) = \mu[\nu\,', \nu_{\max}] = \sup_{\nu' \leq \nu \leq \nu_{max}} \rho(\nu)$, where ν' is the solution of the equation $\gamma_0 \dfrac{\beta(\nu)}{\gamma(\nu)} = \alpha$. Remember that the fixed point of $H(\alpha)$ is the same as that of $FEV[\gamma_0 R_0(\nu)]$.

Case (b): Again, through a direct calculation we conclude that

$$H(\alpha) = \begin{cases} 1 & if \quad 0 \leq \alpha \leq \gamma_0 \dfrac{\beta(\bar{\nu})}{\gamma(\bar{\nu})} \\ \rho(\nu') & if \quad \gamma_0 \dfrac{\beta(\overline{\nu})}{\gamma(\overline{\nu})} < \alpha \leq \gamma_0 \dfrac{\beta(\bar{\nu}+\delta)}{\gamma(\bar{\nu}+\delta)} \\ 0 & if \quad \gamma_0 \dfrac{\beta(\bar{\nu}+\delta)}{\gamma(\bar{\nu}+\delta)} < \alpha \leq 1 \end{cases}.$$

It is easy to see that, if $\delta > 0$, H is a continuous and decreasing function, and in this case, we have that $FEV[\gamma_0 R_0(\nu)]$ is equal to the fixed point of H.

Again, a direct calculation yields $\dfrac{\beta(\bar{\nu})}{\gamma(\bar{\nu})} < \dfrac{FEV[\gamma_0 R_0(\nu)]}{\gamma_0} < \dfrac{\beta(\bar{\nu}+\delta)}{\gamma(\bar{\nu}+\delta)}$ or $R_0(\bar{\nu}) < R_0^f < R_0(\bar{\nu}+\delta)$.

Case (c): As in the previous case, we concluded that it is true that $\dfrac{1}{\gamma(\bar{v})} < R_0^f < \dfrac{1}{\gamma(\bar{v}+\delta)}$ and this guarantees that the disease will invade since $R_0^f > \dfrac{1}{\gamma(\bar{v})} > 1$.

Let us now compare R_0 and R_0^f (for a detailed analysis, see (Barros *et al.*, 2001 and 2002)).

• *Comparison between* R_0 *and* R_0^f

Our analysis here deals with the three cases studied in the previous section related to the three classifications for the amount of infection: low, medium and high parasite load.

In any of these cases, we have

$$\frac{\beta(\bar{\nu})}{\gamma(\overline{\nu})} < \frac{FEV[\gamma_0 R_0(\nu)]}{\gamma_0} < \frac{\beta(\bar{\nu}+\delta)}{\gamma(\bar{\nu}+\delta)}$$

or

$$R_0(\bar{\nu}) < R_0^f < R_0(\bar{\nu}+\delta).$$

As the function $R_0(\nu) = \dfrac{\beta(\nu)}{\gamma(\nu)}$ is crescent and continuous, based on the *Intermediate Value Theorem* (Stewart, 1999), there is only one $\widehat{\nu}$, with $\overline{\nu} < \widehat{\nu} < \overline{\nu}+\delta$, so that:

$$R_0^f = R_0(\widehat{\nu}) > R_0(\bar{\nu}). \tag{9.27}$$

This means that there is such an amount of infection $\widehat{v}$ where R_0 (classical) and the R_0^f (fuzzy) coincide. Moreover, the medium value of the number of secondary cases (R_0^f) is higher than the number of secondary cases due to the medium amount of infection ($R_0(\bar{v})$).

The values ν^*, $\bar{\nu}$ and $\widehat{\nu}$ will be used to get a hint of possible epidemic control strategies.

Epidemic control strategies

The system of equations (9.21) is the classical mathematical model to study disease of SIS type in a homogeneous population. Although we can still use such a system of equations to model the evolution of a disease in a heterogeneous population, such as the one presented in our model where the individuals are distinguished by the amount of infection and, consequently, they present different rates of contact $\beta(\nu)$ and of recovery $\gamma(\nu)$. For that, we should understand (9.21) as a family of systems depending on the parameter ν. However, if we intend to simplify it in the sense of "replacing" that family of systems by a unique system of equations, with the same outcomes (in our case, the same number of of secondary cases) that the family as a whole, the above analysis indicates that, among the different systems of families, there is one which performs this role, namely, that which parameters are $\beta = \beta(\widehat{\nu})$ and $\gamma = \gamma(\widehat{\nu})$ and not that which represents the individuals' system with medium amount of infection $\overline{\nu}$, as it seems intuitively. Moreover, according to (9.27), to elect $R_0(\bar{\nu})$ as an indicator of disease control forces us to evaluate the correct parameter for the population as a whole, that is $R_0(\widehat{\nu})$.

To justify even more the legitimacy of system (9.21) with the parameter $\widehat{\nu}$, to describe the dynamics of the disease in the population as a whole, we will study the control of the disease in the population through $R_0(\widehat{\nu}) = R_0^f$:

- For the case where the amount of infection is *low*, we have $\widehat{\nu} < \overline{\nu}+\delta \geq \nu_{\min}$. Therefore $R_0(\widehat{\nu}) = 0$ and the disease will not establish itself.

- For the case the amount of infection is *high*, we have $\widehat{\nu} > \overline{\nu} > \overline{\nu}+\delta \geq \nu_M$. Therefore, $R_0(\widehat{\nu}) = \dfrac{1}{\gamma(\overline{\nu})} > 1$, indicating that the disease will invade.

- For the case of medium amount of infection we have:
  - if $\nu^* > \widehat{\nu}$ then $R_0(\widehat{\nu}) = \dfrac{\beta(\widehat{\nu})}{\gamma(\widehat{\nu})} < \dfrac{\beta(\nu^*)}{y(\nu^*)} = 1$, indicating that the disease will not invade; and
  - if $\nu^* < \widehat{\nu}$ then $R_0(\widehat{\nu}) = \dfrac{\beta(\widehat{\nu})}{\gamma(\widehat{\nu})} > \dfrac{\beta(\nu^*)}{y(\nu^*)} = 1$, indicating that the disease will invade.

Finally, in Barros et al. (2002) it is shown that R_0^f is the positive solution of an equation of second degree, with characteristics that allow us to deduce:

1. decrease (increasing $\nu_{\min}$, consequently increasing ν^*) of the population susceptibility, what can be done through improvements in life quality of the studied population; and
2. decrease of the medium amount of infection, by the use of drugs, for example, and quarantine (decreasing δ) of the infected individuals.

These conditions are the same as that obtained to the SI fuzzy model.

Finishing this chapter, we now apply the notion of fuzzy differential inclusion to study the effect of heterogeneities in epidemic models.

9.4 Fuzzy Differential Inclusion

In this section we propose again a model in which individual heterogeneities in infectiousness is considered, assuming that the infectious capacity is a function of the parasite load. The model here analyzed is mathematically distinct from those previously studied and are treated by differential inclusion (Krivan & Colombo, 1998; Diamond, 1999). The fact of considering the contact rate as a fuzzy set allows us to use different mathematical tools for representing the model's dynamics. In the previous sections we applied differential equations and here we introduce the notion of differential inclusion, as proposed by Hüllermeier (1997). An interesting study about solutions of fuzzy differential inclusion was done by Diamond (1999).

The differential inclusion theory will be applied to obtain a solution for the SI fuzzy model. This solution is compared with the solution from fuzzy differential equation and with the number of infected individuals obtained using the deterministic model.

9.4.1 The *SI* Fuzzy Model by Fuzzy Differential Inclusion

The great difference between fuzzy differential inclusion and fuzzy differential equation is in the form that its solutions are constructed. For the differential equation, the solution of a problem is given by the collection of all trajectories of equations of the type:

$$\frac{dI}{dt} = \beta(\nu)I(1-I). \tag{9.28}$$

On the other hand, in the fuzzy differential inclusion, the solution is constructed level by level for each instant of time t.

Since β is a fuzzy set, we expect that, at each instant, the number of infected individuals is a fuzzy set too. The uncertainty in the fuzzy model is considered in the parameter ν, because we only know that $\nu \in [0, \nu_{\max}]$. In order to obtain the number of infected individuals (at each instant) as a fuzzy set, we will use the theory of differential inclusions. Krivan and Colombo (1998) suggest the substitution of the equation (9.28) by the *parametrized differential inclusion*

$$\frac{dI}{dt} \in \{\beta(\nu)I(1-I),\ \nu \in [0, \nu_{\max}]\}, \tag{9.29}$$

whose solution is a collection of all solutions of (9.28), where $\nu(t) \in [0, \nu_{\max}]$ is a measurable function (Aubin & Cellina, 1984; Krivan & Colombo, 1998). In this way, the inclusion (9.29) is a common case of unknown but limited noise, which has to be analyzed with the following control system:

$$\frac{dI}{dt} = \beta(\nu)I(1-I),\ I(0) = I_0,\ \nu(t) \in [0, \nu_{\max}]. \tag{9.30}$$

The set of all solutions of the differential inclusion (9.29) coincides with the set of all solutions of the control systems (9.30). However, the attainable set of (9.29) at the instant $t > 0$, defined by $R(t) = \{I(t) : I \text{ is a solution of (9.29)}\}$, is an interval (Aubin & Cellina, 1984). In our case, this attainable set is given by:

$$R(t) = [I_-(t), I_+(t)],$$

where $I_-(t)$ and $I_+(t)$ are, respectively, the solutions of

$$\begin{cases} \dfrac{dI}{dt} = \min\{\beta(\nu)I(1-I),\ \nu \in [0, \nu_{\max}]\},\ I(0) = I_0 \\ \dfrac{dI}{dt} = \max\{\beta(\nu)I(1-I),\ \nu \in [0, \nu_{\max}]\},\ I(0) = I_0 \end{cases}$$

or

$$\frac{dI}{dt} = I(1-I) \qquad \textit{and} \qquad \frac{dI}{dt} = 0, \qquad I(0) = I_0.$$

Then, $I_-(t) = I_0$ and $I_+(t) = \dfrac{I_0 e^t}{S_0 + I_0 e^t}$.

Thus,

$$R(t) = \left[I_0, \frac{I_0 e^t}{S_0 + I_0 e^t}\right].$$

This means that, for each $t > 0$, the number of infected individuals belongs to the interval $\left[I_0, \dfrac{I_0 e^t}{S_0 + I_0 e^t}\right]$.

From the hypothesis of β as a fuzzy set, that is, a membership function of some fuzzy subset whose domain is the set of the parasite load values, Hüllermeier (1997) proposes that the inclusion (9.29) can be seen as a fuzzy differential inclusion parametrized by ν, whose solution is, at each instant t, the fuzzy set $I(t)$, whose $\alpha - levels$, $[I(t)]^\alpha$, are given by the attainable sets $R^\alpha(t)$ of the differential inclusion

$$\begin{cases} \dfrac{dI}{dt} \in \{\beta(\nu)I(1-I),\ \nu \in [\beta]^\alpha\} \\ \\ I(0) = I_0 \end{cases}, \tag{9.31}$$

that is equivalent to

$$\begin{cases} \dfrac{dI}{dt} \in \{\beta(\nu)I(1-I)\} \\ \\ I(0) = I_0 \end{cases}, \tag{9.32}$$

with $\nu \in [\alpha(\nu_M - \nu_{\min}) + \nu_{\min}, \nu_{\max}]$.

In the same way that $R(t)$ was obtained above, we can conclude that

$$R^\alpha(t) = \left[\frac{I_0 e^{\beta(\alpha(\nu_M - \nu_{\min})+\nu_{\min})t}}{S_0 + I_0 e^{\beta(\alpha(\nu_M - \nu_{\min})+\nu_{\min})t}}, \frac{I_0 e^t}{S_0 + I_0 e^t}\right] \tag{9.33}$$

$$= \left[\frac{I_0 e^{\alpha t}}{S_0 + I_0 e^{\alpha t}}, \frac{I_0 e^t}{S_0 + I_0 e^t}\right], \qquad 0 \le \alpha \le 1.$$

Note that $R^0(t) = R(t)$.

Now, if we want to adopt a crispy curve to represent the number of infected individuals we need to choose a defuzzification method. Here we choose the classical expectancy ($E[I(\nu, t)]$) and, just to illustrate we will again analyze the cases of *low*, *medium* and *high* amount of pathogens.

So, we have (Barros *et al.*, 2004):

If $\overline{\nu} + \delta < \nu_{\min}$ (low amount of pathogens) then $E[I(\nu, t)] = I_0, \forall t > 0$.

If $\overline{\nu} - \delta > \nu_M$ (high amount of pathogens) then $E[I(\nu, t)] = \frac{I_0 e^t}{S_0 + I_0 e^t}, \forall t > 0$.

and consequently, we have

$$I(\overline{\nu} - \delta, t) \le E[I(\nu, t)] \le I(\overline{\nu} + \delta, t).$$

Then, $E[I(\nu, t)] \in R(t)$, where $R(t)$ is the attainable set of (9.29). Since $R(t)$ is an interval, again by the *Intermediate Value Theorem* (Stewart, 1999), there exists only one $\nu = \nu(t) \in [\overline{\nu} - \delta, \overline{\nu} + \delta]$, for which we have:

$$I(\nu(t), t) = E[I(\nu, t)].$$

In this way, if we represent the phenomenon by some crispy curve, $E[I(\nu,t)]$ would be the candidate, because it represents the average weighted by $\rho(\nu)$ to each $t>0$. However, this curve is not a solution of the initial model (9.28), since $E[I(\nu,t)] = I(\nu(t),t)$, where $\nu(t)$ is not constant.

To end this section we would like to highlight that the deterministic model (without uncertainty) indicates the adoption of $\overline{\nu}$ for the parasite load and, in this case, the number of infected individuals follows the trajectory $I(\overline{\nu},t) = \dfrac{I_0 e^{\beta(\overline{\nu})t}}{S_0 + I_0 e^{\beta(\overline{\nu})t}}$, whose membership in the fuzzy solution (9.33) is

$$u_{I(t)}(I(\overline{\nu},t)) = \frac{1}{t}\ln\left(\frac{S_0}{I_0}\frac{I(\overline{\nu},t)}{1-I(\overline{\nu},t)}\right) = \beta(\overline{\nu}), \qquad for\ all \qquad t>0.$$

Then, we can say that the deterministic solution is the one that has the higher possibility of occurring (it is preferred) since $\overline{\nu}$ is the amount of pathogen with the higher chance of occurring.

Epidemic control strategies by differential inclusion

In Barros *et al.* (2004), as in the SI fuzzy equations, it is shown that:

- if $S_0 > I_0$, and while $t \le \ln\dfrac{S_0}{I_0}$, then $E[I(\nu,t)] > I(\overline{\nu},t)$. From $t = \ln\dfrac{S_0}{I_0}$, there is an instant $\overline{t}$ for which $E[I(\nu,t)] = I(\overline{\nu},t)$. For $t > \overline{t}$, we have $E[I(\nu,t)] < I(\overline{\nu},t)$, indicating that the deterministic model underestimates the number of infected individuals at the beginning and overestimates it from $\overline{t}$.
- if $S_0 \le I_0$ then $E[I(\nu,t)] \le I(\overline{v},t), \forall t>0$, the deterministic model overestimates the number of infected individuals.

Then, at the beginning of the epidemic $\left(t < \ln\dfrac{S_0}{I_0}\right)$ and $S_0 >> I_0$, we have

$$I(\overline{\nu},t) \le E[I(\nu,t)] \le I(\overline{\nu}+\delta,t).$$

Therefore, $\nu(t) \in [\overline{\nu}, \overline{\nu}+\delta]$. Since $E[I(\nu,t)] = I(\nu(t),t)$ increases when $\nu(t)$ increases too, we have that the higher the $\overline{\nu}$, the higher the $E[I(\nu,t)]$; the higher the δ, the higher the $E[I(\nu,t)]$ and the higher the $\nu_{\min}$, the higher the $E[I(\nu,t)]$.

Then, it is possible to interfere in the disease evolution in two ways:

1) Increasing the value of $\nu_{\min}$. This is a consequence of increasing the resistance of susceptible individuals (diminishing the susceptibility) and it can be obtained, for example, through vaccination, sanitation, etc, indicating that the parameter $\nu_{\min}$ is related with the susceptible individuals; and

2) Another option is to diminish $E[I(\nu,t)]$ by reducing δ. The reduction of δ could be made with control policy related to the infected population, for example,

by quarantine isolation. The reduction of $\bar{\nu}$ is related to the treatment as, for example, with drugs.

We, therefore, have two possibilities of control strategies: (a) an action with all the population and (b) to act directly to the infected individuals. Obviously, a combination of both strategies would have a better efficacy to prevent and to control the disease. We would like to highlight that the control strategies proposed for both fuzzy differential equation and fuzzy differential inclusion are qualitatively the same.

In this chapter we presented some structures to treat dynamical epidemiological systems. In spite of the calculations, which can be quite complex in some situations, the examples described here have demonstrated how important it is to discuss the role that fuzzification can play from the dynamical point of view. The differences in R_0 values, comparing the classic and fuzzy approaches, illustrate this importance.

From the calculations point of view both approaches, fuzzy differential equation and fuzzy differential inclusion, presented an important limitation: the number of parameters that can be fuzzified. In fact, the calculations become almost impossible if we wish to defuzzify more then one parameter in the equations. To circumvent this situation an alternative approach is presented in the next chapter.

10 Classical Dynamical Systems with Fuzzy Rule-Based Parameters

Along chapters 8 and 9 we presented several approaches to treat the uncertainties in dynamical systems applied to epidemic problems. We have also discussed the limitations and advantages of those techniques comparing them to each other. While chapter 8 is devoted to dynamical systems based only on fuzzy linguistic rules structure, chapter 9 presents different mathematical ways to apply fuzzy differential equations, considering the uncertainties of variables or parameters of a classical differential equation system, based on the fuzzification of them using fuzzy sets. In the first approach the most important limitation is the explosion of number of rules and dependence of the expert knowledge, which is particularly true in epidemic studies. In contrast, the largest difficulty found in chapter 9 approaches is the complexity of the calculations, particularly if we want to fuzzify more than one parameter. In fact, to analyze the FEV structure considering two or more parameters as fuzzy sets is a really hard task, from the mathematical point of view.

In this chapter we present an alternative to elaborate fuzzy dynamical systems by means of another methodology. The idea is to use the classic differential equations, and consequently their mathematical theorems and solutions available, and to find the equations parameters through fuzzy rules modeling. So, the differential equations and the fuzzy linguistic models are applied together. Thus, in a certain way, this approach can be seen as a mixing of the approaches exposed in the two previous chapters. The great advantage of this methodology when compared with the approaches presented in chapter 9 is that it allows to consider the uncertainties of more than one parameter, aggregating them by means of a fuzzy rule-based model.

To illustrate the methodology we present two examples. In the first example it is shown a HIV model for dynamical behavior between non-symptomatic and symptomatic seropositive individuals (Jafelice *et al.*, 2004a); in the second example this structure is used to study the influence of HIV epidemic in the expectancy of life in a group of seropositive individuals (Jafelice *et al.*, 2004b).

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 207–223, 2008.
springerlink.com

10.1 Fuzzy Modeling in Symptomatic a HIV Infected Population

Jafelice *et al.* (2004a) proposed a model to study the the evolution of HIV positive individuals. In this work the focus is on the nature of the transference rate, λ, of HIV to full blown AIDS. Expert knowledge indicates that this transference rate is uncertain and depends strongly on the viral load and on the $CD4+$ level of infected individuals (see section 7.6.1 in chapter 7). Jafelice *et al.* (2004a) propose the transference rate as a linguistic variable of the viral load and $CD4+$ level values. In this case the dynamical model results in a fuzzy model that preserves the biological meaning and nature of the transference rate λ. Its behavior fits the natural history of HIV infection reported in the medical literature. A comparison between the fuzzy model and a classical model using data available in the literature was also done.

The model used here to predict the number of HIV positive individuals is the same that proposed by Anderson (1988). However, the transference rate, λ, is viewed as a linguistic variable, whose values are fuzzy sets that depend on the viral load v and $CD4+$ level.

10.1.1 Classical HIV/AIDS Models

The classical Anderson's model (1988) is a macroscopic model for HIV/AIDS. It is described by:

$$\begin{cases} \dfrac{dx}{dt} = -\lambda(t)x & x(0) = 1 \\ \dfrac{dy}{dt} = \lambda(t)x = \lambda(t)(1-y) & y(0) = 0 \end{cases}, \tag{10.1}$$

where $\lambda(t)$ is the transference rate between infected individuals and individuals with AIDS; x is the proportion of infected individuals that does not develop AIDS (asymptomatic); and y is the proportion of the individuals that develop AIDS (symptomatic). Anderson (1988) assumes $\lambda(t) = at$, $a > 0$. Thus the solution of (10.1) becomes:

$$\begin{cases} x(t) = e^{-\frac{at^2}{2}} \\ y(t) = 1 - e^{-\frac{at^2}{2}} \end{cases}. \tag{10.2}$$

Peterman *et al.* (1985) present data related to 194 cases of blood transfusion-associated AIDS. From Peterman *et al.* data Murray (1990) shows that Anderson's model (10.1) can be fitted to find the value for the parameter a. The rate of increase dy/dt of AIDS patients as a function of time, provided by model (10.1), is shown by the continuous curve of figure 10.1.

Nowak and Bangham (1996) introduce three microscopic models for HIV infection dynamics in the individuals organism, with no anti-retroviral therapy.

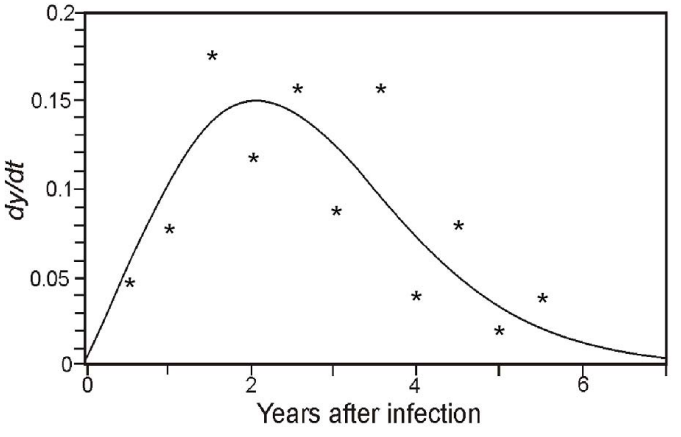

Fig. 10.1. The rate of change in the proportion of the individuals who develop AIDS after infected with HIV (through blood transfusion) at time $t = 0$. The data (from Peterman *et al.*, 1985) shows a best-fit $a = 0.237\text{yr}^{-1}$ solution for model (10.1) (Murray, 1990)

Two of these models will be explained in this section. The first is a model for the interaction between replicating virus and host cells. In this case, there are three variables: uninfected cells n, infected cells i, and free virus particles v. Infected cells are produced from the interaction between uninfected cells and free virus at rate βnv, and die at rate bi. Free viruses are produced from infected cells at rate ki and decline at rate sv. Uninfected cells are produced at a constant rate, r, from a pool of precursor cells, and die at rate an (see figure 10.2 of Nowak, 1999).

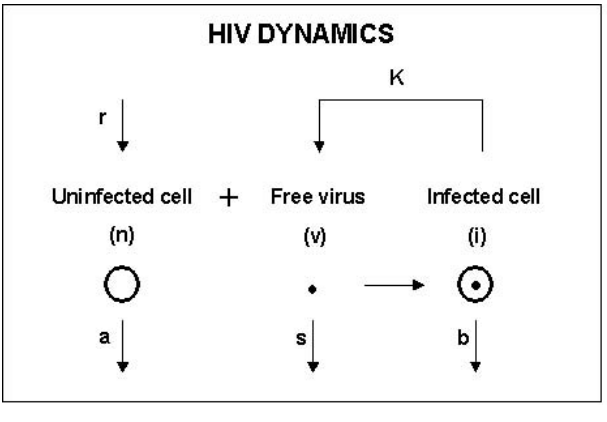

Fig. 10.2. Microscopic model of HIV virus dynamics (modified from Nowak, 1999)

These assumptions are described by the following system of differential equations:

$$\begin{cases} \dfrac{dn}{dt} = r - an - \beta nv \\ \dfrac{di}{dt} = \beta nv - bi \\ \dfrac{dv}{dt} = ki - sv \end{cases} . \tag{10.3}$$

Recently, Filter *et al.* (2005) presented a mathematical method for estimating all the parameters of the three-dimensional (3-D) model (10.3) of the HIV infection. An application of vaccine readiness was deduced from the estimation of the viral load set point and the setting time for patients from a South African cohort. Another recent contribution is the study of Ouattara *et al.* (2008) where the HIV dynamic is described by means of (3-D) model (10.3) adding to the first equation a $CD4+$ T cells proliferation term. The authors studied also the influence of the *Highly Active AntiRetroviral Therapy* (HAART) in the parameters of the model (10.3), analyzing the therapeutic failures based on mathematical modeling of the HIV infection (Ouattara *et al.*, 2008).

The second model includes immune responses against infected cells, and extends system (10.3) by adding an equation to describe the immune response against infected cells:

$$\begin{cases} \dfrac{dn}{dt} = r - an - \beta nv \\ \dfrac{di}{dt} = \beta nv - bi - piz \\ \dfrac{dv}{dt} = ki - sv \\ \dfrac{dz}{dt} = ciz - dz \end{cases} . \tag{10.4}$$

The variable z denotes the magnitude of the CTL (cytotoxic T lymphocyte)-that is, the abundance of virus-specific CTLs. The rate of CTL proliferation in response to antigen is given by ciz. In the absence of stimulation, CTLs decay at rate dz. Infected cells are killed by CTLs at rate piz. Figure 10.3 shows the solution of (10.4) using the following parameters: $r = 0.3$, $a = 0.1$, $\beta = 1$, $b = 0.01$, $p = 0.03$, $k = 0.5$, $s = 0.01$, $c = 0.01$ and $d = 0.01$; and the following initial conditions: $n(0) = 0.99$, $i(0) = 0.01$, $v(0) = 0.1$, $z(0) = 0.01$, $t_{initial} = 0$ time units and $t_{final} = 500$ time units (Caetano & Yoneyama, 1999; Jafelice *et al.*, 2004a).

In logarithmic scale, the uninfected cells of $CD4+$ show a rapid decline in the first weeks and a slow recovery when the number of lymphocytes is close to the maximum. The increase in the number of lymphocytes is related to the presence of infected cells and the virus replication mediated by them (see figure 10.3).

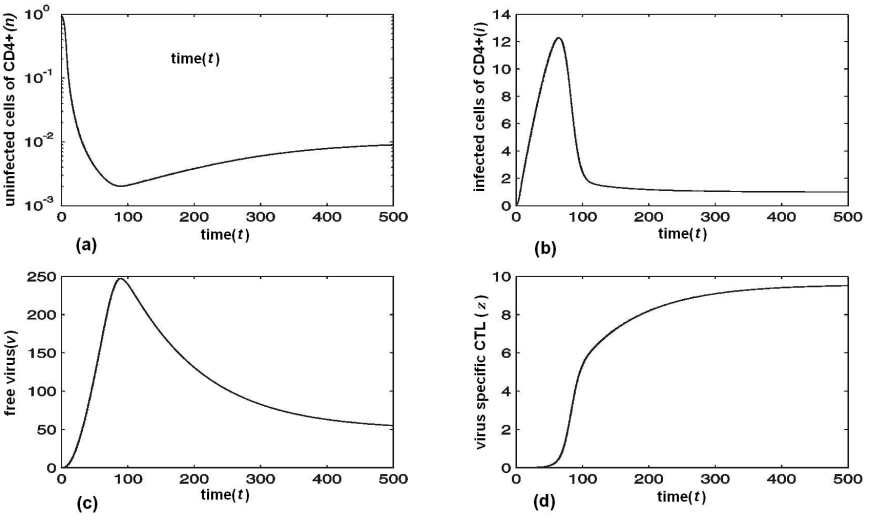

Fig. 10.3. Solution of the microscopic HIV model (Jafelice *et al.*, 2004a)

Comparing the solution of system (10.4) shown in figure 10.3 with the plots shown in figure 7.11, in chapter 7, it is possible to note similarities between the behavior of the uninfected cells of $CD4+$, the free virus, and the virus-specific CTLs with the $CD4+$ level, the HIV virus, and the HIV antibodies, respectively. In figure 7.11 it can also be observed that, during the asymptomatic phase, the variation of uninfected cells of $CD4+$ is small. Therefore, it may be assumed that $dn/dt \cong 0$ which means, from (10.4), that $n(v) \cong \dfrac{r}{a + \beta v}$.

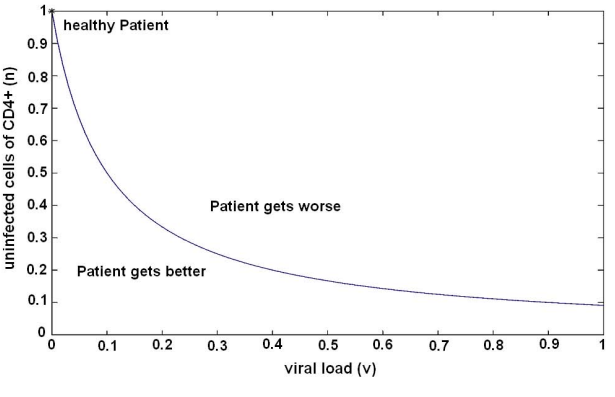

Fig. 10.4. Uninfected $CD4+$ cells *versus* viral load (v) (Jafelice *et al.*, 2004a)

Moreover, after straightforward calculations, we note that (see figure 10.4):

- If $n(v) = \dfrac{r}{a+\beta v}$ then $\dfrac{dn}{dt} = 0$. Thus, $n(t)$ is constant;
- If $n(v) < \dfrac{r}{a+\beta v}$ then $\dfrac{dn}{dt} > 0$. Therefore, $n(t)$ is increasing. This means that the infected individual is recovering because the number of cells of $CD4+$ is growing; and
- If $n(v) > \dfrac{r}{a+\beta v}$ then $\dfrac{dn}{dt} < 0$. Therefore, $n(t)$ is decreasing. This means that the infected individual is worsening because the number of $CD4+$ cells is diminishing.

10.1.2 A Fuzzy Rule-Based Model to Estimate λ

When HIV reaches the bloodstream, it attacks mainly the lymphocyte T of the $CD4+$ type. The amount of $CD4+$ cells in the peripheral blood has prognostic implications for the infection evolution by HIV. Currently, the concentration of immune competent cells is the most clinically useful and acceptable assessment of the treatment of infected individuals with HIV, although it is not the only one.

The identification of the disease's stages and its respective treatment is based on the relationship between viral load and $CD4+$ level. The control of the viral load and $CD4+$ cells level can interfere in the control of the transference rate λ. Thus, the conversion from an asymptomatic individual to a symptomatic individual depends on the individual characteristics, as measured by the viral load, v, and level of $CD4+$, c. Therefore, it is suggested the following model:

$$\begin{cases} \dfrac{dx}{dt} = -\lambda(v,c)x & x(0) = 1 \\ \dfrac{dy}{dt} = \lambda(v,c)x = \lambda(v,c)(1-y) & y(0) = 0 \end{cases}. \tag{10.5}$$

The difference between the model suggested in (10.5) and the classical model (10.1) is that in (10.5) the parameter $\lambda = \lambda(v,c)$, that is, λ is not a constant value. This assumption has a clear biological meaning and thus is a more reliable characterization of λ. From the mathematical point of view, (10.5) can be seen as a parametric family of systems. It seems reasonable that λ, and consequently the population of infected individuals y, could be controlled via v and c. From (10.5) we have, for $t > 0$:

$$\begin{cases} x(t) = e^{-\lambda(v,c)t} \\ y(t) = 1 - e^{-\lambda(v,c)t} \end{cases}. \tag{10.6}$$

The estimation of the transference rate $\lambda = \lambda(v,c)$ is based on linguistic medical information in the form of fuzzy if-then rules. Therefore, it was

adopted a fuzzy rule-based modeling approach assuming, as in the case of medical knowledge, that the viral load, v, the level of $CD4+$, c, and the transference rate, λ, are linguistic variables denoted by V, $CD4+$ and Λ, respectively. Viral load V has its values in the set of terms $\{low, medium, high\}$, $CD4+$ in $\{very\ low, low, medium, high\ medium, high\}$, and transference rate Λ in the terms set $\{weak, medium\ weak, medium, strong\}$. The $CD4+$ level between 0.2 and 0.5 cells/ml was divided into two ranges because it relates to an important phase of the transference from asymptomatic to symptomatic individuals. The membership functions that specify the meaning of the linguistic variables are shown in figures 10.5, 10.6 and 10.7 for *viral load*, $CD4+$ *level*, and *transference rate*, respectively. Table 10.1 presents the rule base that encodes the relationships between c, v, and λ suggested by expert medical knowledge.

Note that, in chapter 7 we presented a fuzzy linguistic model, of Mamdani type, to find the clinical progression considering as input variable $CD4+$ and *viral load*, both of them elaborated based on the absolute measures of variables (see

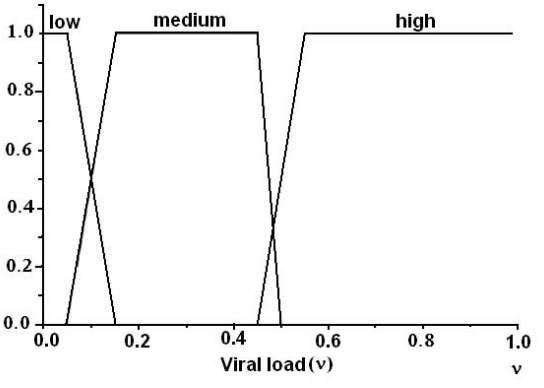

Fig. 10.5. Membership functions for viral load (V) (Jafelice *et al*, 2004a)

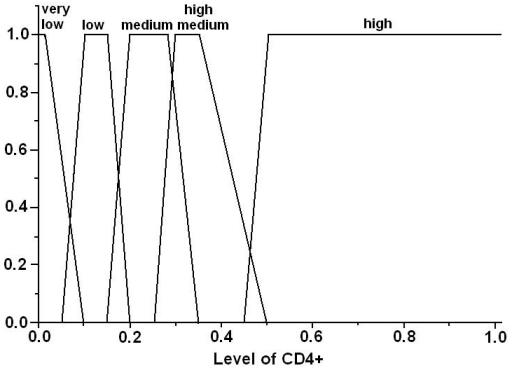

Fig. 10.6. Membership functions for $CD4+$ level (Jafelice *et al*, 2004a)

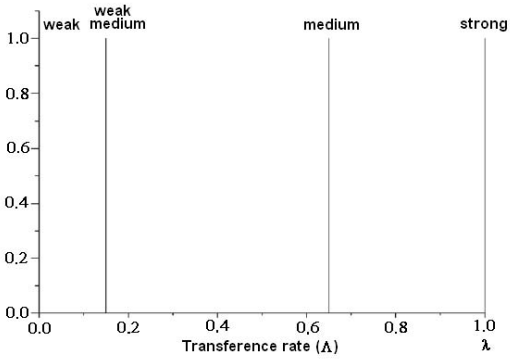

Fig. 10.7. Membership functions for transference rate (Jafelice *et al*, 2004a)

Table 10.1. Rules that encodes the relationships between c, v, and λ (Jafelice *et al.*, 2004a)

Rule	IF V is	AND CD4+ is	THEN Λ is
1	*low*	*very low*	*strong*
2	*low*	*low*	*medium*
3	*low*	*medium*	*medium*
4	*low*	*high medium*	*weak medium*
5	*low*	*high*	*weak*
6	*medium*	*very low*	*strong*
7	*medium*	*low*	*strong*
8	*medium*	*medium*	*medium*
9	*medium*	*high medium*	*weak medium*
10	*medium*	*high*	*weak*
11	*high*	*very low*	*strong*
12	*high*	*low*	*strong*
13	*high*	*medium*	*medium*
14	*high*	*high medium*	*medium*
15	*high*	*high*	*medium*

their membership functions in chapter 7). However, in the fuzzy rule-based model considered here the output variable is the transference rate from asymptomatic to symptomatic groups. Thus the membership functions of input variables were defined based on more adequate measures. In addition, the fuzzy model used in this case is the Takagi-Sugeno-Kang type.

Thus, the solution of (10.5) is given by (10.6) and $\lambda(v, c)$ is obtained from the rule-based fuzzy system. The next section shows how to obtain an analytical expression for λ as a function of v and c using fuzzy inference and a defuzzification method.

Transference rate λ

Given the fuzzy rule base above it is possible to compute the value of $\lambda = \lambda(v, c)$ for values of viral load v and the respective values of $CD4+$ levels. Note that in this case the fuzzified parameter of the differential equation, λ, depends on two variables.

We identified the uninfected cells n with the $CD4+$ level because the uninfected cells of figure 10.3 have similar behavior as the $CD4+$ T lymphocyte of figure 7.11 and the blood test does not differentiate uninfected cells n from infected cells i. The blood test identifies $CD4+$ level only. Therefore, once $CD4+$ is proportional to n, it was assumed that:

$$c(v) = \frac{r}{a + \beta v}. \tag{10.7}$$

Moreover, when we project the defuzified transference rate curve in the transference rate versus viral load plane, and approximate the projection by straights lines, the approximation of λ becomes (see Jafelice *et al.*, 2004a):

$$\lambda(v) = \begin{cases} 0 & if\ 0 < v < v_{min} \\ \dfrac{v - v_{min}}{v_M - v_{min}} & if\ v_{min} \leq v < v_M \\ 1 & if\ v \geq v_M \end{cases}. \tag{10.8}$$

Note that there is a need for a minimum amount of virus v_{min} for disease evolution and that, above a certain amount of virus v_M, the chance for evolution becomes high. The amount of virus is always limited by v_{max}. Thus, the solution of (10.5) is given by (10.6) where $\lambda(v)$ is given by (10.8).

Note that, in this case, the rate λ is a function of v. So, the functions in (10.6) are a solution family of (10.5). To obtain just one value for each t, the $FEV(y)$ value was calculated, as done in chapter 9 (see also chapter 4). For this, the distribution of the value of the viral load, V, is needed.

Medical knowledge suggests that viral load can be reasonably represented by a triangular fuzzy set, whose analytical form is (see figure 9.3, in chapter 9):

$$\rho(v) = \begin{cases} 1 - \left| \dfrac{v - \bar{v}}{\delta} \right| & if\ v \in [\bar{v} - \delta, \bar{v} + \delta] \\ 0 & if\ v \notin [\bar{v} - \delta, \bar{v} + \delta] \end{cases}. \tag{10.9}$$

Parameters $\bar{v}$ and δ are the modal value and spread of the fuzzy set, respectively. This can be viewed as a way to express viral load values through a fuzzy number, a particular kind of a linguistic variable whose values are normal. Fuzzy number is a way to quantify imprecise valuations such as *the viral load is around* $\bar{v}$. Precise viral load values for populations are rare, but fuzzy numbers capture the inherent imprecision typical of biological variables, such as viral load and $CD4+$ level. So, given a distribution of the viral load and the transference rate function, it is possible to find the expected value of the symptomatic individuals for a time horizon.

Fuzzy expectancy of the symptomatic individuals

Fuzzy expectancy is the concept used here to determine the expected value of the proportion of symptomatic individuals $y(t) = 1 - e^{-\lambda(v)t}$ for each time $t > 0$. Fuzzy expectancy uses the idea of fuzzy measure and fuzzy integrals (see chapter 4).

Considering a procedure analogous to that adopted in section 9.3.1 (in chapter 9), results in three cases of interest (Jafelice *et al.*, 2004a):

Case a) *Viral load is low*: It is assumed $v_{min} > \bar{v} + \delta$, and it is find $FEV[y] = 0$.

Case b) *Viral load is high*: It is assumed $v_M \geq \bar{v} - \delta$ and $\bar{v} + \delta \geq v_{max}$, and it is find $FEV[y] = 1 - e^{-t}$.

Case c) *Viral load is medium*: In this case it is assumed $\bar{v} - \delta > v_{min}$ and $\bar{v} + \delta < v_M$, and we obtain the following inequality:

$$1 - e^{\left(\frac{-\bar{v}+\delta+v_{min}}{v_M - v_{min}}\right)t} < FEV[y] < 1 - e^{\left(\frac{-\bar{v}-\delta+v_{min}}{v_M - v_{min}}\right)t} \tag{10.10}$$

for

$$\bar{v} - \delta < v(t) < \bar{v} + \delta,$$

where $v(t) = v_{min} + \left[\frac{-ln(1-\alpha)}{t}\right](v_M - v_{min})$, with $\alpha = FEV[y]$.

We denote by $FEV_{low} = 1 - e^{\left(\frac{-\bar{v}+\delta+v_{min}}{v_M - v_{min}}\right)t}$ the optimistic proportion of the symptomatic individuals and by $FEV_{upper} = 1 - e^{\left(\frac{-\bar{v}-\delta+v_{min}}{v_M - v_{min}}\right)t}$ the pessimistic proportion of the symptomatic individuals. Therefore,

$$FEV_{low} < FEV[y] < FEV_{upper}. \tag{10.11}$$

In addition, we have the following proposition.

Proposition 10.1. *For each $t > 0$, there exists a unique $v(t) \in (\bar{v} - \delta, \bar{v} + \delta)$ for which*

$$FEV[y] = 1 - e^{\left(\frac{-v(t)+v_{min}}{v_M - v_{min}}\right)t}. \tag{10.12}$$

(for proof see Jafelice et al., 2004a).

It is important to highlight that the fuzzy expected value is not a solution of (10.5). What proposition 10.1 shows is that, for each time t, there is a solution of (10.5), whose value at t is the same as $FEV[y]$ at t. Actually, it is not difficult to verify that $FEV[y]$ is differentiable and satisfies the following differential equation with the time dependent parameter $v(t)$:

$$\frac{dy}{dt} = \left[\frac{v(t) - v_{min}}{v_M - v_{min}} + \frac{t}{v_M - v_{min}}\frac{dv(t)}{dt}\right](1 - y) \tag{10.13}$$

Interestingly, the fuzzy expected value $FEV[y]$ of an autonomous differential equation is a solution of a corresponding non-autonomous differential equation.

In the next section it is verified how the fuzzy expected value of symptomatic individuals compares with the solution suggested in (Murray, 1990) using real data.

10.1.3 Fuzzy Expectancy of Symptomatic Individuals and Real Data

As it was discussed above, Murray (1990) presents the model (10.1) fitted to find a from the data of Peterman *et al.* (1985). In this section these data were used to derive a fuzzy model and to compute the fuzzy expected value of the infected individuals (see figure 10.1). It is assumed that, initially, the fraction of infected asymptomatic individuals x is 1 (maximum), and that the fraction of AIDS symptomatic individuals y is null. Note that $FEV[y]$ (see proposition 10.1) depends on the parameters $v(t)$, v_{min} and v_M. Since the values of $y(t)$ can be obtained from the data shown in figure 10.1 and, from proposition 10.1 and equation (10.6), it is possible to find the values of v for each t. From these values of v a least-squares fitting provides the following estimates for $v(t)$, (see figure 10.8):

$$v(t) = 0.067t + 0.036. \tag{10.14}$$

Jafelice *et al.* (2004a) adapted the values $v_{min} = 0.046$ and $v_M = 0.56$ and compute $FEV[y]$.

Recall that the transference rate that fits the classic model (10.1) to the data is $\lambda(t) = 0.237t$ (Murray, 1990).

In this case, it is possible to find an explicit view of the time behavior of the transference rate if the value of $v(t)$, given by (10.14), is replaced in (10.8). The

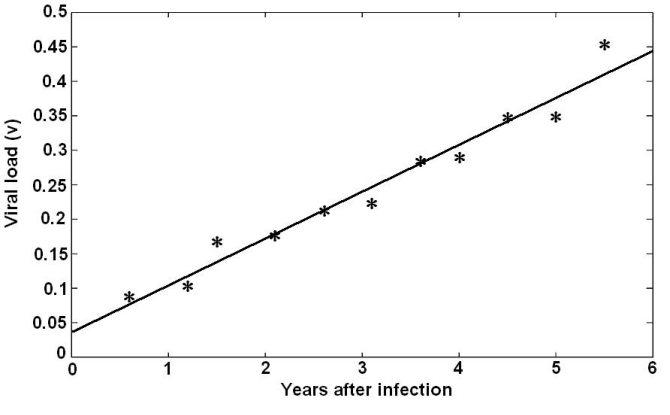

Fig. 10.8. Least squares fit of $v(t)$ (Jafelice *et al.*, 2004a)

result, given by (10.15) below, is an estimate of the transference rate λ of (10.5), shown in figure 10.9 for $0 < t < 6$,

$$\lambda_{v,t} = \begin{cases} 0 & if \quad 0 < v < 0.046 \\ 0.12t - 0.019 & if \quad 0.046 < v < 0.56 \end{cases} . \tag{10.15}$$

As it can be noted from figure 10.9, the estimates of the transference rate as suggested by the fuzzy model fits real data much more accurately than the one provided by the classical model.

Given the estimated values of λ (figure 10.9), from proposition 10.1 it is easy to compute $FEV[y]$ for each $t > 0$. Figure 10.10 shows the plots of the upper and lower bounds for $FEV[y]$, the estimated values of $FEV[y]$ together with

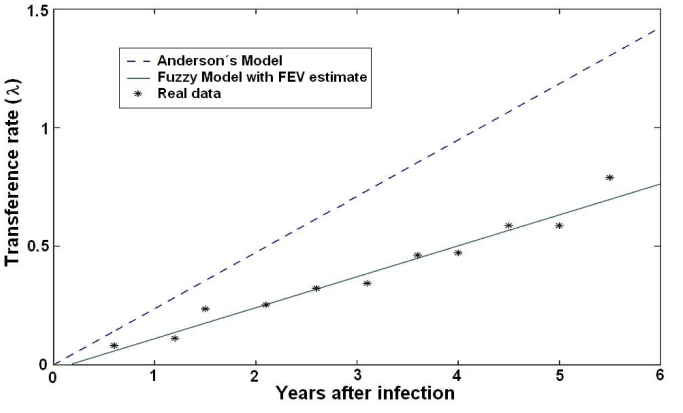

Fig. 10.9. Time behavior of the transference rate (Jafelice *et al.*, 2004a)

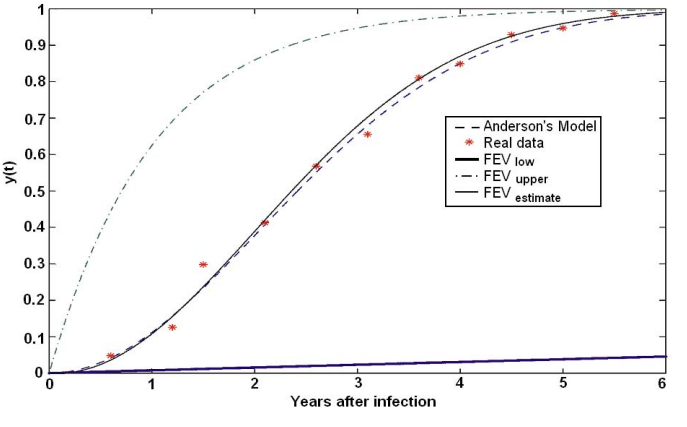

Fig. 10.10. Bounds for $FEV[y]$, the fuzzy expected value of the symptomatic individuals, Anderson's model, and real data (Jafelice *et al.*, 2004a)

the values given by Anderson's model (1988) and the data reported in (Murray, 1990). Clearly, $FEV[y]$ does provide a solution that is closer to actual data than Anderson's model. Moreover, the fuzzy model uses information from medical experts and biological principles to determine estimates of transference rate.

The main difference between the classical model (10.1) and the fuzzy model (10.5) is that the fuzzy model exploits parameter uncertainty whereas classical model does not. In a sense, the classical model is a particular case of fuzzy models. In addition, we saw that Anderson's model is derived from a best fit to data while the fuzzy model is constructed from biological principles. In other words, the fuzzy model provides a clear and meaningful characterization of parameter λ, since it is compatible with the available medical knowledge and perception of its values. The adherence of the transference rate of (10.5) to real data reported in the literature, when $v = v(t)$, is very significant. It gives the medical science better estimates to control the AIDS evolution.

10.2 Fuzzy Model to Compute the Life Expectancy of HIV Infected Individuals

This section is based on the Jafelice and collaborators (2004b) work where it was studied the impact of AIDS on the life expectancy of a population, considering a classical population model complete with expert knowledge information via a fuzzy rule-based system. In this model, it was assumed that AIDS has a direct influence on the mortality rate of a population and a fuzzy rule base system was derived to capture this influence using as variables the viral load, the $CD4+$ level and a parameter that modify the transference rate. They determined the average number of individuals and the life expectancy for specific groups without anti-retroviral therapy. Thus, the aim was to study the life expectancy of an HIV infected individuals using a conventional population model with a mortality rate derived from expert knowledge.

We assume that in a group of people deaths occur due to other causes in addition to HIV infection. The model assumes neither birth nor migration. The number of individuals, $n(t)$, living at time t satisfies the following differential equation:

$$\frac{dn}{dt} = -(\lambda_1 + \phi(v, c)\lambda_2)n, \tag{10.16}$$

where λ_1 is the natural mortality rate, $\phi(v, c)$ indicates the influence of AIDS in the mortality rate, and λ_2 is a constant, characteristic value of the population and depends on its social and environment behavior. The solution of (10.16), $n(t)$, is given by:

$$n(t) = n(0)e^{-(\lambda_1+\phi(v,c)\lambda_2)t}, \quad t > 0. \tag{10.17}$$

As we have seen, the population model above suggests a straightforward mechanism to describe the population behavior. However, in practice, precise values of the influence of AIDS on the mortality rate are rare, but experts do know how to evaluate the influence of AIDS from their perception of the

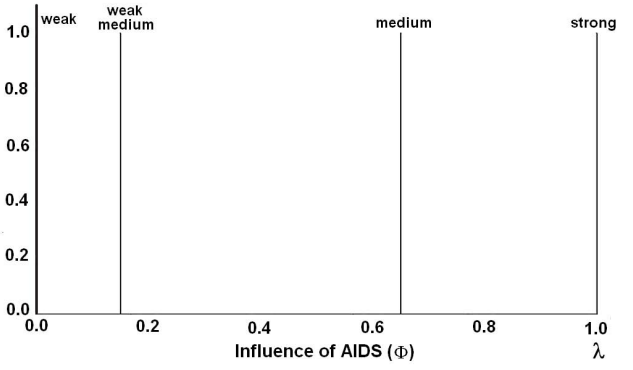

Fig. 10.11. Membership functions for influence of AIDS, Φ (Jafelice *et al.*, 2004b)

Table 10.2. Fuzzy rules for the fuzzy rule-based model to estimate the influence of AIDS (Jafelice *et al.*, 2004b)

	V		
CD4+	**Low**	**Medium**	**High**
Very Low	Strong	Strong	Strong
Low	Medium	Strong	Strong
Medium	Medium	Medium	Medium
High Medium	Weak Medium	Weak Medium	Medium
High	Weak	Weak	Weak

relationships between c, v, and ϕ. In this section, the value of the influence of AIDS, $\phi = \phi(v, c)$, is obtained from linguistic medical information translated into a set of fuzzy if-then rules. In other words, we adopt a fuzzy rule-based modeling approach assuming, as it is the case with medical knowledge, that the viral load, v, the level of $CD4+$, c, and the influence of AIDS on the mortality, ϕ, are linguistic variables denoted by V, $CD4+$ and Φ, respectively. Viral load V has its values in the set terms $\{low, medium, high\}$, $CD4+$ in $\{very\ low, low, medium, high\ medium, high\}$, and influence of AIDS in $\{weak, medium\ weak, medium, strong\}$. The membership functions that specify the meaning of the linguistic variables are the same that is shown in figures 10.5, 10.6 and 10.11 for viral load, $CD4+$ level, and influence of AIDS, respectively. The rule base that encodes the relationship between c, v, and ϕ, as suggested by expert medical knowledge, is summarized in table 10.2.

10.2.1 The Influence of AIDS on the Mortality Rate

First, we note that the set of differential equations (10.4) gives a microscopic dynamic model for HIV infection dynamics within the organism of an individual, assuming no anti-retroviral therapy. When in equilibrium, this model suggests an important relationship between the viral load, v, and the level of $CD4+$, c:

$$c(v) = \frac{r}{a + \beta v}. \tag{10.18}$$

Thus, (10.18) shows that *viral load* and $CD4+$ level are not independent variables. The parameters values used here are $r = 0.3$, $a = 0.1$ and $\beta = 1$.

Second, from the fuzzy rule base of table 10.2 we can compute $\phi(v, c)$ for given values of viral load v and respective values of $CD4+$ level. We propose that $\phi(v)$ is given by:

$$\phi(v) = \begin{cases} 0 & if \quad 0 < v < v_{min} \\ \dfrac{v - v_{min}}{v_M - v_{min}} & if \quad v_{min} \leq v < v_M \\ 1 & if \quad v \geq v_M \end{cases} . \tag{10.19}$$

Note that there is a need for a minimum amount of virus v_{min} for the disease's evolution and that, above a certain amount of virus v_M, the chance for evolution becomes high. The amount of virus is always limited by v_{max} .

Replacing $\phi(v, c)$ by (10.19) in (10.17) we find (10.20); that is, the number of individuals at time t becomes:

$$n(t) = n(0)e^{-(\lambda_1 + \phi(v)\lambda_2)t}, \quad t > 0. \tag{10.20}$$

10.2.2 Average Number and the Life Expectancy of Individuals

Assuming that viral load can be reasonably represented by a triangular fuzzy set given by equation (10.9). Thus, for each instant t, the average number of individuals, $n(t)$, and the life expectancy, E, are determined by:

$$\langle n(t) \rangle = n(0) \int_R e^{-(\lambda_1 + \phi(v)\lambda_2)t} \frac{\rho(v)}{\delta} dv \tag{10.21}$$

$$E = \int_R \frac{1}{\lambda_1 + \phi(v)\lambda_2} \frac{\rho(v)}{\delta} dv, \tag{10.22}$$

where $\dfrac{\rho(v)}{\delta}$ is the distribution of the viral load.

Considering the HIV positive individuals with viral load, v, *low*, *medium* and *high*, whose fuzzy sets have the same shape as presented in the previous sections, we can study the life expectancy through the FEV calculation:

Case a) *Viral load is low*: We assume $v_{min} > \bar{v} + \delta$, and get:

$$\langle n(t) \rangle = n(0)e^{-\lambda_1 t} \quad and \tag{10.23}$$

$$E = \frac{1}{\lambda_1}. \tag{10.24}$$

Case b) *Viral load is high*: We assume $v_M \leq \bar{v} - \delta$ and $\bar{v} + \delta \leq v_{max}$. Thus,

$$\langle n(t) \rangle = n(0)e^{-(\lambda_1+\lambda_2)t} \quad and \tag{10.25}$$

$$E = \frac{1}{\lambda_1 + \lambda_2}. \tag{10.26}$$

Case c) *Viral load is medium*: In this case, we assume $\bar{v} - \delta > v_{min}$ and $\bar{v} + \delta < v_M$. Therefore,

$$\begin{aligned}\langle n(t) \rangle = \; & \frac{n(0)e^{-\lambda_1 t}}{\delta^2 t^2 \lambda_2^2}\Bigg[- 2(v_M - v_{min})^2 e^{-\left(\frac{\bar{v}-v_{min}}{v_M - v_{min}}\right)\lambda_2 t} \\ & + (v_M - v_{min})^2 e^{-\left(\frac{\bar{v}-\delta-v_{min}}{v_M - v_{min}}\right)\lambda_2 t} + (v_M - v_{min})^2 e^{-\left(\frac{\bar{v}+\delta-v_{min}}{v_M - v_{min}}\right)\lambda_2 t}\Bigg]\end{aligned} \tag{10.27}$$

and

$$\begin{aligned}E = \; & \frac{v_M - v_{min}}{\delta^2}\Bigg\{ - \left[\frac{v_{min}}{\lambda_2} - (v_M - v_{min})\frac{\lambda_1}{\lambda_2^2}\right] ln\Bigg[(\bar{v} - \delta - v_{min})\lambda_2 \\ & + (v_M - v_{min})\lambda_1\Bigg] + \frac{\bar{v} - \delta}{\lambda_2} ln\Bigg[(\bar{v} - \delta - v_{min})\lambda_2 + \; (v_M - v_{min})\lambda_1\Bigg] \\ & + 2\left[\frac{v_{min}}{\lambda_2} - (v_M - v_{min})\frac{\lambda_1}{\lambda_2^2}\right] ln\Bigg[(\bar{v} - v_{min})\lambda_2 \\ & + (v_M - v_{min})\lambda_1\Bigg] - \frac{2\bar{v}}{\lambda_2} ln\Bigg[(\bar{v} - v_{min})\lambda_2 + (v_M - v_{min})\lambda_1\Bigg]\Bigg\} \\ & - \frac{v_M - v_{min}}{\delta}\Bigg\{\left[\frac{v_{min}}{\lambda_2} - (v_M - v_{min})\frac{\lambda_1}{\lambda_2^2}\right] ln\Bigg[(\bar{v} + \delta - v_{min})\lambda_2 \\ & + (v_M - v_{min})\lambda_1\Bigg] - \frac{\bar{v} + \delta}{\lambda_2} ln\Bigg[(\bar{v} + \delta - v_{min})\lambda_2 + (v_M - v_{min})\lambda_1\Bigg]\Bigg\}.\end{aligned} \tag{10.28}$$

Figure 10.5 shows the average number of individuals for each population studied, considering $n(0) = 24,500,000$, $\lambda_1 = 1/58$ and $\lambda_2 = 1/38 - 1/58$ (Jafelice *et al.*, 2004b). The United Nations (UN), estimate a number of $24,500,000$ HIV positive people in Africa; therefore, we assume this value as the one for $n(0)$. Other important information found in UN is that '*The average life expectancy has diminished by approximately 20 years; in Moçambique nowadays it is 38 years*'. Hence, we assume natural mortality rate $\lambda_1 = 1/58$, the inverse of natural life expectancy, and $\lambda_2 = 1/38 - 1/58$. The values $v_{min} = 0.05$ and $v_M = 0.6$

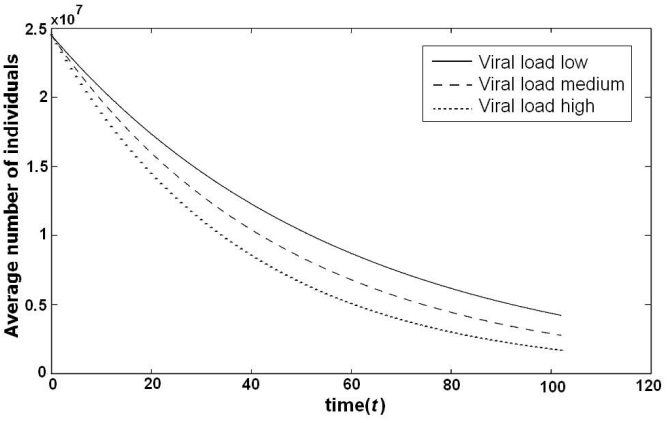

Fig. 10.12. Average number of individuals (Jafelice *et al.*, 2004b)

were obtained from the rule base and from the values of $\phi(v, c)$. With this data, the life expectancy for each group is found as follows:

Case a): Viral load is *low* - $E = 58$years;
Case b): Viral load is *high* - $E = 38$ years; and
Case c): is *medium* - $E \cong 47$years.

These values are in accordance with the ones reported by UN. We used the data from Africa because the model suggested in this section does not consider treatment with anti-retroviral therapy, as it is the case of Africa.

Finally we want to address the importance that rule-based models can have in the fuzzy dynamical systems development. Both here and in chapter 8 they allow to dealing with more complex systems since the rule-based structure can provide the aggregation, in a easier way, of several input variables. In addition, they allow the expert knowledge be applied directly in the dynamical structure. In the next chapter we present another way to consider this expert knowledge in a dynamical context and with a simulation approach.

11 Fuzzy Reed-Frost Model

In the previous chapters we presented several ways to treat dynamical systems from the fuzzy sets theory point of view, considering particular epidemic scenarios. We have discussed some approaches highlighting the advantages and limitations of each technique. However, all structures presented were macro-epidemic-models, that is, the epidemic spreading depends on the global variables and does not consider what happens inside each individual in the population. The only exception is the fuzzy modeling in symptomatic HIV infected population, presented in chapter 10, in which the fuzzy transference rate developed was based on the Nowak's micro model (Nowak, 1999).

Models whose epidemic spreading consider individual aspects in their dynamics are rare and are called micro-epidemic-models. This chapter is devoted to this kind models and two micro-epidemic-models are presented, where the individual's symptomatology are considered into the epidemic spreading.

The simplest stochastic epidemic model is the so-called Reed-Frost model (Abbey, 1952). In this model there is a chance that a susceptible individual become infected when he/she contacts an infected individual. The dynamics of spreading is determined by a Markovian process and it depends on a probabilistic parameter. Since the Reed-Frost model is the simplest epidemic model available, it is interesting to investigate the fuzzification possibilities of its structure, and discuss those possibilities from a fuzzy dynamical point of view. This is the aim of this chapter.

11.1 The Classical Reed-Frost Model

The Reed-Frost model was proposed by L. J. Reed and W. H. Frost in a series of lectures held at Johns Hopkins University (Abbey, 1952). It is a particular case of a chain-binomial model, in which it is assumed that each infected individual infects susceptible individuals independently, and that individuals are under the same contact rate with each other. If we represent by p the probability of a contact between a susceptible and an infected individual resulting in a new case, we have that, at time t, the probability that a susceptible individual does become

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 225–251, 2008.
springerlink.com

infected, C_t, is equal to the probability of at least one infectious contact, that is,

$$C_t = 1 - (1-p)^{I_t}, \quad t > 0, \tag{11.1}$$

where I_t is equal to the number of infected individuals at time t. Time t is assumed to be a discrete variable, and an individual's classification, as either susceptible, infected or resistant, can only change when time changes from t to $t+1$. Other assumptions are that the probability of an infectious contact is fixed along the epidemic course and is the same for all individuals.

The Reed-Frost model (11.1) can be used to describe the spread of any infectious disease affecting closed, uniformly-mixed groups. The group has a constant and small size N, is homogeneous, both from the susceptibility and infectivity points of view (see Bailey, 1975), with individual members spending a significant and constant part of the day in close contact. From the infection point of view, the infectious period is assumed to be short compared to the incubation period which, in turn, is taken as constant.

The Reed-Frost model assumes that individuals are classified according to their disease status: susceptible and infected and, in some cases, also resistant or immune. No error involved in the classification process, such as a truly infected individual being classified as susceptible, is considered in the model. For a great number of infectious diseases, however, such a diagnostic test is neither readily nor easily available: examples are influenza and several other viral and bacterial infections. The corresponding diagnostic process involves uncertainty, and is based upon a set of clinical characteristics, often subjective, which we call *signals*. It is then important to consider, in the epidemic model, the uncertainty involved in the classification process.

The homogeneity assumption is unlikely to hold in real epidemics, especially in large groups (see for example Becker, 1979). In certain cases, the assumption of time-invariant susceptibility/infectivity levels does not hold either. Each individual may have a varying susceptibility level to infections, depending on physical and psychological factors, even within a short-lasting epidemic. Infectivity levels may also vary according to similar factors. Indeed, the capacity of an infected individual to produce an infectious contact may depend upon the set of signals developed. Some signals, such as sneezing and coughing in influenza-type infections, may increase the probability that a contact be infectious, while others, such as fever, may decrease it by making the host less prone to contacts.

There have been several attempts to generalize the Reed-Frost model so as to consider a non-homogeneous group, either from the susceptibility or from the infectivity points of view (Maia, 1952; Scalia-Tomba, 1986; Lefèvre & Picard, 1990; Picard & Lefèvre, 1991). In all these, the homogeneity assumption is relaxed by dividing the main group into subgroups, and considering that there is homogeneous mixing within each subgroup. Subgroups are closed and individuals remain within the same subgroup for the entire duration of the epidemics, which means that an individual's susceptibility and infectivity levels are taken as constant throughout the epidemic course. However, it would be interesting if

the individual's heterogeneities were treated without the separation of the population in subgroups, incorporating it directly into the dynamics of the system.

In the next section we present in detail a stochastic Reed-Frost generalization, developed by Menezes and collaborators (2004), that considers the clinical signals involved in the classification process in the study of the epidemic course. Those clinical signals may include symptoms, results from laboratory and physical examinations. They assumed that, after being infected, no resistance is gained and the individual becomes susceptible again (SIS model). The individual's infectivity is modeled as a function of the signals, therefore allowing for time-dependent, heterogeneous infectivity. In this work susceptibility levels are kept constant. Since this model involves only random variables it is possible to obtain expressions for the epidemic basic reproduction number and its probability function.

The structure of this stochastic Reed-Frost model can be completely generalized for a fuzzy approach, where the signals' information is evaluated in a different way. In section 11.3 we present a fuzzy version for this model and for comparing the models' possibilities we present a real problem. Simulations are performed and the models are compared with real data.

11.2 The Stochastic Reed-Frost Model

In the model proposed by Menezes *et al.* the clinical signals are recorded and are taken into account in the epidemic course via a signal summary, both as part of the classification process and to define the probability of an infectious contact (Menezes *et al.*, 2004). It is assumed that, the higher the signal summary, the higher the probability that a contact be infectious. The probability that an individual has at least one infectious contact, which is the core of the Reed-Frost model, is then computed taking into account the heterogeneous infectivity in the group.

The model can include both signals linked to an increased infectiousness and signals linked to a decreased infectiousness. Both types of signals enter the signal summary, affecting it in opposite directions. Distinct signals can have different weights in the summary, reflecting the impact they are believed to have on both the classification process and on the infectious contact probability.

A probability distribution is assigned to the signal summary, conditioning on the previous probability of at least one infectious contact. This distribution is a mixture of the one given the individual is infected, with the one given the individual is susceptible. In this approach the classification is seen as a probabilistic step conditioned on the signal summary. The probability of an infectious contact is taken as a deterministic, polynomial function of the signal summary.

In this formulation a generalized Reed-Frost model is constructed taking a susceptible individual as reference. It is first assumed that, at time t, each individual i has a true health status represented by $\eta_{i,t}$, which takes value 1 if the individual is infected at t, and 0 if the individual is susceptible. Thus, the number of infected individuals at time t is given by,

$$I_t = \sum_{i=1}^{N} \eta_{i,t}. \tag{11.2}$$

Each individual has one or more clinical signals, which can be summarized by one variable $D_{i,t}$, taking values between 0 and 1. At time t, the probability $P_{il,t}$ that a contact between a susceptible individual i and an infected individual l results in a new case is a function of the signals of the infected individual only, $D_{l,t}$, as a consequence of the homogeneous susceptibility assumption. It is assumed in particular that this function can be written as a polynomial of degree M. That is,

$$P_{il,t} \equiv P_{l,t} = \sum_{j=1}^{M} \varphi_j D_{l,t}^{j}, \tag{11.3}$$

where $0 \leq \varphi_j \leq 1$ and $\sum_j \varphi_j = 1$, that is, $P_{l,t}$ is a convex combination of $D_{l,t}^{j}$, guaranteeing that $P_{l,t} \in [0,1]$ for all l, t. Then the probability that a susceptible individual has, at time t, at least one infectious contact defines the stochastic Reed-Frost model as

$$C_t = 1 - \prod_{l=1}^{N} (1 - P_{l,t})^{\eta_{l,t}}. \tag{11.4}$$

Note that C_t here can be interpreted as the probability that an individual be infected at time $t+1$, as in the classic Reed-frost model, and it is possible to write $C_t = P\eta_{i,t+1} = 1$.

In some cases $\eta_{i,t}$ is unknown, so individuals have to be diagnosed as either infected or susceptible. This consists in a classification procedure which takes into account the clinical signals or, for simplicity, the signals summary $D_{i,t}$, and is defined outside the model, probably by experts. Let $G_{i,t} = 1$ indicate that the individual i is diagnosed as infected at t, and $G_{i,t} = 0$ indicate that the individual is diagnosed as susceptible. So, the number of individuals diagnosed as infected at time t is an estimation of the number of infected individuals at T, and is given by

$$\widehat{I}_t = \sum_{i=1}^{N} G_{i,t}. \tag{11.5}$$

In this way, the probability that a contact between a susceptible individual i and an infected individual l results in a new case is defined by (11.3) and, in this case, (11.4) is estimated as

$$\widehat{C}_t = 1 - \prod_{l=1}^{N} (1 - P_{l,t})^{G_{l,t}}. \tag{11.6}$$

Thus, $\widehat{C}_t$ here is the estimated probability that an individual be infected at time $t+1$.

It is important to highlight that this generalization of Reed-Frost model, taking into account the individual's heterogeneities, has a particular probability

structure, which allows some analytical calculus be performed. As it will be mentioned later, these calculations supply interesting results from epidemiological point of view. Nevertheless, in the fuzzy Reed-Frost model the analytical calculus is not easy and represents a really difficult task.

Another interesting point is that this stochastic approach of the Reed-Frost model can be used in two related contexts. One is that of a retrospective study, in which patient's health status are observable and modeled as random variables. The objective of such a study is typically to estimate the parameters of the signals'distributions, and it involves relations (11.2-11.4). The other is that, once these parameters estimates are available, the approach can be used in a prospective way, where the true $\eta_{i,t}$ are not known, due either to time or cost constraints. Such a study could include, as an objective evaluating, the function $P(\cdot)$, and it involves expressions (11.3), (11.5) and (11.6). This consists in recording the patients'clinical signals over a certain period of time, and then estimating their true health status using the model and the classification processes.

11.2.1 The Probabilistic Structure

Each infectious disease produces clinical signals with varying degrees of severity, which depend upon both the pathogen and the individual's variability. Susceptible individuals may also present some of these signals, for reasons other than the infection considered, but it is expected that they do so with a lower severity than if they were infected.

For the *retrospective study*, the true health status $\eta_{i,t}$ is a binary variable. For $t > 1$, and given all the epidemic information up to $t-1$, $P\eta_{i,t} = 1$ is equal to the probability of having at least one infectious contact at $t-1$, C_{t-1}, which is the same for all individuals due to the homogeneous mixing assumption. For $t = 1$, it is defined $P\eta_{i,t} = 1 \equiv \theta$ as the *a priori* probability that any individual is infected at the epidemic onset, which must be evaluated via populational measurements, for example, the estimated prevalence of the pathogen or the disease in this population.

For a given pathogen, the clinical summary for any infected individual is represented by X_I, and any susceptible individual is represented as X_S. Given an individual's health status η, X_I and X_S can be seen as random variables, intrinsically linked to the pathogen, and their distributions remaining unaffected by the epidemic progress. In this work it was assumed that they take values within the interval $[0, 1]$ with a distribution within the Beta family, as follows:

$$\begin{aligned} X_I &\sim \text{Beta}\,(\alpha_I, \beta_I)\,, \\ X_S &\sim \text{Beta}\,(\alpha_S, \beta_S)\,. \end{aligned} \tag{11.7}$$

Note that, with the choice of the distribution Beta all moments $E(X_I^k)$, $E(X_S^k)$ of X_I and X_S are finite (for all $k = 1, 2, ...$). In fact, for any random variable X with Beta distribution with parameters (α, β), its probability density function is

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 \leq x \leq 1, \alpha > 0, \quad \beta > 0,$$

where $\Gamma(x)$ is the Gamma function, defined as

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du, \quad x > 0.$$

Note that $\Gamma(x) = x\Gamma(x-1)$, for all $x > 1$. The expected value of X^k, for any $k > 0$, integer, is

$$E\left(X^k\right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+k)} \times \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} = \frac{\prod_{m=0}^{k-1}(\alpha+m)}{\prod_{m=0}^{k-1}(\alpha+\beta+m)}. \tag{11.8}$$

In particular,

$$E\left(X\right) = \frac{\alpha}{\alpha+\beta}, \tag{11.9}$$

$$E\left(X^2\right) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}, \tag{11.10}$$

and

$$var\left(X\right) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{11.11}$$

So, it is possible to define $\mu_I = E(X_I)$, $\mu_S = E(X_S)$, $\Delta = X_I - X_S$ and $\delta \equiv E(\Delta) = \mu_I - \mu_S$, which means the difference between the mean clinical summary of an infected and a susceptible individual.

For this approach, the observed clinical summary for individual i at time t, $D_{i,t}$, is equal to X_I if the individual is infected and it is equal to X_S if the individual is susceptible. So, it is possible to write:

$$D_{i,t} = X_I\eta_{i,t} + X_S\left(1-\eta_{i,t}\right) = X_S + \Delta\eta_{i,t}, \tag{11.12}$$

for all $i = 1, \ldots, N$ and all $t \geq 1$. For $t > 1$, the conditional mean of $D_{i,t}$, given by C_{t-1}, is given by

$$E\left(D_{i,t}|C_{t-1}\right) = \mu_S + \delta C_{t-1}. \tag{11.13}$$

Similarly, for $t = 1$ we have,

$$E\left(D_{i,t}|\theta\right) = \mu_S + \delta\theta. \tag{11.14}$$

In other words, the expected value of the signal summary at time t is a convex combination of the mean signals for infected and susceptible individuals, based upon the probability of being truly infected. Given C_{t-1}, the number of infected, I_t, defined by (11.2), is a sum of conditionally independent binomial variables, all with the same probability of success: C_{t-1} for $t > 1$, and θ for $t = 1$. Thus,

I_t is in this case a binomial random variable with probability C_{t-1} and sample size N.

Since the probability of an infectious contact $P_{l,t}$ is a deterministic function of $D_{l,t}$, it is a constant when $D_{l,t}$ is given. Unconditionally, a probability distribution is effectively assigned to each probability of an infectious contact, $P_{l,t}$, for each individual l at each generation t, $l = 1, \ldots, N$, $t > 0$. The probability of at least one infectious contact, C_t, does not have a well-known probabilistic distribution, but its conditional expected value, given the clinical summaries $D_{i,t}$, can be computed.

In the above, the calculations are performed in terms of μ_S and δ instead of μ_S and μ_I. This separates the contribution of the signal's distribution, which can be treated as unaffected by the disease spread, from the probability that each individual is infected. Therefore, it incorporates treatment effects naturally as a reduction of the difference between mean signals, $\delta = \mu_I - \mu_S$.

In a *prospective study* context, most of the probability structure introduced above applies, but the diagnostic uncertainty must be included. When the $\eta_{i,t}$ are unknown, patients must be classified as either infected or susceptible. The classification is represented by $G_{i,t}$ and defined as a simple process: given the clinical summaries $D_{i,t}$, each individual is independently classified as infected with probability:

$$P\{G_{i,t} = 1 | D_{i,t}\} = D_{i,t}. \tag{11.15}$$

The probabilistic structure for $D_{i,t}$ and $G_{i,t}$ can be interpreted as defining a conditional Binomial distribution for $G_{i,t}$, given the probability of success $D_{i,t}$ which itself has a conditional Beta distribution, given the health status $\eta_{i,t}$.As a consequence, we have that $E(G_{i,t}|D_{i,t}) = D_{i,t}$.

The conditional probability that an individual is classified as infected in generation t, given the probability of at least one infectious contact in the previous generation C_{t-1}, is given by:

$$\begin{aligned} P\{G_{i,t} = 1 | C_{t-1}\} &= E\left[P\{G_{i,t} = 1 | D_{i,t}\} | C_{t-1}\right] \\ &= E\left[D_{i,t} | C_{t-1}\right] = C_{t-1}, \quad t > 1. \end{aligned} \tag{11.16}$$

where either (11.13) and (11.14) can be used to re-express this as a function of the signals.

The number of patients diagnosed as infected, $\widehat{I}_t$, defined by (11.5), is again a sum of conditionally independent binomial variables, given $D_{i,t}$, but since each one of these has a different probability of success, the distribution of I_t is not the usual binomial.

The basic reproduction number

As defined in chapter 3 the basic reproduction number, R_0, is the number of secondary infections resulting from a single case in an entirely susceptible group, during its infectious period. In the stochastic Reed-Frost presented, $R_0 = I_2$, given that $I_1 = 1$. Note that this definition is coherent with the Diekmann *et al.* (1990) definition of the next generation operator.

In a *retrospective study*, the expected value of R_0 is given by:

$$E(R_0) = E(I_2|I_1 = 1) = \sum_{j=1}^{N} E(\eta_{j2}|I_1 = 1). \tag{11.17}$$

By conditioning on C_1, we get $E\left(\eta_{j2}|I_1 = 1\right) = E\left[E(\eta_{j2}|C_1)|I_1 = 1\right] = E(C_1|I_1 = 1)$ for all $j = 1, \ldots, N$. This means that,

$$E\left(R_0\right) = NE\left(C_1|I_1 = 1\right). \tag{11.18}$$

The same result is obtained by considering that, give C_1, I_2 has Binomial distribution mean NC_1. Using the definition of I_1, we can re-write $E(C_1|I_1 = 1)$ as

$$\sum_{j=1}^{N} E\left\{C_1|\eta_{j,1} = 1, I_1 = 1\right\} P\left\{\eta_{j1} = 1|I_1 = 1\right\}. \tag{11.19}$$

All individuals are equally likely to be the first one infected, so

$$P\left\{\eta_{j1} = 1|I_1 = 1\right\} = 1/N.$$

Moreover, given that $\eta_{j,1} = 1$ and all other $\eta_{k,1}$ are equal to zero, we have from (11.4) that $C_1 = P(D_{j,1}) = P(X_I)$, since this individual is infected. Therefore, we can re-write (11.19) as

$$\frac{1}{N} \sum_{j=1}^{N} E\left\{P(X_I)\right\} = E\left\{P(X_I)\right\}, \tag{11.20}$$

and thus

$$E(R_0) = NE\left\{P(X_I)\right\} = N \sum_{k=1}^{M} \varphi_k E(X_I^k). \tag{11.21}$$

In the simple case where $P(D) \equiv D$, we get

$$E(R_0) = N\mu_I = N(\mu_S + \delta). \tag{11.22}$$

In a *prospective study* a diagnostic is estimated with a certain margin of error, and similar arguments to the retrospective study can be used to derive an expression for the expected value of R_0. Indeed, we have, in this case,

$$E(R_0) = E\left(\widehat{I}_2|I_1 = 1\right) = \sum_{j=1}^{N} E\left(G_{j,2}|I_1 = 1\right), \tag{11.23}$$

where it is clear that we must be sure about the first recorded case being indeed an infection. By conditioning on $D_{j,2}$, we get $E\left(G_{j,2}|I_1 = 1\right) = E\left(D_{j,2}|I_1 = 1\right)$, for all $j = 1, \ldots, N$. Now, conditioning on C_1, we get $E\left(D_{j,2}|I_1 = 1\right) = \mu_S + \delta E\left(C_1|I_1 = 1\right)$, for all j. This means that,

$$E(R_0) = N\left\{\mu_S + \delta E\left(C_1 | I_1 = 1\right)\right\}. \tag{11.24}$$

In the retrospective study we saw that $E\left(C_1 | I_1 = 1\right) = E\left\{P(X_I)\right\}$, and the same still holds in the prospective study, as conditioning on having one infected individual at $t = 1$ there is no diagnostic uncertainty at $t = 1$. Thus, we find

$$E(R_0) = N\left\{\mu_s + \delta E P(X_I)\right\}. \tag{11.25}$$

In particular, when $P(D) \equiv D$, we get:

$$E(R_0) = N\left\{\mu_s + \delta(\mu_s + \delta)\right\}. \tag{11.26}$$

The main goal with the evaluation of R_0 and related functions is to define criteria yielding clues as to the long-term disease establishment. While the expected value gives some clues, it is also important to evaluate how likely the value of R_0 is to spread around it. The traditional statistical approach is to construct a confidence interval for R_0, and check if the value contains those leading to long-term disease establishment, in this case any value greater than, or equal to, 1. In the present problem, however, this is of limited use: R_0 being a random variable assuming only nonnegative integer values $0, 1, 2, \cdots$, all of these values have positive probability mass and, thus, any confidence intervals is likely to include the value $R_0 = 1$ at least. A more useful measurement seems to be the probability that $R_0 \geq 1$. While taking uncertainty into account, this is perhaps more useful: if, under certain conditions, it is known that the probability that $R_0 \geq 1$ is about 20%, then only 1 in 5 independent initial cases are likely to propagate the disease. Considering this discussion, it is interesting to evaluate the R_0 uncertainty in both retrospective and prospective studies.

In order to evaluate the uncertainty around the R_0, in the *retrospective study*, it is necessary to evaluate its probability distribution. We have that:

$$\begin{aligned} P\left\{R_0 = z\right\} &= P\{I_2 = z | I_1 = 1\} \\ &= E\left(PI_2 = z | C_1 | I_1 = 1\right) \\ &= E\left(n_z C_1^z (1 - C_1)^{N-z} | I_1 = 1\right), \end{aligned} \tag{11.27}$$

where it is used the fact that, given C_1, I_2 has a Binomial distribution, with n_z representing the binomial coefficient equal to $N!/[z!(N-z)!]$. If we then condition on the true health status $\eta_{j,1}$, we can write:

$$\begin{aligned} E\left(C_1^z (1 - C1)^{N-z} | I_1 = 1\right) &= \sum_{j=1}^{N} E\left\{C_1^z (1 - C1)^{N-z} | \eta - j, 1, I_1 = 1\right\} \frac{1}{N} \\ &= E\left\{P(X_I)^z (1 - P(X_I)^{N-z})\right\}. \end{aligned} \tag{11.28}$$

In particular for $z = 0$, we get:

$$P\left\{R_0 = 0\right\} = E\left\{(1 - P(X_I))^N\right\}. \tag{11.29}$$

In the *prospective study*, using the conditional probability properties, it is possible to write:

$$P\{R_0 = z\} = P\left\{\widehat{I}_2 = z | I_1 = 1\right\} = P\left\{\sum_{j=1}^{N} G_{j,2} = z | I_1 = 1\right\}$$
$$= E\left[P\left\{\sum_{j=1}^{N} G_{j,2} = z | \{D_{i,2}\}\right\} I_1 = 1\right], \tag{11.30}$$

where the fact that, given $D_{i,2}$, the random variables $G_{i,2}$ are conditionally independent of $I_1 = 1$ is used. Since each $G_{i,2}$ is a binary random variable, $\sum_{j=1}^{N} G_{j,2} = z$ occurs whenever z of the $G_{i,2}$ are exactly equal to 1. Define J_z as exactly a set of z indices, ranging from 1 to N. That is, J_z is exactly a subset of z elements of the discrete set $1, 2, \cdots, N$. Note that there exist $n_z = N!/[z!(N-z)!]$ such subsets. Let these be represented by $J_{z,1}, J_{z,2}, \cdots, Jz, n_z$. Given $D_{i,2}$, the conditional probability that only the variables within the subset $G_{i,2} \in J_{z,l}$ are equal to 1, is equal to

$$\left[\prod_{j \in J_{z,i}} P\{G_{j,2} = 1 | D_{j,2}\}\right] \quad \left[\prod_{j \notin J_{z,i}} P\{G_{j,2} = 0 | D_{j,2}\}\right]$$
$$= \left[\prod_{j \in J_{z,i}} D_{j,2}\right] \left[\prod_{j \notin J_{z,i}} \{1 - D_{j,2}\}\right]. \tag{11.31}$$

When taking the conditional expectation given C_1, we use the fact that the $\{D_{i,2}\}$ are conditionally independent and we get that (11.31) is

$$\left[\prod_{j \in J_{z,i}} E(D_{j,2} | C_1)\right] \left[\prod_{j \notin J_{z,i}} E(1 - D_{j,2} | C_1)\right],$$

which means that, from (11.27), it becomes,

$$P\{R_0 = z\} = E\left[n_z (\mu_S + \delta C_1)^z (1 - \mu_S - \delta C_1)^{N-z} | I_1 = 1\right]. \tag{11.32}$$

Note that we have shown that, given C_1, R_0 has a conditional Binomial distribution with mean $N(\mu_S + \delta C_1)$. For $z = 0$ the right-hand side of (11.32) becomes $E\left[(1 - \mu_S - \delta C_1)^N | I_1 = 1\right]$ and, given that only one individual is infected, $C_1 = P(X_I)$ as before. Thus, we can write

$$P\{R_0 = 0\} = E\left\{[1 - \mu_S - \delta P(X_I)]^N | I_1 = 1\right\}$$
$$= \sum_{k=0}^{N} \frac{N!}{k!(N-k)!} (-1)^k E\left\{[\mu_S + \delta P(X_I)]^k | I_1 = 1\right\}. \tag{11.33}$$

Now, we use that

$$(\mu_S + \delta P(X_1)) = (\mu_S + \delta)\left\{1 - \frac{\delta\left[1 - P(X_I)\right]}{\mu_S + \delta}\right\}, \tag{11.34}$$

to re-express (11.33) as

$$\sum_{k=0}^{N}\sum_{l=0}^{k}\frac{N!}{k!(N-k)!}\frac{k!}{l!(k-l)!}(-1)^{k+l}\delta^l\left(\mu_S+\delta\right)^{k-l}E\left\{\left[1-P(X_I)\right]^l|I_1=1\right\}. \tag{11.35}$$

Thus, we get

$$P\left\{R_0=0\right\}=\sum_{k=0}^{N}\sum_{l=0}^{k}\sum_{m=0}^{l}\frac{N!(-1)^{k+l+m}\delta^l\left(\mu_S+\delta\right)^{k-l}}{(N-k)!\,(k-l)!m!\,(l-m)!}E\left\{\left[P(X_I)\right]^m\right\}. \tag{11.36}$$

In order to evaluate the uncertainty, it is interesting to calculate the probability $P\{R_0 \geq 1\}$, which is easily evaluated from the previous formulation since $P\{R_0 \geq 1\} = 1 - P\{R_0 = 0\}$. We assumed, for simplicity, that $P(D) \equiv D$, and we evaluated this probability in both retrospective and prospective studies, along with $E(R_0)$, for a range of values of μ_I and δ, assuming also $N = 5$. There is no need to evaluate the functions for other values of N as, from the theoretical point of view, N is just a scale factor. In practice, for larger values of N the function defining the probability of an infective contact given the signal, $P(D)$, is likely to be other than the identity, reflecting lower contact rates between the individuals.

For the *retrospective study*, $E(R_0)$ and $P\{R_0 = 0\}$ are computed using (11.21) and (11.29), respectively, whilst for the *prospective study* we used (11.25) and (11.33). In figure 11.1(a) we can see $E(R_0)$ as a function of δ in both kinds of study, assuming that the variances for both X_S and X_I are fixed and equal to 0.06, $\mu_S = 0.1$ and μ_I varies from 0.13 to 0.83 by 0.1 of step size. First, we note that the expected value in the retrospective study is an upper bound for the prospective study. This suggests that the uncertainty involved in the diagnostic process implies an underestimation of $E(R_0)$.

We can also note from figure 11.1(a) that the two quantities are similar for values of $\delta = \mu_I - \mu_S$ in the extremes of the range considered, differing more for intermediate values. This was expected because, when $\mu_I \to 1$, $\mu_S + \delta\mu_I$ tends to $\mu_S + \delta = \mu_I$ and, thus $E(R_0)$ in the prospective study (11.25) tends to $N\mu_I$, which is $E(R_0)$ in the retrospective study, (11.21). On the other hand, when $\mu_I \to \mu_S$, we also have $\delta \to 0$ and $\mu_s + \delta\mu_I \to \mu_s$, and thus in both studies we have $E(R_0) \to N\mu_S$.

However, note that in spite of the two expectancies converging to the same value as δ decreases, there is a *discordance* region where one of them is above 1, while the other is below 1. Clearly, it is an important result from the epidemiological point of view.

In figure 11.1(b) we have the computed probabilities of R_0 being greater than zero, as a function of δ, for the same parameter values considered. Differently

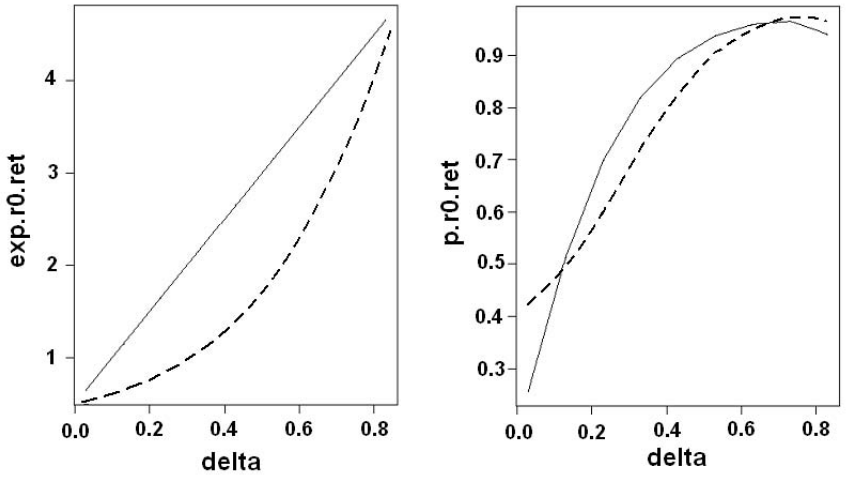

Fig. 11.1. Functions of retrospective study (solid line) and prospective study (dashed line) for several values of δ. (a): Expected R_0 and, (b): probability that $R_0 \geq 1$ (Menezes *et al.*, 2004).

from the expected values, the probabilities for one study are not consistently greater than those for the other study. Perhaps the most important aspect highlighted by this figure is the fact that $P\{R_0 \geq 1\}$ may solve apparent discordances between the expected values from different studies.

Comparing the stochastic model with the classic Reed-Frost

The classic Reed-Frost model can be seen as a particular case of both Reed-Frost models presented before. Let us consider the more general model for a prospective study, and the result follows for the retrospective study model. Suppose that $\mu_S = 0$ and $\mu_I = \delta = 1$, so that the variances of X_I, X_S are both equal to zero (see equations 11.9-11.11). This implies that $D_{i,t} = X_I = \mu_I$ for all i, t. Then, from (11.15) $P\{G_{i,t} = 1 | D_{i,t}\} = 1$ for an infected individual, and 0 otherwise, with no uncertainty, implying that $\sum_l G_{l,t} = I_t$ in this case. Suppose also that $M = 1$ in (11.3), meaning that $P_{l,t} = \varphi \equiv P$ for all infected individuals, whilst $P_{l,t} = 0$ for all susceptible individuals, for all t. Then equation (11.4) becomes:

$$C_t = 1 - (1 - P)^{\sum_l G_{l,t}} = 1 - (1 - P)_t^I \,, \tag{11.37}$$

which is the equation for the classical Reed-Frost model.

Conditions determining long-term establishment of the disease

The probabilistic structure of the model proposed by Menezes *et al.* allows the evaluation of the conditions that determine the long-term establishment of the

disease. Conditions under which the disease establishes itself in the group can be obtained in terms of quantities of interest, by determining when $E(R_0)$, as a function of these quantities, is greater than, or equal to, 1. Assuming for simplicity $P(D) \equiv D$, in a *retrospective study* it is possible to use the expression (11.22) to say that the disease establishes itself in the population whenever

$$\mu_S + \delta \geq \frac{1}{N}. \tag{11.38}$$

In a *prospective study*, it is possible to use (11.26) to say that the disease establishes itself in the population whenever

$$\mu_S (\delta + 1) + \delta^2 \geq \frac{1}{N}. \tag{11.39}$$

In this context, it is interesting to consider $P\{R_0 \geq 1\}$ as a stochastic way of evaluating how likely the disease is of establishing itself in the long term. For that, we simply compare the value obtained for $P\{R_0 \geq 1\}$ to a pre-specified threshold; when the probability is below the threshold, the disease may be said to be unlikely to establish itself. This threshold may vary according to disease and context.

Since the classic Reed-Frost model can be seen as a particular case of the stochastic model, it is also possible to evaluate the conditions that determine the long-term establishment of the disease. As discussed above, we fall into the classic Reed-Frost model when $\mu_S = 0$ and $\mu_I = \delta = 1$. In this case, from (11.39) the epidemic establishes itself whenever

$$E(C_1 | I_1 = 1) \geq \frac{1}{N} \quad \text{or} \quad P \geq \frac{1}{N}, \tag{11.40}$$

which implies that, for large N, any epidemic with non-negligible probability of an infectious contact, P, establishes itself in the population. It is important to point out, however, that this remark has a purely theoretical interest: in practice, when N is large the homogeneous mixing assumption rarely holds.

Applications to intervention strategies design

An important practical application of epidemic models is that of intervention strategies design. An intervention may involve simply a change in risk behavior, thus changing $P(D)$, or it may involve treatment, which might affect both $P(D)$ and infected individuals' signals summary distribution. The signals summary distribution for susceptible individuals, including the mean μ_S, is assumed not to be affected, as it represents the population distribution of the signals summary under study, due to causes other than disease. For simplicity it is assumed that $P(D) = \varphi D$.

First, let us consider the impact of a risk-behavior reducing intervention, which can be represented by a change from $P(D)$ to $P^*(D) = \varphi^* D$, where $\varphi^* < \varphi$. In a *retrospective study*, the post-intervention expected R_0 is given by $E(R_0^*) = NE[P^*(X_I)] = N\varphi^* \mu_I$, from (11.21). A desirable intervention yields $E(R_0^*) < 1$, which is guaranteed to hold if $\varphi^* < (N\mu_I)^{-1}$.

In a similar way, such intervention can be designed to guarantee that

$$1 - P\left\{R_0^* = 0\right\} < p_0, \tag{11.41}$$

where p_0 is a pre-specified threshold. Indeed, using (11.29) we get that (11.41) is satisfied by designing the intervention so that

$$\sum_{k=1}^{N} \frac{N!}{k!(N-k)!}(-1)^k(\varphi^*)^k E\left(X_I^k\right) > 1 - p_0 \tag{11.42}$$

holds.

Similar conditions can be obtained to evaluate *a priori* the intervention impact in a *prospective study*. Indeed, using (11.25) we can conclude that the intervention will generate an expected basic reproduction number smaller than 1 whenever:

$$\left[1 - \varphi^* E\left(X_I^2\right)\right] \sum_{k=0}^{N-1} \frac{N!}{k!(N-k)!}(-1)^k(\varphi^*)^k E\left(X_S^{2k}\right) > \frac{1}{\delta}\left(\mu_I - \frac{1}{N}\right). \tag{11.43}$$

No analytical expressions are available for the roots of this polynomial on φ^* for general N and, thus, in practice this condition can be used mainly to check whether or not a specific value of φ^* satisfies it.

The post-intervention probability of long-term disease establishment can be evaluated by replacing $P(X_I)$ by $P^*(X_I) = \varphi^* X_I$ in (11.36), in the same way as before.

Again, the impact of treatment affecting only the signals distribution can be evaluated prior to introduction via replacing δ by δ^* in expressions for $E\left(R_0\right)$ and $P\left\{R_0 \geq 0\right\}$. For example, in a retrospective study such a treatment generates on average less than 1 new infected cases for each first case whenever

$$\delta^* < \frac{1}{N\varphi} - \mu_S. \tag{11.44}$$

A treatment can also affect both δ and $P(D)$. Expressing the treatment impact in the same way as before, the condition guaranteeing that on average less than 1 new infected cases are generated for each first case is

$$\varphi^*\left(\mu_S + \delta^*\right) \leq \frac{1}{N}. \tag{11.45}$$

The treatment impact on the probability of long-term disease establishment, as well as in the prospective study case, can be evaluated in the same way.

The impact of a treatment affecting both signals summary distribution and the probability of an infectious contact can be evaluated, by combining the ideas above.

11.2.2 Theoretical Remarks

The stochastic Reed-Frost model presented handles varying infectivity levels by assuming that infectivity is determined by observable and quantifiable clinical

signals which, in turn, are assumed as having a specific probability distribution. Therefore, this model takes into account heterogeneous infectivity by assigning a probability distribution to each individual's infectious contact probability, considering also time-varying individual infectiousness. In addition, the proposed approach can be used in intervention design, which consists in one of the most interesting features of this model.

Disease studies to which the proposed model can be applied include all fast-propagating infectious diseases, in the sense that the disease propagates at a faster rate than its diagnostic and control can be performed. Examples of such diseases are influenza and meningitis. Applications also include several kinds of confinement, starting from the reinforced confinement of hospital wards and including classrooms at winter time, when weather imposes confinement.

In chapters 8, 9 and 10 we presented some fuzzy dynamic structures and discuss how it would be complex to develop them. In the next section we present a fuzzy version for the Reed-Frost model that consists of a simple way of elaborating fuzzy dynamic systems (Ortega *et al.*, 2008a). As in the stochastic approach, this fuzzy model is an epidemic model based on the microscopic information, the individual's clinical signals, and consider a fuzzy relation to evaluate the individual's infectiousness, performing a fuzzy decision process where the infectiousness degree is applied directly in the epidemic dynamic.

11.3 Fuzzy Reed-Frost Model

As in the previous approach, modeling the fuzzy Reed-Frost dynamics based on the signals scenario is based on the idea that there is an association between the intensity of the signals present in an infected individual and the possibility of an infectious contact with this individual. As it can be observed, the fuzzy approach has not a probabilistic connotation as intense as in the stochastic Reed-Frost. In this formulation it is assumed that each individual i has a health status, susceptible or infected represented by $G_{i,t}$. The binary variable $G_{i,t}$ takes value 1 if the individual i is infected at t, and 0 if the individual is susceptible. In this way, the number of individuals infected at t, in a group with size N, is given by the equation (Ortega *et al.*, 2008a):

$$I_t = \sum_{i=1}^{N} G_{i,t}. \tag{11.46}$$

In general, the diagnostic process is based upon the set of signals present in the individual under analysis. This signals set can be summarized by one variable $ID_{i,t}$, taking normalized values into the interval $[0, 1]$, since it must be a fuzzy measure. In addition, these clinical signals usually vary their severity depending on both the disease and on the individual variability. Furthermore, for reasons other than the infection considered, these signals can also be present in susceptible individuals. In this case, it is expected that its expression should be less intense than in the presence of the infection.

Since the clinical signals expression is different for infected and susceptible individuals, we assumed two probability distributions, depending on the parameters of either the susceptible or the infected populations. We represent by X_I the signal for any infected individual, and by X_S the signal for any susceptible individual. So, given an individual's health status, $G_{i,t}$, X_I and X_S are random variables intrinsically linked to the pathogen; therefore, their distributions remain unaffected by the epidemic course. Since they take values within the interval $[0, 1]$ we assume the same probability distributions used in the stochastic approach (see equations 11.7).

Thereby, at time t, the possibility $P_{jl,t}$ that a contact between a susceptible individual j and an infected individual l results in a new case is a function of the signals of the infected individual only, $ID_{l,t}$, as a consequence of the susceptibility homogeneity assumption. We assume in particular that this function is

$$P_{il,t} \equiv P_{l,t} = \varphi ID_{l,t}^{\omega}, \tag{11.47}$$

where φ and ω are parameters of the model and should be chosen in a way to guarantee that $P_{l,t} \in [0, 1]$ for all l, t. Then the epidemic dynamics in this fuzzy Reed-Frost model is also given by:

$$C_t = 1 - \prod_{l=1}^{N} (1 - P_{l,t})^{G_{l,t}} \tag{11.48}$$

and C_t here can be interpreted as the possibility that an individual be infected at time $t + 1$, similarly to the classic Reed-Frost model, and will be used to generate the health status of the individuals in time $t + 1$.

11.3.1 The Possibilistic Structure

The main difference between the proposal by Menezes *et al.* and the fuzzy approach consists in the structure of the summary of signals, which are performed by a random variable in the former and by a possibility measure in the latter. In the fuzzy case, the individual's infectiousness is calculated through a membership degree based on the $\max - \min$ composition (see section 2.5 in chapter 2).

Consider a set of signals S and the matrix representation of a fuzzy relation. Thus, $S_l = [s]_{1\times k}$ is the array of k signals of the individual l, $I_l = [i]_{k\times q}$ is the matrix that associates each signal to the infective statement and $DI_l = [di]_{1\times q}$ is the membership degree of the individual l in the fuzzy set $Infected$, interpreted here as the degree of infectiousness, found by the fuzzy composition given by:

$$DI = S \circ I, \tag{11.49}$$

whose fuzzy composition $\circ$ is the $max - min$ composition defined by:

$$DI(di) = \max_{s\in S} \left[\min(S(s), I(s, i))\right]. \tag{11.50}$$

As an example, consider the set of signals $S = [fever, cough]$, i.e., s_1 is fever and s_2 is cough, and an individual who presents fever degree $s_1 = 0.7$ and

cough degree $s_2 = 0.4$. The matrix I that relates signals and infectiousness is $I = [i_{fever}, i_{cough}]$, where i_{fever} is the relationship degree between the symptom *fever* and *infectiousness status* and i_{cough} is the relationship degree between the symptom *cough* and *infectiousness status*. So, an individual that has degree of fever, s_{fever}, and degree of cough, s_{cough}, belongs to the infectiousness fuzzy set with degree given by:

$$DI = \max\{\min[s_{fever}, i_{fever}]; \min[s_{cough}, i_{cough}]\}. \tag{11.51}$$

We assume that each individual i has k signals, whose levels represented by membership degree in each fuzzy subset of clinical signal (like fever, cough) s_{i1}, s_{i2},...,s_{ik}. So, these levels are numbers between 0 and 1, with $s_{i1} = 0$, indicating that the clinical signal 1 is absent in patient i, and $s_{i1} = 1$, indicating that patient i presents the clinical signal 1 with maximum level (or severity). The infectiousness degree is computed for all individuals and the heterogeneity is considered in the epidemic dynamics through the signal influence on the possibility p (the possibility of an infective contact between a susceptible and an infected individual) and, consequently, C_t (the risk of a susceptible individual becoming infected). The new individual set of signals at time $t+1$ is found from C_t, by equation (11.48), and the epidemic spreading is built through the follow up of the number of infected individuals generated.

It is important to highlight that the use of the $max - min$ composition of fuzzy relations consists in an arbitrary choice and, as it is common in fuzzy models, other possibilities could be explored since these compositions are generated through fuzzy operators for disjunction and conjunction manipulations (see chapter 2). In addition, the relational matrix that joins the signals and infectiousness can be elaborated through experts opinion, which allows the introduction in the model of informations that otherwise are not available.

Although the fuzzy approach does not have an explicit probability structure, all calculations performed in section 11.2 for the stochastic proposal is possible to be developed here. This is the reason for which the calculations of this stochastic model were presented. However, the calculations in the fuzzy approach demand greater care because they are more complex. The calculus used in the stochastic model are based on the mathematical manipulations over conditional probabilities and, therefore, consider random variables. However, in the fuzzy model the measure generated through the $max - min$ composition does not produce a pure random variable, as in the probabilistic context, but a possibility measure, where the $\sigma - additive$ property does not always hold (see section 4.1 in chapter 4). Nevertheless, these calculations would allow an analysis of fuzzy R_0 and the possibility that R_0 to be greater than 1. This analysis could provide different results than that obtained through classic structures (stochastic), which has already been shown in several works (see chapters 9 and 12).

11.3.2 Theoretical Remarks

Both fuzzy and stochastic proposals allow several variants. Consider, for instance, the possibility/probability of an infectious contact, which is assumed to

be a function of the signals. This function can have any polynomial form, and as such can potentially include any desired function. For instance, by assigning Beta distributions to the individual signals, not only a flexible distribution family, but also one for which all moments are available, with no limitation on the polynomial degree. Besides, it can be generalized to take the possibility/probability of an infectious contact as probabilistic, rather than deterministic, as a function of the signals. Other variants of these models can be obtained by considering more sophisticated classification procedures, which effectively suggests separating the clinical signals effect on different aspects of the epidemic.

The fuzzy model, as proposed here, is more in accordance with the prospective approach of the stochastic Reed-Frost model. However, it is entirely possible to apply this fuzzy structure for retrospective studies. Furthermore, if there are data available they can be used to improve the fuzzy relational matrix provided by the experts.

Some differences can be pointed between fuzzy and stochastic structures. In the former all signals information related to the possibility function can be performed through fuzzy relational matrix, while in the latter it should be done through the probabilistic function. Clearly, from the interdisciplinary point of view, it is easer to understand the fuzzy relational approach than the mathematical formalism of polynomial functions. In the same way, the heterogeneity of susceptible individuals can be more easily considered in the fuzzy structure. This can be made by simply considering a fuzzy relational matrix that cross informations about the immunological characteristics (as informations about the child's history, family and personal antecedents, breastfeeding, re-infections etc.) and the degree of susceptibility for the infection. In this sense, the fuzzy relational matrix can supply a fuzzy measure of the individual's protection for certain infection, taking into account the aspects of the identification uncertainties, commonly present in a real epidemic process. In addition, both fuzzy measures for the susceptibility and infectiousness individual's degree, can be elaborated based on the experts opinion.

In order to study the behavior of fuzzy and stochastic models from a theoretical and applied points of view, we present, in the next section, simulations of both models and compare them considering a real epidemic data of viral infections.

11.4 Simulations

The simulations carried out consider an infection scenario and its main objective is to analyze, from the theoretical point of view, the behavior of both models. Besides, the performance of each model was evaluated in a quantitatively and qualitatively way by comparing them with real data (Ortega *et al.*, 2008a).

During the entire 2003 year, all children of a daycare, corresponding to roughly 120 children, with age varying from 1 month to 6 years old, were followed up in São José do Rio Preto, São Paulo, Brazil. The objective of this work, among other things, was to study the circulation of viruses for respiratory infections. All

daycare's children with cold symptoms had nasopharyngeal aspirates collected and analyzed with multiplex technique. Therefore, it was possible to determine the true health status of each child. Also, the epidemiological data were collected for all children in the study, independently of the symptomatic status. All children stayed at the daycare during the whole day, what can be considered a quasi-closed group. Although the children are usually distributed in small groups, there are periods along the day that they interact with each other, as in the meals time and in the playful moments. In the same way, there is also interaction among the teachers during the workday. These characteristics, added to the fact that the respiratory infections can be configured as infections of long reach, allow us to consider that the data and the study conditions are in agreement with the model's assumptions.

In order to find the fuzzy relations between signals and infectiousness, four experts in childhood diseases supplied the relational matrices considering the more important clinical signals for infections by viruses. The matrix with the fuzzy relations between signals and infectiousness degree was found by the median of that four experts values. The signs considered for the infections and their respective fuzzy relations values were: fever (0.85), cough (0.85), coryza (0.85); sneezing (0.70) and wheezing (0.60). In this simulations we assumed homogeneous susceptibility and an infected individual was considered immunized to new infections during 3 weeks, which was the minimum period for re-infection observed. So, in the model re-infection is possible, once the protection period is observed.

The model has basically three parameters, which are presented in the equations of the dynamics structure: φ, the polynomial's coefficient; ω, the polynomial's power; and θ, the prior probability of infected status. In addition, the size of the population N was maintained constant since small variations in its value do not affect the result of the model. We assumed $N = 120$, which is around the monthly average of the number of children in daycare.

In order to generate the signals of susceptible and infected individuals we elaborated Beta distributions considering the prevalence of the symptoms of viral infections in the population. The signals prevalence were classified in five categories as follows: very low, when the most probable prevalence is roughly 10%; low, when this prevalence is roughly 30%; medium, when the prevalence is about 50%; high, when it is around 75%; and very high, when the expected prevalence is around 90%. Figure 11.2 presents all distributions used and their respective α and β parameters.

As discussed previously, depending on the signal considered, it is possible that an uninfected individual presents signals in some intensity. However, it is not expected that this happens with great frequency in the population. In other words, it is expected that the majority of the susceptible individuals should be not symptomatic. So, it was assumed that all signals of the susceptible individuals have very low prevalence. To determine the prevalences for the signals of infected individuals in the viral infection scenario simulated, it was considered the prevalences observed in the daycare children during the time of the study.

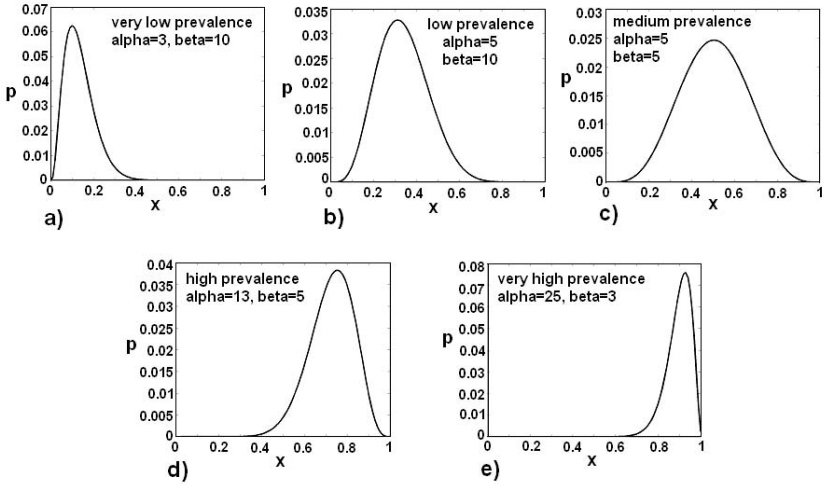

Fig. 11.2. Beta distributions used for the categories prevalences: a) *Very Low*, with $\alpha = 3$ and $\beta = 20$ parameters; b) *Low*, with $\alpha = 5$ and $\beta = 10$; c) *Medium*, with $\alpha = 5$ and $\beta = 5$; d) *High*, with $\alpha = 13$ and $\beta = 5$; and e) *Very High*, with $\alpha = 25$ and $\beta = 3$.

So, based on this observation it was assumed the following signals prevalences: fever, sneezing and wheezing are *Very Low*, cough is *High* and corysa is *Very High*. Note that the Beta distributions defined in figure 11.2 can be used to describe the prevalences of several signals, in different contexts.

Since the simulation of both models involve random process, each simulated condition was repeated 150 times, aiming to find the results through statistical analysis. As expected, the simulations of the models showed that there is a great diversity of dynamical behavior, depending on the parameters values. In some areas of the parameters space the fuzzy and stochastic models are equivalent (for example to small values of φ and ω, with fixed θ). However, there are areas in the phase space where the models present quite different behaviors (see figure 11.3).

In order to analyze the differences and similarities between the fuzzy and the stochastic models in a more detailed way, a diagram was elaborated varying all parameters of the model and considering the dynamical equilibrium provided by both models. As can be noted in figures below, the diagram presents areas in which the epidemic responses of the models completely agree and areas where they have not similar behavior. In fact, there are no abrupt transition between the regions and frontiers between the regions in this diagram could be considered as fuzzy limitations. However, it is possible to define two crisp states: a so-called *concordant* area, where the systems present very similar behavior in the majority of the points; and a so-called *discordant* area, where the systems present quite different behavior for the majority of the points. The *concordant* area is characterized by the presence of the few types of epidemic response, that is, where the epidemic does not hold or it is endemic for both models. On the other

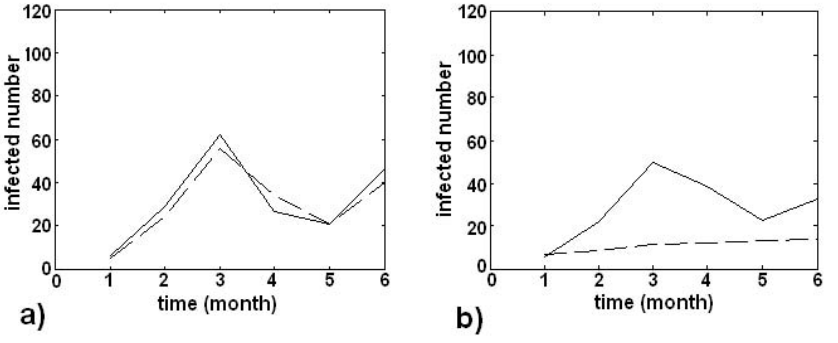

Fig. 11.3. Behavior of the Infected Number in time for the fuzzy model (solid line) and stochastic model (dashed line) for the parameters: a) $\varphi = 0.05$ and $\omega = 0.1$, corresponding to the parameter values in which there is equivalence between the models; and b) $\varphi = 0.05$ and $\omega = 1.5$, corresponding to the parameter values in which there is no equivalence between the models.

hand, in the *discordant* area there are several concomitant epidemic behaviors (endemic, strong epidemic, etc).

Figure 11.4 shows that for small values of initial proportion of infected individual (parameter $\theta \leq 0.04$) there are only three regions in the diagram: 1) a *concordant* area, for small values of φ parameter ($\varphi < 0.01$), where the epidemic does not hold for both models; 2) a *discordance* area, where the fuzzy model always present endemic response and the stochastic model presents both no epidemic and endemic responses; and 3) a *concordant* area, where both models present an endemic behavior (this concordance area is maintained for values of $\varphi \geq 0.07$). In none of these regions of the parameters space, it was observed strong epidemics. This is due to the small values of *theta* parameter, which is responsible for the starting of the infection process in the population.

Varying the values of θ, we can note that the regions in the diagram is modified: a fourth region appears in the map, corresponding to a *discordant* area. Figure 11.5 shows that, for $\theta = 0.05$ this new discordant region start for $\varphi \geq 0.8$ and small ω values. Moreover, this region increases according to the θ value. This should expected since high levels of θ implies in high virus circulation and, by the models assumptions, the signals are more intense. In addition, due to the properties of the $max - min$ composition of fuzzy relations and the summary of the signals, the epidemic course tends to be stronger in the fuzzy model than in the stochastic approach. This occurs because the differences between the values of possibility and probability of infectious contact is more expressive in this situation. Therefore, in this region both models can results in no epidemic response (particularly for high θ values, because the number of susceptible individuals are very low), weakly, moderate or strong epidemic behaviors, but they do not agree for the majority of the parameters set.

In order to evaluate the models performance when faced with real data, we explored the parameters space seeking to find epidemic behaviors that were

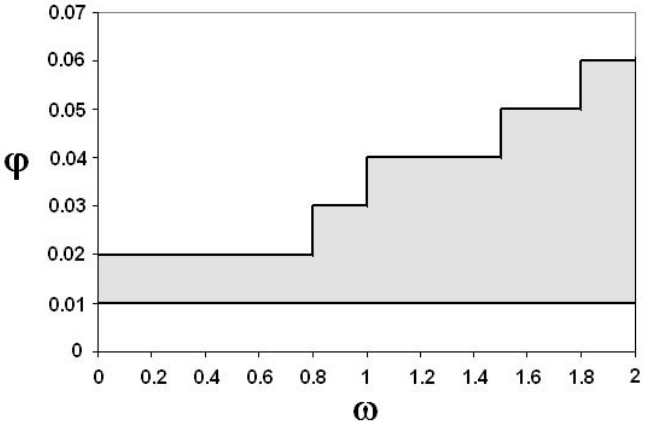

Fig. 11.4. Diagram comparing the epidemic responses of both fuzzy and stochastic models, for $\theta = 0.02$, where two regions are characterized: a concordant region (white area in the graph) and a discordant region (gray area in the graph). For values of φ less than 0.01 the epidemic does not hold for both models. For values of φ greater than 0.06 both models present endemic behavior (Ortega *et al.*, 2008a).

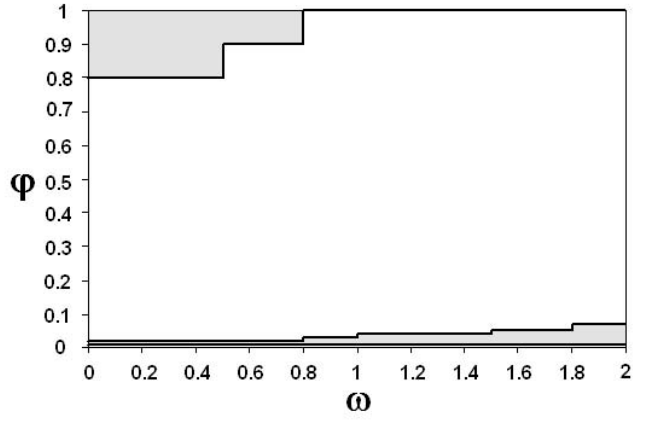

Fig. 11.5. Diagram comparing the epidemic responses of both fuzzy and stochastic models, for $\theta = 0.05$, where four regions are characterized: 1) a concordant region (white area in the inferior part of the graph) in which the epidemic does not hold for both models (for values of φ less than 0.01); 2) a discordant region (gray area in the inferior part of the graph) in which the fuzzy approach provides endemic behavior and in the stochastic model the epidemic does not hold (for small values of φ); 3) a concordant region (white area) in which both models present endemic behavior (for φ values between 0.01 and 0.8); and 4) a discordant region (gray region in the superior part of the graph) in which both models provide no endemic, weakly, moderate or strong epidemic behaviors, but they do not agree for a fixed parameters set (φ values greater than 0.8) (Ortega *et al.*, 2008a).

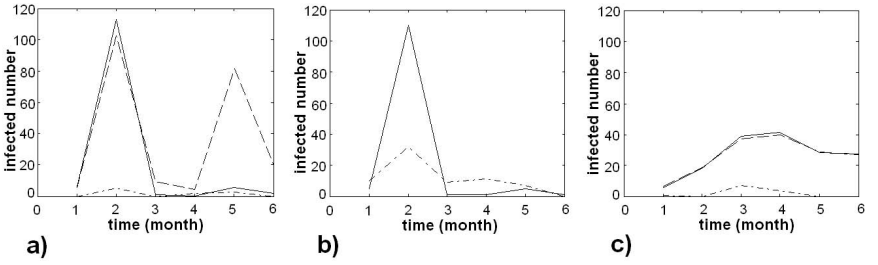

Fig. 11.6. Qualitative comparison between fuzzy (solid line) and stochastic (dashed line) models with real data (dashed dot line), considering the daycare infections in the first semester: a) for RSV infection, in which the results provided by the stochastic model was worse than the one of the fuzzy approach; b) for picornavirus, in which the behavior of the fuzzy and stochastic models are identical; and c) for meta-influenza B infection.

comparable to the daycare infections curves. As the number of children varies in time, particularly between the first and the second semesters due to the holidays, the dynamical simulations were performed to a period of half year (each simulation step corresponding to a month). Figure 11.6 illustrate some examples of these results. We can see in this figure that, for some areas in the parameters space, the models were able to supply a behavior qualitatively similar to the real data. However, the quantitative agreement were not so good when we consider the total number of infected individuals in time.

It is possible to note in figure 11.6a that the models, as well as the real data of the RSV infection in the daycare, present a double peak. In addition, the moment that these peaks occurs were the same in the models and in the real data. However, the maximum number of infected is very large in the models when compared with the data. In this case the fuzzy model presents a slightly better results than the stochastic one, since it provided an attenuation of the infection with time (second small peak). In figure 11.6b we show a comparison between the fuzzy model and the real values of the infection by picornavirus. In this case, the fuzzy and the stochastic results were almost identical. The models supplied a peak of infection in the second month and a second attenuated peak in the fifth month, finishing the epidemic in the 6_{th} month. But in the real data a second peak occurs in the fourth month and it was not so expressive. Figure 11.6c shows another example considering the influenza B infection, where a qualitative similarity between the models and the real data can be analyzed.

Although it is interesting to compare the models performance with the real number of infected distribution, it is most informative to consider all dynamics behavior provided by the models. In other words, it is important to study the dynamic equilibrium of the system, after discarding the transient phase. Thereby, looking for a more accurate quantitatively analysis we study the steady-state of the models and compare it with the average of the number of infected children in daycare.

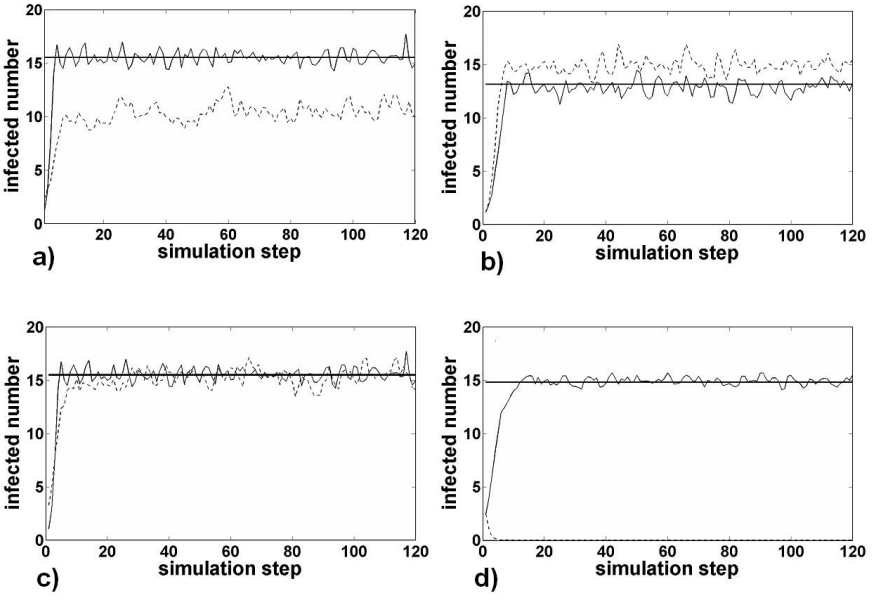

Fig. 11.7. Steady-state analysis comparing the fuzzy (solid line) and stochastic (dashed line) dynamical equilibria with the average of the real data (horizontal line): a) total viral infection, in which the annual average of data was equal to 15.50, and the parameters $\theta = 0.01$, $\varphi = 0.03$ and $\omega = 1$ in both models; b) picornavirus infection, in which the annual average of the data was equal to 13.15, and the parameters $\theta = 0.01$, $\varphi = 0.02$ and $\omega = 1.5$ in both models; c) the same viral infection presented in figure (a) but with the models run with different parameters (for the fuzzy approach it were used parameters $\theta = 0.01$, $\varphi = 0.03$ and $\omega = 1$, and for the stochastic model it were used parameters $\theta = 0.02$, $\varphi = 0.04$ and $\omega = 1.2$); and d) picornavirus infection during the second semester in which the average was equal to 13.15, and parameters $\theta = 0.02$, $\varphi = 0.02$ and $\omega = 1.9$ in both models (case in which the epidemic does not hold in the stochastic approach) (Ortega *et al.*, 2008a).

Figure 11.7 shows four examples that illustrate the results found for the steady-state of the models, comparing them with the average of the real data. In figure 11.7a it is shown the total viral infection in both first and second semester, in which the annual average of infected number was equal to 15.50 cases, and the models' performance, fixing the parameters $\theta = 0.01$, $\varphi = 0.03$ and $\omega = 1$. For this parameters set the fuzzy model showed a better performance than the stochastic approach. The average number of infected in the dynamical equilibrium of the fuzzy and stochastic models are equal to 15.52 and 10.66, respectively.

In figure 11.7b it is shown the real data for picornavirus infection in which the annual average was equal to 13.15 cases, and the models' performance, fixing the parameters set $\theta = 0.01$, $\varphi = 0.02$ and $\omega = 1.5$. Although the stochastic and

fuzzy behaviors, in this case, were more similar than in the prior case, the fuzzy performance was again the best. The fuzzy model provided an average number of infected equal to 12.84, contrasting with the average of 14.94 supplied by the stochastic one.

Considering the same situation as in figure 11.7a and exploring the parameters space searching for a parameters set that provide the best result of the stochastic model we find figure 11.7c. It can be noted in this figure that the fuzzy result, shown in the scenario (a) above, presents a performance as good as the stochastic approach (parameters values of stochastic model were $\theta = 0.02$, $\varphi = 0.04$ and $\omega = 1.2$). The annual average of the stochastic model for number of infected was 15.38 in this case.

Figure 11.7d illustrates how different the model's dynamical behavior can be. In this figure it is compared the average of the number of infected with picornavirus (equal to 13.15) with the fuzzy and stochastic dynamics. Note that, while the fuzzy system reaches a non-trivial steady-state, resulting in an average number of infected equal to 14.93, in the stochastic approach the epidemic does not hold (parameters $\theta = 0.02$, $\varphi = 0.02$ and $\omega = 1.9$ in both models).

11.5 Discussion

From the theoretical point of view, several notes were made in sections 11.2.2 and 11.3.2. However, from the point of view of simulation it is interesting to remark that both fuzzy and stochastic Reed-Frost models can provide a diversity of epidemic behaviors, depending on the parameters set. As observed in the parameters space diagrams, there are regions in which they completely agree with each other and regions where they provide quite different results. This illustrates the fact that different mathematical structures can result in different results.

During the year in that the daycare's children were followed up it were realized a total of 255 exams, in which 186 were diagnosed as respiratory infections. However, in spite of the intense epidemiological work, the number of infected children per month is reasonably small to allow the the performance of the dynamical model more deeply. Besides, the data are insufficient for simulating a retrospective version of the model. In other words, it is not possible to introduce in the model, via likelihood analysis for instance, the information contained in the real data and later to evaluate your predictive ability. This partly explains the poor quantitative results of both models when confronted with the distribution of the number of infected individuals in time, observed in figure 11.6. Moreover, the great number of infected individuals provided by the models, and showed in figure 11.6, can be due to the homogeneous mixing assumption, one of the basis of the classical Reed-Frost model, which in this case perhaps does not apply.

Another aspect to be considered is the seasonal characteristics of the circulation of those viruses in the population and also their relationship with the climatic variation. Although the proposed models do not contain in their

structure any aspect of seasonal or climate conditions, they were able to recover the specific virus type. Although the fuzzy approach allows it through relational matrix, it was considered the only global aspects about respiratory infections virus. This highlight the good results supplied by the models, once they worked well independently of the type of virus. The worst results were obtained for the analysis of the distribution of the number of infected individuals with influenza A and B. However, the number of observations in this case was quite reduced.

Therefore, although the quantitative results of the models are still distant from the real data, the qualitative results are encouraging. Nevertheless, they can be improved by small adaptations in the model aiming to a better fitting of real data. Certainly a variable to be investigated is the function of possibility/probability. Other non-linear functions of the clinical signals could perhaps supply quantitative results more accurately.

By analyzing the dynamical behavior of the steady-state, we can see that the models were able to describe the average of the number of infected individuals for all virus type, considering both annual and half-yearly data. In some cases the fuzzy approach provides a slightly better results. In addition, the situations presented in figure 11.7 are in accordance with that showed by the parameters space analysis, in the sense that in the case in which the set of parameters corresponding to a *concordant* area (figures 11.7a and 11.7c) the fuzzy and stochastic dynamic equilibrium were similar. On the other hand, for the parameters set corresponding to the *discordant* area (figures 11.7b and 11.7d) they do not present equivalent results. In figure 11.7b both of them present endemic situation, but with different average values. However, in figure 11.7d, while in the fuzzy model we have an endemic state, in the stochastic approach the epidemic does not hold. This also illustrates the fact that the transitions in the space of the parameters occur in a fuzzy rather than in a crisp way.

In addition, it is important to point out that both models presented here do not consist simply of generalizing the classical Reed-Frost model formulation. Actually, by including several heterogeneities in the model, the existing differences among individuals is naturally incorporated, making it applicable to real epidemic scenaria. For, on the one hand, there are several infections, like influenza, which, besides being transmitted among small groups of individuals, produce highly heterogeneous clinical pictures, on the other hand, the huge amount of genetic information provided by the emerging field of genomics (and proteonomics) generates clinical information that may sharply distinguish individuals. These tailor-made diagnostic techniques make obvious the necessity of new tools to deal with heterogeneities. Clearly, both approaches can be relaxed in their hypotheses, as the homogeneous mixing assumption for instance, becoming more powerful models. Obviously, in this case, they would lose their identification with the classical Reed-Frost model.

Finally, we would like to point out the importance of this work for the area of epidemic modeling, where the scarceness of information usually makes the elaboration of models that involve the individual aspects (micro) in the epidemic

process (macro) unfeasible. Models of this type are rare in epidemiology and their analysis allows a better understanding of the factors that may contribute to the force of the infection during an epidemic. Maybe this is the most important contribution of these epidemic approach, since they consider the individual's heterogeneities in a simple dynamical structure. In chapter 12 we continue the presentation of the dynamical systems in epidemiology, but considering now the mixing of approaches, in a hybrid model context.

12 Hybrid Models in Epidemiology

This book is devoted to the applications of fuzzy sets theory in epidemiology and correlated areas. However, as it is observed in other fields, it is becoming more and more common the use of the multiple and combined mathematical tools, aiming the treatment of complex problems in biomedical sciences. Usually the mixing of these different approaches involves classical mathematics and artificial intelligent theories, such as fuzzy systems, neural networks, evolutionary computation, expert systems, cellular automata, and so on. This kind of modeling, where several approaches are put working together, is called *hybrid models*.

The most applied artificial intelligence approach is the area of neural networks, followed by evolutionary computation and fuzzy sets theory. The main features of the neural network systems are its learning capacity and recognition of patterns, what turns it into a quite powerful tool in knowledge extraction and its great technological applicability. However, it usually requires a large data set and high performance computations. Besides, its structure does not allow the insertion of expert knowledge and the neural network architecture and synaptic matrix are not interpretable. In the *MEDLINE* database we found around 27,000 articles with Artificial Intelligence in biomedical fields, of which 14,000 refer to the Neural Network approach, the first one is reported to the 1960s (Greene, 1962). In medicine the first work reported in *MEDLINE* is the *Control theory applied to neural networks illuminates synaptic basis of interictal epileptiform activity* published in 1986, in *Advances in Neurology* (Johnston & Brown, 1986).

Evolutionary computation is one of the largest promises of the technological areas, particularly in software developments. Its most important characteristic is its adaptation capacity, which also results in a learning ability in some way. Due to its features, evolutionary computation have been widely applied in virtual lives, immunological computers systems and robotics. However, as in neural network systems, it demands a heavy data analysis and hard computation understanding. In the *MEDLINE* database we found around 200 scientific articles about evolutionary computation and more than 8,000 works about genetic algorithms.

Fuzzy logic, in contrast, has as its main feature the capacity to imitate human reasoning, particularly concerning decision making process, and to deal with

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 253–276, 2008.
springerlink.com

identification uncertainty. Frequently, fuzzy systems involve simple mathematical and computational structures and are easily interpretable. In contrast, it does not have adaptation or learning capacity. Perhaps for the fact fuzzy logic is the most recent theory of that three cited above, it is less mentioned in the *MEDLINE* database. In fact, less than 2,000 works reporting to fuzzy logic appear in *MEDLINE*. The first of them was published in the 1970s by Hiramatsu and collaborators (1974), in Japanese, with English title *Applications of the fuzzy logic to medical diagnosis*.

Due to the fact that neural networks, evolutionary computation and fuzzy logic have well defined features that are complementary to each other, it is natural the combination of those tools (see also chapter 13). These combinations usually result in powerful and advantageous hybrid models. Indeed, most of the hybrid models in engineering and bioengineering areas are resulting of these combinations, compounding the so-called neuro-fuzzy systems, the adaptive neural networks or fuzzy-genetic systems, as some examples (Teodorescu *et al.*, 1999a; Teodorescu *et al.*, 1999b; Szczepaniak*et al.*, 2000).

In the fuzzy systems context, it is common the use of both neural networks and evolutionary computation to build linguistic models, since both of them are suitable to extract information available in a data set. As discussed in chapter 8 these artificial intelligence tools can provide membership functions, define input and/or output variables and even the fuzzy rules, playing the expert's role (Wang & Mendel, 1992; Jang, 1993; Peña-Reyes & Sipper, 1999). A good and didactic example of genetic algorithm applied to build fuzzy linguistic model is the work of Peña-Reyes and Sipper (1999), where a breast cancer diagnosis is developed. In this paper the authors compare the models elaborated by an experts panel with that provided by the computational intelligent system, facing both of them with real data. In the *MEDLINE* database we can found more than 200 articles considering the hybrid models such as neuro-fuzzy systems and fuzzy genetic algorithms. The applications of hybrid models on biomedicine are recent and the first works are reported to the 1990s (Barillot *et al.*, 1993; Kuncheva & Andreeva, 1993, Kwok *et al.*, 2003 and 2004).

In epidemiology, applications based on neural networks and evolutionary algorithms are not common. As discussed along this book, modern modeling epidemiology is still concentrated, for the most part, in the use of classical mathematical modeling, based on differential equations, and stochastic approaches. However, to exemplify what happens in the medical area, the publications of neural networks and genetic algorithms techniques in epidemiology are more intense than the fuzzy logic. In fact, if we consider the *MEDLINE* database we found around 300 articles considering the applications of neural networks or genetic algorithms in epidemiology. However, only around 15 articles of *MEDLINE* truly consider the applications of fuzzy logic in epidemiology (Merilan & Roe, 1993; Hammad *et al.*, 1996; Bolotin, 2004).

Considering the discussion above, it is expected that hybrid models with fuzzy logic techniques in epidemiology are rare and consider basically the mixing of classical and soft computing techniques. In this chapter we present two

examples of hybrid systems applied in epidemiology. The first of them consists in an aggregation of Bayesian statistic test with fuzzy hypothesis and fuzzy linguistic model, in a decision making support context, for the optimal age for vaccination against measles (Ortega *et al.*, 2008b). In the second work, linguistic fuzzy models are combined with numerical calculus to elaborate a predator-prey model to study the interaction between aphids and ladybugs in citriculture. This study is particularly important due to its economical appeal, since it allows the understanding of the dynamic process of the Citrus Sudden Death and elaborating control strategies (Peixoto *et al.*, 2008a).

As it will be noted along the next sections, both models have interesting features from the theoretical and epidemiological points of view, which confers originality and creativity to them. In the model for the estimation of optimal age for vaccination against measles the dynamic information is built from the non-dynamic age structure, so avoiding the complex calculus commonly involved in fuzzy dynamic approaches (as presented in chapters 9 and 10). In this model, the distribution of maternally derived antibodies seroprevalence and children seroconversion rate to the vaccine are evaluated through classical and fuzzy Bayesian test. This statistical test decision is aggregated, in a fuzzy rule-base structure, with the risk of adverse effects associated to the vaccine and with the undesirability of acquiring the disease, estimating the vaccine recommendation for a determined age (Ortega *et al.*, 2008b).

In the predator-prey model it is built a dynamical system based only on the experts' information through the fuzzy rule-base structure, since there was no sufficient information about the phenomenon to build classical differential equations (Peixoto *et al.*, 2008a). So, in contrast with the approaches presented in chapters 9 and 10, in this case the identification uncertainties of the states variables and parameters are not based upon the fuzzification of the classical differential equations. This theoretical detail may seems simple. However it illustrates a fundamental difference among those two approaches, particularly in respect to the applicability of that models to real problems. While in the former it is necessary the estimation of epidemic parameters to process the system, in the latter the dynamical expert-rule-based structure can be used to provide these parameters, which are usually not available. Additionally, this dynamical expert-rule-based model can be used to provide the classical equations that best describe the studied phenomenon, transforming, thereby, the experts' knowledge into mathematical equations. This approach was applied by Peixoto and collaborators associated with a cellular automata environment to study the spatial and temporal dynamics of citrus sudden death, which consists also in a hybrid model (Peixoto *et al.*, 2008b).

12.1 A Fuzzy Model for the Estimation of Optimal Age for Vaccination Against Measles

As briefly discussed in chapter 6, measles is one of the most infectious and lethal diseases, being responsible for 10% of global mortality from all causes

among children aged less than 5 years, which represents approximately 1 million deaths annually (CDC, 1998). In spite of a very effective vaccine, which averted approximately 1.67 million measles associated deaths in 1996, measles still is the first cause of deaths preventable by vaccines in children.

Previous studies have already demonstrated the necessity of vaccinating between 94% and 98% of the susceptible children to avoid measles outbreaks (Reed, 1999). However, the optimal age to vaccinate children in a routine immunization calendar is still dependable on a set of variables, characteristic of the target population, like the seroconversion rate of children below 1 year of age, the presence of maternally derived antibodies, the serostatus of mothers, among others (Zanetta *et al.*, 2002).

Mathematical models have been proposed and applied for the estimation of optimal age to vaccinate children, not only against measles (Hethcote, 1988; Zanetta *et al.*, 2002), but also against rubella (Massad *et al.*, 1994b). Those models are of deterministic structure and, despite some attempts to include age-dependence of the force of infection parameter (Amaku *et al.*, 2003), all the available calculations of the optimal age to vaccinate, to the best of our knowledge, have assumed a constant, age-independent force of infection. In addition, the variables which determine the optimal age to vaccinate are usually difficult to determine and usually only fragmentary information on them are available. Therefore, we will present in this example, an alternative approach, which combines two powerful techniques to dealing with subjective and/or imprecise information: the Bayesian approach and the fuzzy sets theory.

Several articles can be found in the literature combining the Bayesian approach with ideas from fuzzy sets theory (Okuda *et al.*, 1978; Tanaka *et al.*, 1979; Uemura, 1991; Viertl, 1996; Gil *et al.*, 1985; Casals, 1993; Delgado *et al.*, 1979; Taheri & Behboodian, 2001). In this work we applied the method proposed by Taheri and Behboodian (2001), who consider the problem of hypotheses testing when the data (observations) are crisp and the hypotheses are fuzzy, such as: θ is approximately one, θ is very high, etc (Taheri & Behboodian, 2001). It was assumed that the estimation of optimal age to vaccinate is a decision analysis procedure, which involves aims and constraints (see Zimmermann, 1996; and Massad *et al*, 1999) for a discussion on fuzzy decision taking). In this case the aim is to minimize the lifetime expected risk of acquiring measles infection and the constraints include the immunological aspects involving the "taking" of the vaccine, like the seroconversion rates, maternally derived antibodies, the mother serostatus, among others. Logistic aspects of vaccination will not be considered in this work. The results of the fuzzy approach was compared with the classical counterpart.

12.1.1 A Bayesian Approach to Fuzzy Hypotheses Testing

As mentioned above we applied the methods proposed by Taheri and Behboodian (2001), who consider the problem of hypotheses testing when the data are crisp measures and the hypotheses are fuzzy.

As defined by those authors, any hypothesis of the form $H : \theta$ *is* $H(\theta)$ is said to be a *fuzzy hypothesis*, where $H(\theta)$ is a membership function from the space Θ to $[0, 1]$. Examples of fuzzy hypothesis include, θ *is approximately* $1/2$, θ *is very low*, among others.

Let $X = (X_1, ..., X_n)$ be a random sample, with observed value $x = (x_1, ..., x_n)$, where X_i has the probability density function, p.d.f., $f(x_i|\theta)$ with unknown $\theta \in \Theta$, whose prior density is $\pi(\theta)$. Suppose, as in Taheri and Behboodian (2001), that two membership functions $H_0(\theta)$ and $H_1(\theta)$ are given. So, the main problem consists in testing

$$\begin{array}{lllll} H_0: & \theta & is & H_0(\theta), \\ H_1: & \theta & is & H_1(\theta) \end{array} \tag{12.1}$$

on the basis of a Bayesian method.

As it was assumed the problem of finding the best age to vaccinate children against measles as a decision analysis problem, it must be defined the space of possible actions (ages of vaccination) A, and the loss function $L(\theta, a) : \Theta \times A \to R$ where R is the risk space. The loss function specifies the loss when taking action a when the true parameter is θ.

Now, assuming that θ has a prior distribution $\pi(\theta)$ and that $f(x|\theta)$ is the p.d.f of X with fixed $\theta \in \Theta$, then, the posterior density of θ, $\pi(\theta|x)$, is proportional to its priori and to the p.d.f. of X,

$$\pi(\theta|x) \propto \pi(\theta) f(x|\theta). \tag{12.2}$$

The Bayes risk of a decision d, associated with the prior $\pi(\theta)$, is then

$$R(\pi, d) = E\left[R(\theta, d)\right]. \tag{12.3}$$

The aim is, therefore, to minimize the risk by taking the optimal decision d^*, that is,

$$R(\pi, d^*) = \inf_{d \in D} R(\pi, d) \tag{12.4}$$

where D is the space of possible decisions.

Bayes test without a loss function

We will now test the fuzzy hypothesis H_0, accepted in point a_0 and rejected in a_1, based on a random sample from $f(x|\theta)$ with prior density $\pi(\theta)$ for θ.

A Bayes test without loss function rejects H_0 if and only if the posterior density under H_0 is less than the posterior density under H_1, that is,

$$\int \pi(\theta|x) H_0(\theta) d\theta < \int \pi(\theta|x) H_1(\theta) d\theta. \tag{12.5}$$

As we are operating in a fuzzy setting, it is important to define a criterion related to the degree of acceptance of H_0 versus H_1 (Taheri & Behboodian, 2001):

$$\frac{\alpha_0}{\alpha_0 + \alpha_1}, \tag{12.6}$$

where

$$\alpha_0 = \int \pi(\theta|x) H_0(\theta) d\theta \tag{12.7}$$

and

$$\alpha_1 = \int \pi(\theta|x) H_1(\theta) d\theta. \tag{12.8}$$

The ratio α_0/α_1 is called the *posterior odds ratio* of H_0 to H_1, and the ratio $\int \pi(\theta|x) H_0(\theta) d\theta / \int \pi(\theta|x) H_1(\theta) d\theta$ is called the *prior odds ratio* (Taheri & Behboodian, 2001).

Bayes test with loss function

First, as in (Taheri & Behboodian, 2001) we define the following loss functions:

$$L(\theta, a_0) = a(\theta) \left[1 - H_0(\theta)\right] \tag{12.9}$$

and

$$L(\theta, a_1) = b(\theta) \left[1 - H_1(\theta)\right] \tag{12.10}$$

where $a(\theta)$ and $b(\theta)$ are nonnegative functions, depending on our sensitivity to false rejection or false acceptance.

Now, according to theorem 3.1 of Taheri and Behboodian (2001), if (12.9) and (12.10), then the Bayes test accepts H_0 iff

$$\int a(\theta)[1 - H_0(\theta)]\pi(\theta|x) d\theta \leq \int b(\theta)[1 - H_1(\theta)]\pi(\theta|x) d\theta. \tag{12.11}$$

Alternatively, if we call $a(\theta) = C_{II}$ and $b(\theta) = C_I$, where C_I is related to error type I and C_{II} to error type II, then the Bayes test accepts H_0 iff

$$\frac{1 - \int \pi(\theta|x) H_1(\theta) d\theta}{1 - \int \pi(\theta|x) H_0(\theta) d\theta} \geq \frac{C_{II}}{C_I}. \tag{12.12}$$

12.1.2 Estimating the Optimal Age to Vaccinate Against Measles

As mentioned above, several variables determine the optimal age to vaccinate children against measles. In this analysis we choose: *the presence of maternally derived antibodies*; *the children seroconversion rate to the vaccine*; *the risk* (side effects) associated to the vaccine; and an usually *ad hoc* function that express *the undesirability of acquiring the disease* (Zanetta *et al*, 2002).

For each of the first two variables we carried out a fuzzy Bayes test, which provided the degree of decision related to the null of alternative hypotheses. The last two variables were both considered as fuzzy sets. These four variables were applied in the rule-based fuzzy model which will determine the decision process.

A Bayes test for maternally derived antibodies

For this variable we adopted the following fuzzy hypotheses:

$$\begin{array}{ll} H_0: & \theta \quad is\ low, \\ H_1: & \theta \ \ is\ not\ low. \end{array} \tag{12.13}$$

The membership functions for those fuzzy hypotheses are:

$$H_0(\theta) = \begin{cases} 1 & if \quad 0 \leq \theta < 0.15 \\ (0.85 - \theta)/0.7 & if \quad 0.15 \leq \theta < 0.85 \\ 0 & if \quad 0.85 \leq \theta \leq 1 \end{cases}$$

and

$$H_1(\theta) = \begin{cases} 0 & if \quad 0 \leq \theta < 0.15 \\ (\theta - 0.15)/0.7 & if \quad 0.15 \leq \theta < 0.85 \\ 1 & if \quad 0.85 \leq \theta \leq 1 \end{cases}.$$

The shape of the membership functions of the H_0 and H_1 are presented in figure 12.1.

The *a priori* density function for θ is assumed as uniform. Therefore, its *a posteriori* density function is $Beta(y+1, n-y+1)$, that is:

$$\pi\left(\theta | X = x\right) = \frac{(n+1)!}{y!\,(n-y)!}\theta^y\left(1-\theta\right)^{n-y}, \tag{12.14}$$

where n is the number of tested children and y the number of positive to the test, and

$$y = \sum_{i=1}^{n} x_i. \tag{12.15}$$

As mentioned above, the Bayes test accepts H_0 iff:

$$\frac{\left(1 - \int \pi(\theta|x) H_1(\theta) d\theta\right)}{\left(1 - \int \pi(\theta|x) H_0(\theta) d\theta\right)} \geq \frac{C_{II}}{C_I}. \tag{12.16}$$

It was assumed, for the sake of simplicity, $a(\theta) = C_{II} = b(\theta) = C_I = 1$, and the condition for accepting H_0 is simplified to

$$\int \pi(\theta|x) H_0(\theta) d\theta \geq \int \pi(\theta|x) H_1(\theta) d\theta. \tag{12.17}$$

Now, for fixed values of n and y in each age class, the condition (12.17) can, for a positive constant $\kappa(n, y) = \frac{(n+1)!}{y!(n-y)!}$, be rewritten as:

$$\int \theta^y (1-\theta)^{n-y} H_0(\theta) d\theta \geq \int \theta^y (1-\theta)^{n-y} H_1(\theta) d\theta \tag{12.18}$$

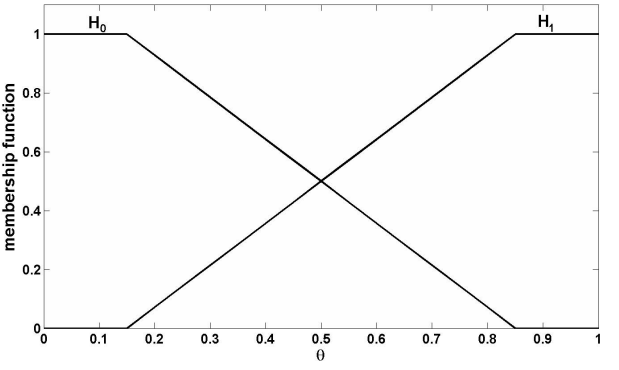

Fig. 12.1. Membership functions for the hypotheses H_0: θ *is low* and H_1: θ *is not low* for the Bayes test of maternally derived antibodies (Ortega *et al.*, 2008b)

which, in terms of the membership functions, takes the form:

$$\begin{aligned}&\int_0^{0.15} \theta^y(1-\theta)^{n-y}d\theta + \int_{0.15}^{0.85} \theta^y(1-\theta)^{n-y}(0.85-\theta)/0.7d\theta \geq \\ &\int_{0.15}^{0.85} \theta^y(1-\theta)^{n-y}(\theta-0.15)/0.7d\theta + \int_{0.85}^{1} \theta^y(1-\theta)^{n-y}d\theta.\end{aligned} \tag{12.19}$$

In addition to the condition for accepting H_0 we define a criterion related to the degree of acceptance of H_0 versus H_1, $\frac{\alpha_0}{\alpha_0+\alpha_1}$, as in equation (12.6).

A Bayes test for vaccine seroconversion

For this variable it was adopted the same fuzzy hypotheses as the previous section (see 12.13), however, the membership functions for this case are:

$$H_0(\theta) = \begin{cases} 1 & if \quad 0 \leq \theta < 0.15 \\ (0.65-\theta)/0.5 & if \quad 0.15 \leq \theta < 0.65 \\ 0 & if \quad 0.65 \leq \theta \leq 1 \end{cases}$$

and

$$H_1(\theta) = \begin{cases} 0 & if \quad 0 \leq \theta < 0.15 \\ (\theta-0.15)/0.5 & if \quad 0.15 \leq \theta < 0.65 \\ 1 & if \quad 0.65 \leq \theta \leq 1 \end{cases}$$

whose shape can be seen in figure 12.2.

The *a priori* and its *a posteriori* density functions are the same as for the previous section. The Bayes condition for accepting H_0 is now:

$$\begin{aligned}&\int_0^{0.15} \theta^y(1-\theta)^{n-y}d\theta + \int_{0.15}^{0.65} \theta^y(1-\theta)^{n-y}(0.65-\theta)/0.5d\theta \geq \\ &\int_{0.15}^{0.65} \theta^y(1-\theta)^{n-y}(\theta-0.15)/0.5d\theta + \int_{0.65}^{1} \theta^y(1-\theta)^{n-y}d\theta.\end{aligned} \tag{12.20}$$

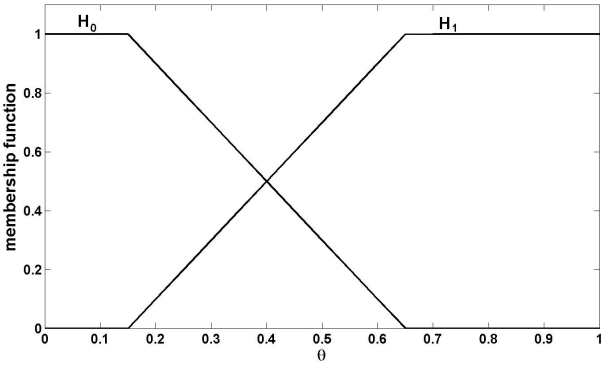

Fig. 12.2. Membership functions for the hypotheses H_0: θ *is low* and H_1: θ *is not low* for the Bayes test of vaccine seroconversion (Ortega *et al.*, 2008b)

And, again, in addition to the condition for accepting H_0 it was defined a criterion related to the degree of acceptance of H_0 versus H_1 , $\frac{\alpha_0}{\alpha_0+\alpha_1}$, as in equation (12.6).

In order to compare with the fuzzy Bayes test approach, a classical Bayes test was also performed. In this case, it was considered the following hypotheses:

$$\begin{aligned} H_0 &: \theta \geq 0.5 \\ H_1 &: \theta < 0.5. \end{aligned} \tag{12.21}$$

Membership degree functions for the Risk of Vaccination and Undesirability Function

The variables *risk of vaccination* and *undesirability function*, related to the vaccine side effects and mortality due to measles, respectively, were not subjected to the hypotheses test but were rather considered as fuzzy sets and included in the fuzzy rule-based model described below. These membership functions were obtained by interviewing experts and by available knowledge in the field of vaccination

Adverse effects following the measles vaccination are generally mild and limited to individuals who are susceptible (Markowitz & Katz, 1994). Of special interest is the occurrence of dysfunction of the central nervous system, such as encephalitis and encephalopathy after vaccination, which have an estimated risk of approximately 1 case per million doses for either of these two events (Landigran & Witte, 1973). These low estimated risks associated with the vaccine are approximately one tenth of the risk of the same effects associated with measles natural infection, which makes the vaccine very safe in relation to the infection. However, it is interesting, for the sake of generality, to include those risks associated with the vaccine in the analysis. It was assumed that the risk of side effects are inversely proportional to the age of vaccination since both the possible effects are graver in lower age children. Therefore, it was assigned the

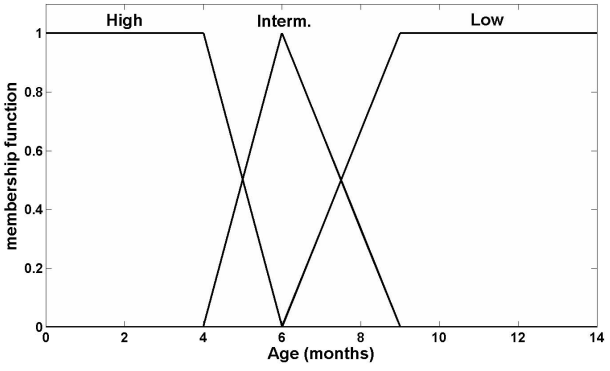

Fig. 12.3. Membership functions for the *Adverse Effects of Vaccination*, that is *Risk* (Ortega *et al.*, 2008b)

following categories of risk, R: *high risk, intermediate risk* and *low risk*, with the corresponding membership degree functions:

$$R\ High: \quad \mu_{R_H}(age) = \begin{cases} 1 & if \quad age < 4\ months \\ -0.5age + 3 & if \quad 4\ months \ \leq age < 6\ months \\ 0 & if \quad age > 6\ months \end{cases},$$

$$R\ Interm.: \quad \mu_{R_I}(age) = \begin{cases} 0.5age - 2 & if \quad 4\ months \ \leq age < 6\ months \\ 1 & if \quad age = 6\ months \\ -1/3age + 3 & if \quad 6\ months \ < age \leq 9\ months \\ 0 & \quad age > 9\ months \end{cases}$$

and

$$R\ Low: \quad \mu_{R_L}(age) = \begin{cases} 0 & if \quad age < 6\ months \\ 1/3age - 2 & if \quad 6\ months \ \leq age < 9\ months \\ 1 & if \quad age \geq 9\ months \end{cases}$$

whose shape can be seen in figure 12.3.

The complications associated with measles infection have been subject of much description and review (Markowitz & Katz, 1994), ranging from otitis media, through pneumonia, encephalitis, culminating in rare cases (0.1 - 1.0/1000 cases) in death. The risk of serious complications and death is increased in young children (Babbott & Gordon, 1954). In order to model the age-dependent undesirability function it was applied the same *ad hoc* approach used in previous publications aimed at the estimation of optimal age to vaccinate for measles (Hethcote, 1988; Zanetta *et al*, 2002) and for rubella (Massad *et al*, 1994b).

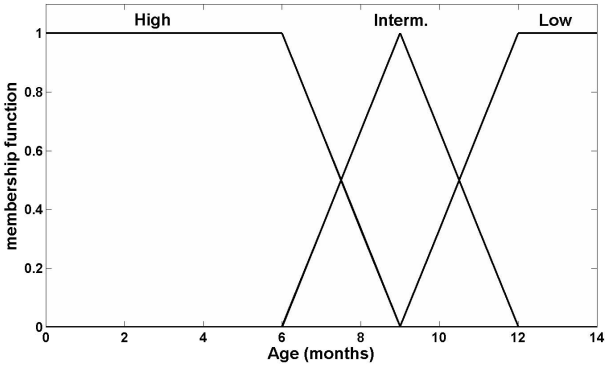

Fig. 12.4. Membership functions for the *Undesirability Function* (Ortega *et al.*, 2008b)

For the undesirability of the vaccine, U, it was assigned the following age-dependent membership functions:

$$U\ High: \quad \mu_{U_H}(age) = \begin{cases} 1 & if \quad age \leq 6\ months \\ -1/3age + 3 & if \quad 6\ months\ < age < 9\ months \\ 0 & if \quad age \geq 9\ months \end{cases},$$

$$U\ Interm.: \quad \mu_{U_I}(age) = \begin{cases} 1/3age - 2 & if \quad 6\ months\ \leq age < 9\ months \\ 1 & if \quad age = 9\ months \\ -1/3age + 4 & if \quad 9\ months\ < age \leq 12\ months \\ 0 & \quad age > 12\ months \end{cases}$$

and

$$U\ Low: \quad \mu_{U_L}(age) = \begin{cases} 0 & if \quad age < 9\ months \\ 1/3age - 3 & if \quad 9\ months \leq age < 12\ months \\ 1 & if \quad age \geq 12\ months \end{cases}$$

whose shape can be seen in figure 12.4.

12.1.3 A Fuzzy Rule-Based Model

The decision making problem of finding the best age to vaccinate children against measles was modeled by a TSK rule-based model (see chapter 7). Each fuzzy rule in the model has four antecedents variables: the outcome of the Bayes test for the *presence of maternally derived antibodies* and for *children seroconversion rate to the vaccine* (H_0 and H_1, respectively, and they acceptance degree), and two fuzzy sets related to the risk (*side effects*) associated to the vaccine and the *undesirability of acquiring the disease*, described above. Note that the variables *risk of side effects* and *undesirability functions* were nor tested. The consequent

of fuzzy rule is the *vaccination status* described as a constant function defined in the interval [0,10], to rank the level of recommendation to vaccinate, classified as *strongly recommended* (10), *recommended* (7) and *not-recommended* (0). These three possibilities were the consequence of the combination of the various antecedent variables, according to their degree of membership functions. An example of a fuzzy rule is:

IF *vaccine seroconversion* is H_0 AND *maternally antibodies* is H_1 AND *undesirability function* is *intermediate* AND *adverse effects* is *low*, THEN *recommendation level to vaccine* is *recommended*.

The combination of all input variables resulted in 36 fuzzy rules, shown in table 12.1.

It is important to note that the *presence of maternally derived antibodies* and *the children seroconversion rate to the vaccine* are singleton sets (see chapter 2). The membership degree of this sets are the acceptance degree for H_0 or H_1, found by Bayesian test described previously. In the sense to clarify the algorithm of the model, figure 12.5 presents a scheme of the methodology applied.

It was assumed that the optimal age to vaccinate against measles is the age that the *recommendation level to vaccinate* is maximum. The basic difference between the fuzzy and classical approaches presented in this model refers to the acceptance degree of *vaccine seroconversion* and *maternal antibodies*. Both degrees are fuzzy under the former approach and crisp under the latter.

In order to demonstrate a practical application of the above theory the model was applied to the data described in (Zanetta *et al.*, 2002). In this work the authors estimated the optimal age for vaccination against measles in the state of São Paulo, Brazil, taking into account mothers' serostata and based on the classical dynamical model. The authors measured the seroprevalence of measles virus antibody of children in the first year of life and their mothers. In addition, they compared maternal antibody decay of two groups of children: those whose mothers were 25 years old or more (mothers born in the pre-vaccination era), and less than 25 years old (mothers born in the vaccination era). Therefore, the 25-year-age cut-off was chosen to distinguish between vaccinated and non-vaccinated mothers. The author's hypothesis was that children born of mothers who are positive to antibodies against measles by natural infection had a higher probability of being seropositive to older ages than those born to mothers who were positive by vaccination. The expected difference would determine a younger age to vaccinated the latter in comparison to the former. Seroconversion rates to the vaccine was also measured for both groups. The optimal age to vaccinated children of each group was estimated by a dynamical model of the type proposed by Hethcote (1988). The data collection comprised 1,216 mothers along with their serostata, and 552 children whose seroconversion was verified.

Table 12.2 present the results of the Bayes test applied for the same population described and studied by Zanetta and collaborators (2002) for maternally derived antibodies. It can be noted that, as expected, children born from the older mothers group remain with high levels of maternally derived antibodies

Table 12.1. Fuzzy rules of the TSK model, where NR is Not Recommended, R is Recommended and SR is Strongly Recommended (Ortega *et al.*, 2008b)

Rule	Vaccine Seroconversion	Maternally Antibodies	Undesirability Function	Adverse Effects	Vaccination Decision
1	H0	H0	Low	Low	NR
2	H0	H0	Low	Interm.	NR
3	H0	H0	Low	High	NR
4	H0	H0	Interm.	Low	R
5	H0	H0	Interm.	Interm.	NR
6	H0	H0	Interm.	High	NR
7	H0	H0	High	Low	R
8	H0	H0	High	Interm.	R
9	H0	H0	High	High	NR
10	H0	H1	Low	Low	NR
11	H0	H1	Low	Interm.	NR
12	H0	H1	Low	High	NR
13	H0	H1	Interm.	Low	R
14	H0	H1	Interm.	Interm.	R
15	H0	H1	Interm.	High	NR
16	H0	H1	High	Low	R
17	H0	H1	High	Interm.	R
18	H0	H1	High	High	NR
19	H1	H0	Low	Low	R
20	H1	H0	Low	Interm.	R
21	H1	H0	Low	High	NR
22	H1	H0	Interm.	Low	R
23	H1	H0	Interm.	Interm.	R
24	H1	H0	Interm.	High	NR
25	H1	H0	High	Low	SR
26	H1	H0	High	Interm.	R
27	H1	H0	High	High	R
28	H1	H1	Low	Low	R
29	H1	H1	Low	Interm.	NR
30	H1	H1	Low	High	NR
31	H1	H1	Interm.	Low	R
32	H1	H1	Interm.	Interm.	R
33	H1	H1	Interm.	High	NR
34	H1	H1	High	Low	SR
35	H1	H1	High	Interm.	R
36	H1	H1	High	High	NR

(H_1 for ages 0 and 1.5 months) for longer period then that born from the younger mothers (H_1 for age 0 and H_0 for 1.5 months).

Table 12.3 shows the Bayes test results for vaccines seroconversion in the three groups. For both variables, *maternally derived antibodies* and *vaccines seroconversion*, the Bayes decision was equal in the classical and fuzzy hypotheses (H_0 or H_1). The main difference consists in the degree of acceptance value, that can be 0 or 1, in the former, and a number in the interval [0,1] for the last.

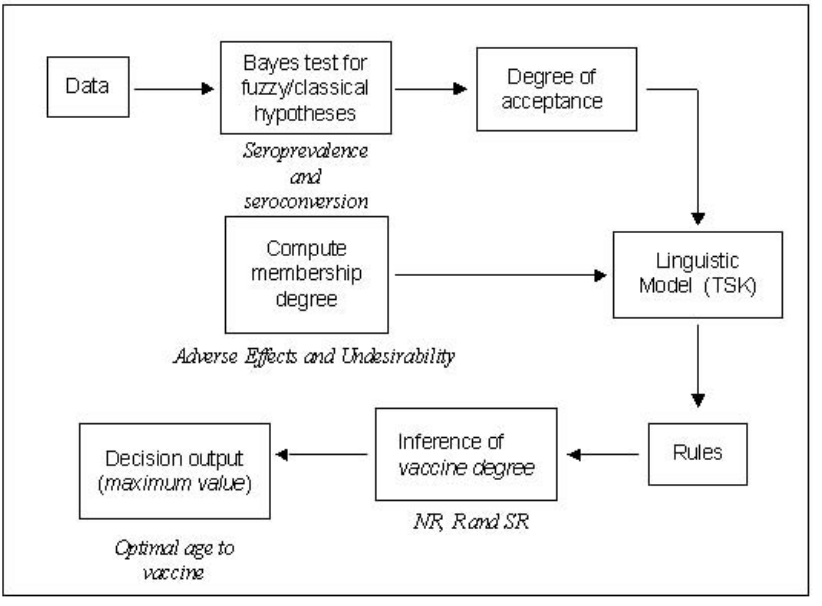

Fig. 12.5. Scheme of the methodology applied (Ortega *et al.*, 2008b)

Table 12.2. Bayes test result for maternally derived antibodies for three groups: 1) mothers older than or equal to 25 years of age; 2) mothers younger than 25 years of age; and 3) all mothers (Ortega *et al.*, 2008b)

Age	**Mothers age** ≥ 25 **years**				**Mothers age** < 25 **years**				**General**			
	n	y	Bayes Test	Degree of acceptance	n	y	Bayes Test	Degree of acceptance	n	y	Bayes Test	Degree of acceptance
0	31	31	H_1	1.0000	60	50	H_1	1.0000	91	81	H_1	1.000
1.5	60	27	H_1	0.6000	63	15	H_0	0.8100	123	42	H_0	0.612
2.5	57	14	H_0	0.7900	53	9	H_0	0.9210	110	23	H_0	0.871
3.5	55	13	H_0	0.8100	55	7	H_0	0.9720	110	20	H_0	0.920
4.5	55	5	H_0	0.9930	55	4	H_0	0.9970	110	9	H_0	0.999
5.5	50	7	H_0	0.9570	53	5	H_0	0.9910	103	12	H_0	0.992
6.5	76	4	H_0	0.9998	78	5	H_0	0.9995	154	9	H_0	1.000
7.5	46	9	H_0	0.8800	42	3	H_0	0.9940	88	12	H_0	0.976
8.5	29	5	H_0	0.8900	38	5	H_0	0.9560	67	10	H_0	0.955
9.5	34	8	H_0	0.7970	32	10	H_0	0.6530	66	18	H_0	0.740
10.5	39	22	H_1	0.8100	38	23	H_1	0.8770	77	45	H_1	0.859
11.5	29	11	H_0	0.5300	30	14	H_1	0.6360	59	25	H_1	0.552
12.5	28	10	H_0	0.5700	30	12	H_1	0.5120	58	22	H_0	0.533

Applying the classical and fuzzy Bayes tests results in the fuzzy rule-based model, to both mothers groups, the recommended level to vaccinate against measles was calculated. The behavior of this recommended degree is presented in figure 12.6.

Table 12.3. Bayes test results for vaccine seroconversion for three groups: 1) mothers older than or equal to 25 years of age; 2) mothers younger than 25 years of age; and 3) all mothers. In this case the y^* values were estimated from seroconversion function presented in Zanetta *et al.*, 2002 (Ortega *et al.*, 2008b)

Age	**Mothers age $\geq$ 25 years**				**Mothers age $<$ 25 years**				**General**			
	n	y^*	Bayes Test	Degree of acceptance	n	y^*	Bayes Test	Degree of acceptance	n	y^*	Bayes Test	Degree of acceptance
0	–	0	H_0	1.000	–	0	H_0	1.000	–	0	H_0	1.000
1.5	–	1	H_0	1.000	–	3	H_0	1.000	–	2	H_0	1.000
2.5	–	3	H_0	1.000	–	6	H_0	1.000	–	4.5	H_0	1.000
3.5	–	6	H_0	0.9999	–	11	H_0	0.996	–	9	H_0	0.999
4.5	–	10	H_0	0.9976	–	19	H_0	0.932	–	15	H_0	0.974
5.5	–	16	H_0	0.9656	–	28	H_0	0.808	–	22	H_0	0.890
6.5	67	25	H_0	0.6760	59	22	H_0	0.6760	126	39	H_0	0.768
7.5	47	21	H_0	0.5730	55	24	H_0	0.5880	102	38	H_0	0.679
8.5	32	18	H_1	0.5840	33	19	H_1	0.6020	65	36	H_1	0.575
9.5	32	20	H_1	0.5550	35	21	H_1	0.6350	67	42	H_1	0.676
10.5	37	32	H_1	0.9650	36	29	H_1	0.9060	73	61	H_1	0.956
11.5	27	22	H_1	0.9080	30	26	H_1	0.9598	57	48	H_1	0.957
12.5	31	23	H_1	0.8230	31	23	H_1	0.8230	62	46	H_1	0.835

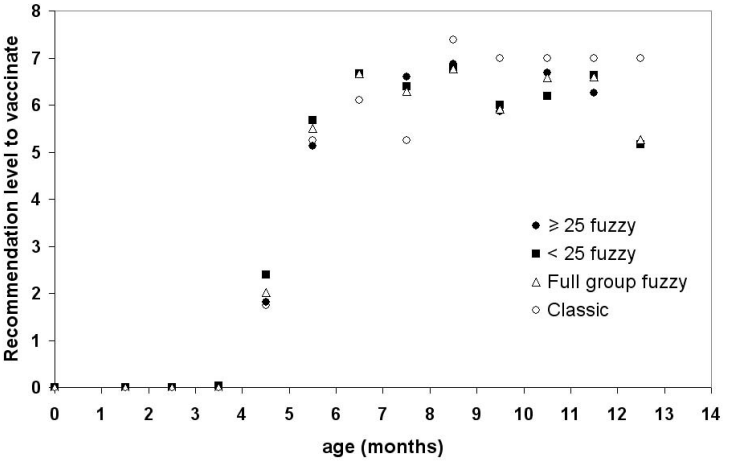

Fig. 12.6. Behavior of Recommendation level to vaccinate with age, and the optimal age to vaccination against measles, by fuzzy and classical approach, considering tree groups: 1) mothers 25 years old or more; 2) mothers less then 25 years old; and 3) full group which there are no distinction between mothers age (Ortega *et al.*, 2008b)

It is possible to note in figure 12.6 that although the optimal age to vaccinate (8.5 months, the maximum value) is the same under both approaches, fuzzy and classical, the behavior of the level recommendation is different from 6.5 months. In the classical approach the recommendation level to vaccinate is level 7 for

all ages in the interval [9.5,12.5]. On the other hand, in the fuzzy approach the function presents a peak at 8.5 months. In addition, it is important to note that the differences between mothers groups were emerged only under the fuzzy hypothesis approach.

12.1.4 Discussion

Due to the structure of the linguistic model, if the Bayes test results are the same for the all mothers groups, no differences will be detected by the classical model, as happened in the present data. Therefore, in this case only the fuzzy approach was able to discriminate between the recommendation levels to vaccinate in the mothers groups. This happens because the fuzzy hypothesis is based on an acceptance degree in addition to the condition for accepting H_0.

The fuzzy hypothesis presented some advantages when compared with its classical counterpart. Firstly, under the fuzzy approach it is not necessary to specify the cut off for θ in the null alternative. Furthermore, it is possible to add available expert knowledge to this procedure, taking into account if necessary a membership function to an alternative hypothesis other than the complementary of the null hypothesis. It is worth noting that the expert knowledge considered under the fuzzy hypotheses testing is completely different than those possible under a pure Bayesian approach. This possibility is introduced in the system through the definition of *a priori* probabilities (0.5 in this case). Finally, under the classical hypotheses testing there is no information about the magnitude of the test result (H_0 or H_1), as under the fuzzy approach. The results presented in this work show that this information was important in order to identify differences between the mothers groups.

Considering the maximum value for the recommended level to vaccinate as the optimal age to vaccinate it is found that, in all groups, under both models, fuzzy and classical, the optimal age is 8.5 months (see figure 12.6). Using the classical dynamical system proposed by Zanetta *et al.* this optimal age was 17 months to mothers 25 years old or more, and 14 months to mothers less than 25 years old. Therefore, the results presented in this work were completely different than those found using the classical dynamical model proposed in (Zanetta *et al*, 2002). This fact reveals the importance to discuss how the mathematical model structure could affect the results in modern epidemiological studies. The theoretical process of developing new analytical tools for dealing with real-world problems is still in its infancy. In this sense, this model could contribute to better understanding how these processes could perform under different scenarios when uncertainty and vagueness are present.

It is also important to consider the differences in the recommended vaccination age obtained by the fuzzy model and the classical dynamic system. Both the classical and the fuzzy model point to the fact that as timing passes, the susceptibility window bellow 1 year age is widening due to a lower antibody concentration of younger mothers. Therefore, as the fuzzy model recommends earlier age of vaccination, it is reasonable to suppose that its result is more reliable then classical ones (Ortega *et al.*, 2008b).

Finally, by merging Bayesian analysis and the theory of fuzzy sets, we illustrated the potential usefulness of this hybrid technique to address important public health and epidemiological problems. In the next section we will present another hybrid model applied in the epidemic context, but with a mathematical structure completely different.

12.2 A Fuzzy Model to Study a Predator-Prey Dynamics

In this section we will present the fuzzy rule-based system elaborated by Peixoto and collaborators (2008a), to study the interaction between aphids (preys) and ladybugs (predators) in citriculture, where the aphids are considered as the transmitter agents of the Citrus Sudden Death (CSD). As discussed previously, the great innovation of this work is the combination of the linguistic fuzzy model and the numerical calculus, aiming at the development of a dynamical system in an indirect way. The dynamics of the predator-prey interaction was built from a linguistic model based on the expert knowledge and simulations were performed in order to found the prey population, the potentiality of the predators, and the phase-plane (that is, the graph of potentiality of the predators *versus* preys number, for a fixed time). From the information presented in this phase-plane, a classic model was fitted and its parameters were found. In this sense, the classical dynamical system able to describe the phenomenon was found through the hybrid technique that will be shown here.

Citrus Sudden Death (CSD) is a disease that affects sweet orange trees grafted on Rangpur lime in the south of the state of Minas Gerais and in the north of the state of São Paulo, Brazil (Bassanezi *et al.*, 2003). It is believed that this disease is caused by a virus transmitted by insects known as aphids (vector). Among the most known predators of aphids in citrus in Brazil, the ladybugs are one of the most outstanding (Morales & Buranr Jr., 1985). In this sense, ladybugs have played an important role in the biological control of the aphids, and consequently of the CSC, in the Brazilian citriculture, what clearly has economical interest (Peixoto *et al.*, 2008a and 2008b).

Thereby, the proposal is to study the interaction between a prey (aphid) and its predator (ladybug) through fuzzy set theory, instead of using the usual differential equations, which characterize the classical deterministic models. Since there are no sufficient information about that phenomena, it is difficult to express the variations as functions of the states. In contrast, qualitative information from experts allows the elaboration of linguistic rules that relate, at least partially, the state variables with their own variations. The study of the variation of these states values in time allows to find, from fitting techniques, what is the classical deterministic model, given by a system of ordinary differential equations, whose solutions coincide with those provided by the fuzzy model. Thus, this approach allows the estimation of the parameters of the differential equations from fuzzy model, which consists in a great advantage because such parameters have biological meaning, like mortality rates, growth rates, and so on.

12.2.1 The Predator-Prey Model

Mathematical models that describe prey and predator relationship are used to study interactions between two populations, when one of them depends on the other for food and for survival. Such dynamic relationship between preys and predators are prominent subjects in Ecology (Edelstein-Keshet, 1987; Murray, 1990).

In short, it is presented below the hypotheses that characterize a predator-prey model, whose trajectories show the following features:

1. The number of prey population and the number of predator population have an oscillatory character;
2. An increase in the prey population is followed (with a delay) by an increase in the predator population;
3. A decrease in the prey population is followed (with a delay) by a decrease in the predator population;
4. If the number of predators is small, the number of preys increases;
5. If the number of predators is large, the number of preys decreases;
6. If the number of preys is large, the number of predators increases; and
7. If the number of preys is small, the number of predators decreases.

So this dynamics is characterized by chained oscillations in both populations: predators and preys. What is most interesting is that these oscillations have the following property: the peak of the prey population will always occur some time before the predator population peak.

According to the information above, it is possible to elaborate a fuzzy rule base that replaces differential equations, which characterize the classic deterministic models that are used to model the dynamics between preys and predators. In fact, the main interest here is to elaborate a predator-prey model that represents the interaction between aphids (preys) and ladybugs (predators) in citriculture. Then, it is presented a short review of the Citrus Sudden Death and the interaction between aphids and ladybugs.

12.2.2 Aphids *versus* Ladybugs

Citrus Sudden Death (CSD) is a disease that has caused serious harm to citriculturists, up to the point of destroying big plantations in the north of the state of São Paulo and in the south of Minas Gerais, in Brazil. CSD is a disease combining canopy/rootstock and it can lead plants on intolerant rootstock to death. Researches have shown that the ducts, which lead nutrients generated by the photosynthesis to the roots, become obstructed and degenerated. Without food, the roots putrefy, the tree decays and dies (Bassanezi *et al.*, 2003).

As discussed before, it is believed that this disease has been caused by a virus transmitted by insects known as aphids and the most important predator of aphids in citrus in Brazil is the ladybug *Cycloneda sanguinea*, that belongs to the *Coleoptera* Order and the *Coccinelidae* Family. Therefore, the constant occurrence of larvae and adults of ladybugs is important to control the aphids

(Hodek, 1973). Thus, the main interest here is to take into account the particularities of the data and reports of experts, since the quality of the predators is of most importance.

12.2.3 Formulation of the Predator-Prey Fuzzy Model

In predator-prey systems, the structure of the predator population changes in time and there are phases where there is no predation at all. As it have been observed, the ladybugs only prey upon aphids in the larva and adult stages. The aphids are captured by their enemies independently of their life phase. Therefore, incorporating this new characteristic into the model will help to understand predator-prey interactions. Moreover, it can be applied to the predation theory for biological control. Facts like these are not considered in the simple models of predation, for example, the classical Lotka-Volterra model. According to Hsin and Yang (2003), simple models are not adequate to study the predator-prey relationship, when the populations involved present different dynamics according to their ages. Each of these predators' larva can consume up to 200 aphids a day, and the adult predators prey, on average, upon 20 aphids a day (Gravena, 2003). Hence the population of predators will consist of larvae and adults. So we should distinguish these subpopulations and their particularities in the predator-prey model.

From the information above, it was consider that predators are differentiated in accordance with their potential of predation, through a membership function of predator class as follows:

$$P_{y_i} = \begin{cases} 1, & if \ \ larvae \\ 0.1, & if \ \ adults \end{cases} \tag{12.22}$$

and the potential of predation of a predators population as being $P_y = p_1 + 0.1p_2$, where p_1 is the number of larvae population and p_2 is the population of adults.

As the aphids are captured by their enemies independently of their life phase, the population of preys is not subdivided, since the quality of being a prey does not depend on its time of life to be classified according to its readiness to escape from their predators.

The input variables of the system are the number of preys and the potentiality of the predators, and the output variable is their variations. However, accurate knowledge about the input variables and their variations is not available. On the other hand, qualitative information from experts, in particular by entomologists, allows the elaboration of rules that relate the variables of state with their own variations. In this case, the fuzzy rule base was given by 30 rules of the type:

IF *the number of preys* is *large* AND *the potential of predation* is *very small*, THEN *the variation of preys increases a little* AND *the variation of the potential of predation increases a lot.*

From the Mamdani inference method and defuzzification by the center of area method, it was obtained the variation rates of the preys and the potential of

predation. In each moment t, the number of preys and the potential of predation are given by the expressions:

$$\begin{cases} x(t) = x(t_0) + \int_{t_0}^{t} x'(s)ds \\ P_y(t) = P_y(t_0) + \int_{t_0}^{t} P'_y(s)ds \end{cases} . \tag{12.23}$$

In order to observe the variation in the number of preys and the potential of predation the numerical simulations were performed from Numerical Integral Methodology (Conte & Boor, 1981). To achieve this, it was considered an initial number of aphids, x_0, and an initial number of potential of predation, P_{y0}, in a branch of a tree, chosen randomly. From the initial conditions, the fuzzy system produces x' and P'_y as outputs. From these two last values, x and P_y is found in each iteration by means of:

$$\begin{cases} x(t_{i+1}) = x(t_i) + \int_{t_i}^{t_{i+1}} x'(s)ds \\ P_y(t_{i+1}) = P_y(t_i) + \int_{t_i}^{t_{i+1}} P'_y(s)ds \end{cases} , \tag{12.24}$$

which is the discrete form of the continuous system (12.23).

Finally, to solve the integral above it was adopted the Trapezoidal Numerical Integration, since the fuzzy system provides x' and P'_y in each iteration t_i. Thus the system (12.24) turns to:

$$\begin{cases} x(t_{i+1}) = x(t_i) + \frac{1}{2}\left[x'(t_{i+1}) + x'(t_i)\right] \\ P_y(t_{i+1}) = P_y(t_i) + \frac{1}{2}\left[P'_y(t_{i+1}) + P'_y(t_i)\right] \end{cases} . \tag{12.25}$$

Using (12.25) and considering $t_i = t_0 + i$ and $t_0 = 0$, it is possible to get the values of x and P_y, and so on, successively.

Summarizing, the simulations of the trajectories produced by the fuzzy model follow the steps below:

- Given the initial population of the preys (x_0) and the initial potential of predation (P_{y0}) as inputs data of the fuzzy rule-based system;
- The fuzzy rule-based system gives the values of the output data: x'_1 e P'_{y1};
- From (12.25), it is found x_1 e P_{y1};
- x_1 and P_{y1} are the new inputs variables of the fuzzy rule-based system in the next step simulation, and so forth.

The evolution of the population of preys and potential of predation given by (12.25) through the fuzzy model over time, as well as, its respective phase-plane, are illustrated in figure 12.7.

It is important to highlight that even without any equations, this approach allows to obtain a phase-plane where the trajectories appear to converge to a

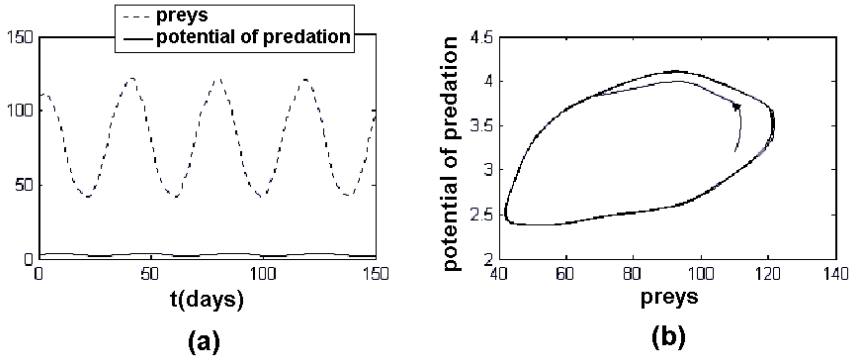

Fig. 12.7. (a) The evolution of the populations in time; and (b) Phase-plane of the fuzzy model, with $x_0 = 110$ and $P_{y0} = 3.2$ (Peixoto *et al.*, 2008a)

cycle with certain regularity. The question now is "is there a system of classical differential equations able to provide a similar phase-plane solution to that obtained by the fuzzy approach?"

12.2.4 Fitting the Holling-Tanner Model

Next it is proposed a classic deterministic model, given by a system of ordinary differential equations, assuming a predator-prey system, whose solutions coincide with those of the fuzzy model described above (Peixoto *et al*, 2008a). In this way, it is assumed that, in the classic model, there is no heterogeneity in the class of predators. Therefore, it is possible to find parameters of the new model, using the phase-plane of the fuzzy model illustrated in figure 12.7. The goal is to compare the predator-prey fuzzy model with the Holling-Tanner Model. In order to achieve this purpose, the parameters of the system given by differential equations were fitted from the fuzzy model. The following predator-prey system of the Holling-Tanner type was considered (Holling, 1959; Tanner, 1975):

$$\begin{cases} \dfrac{dx}{dt} = rx\left(1 - \dfrac{x}{K}\right) - \dfrac{mxy}{D+x} \\ \dfrac{dy}{dt} = sy\left(1 - h\dfrac{y}{x}\right) \\ x(0) > 0,\ y(0) > 0, \end{cases} \tag{12.26}$$

where $x(t)$ and $y(t)$ denote prey and predator densities, respectively, as functions of time, and $r, m, s, h, D, K > 0$.

In system (12.26) it was assumed that:

- The prey population grows logistically with carrying capacity K and intrinsic growth rate r in the absence of predation;

- The predator consumes the prey, according to the functional response $p(x) = \dfrac{mx}{D+x}$ and grows logistically with intrinsic growth rate s and carrying capacity proportional to the population size of prey. In (12.26), the functional response $p(x)$ is classified into type II (Svirezhev & Logofet, 1983);
- The parameter h is the number of prey required to support one predator at equilibrium when y equals x/h;
- m is the per capita maximum predator consumption rate, that is, the maximum number of preys that can be captured by a predator in each time unit; and
- h is a measure of the food quality that the prey provides for conversion into predator births.

This choice is justifiable, because:

- The prey population grows logistically and, in a branch of an orange tree it attains its carrying capacity in the absence of predation;
- The population of ladybugs, in the branch of an orange tree attains its carrying capacity proportional to the population of preys; and
- According to Morales and Buranr Jr. (1985), the number of aphids captured per day by *Cycloneda sangunea* (adults, male and female), corresponds to the functional response of Holling's Type II.

Some parameters may be obtained:

- Since the adult ladybug consumes, on average, 20 aphids a day, then $D = 10$;
- A population of 200 aphids per branch is considered large, that is why it was considered $K = 200$ as the carrying capacity of the prey population;
- Considering that an adult aphid generates up to 5 new nymphs a day, it shall be taken $r = 2$; and
- If the ladybug population duplicates in 1.03 weeks and only the females reproduce, it was assumed that $s = 0.3$.

The other parameters, m and h, were fitted according to the data, (x, y), generated by the fuzzy model, and substituting into equations (12.26) it was found the following system of equations:

$$\begin{cases} \dfrac{dx}{dt} = 2x\left(1 - \dfrac{x}{200}\right) - \dfrac{30.625xy}{10+x} \\ \dfrac{dy}{dt} = 0.3y\left(1 - 22.142857\dfrac{y}{x}\right) \\ x(0) > 0,\ y(0) > 0, \end{cases} \tag{12.27}$$

where x is the number of preys and y is the number of predators.

Figure 12.8 presents an example of fitting for the two models. We can note in figure 12.8b the limit cycle commonly expected in the classical predator-prey dynamical system. We can also observe that both two initial conditions tend to this limit cycle: $x_0 = 110$ for prey, and $y_0 = 3.2$ for predator, which is a start

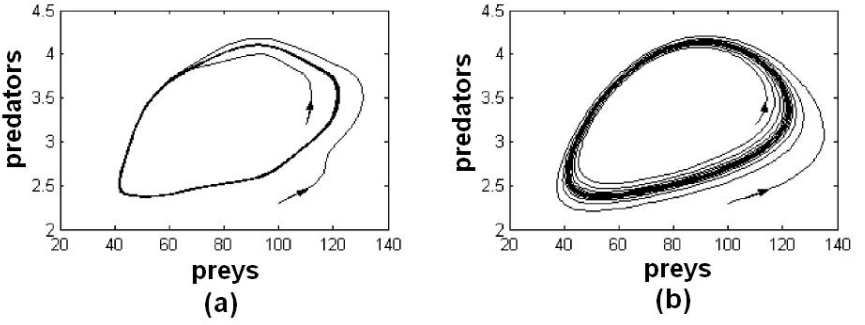

Fig. 12.8. (a) Phase-plane of the fuzzy model and (b) phase-plane of the deterministic model; considering two initial conditions: $x_0 = 110$ for prey, $y_0 = 3.2$ for predator and $x_0 = 100$ for prey, $y_0 = 2.3$ for predator (Peixoto *et al.*, 2008a)

point inside the limit cycle; and $x_0 = 100$ for prey, and $y_0 = 2.3$ for predator, which is a start point outside the limit cycle. The same behavior was found from fuzzy model (see figure 12.8a).

Note that it is possible to fit the curves in order to find suitable parameters for deterministic models. This is achieved by using the phase-plane curve from the fuzzy model. The great advantage of obtaining parameters for differential equations given by (12.26) is the fact that it allows a stability analysis of the system.

System (12.27) was analyzed in order to find its critical points, that is, the pair of values x and y that turns the derivatives null and kept the system in equilibrium, without changing the values of x and y. One can note that the system above has a possible pair: $(77.5, 3.5)$.

It is important to study what happens when the initial populations x_0 and y_0 are very near the critical populations, that is, (x, y) is near $(77.5, 3.5)$.

The system (12.26) is the same one to which Tanner (1975) wrote the stability analysis. Following Tanner's procedure, we may consider isoclines equations of the prey and predator populations. The peak of the prey critical line of (12.26) is at $(K - D)/2$. It results in:

1. If K is small, then the critical point to the right of the peak of the prey critical line is a stable focus for all values of s/r;
2. If K is large, the critical point is left to the peak of the prey critical line, and s/r is larger than a boundary value determined by rh/m, then the critical point is an stable focus;
3. If K is large, the critical point is left to the peak of the prey critical line, and s/r is less than a boundary value determined by rh/m, then the critical point is focus of a limit cycle; and
4. If K is infinite and s/r is smaller than rh/m, then the critical point is an unstable focus.

To analyze the stability of the critical point $(77.5, 3.5)$, it was used the result above. The isocline equations are given by:

$$\begin{cases} y = \dfrac{2}{30.625}(10+x)\left(1-\dfrac{x}{200}\right) \\ y = \dfrac{x}{22.142857} \end{cases} \tag{12.28}$$

isoclines of the prey and predation population, respectively, referring to the system of equations (12.27).

The maximum of prey isocline is:

$$\frac{K-D}{2} = 95 > x^* = 77.5, \tag{12.29}$$

and thus, the critical point is on the right side of the maximum. Yet,

$$\frac{rh}{m} = 1.446 > \frac{s}{r} = 0.15. \tag{12.30}$$

From (12.29) e (12.30), the critical point $(77.5, 3.5)$ is the focus of a cycle.

12.2.5 Discussion

In this section a hybrid fuzzy logic approach, proposed by Peixoto *et al* (2008a), was applied to Ecology/Epidemiology. Primarily, they have been able to model the *ladybug-aphids* dynamics without using explicit differential equations. They used only intuitive hypotheses of the predator-prey interaction and data from experts. This work illustrates that the fuzzy sets theory can contribute in an important way to the construction of mathematical models, mainly when some parameters of the differential equations are not available.

Without doubt, the great advantage of obtaining the parameters of differential equations lies in the fact that it allows the stability analysis of the system can be carry out.

Finally, we would like to highlight some advantages of using fuzzy rule-based models as opposed to deterministic models, in the particular scenario of dynamical systems presented:

- Several differential equations parameters of the predator-prey type systems are not available;
- In the fuzzy model, a rule base was used instead of systems given by equations, eliminating the difficulty of obtaining the parameters. In addition, these parameters can be obtained, if needed, through curve fitting procedure from the solutions obtained by the fuzzy rule-based models; and
- The input and output sets of the fuzzy rule-based systems can be easily constructed with the help of experts in the field, that is, a expert will know when the population of a particular species is small, large, and so forth.

The fuzzy dynamical methodology described in this section has been widely applied by Barros and collaborators to describe the dynamic of diseases with direct transmission (Barros *et al.*, 2007).

13 ...and Beyond: Fuzzy Logic in Medical Diagnosis

The purpose of this chapters is to provide a review and commentary on the current state of fuzzy logic applications in medical diagnosis. A symposium on fuzzy diagnostic and therapeutic decision support, organized by Adlassnig (2000) may be considered a watershed in the application of fuzzy logic in medical problems. For a deeper discussion on fuzzy logic in medicine see Szczepaniak *et al.* (2000).

Doctors have always been fascinated by diagnosis and the means by which it can be reached (see figure13.1), but the purpose of studying diagnostic logic has simply been to improve thought processes (Macartney, 1987). More recently, however, a second purpose is becoming more important: the design of expert systems and computer modeling able to perform medical diagnosis. In addition, the paradigm shift represented by the emergence of Evidence Based Medicine, the conscientious, explicit and judicious use of current best evidence in making decision about the care of individual patients (Sackett *et al.*, 1997) is unearthing new problems related to the logic behind diagnosis.

Medical diagnosis has been defined as "the crucial process that labels patients and classifies their illnesses, that identifies their likely prognosis, and that defines the best treatment available" (Sackett *et al.*, 1991). It is, actually, a complex process characterized by uncertainty in many stages (Bellamy, 1997).

The act of clinical diagnosis is, therefore, a process of classification, that is, an effort to recognize the class to which a patient's illness belongs (Sackett *et al.*, 1991). In a broader context, the clinical practice should be focused on the five clinical objectives, described by Sackett *et al.* (1997):

1. achieving a diagnosis;
2. estimating a prognosis;
3. deciding on the best therapy;
4. determining harm (related to item 3); and
5. providing care of the best quality.

Therefore, we may think of diagnosis proceedings from symptoms and signs (and laboratory tests) to focus on the documentation of maladaptive alterations in structure, function, and/or response to stimuli (Sackett et al, 1991).

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 277–310, 2008.
springerlink.com

Fig. 13.1. Doctor Rene Laennec ausculting a patient (from Porter, 1996)

Alternatively, diagnosis can proceed from symptoms and signs (and laboratory tests) to focus on prognosis. Finally, diagnosis may focus on a therapeutic trial of identifying the target disorder on the basis of its response to specific therapy (Sackett *et al.*, 1991).

13.1 The Diagnostic Process

Several attempts have been made to identify the possible cognitive pathways that lead to diagnosis. In this chapter we focus on Sackett *el al.* (1991) description of the four strategies of clinical diagnosis.

The first strategy is called *pattern recognition*, and is based on *gestalt* methods. It is defined as the "instantaneous realization that the patient's presentation conforms to a previously learned pattern of disease" (Sackett *et al.*, 1991). It is usually sensorial and reflexive. Doctors do it but cannot explain to others why or how they do it. This strategy is applied by experienced clinicians and is often described as intuitive. Normally, the diagnosis is performed at first sight (or any other sensorial input from the patient) of the patient. An example of such an approach is the diagnosis of parkinsonism in which the doctor labels the patient rather quickly just by watching her gait or by hearing his/her speech.

The second strategy is called *multiple-branching* or *arborization* strategy of diagnosis. It is defined as "the progression of the diagnostic process down but

one of a large number of potential, presents paths by a method in which the response to each diagnostic inquiry automatically determines the next inquiry to be carried out and, ultimately, the correct diagnosis" (Sackett *et al.*, 1991). It is very logical and it is based on algorithms and flow charts and it is particularly useful when trials, not treatment, is the objective. This method is also very useful when diagnosis is delegated from physicians to nurses or paramedical staff.

The third strategy is called *strategy of exhaustion*. It is defined as the "painstaking, invariant search for all medical facts about the patient, followed by sifting through the data for the diagnosis" (Sackett *et al.*, 1991). This method is very time consuming and it is performed in two stages: first, the collection of all the potentially pertinent data, and second, the searching through it for the diagnosis. The strategy of exhaustion is the method of the novice and it is abandoned with experience.

The fourth strategy is called the *hypothetic-deductive* strategy and it is the one used by virtually all doctors, virtually all the time. It is the "formulation, from the earliest clues about the patient, of a *short list* of potential diagnosis or actions, followed by the performance of those clinical and complementary maneuvers that will best reduce the length of the list" (Sackett *et al.*, 1991). The hypothetic-deductive approach has been considered the most appropriate diagnostic process in the sense that it is time saving and it has the greatest accuracy.

In summary, diagnostic approaches can usefully be described as one or a combination of four types: the pattern recognition approach of the seasoned clinical, the multiple-branching method of the delegate, the exhaustion method of the novice, and the most widely used strategy, the hypothetic-deductive approach (Sackett *et al.*, 1991).

13.2 Computer Models and Expert Systems for Medical Diagnosis

As mentioned above, the logical analysis of the diagnostic process makes the medical diagnosis amenable to be mimicked (and if possible, surpassed) by computer systems. This sort of approach is based on the assumption that diagnosis is a highly desirable end and that doctors want a system that will do better than the best clinician.

The development of computer models of medical diagnosis as applied to the construction of expert systems is part of *Artificial Intelligence*. Before we describe the expert systems already developed for medical diagnosis, however, let us first discuss the mathematical foundations of the available computer models.

13.2.1 Mathematical Models in Medical Diagnosis

The vast capacity of modern computers to process and store data and to carry out, almost instantaneously, complex logical manipulations has encouraged the description of the diagnostic process in mathematical terms. This formalization

of the diagnostic logic, in turn, has allowed the development of computer programs that can aid physicians in their attempts to solve their patients' problems accurately and safely through the application of ever-expanding medical knowledge (Miller *et al.*, 1981).

The mathematical and statistical techniques already applied in classical expert systems include decision trees and other logical schematics, likelihood ratio, Bayes'theorem, discriminant analysis, and cluster analysis. In addition, subsidiary mathematical techniques has been developed for helping machine-learning systems, like the backpropagation algorithm of connectionist networks. Some of the above techniques are aimed at the classificatory aspect of diagnosis, like discriminant analysis and cluster analysis. In this chapter we deal only with those mathematical and statistical aspects of *uncertainty*.

In what follows we, therefore, briefly describe the five main "classical" techniques for dealing with uncertainty, namely, *Bayesian reasoning*, *Bayesian belief networks*, the *Dempster-Schaffer theory of evidence*, the *Stanford certainty factor algebra*, and *Causal networks*.

Bayesian reasoning

Using probability, we can determine, often from *a priori* argument, the chances of events occurring. In knowledge-based problem solving, like in medical diagnosis, we often find ourselves reasoning with limited knowledge and incomplete information. Several techniques have been designed to deal with such a limited knowledge and information and, before we start the discussion of fuzzy reasoning we describe some of the techniques of probabilistic reasoning.

Bayesian reasoning is based in formal probability theory and is used extensively in several current areas of research, including pattern recognition and classification, both of paramount importance in diagnostic applications.

The Bayesian approach is based on *prior probabilities*, the unconditioned probability assigned to an event in the absence of knowledge supporting its occurrence or absence, and *posterior probability*, the condition probability of an event given some evidence. The usual notation for prior probability is $p(event)$ and for posterior probability is $p\,(event|evidence)$. So, for instance, the prior probability of a person having a disease is the number of people with the disease divided by the number of people in the domain of concern. The posterior probability of a person having a disease d with symptom s is given by:

$$p\,(d|s) = \frac{|d \cap s|}{|s|}, \tag{13.1}$$

where the bars brackets meaning the number of elements in that set. Therefore, the posterior probability given by equation (13.1) is the number of people having both (intersection) the disease d and symptom s divided by the total number of people having the symptom s. Equation (13.1) can also be written:

$$p\,(d|s) = \frac{p(d) \times p(s|d)}{p(s)} \tag{13.2}$$

also known as *Bayes equation* or *Bayes theorem*.

In clinical terms, the prior probability $p(d)$ is called the *prevalence* of the disease, and the posterior $p(s|d)$ the *sensitivity* of the diagnostic test that revealed s (by the way, s may be a symptom, a sign, a laboratory test or any other diagnostic information). In real situations, however, rarely a final diagnostic is reached by a single symptom (sign, and/or laboratory test). We, therefore, need a generalized version of equation (13.2), a form of Bayes with multiple symptoms (signs and/or laboratory tests):

$$p\left(d|s_1\&s_2\&...\&s_n\right) = \frac{p(d) \times p\left(s_1\&s_2\&...\&s_n|d\right)}{p(s_1\&s_2\&...\&s_n)}. \tag{13.3}$$

Now, for m diseases and n symptoms (signs and/or laboratory tests) there will be about ($m \times n^2$ conditional probabilities) plus (n^2 symptom probabilities) plus (m disease probabilities), or about $m \times n^2 + n^2 + m$ pieces of information to collect. In a realistic medical system with 100 diseases and 1000 symptoms (signs and/or laboratory tests), this value is $(100 \times 1000^2 + 1000^2 + 100) = 1.01 \times 10^8$, that is, over 100 millions! This simple calculation illustrates the difficulties involving such an approach.

Bayesian belief networks

Bayesian belief networks relax several constraints of the full Bayesian approach, basing on three assumptions. The first is that the modularity of the problem domain may allow us to relax many of the dependence/independence constraints required for Bayes approach. The second assumption is that the links between the nodes of the belief network are represented by conditional probabilities. Thus for nodes A and B of the network, the link between A and B, denoted by $A \rightarrow B(c)$, reflects the evidences $A's$ support for the belief in B with some confidence c, sometimes called *causal influence measure* (Luger & Stubblefield, 1998). The third assumption is that coherent pattern of reasoning may be reflected as paths through cause/symptoms relationships. Causes can influence the likelihood of their symptoms and the presence of a symptom can affect the likelihood of all its possible causes. To create a belief network we must make a clear distinction between these two kinds of potential influence, and then select the path our reasoning will take through the network.

In a very simple case, let us consider a possible serial relationship of A on B and B on C (Luger & Stubblefield, 1998):

$$A \longleftrightarrow B \longrightarrow C \tag{13.4}$$

with $A \rightarrow B(c_1)$ and $B \rightarrow C(c_2)$, where c_1 and c_2 are the causal influence measures. If there is no evidence supporting B then A and C are called *d-separated* and independent. If there is evidence of B then C cannot support A, although evidence for B can support A.

As mentioned by Luger and Stubblefield (1998), Bayesian belief networks seem to reflect how humans reason in complex domain where some factors are known and related *a priori*.

A clinical example of such a complex relationship would be the infections of the urinary tract in diabetic patients. It is known that diabetes increases the susceptibility to urinary infections and these may by themself cause a worsening in the diabetic condition by causing a disequilibrium of the clinical state. So diabetes, A, is correlated with urinary infections, B, which causes a set of symptoms, C, unrelated to the diabetic condition.

The Dempster-Schafer theory of evidence

Uncertainty often results from a combination of missing evidence, the inherent limitations of heuristic rules and the limitation of our own knowledge. The Dempster-Schafer theory of evidence considers sets of propositions and assigns to each of them an interval [*belief*, *plausibility*] within which the degree of belief for each proposition must lie. This *belief* measure, denoted bl, ranges from zero (no evidence) to one (certainty). Its complement is called *plausibility*, and is denoted pl. So, for a proposition a we have:

$$pl(a) = 1 - bl(\neg a), \tag{13.5}$$

where $\neg a$ means *not a*. Plausibility also ranges between zero and one and reflects how evidence of not a, $\neg a$, relates the possibility for belief in a. Dempster-Schafer address the problem of measuring certainty by asking for a fundamental distinction between lack of certainty and ignorance. Belief functions allow us to use our knowledge to bound the assignment of probabilities to events in the absence of exact probabilities. The Dempster-Schafer theory is based on the idea of obtaining degrees of belief for one question from subjective probabilities for related questions and the use of a rule for combining the degrees of belief when they are based on independent items of evidence.

Let us suppose that we have a diagnosis domain H containing some diagnostic hypothesis that a patient has tuberculosis (T), pneumonia (P), or common cold (C), that is, $H = \{T, P, C\}$. We have to associate measures of beliefs with the hypotheses set within the domain H. Evidence need not support individuals hypothesis exclusively. So, for instance, the presence of fever would support our three hypotheses simultaneously. On the other hand, evidence in favor of some hypothesis may affect belief in others.

The next step is to define a *probability density function*, d, for all subsets of the set H, where $d(h_i)$ represents the belief that is currently assigned to each h_i of H (where in this case $\sum d(h_i) = 1$). If H has n elements then there are 2^n subsets of H. Since many of the subsets will never occur, it is possible to deal with the remaining subsets. The plausibility of H is:

$$pl(H) = 1 - \sum d(h_i), \tag{13.6}$$

where the h_i are the sets of hypotheses that have some supporting belief. Whenever we start a diagnosis, it is often the case that we have no information about any hypothesis, then $pl(H) = 1$.

In the our diagnostic example the set H is composed by 3 elements, resulting in 8 subsets found by the combinations of those elements. Suppose our first evidence is that the patient has cough, and that this supports $\{T, P\}$ at 0.8. If this is our only hypothesis, then $d_1\{T, P\} = 0.8$ and $d_1(Q_1) = 0.2$ to account for the remaining distribution of belief, that is, all other possible beliefs across H (and not our belief in the complement of $\{T, P\}$). Therefore, Q_1 is the subset composed by all possible subsets of H except the set $\{T, P\}$. We next proceed by amplifying our investigation space and have now that the patient also has headache, which has the support level of $\{P, C\}$ to 0.6, and so we have $d_2\{P, C\} = 0.6$ and $d_2(Q_2) = 0.4$, where Q_2 is the subset found by all possible subsets of H expect the set $\{P, C\}$. These two beliefs may now be combined by the Dempster's rule (Luger & Stubblefield, 1998) in order to find another belief measure, d_3:

$$d_3(Z) = \frac{\sum_{X \cap Y = Z} d_1(X) d_2(Y)}{1 - \sum_{X \cap Y = \emptyset} d_1(X) d_2(Y)}. \tag{13.7}$$

where the belief in the hypothesis Z, or $d_3(Z)$, is the sum of the products of the hypothetical situations $d_1(X)$ and $d_2(Y)$, whose co-occurrence supports Z, that is, $X \cap Y = Z$. Since there are situations in that $X \cap Y = \emptyset$, so the sum of the confidences must be normalized by one minus the sum of these values (Luger & Stubblefield, 1998).

So that, considering the two evidences of our example, d_1 and d_2, and computing all possible ways of intersecting X and Y hypotheses, we can find table 13.1.

Table 13.1. Possible hypotheses and their belief values through the Dempster's rule (13.7) for the diagnostic example

Belief measure on evidence d_1	**Belief measure on evidence** d_2	**Belief measure on evidence** d_3
$d_1\{T, P\} = 0.8$	$d_2\{P, C\} = 0.6$	$d_3\{P\} = 0.48$
$d_1\{Q_1\} = 0.2$	$d_2\{P, C\} = 0.6$	$d_3\{P, C\} = 0.12$
$d_1\{T, P\} = 0.8$	$d_2\{Q_2\} = 0.4$	$d_3\{T, P\} = 0.32$
$d_1\{Q_1\} = 0.2$	$d_2\{Q_2\} = 0.4$	$d_3\{Q_3\} = 0.08$

As there are no sets $X \cap Y$ that are empty, the denominator of equation (13.7) is 1. So, in this case, the sum of all d_3 belief values in the third column of table 13.1 is equal 1. We may, therefore, assign a belief of 0.48 that the patient in our example has pneumonia.

The great advantage of this Dempster-Shafer approach is that if another evidence is acquired, d_4, the set of new possible hypothesis and their belief values can be computed through the combination of the *probability density functions* d_3 and d_4, by the application of the Dempter's rule over them. The

Dempster-Shafer approach is a very useful tool when the stronger Bayesian conclusions may not be justified (Luger & Stubblefield, 1998).

13.2.2 The Stanford Certainty Factor Algebra

Unlike Bayesian approaches, which attempt to measure the probability with which evidence supports a conclusion, certainty theory (Buchanan & Shortliffe, 1984) attempts to measure the confidence merited by a given heuristic. This heuristic approach is based on estimates of the confidence we are justified in having experts conclusions. Those estimates are weighted with other heuristics derived from the experts experiences and comprise terms like "very probable", "almost certainly" or "possible". Certainty theory is an effort to formalize this heuristic approach to reasoning with uncertainty.

The Stanford certainty theory creates confidence measures and some simple rules for combing these confidences. It splits "confidence for" from "confidence against". These two measures constrain each other in the sense that a given piece of evidence is either for or against a particular hypothesis.

The confidence measures of the Stanford certainty factor tradition are a human subjective estimate of symptom/cause probability measure. Although it is defined in a formal algebra, the meaning of the certainty measures is not as rigorously founded as is formal probability theory. Its measures are *ad hoc* in the same sense that a human expert's confidence in his/her results is approximate, heuristic, and informal. The most important application of this theory on medical diagnosis is the famous MYCIN program described below.

Causal Networks

In causal models relationships are depicted as links between nodes in a graph or a network. This approach consists in mapping of observations onto a network of nodes and the linking the nodes in a causally coherent pattern.

A complete causal pathway from a start node to a terminal node represents a complete disease process. Confirmation of a state is derived either from associated observations, or indirectly through the causal link to another state for which there is some evidence.

The construction of the causal network consists in five operators (Luger & Stubblefield, 1998): *aggregation*, *elaboration*, *decomposition*, *summation* and *projection*. This approach has been applied in several diagnostic problems like glaucoma (Weiss *et al.*, 1978a and 1978b) and acid base and electrolyte imbalances (Patil *et al.*, 1981).

13.2.3 Expert Systems for Medical Diagnosis

Expert systems applied to medical diagnosis are part of the so called *clinical decision-support systems* (Shortliffe, 2001), which can be defined as any computer program designed to health professionals make decisions. They can be

divided basically into three categories: *tools for information management*, *tools for focusing attention* and *tools for patient-specific consultation*. We will stick with the latter.

The first computer softwares aiming medical diagnosis were the de Dombal's system for diagnosis of abdominal pain (de Dombal *et al.*, 1972) and Shortliffe's system for selection of antibiotic therapy (Shortliffe, 1976).

De Dombal's system is based on Bayesian reasoning and it is so successful that it is still in use in the United Kingdom. Using surgical or pathologic diagnosis as the gold standard, the group leaded by de Dombal have emphasized the importance of deriving conditional probabilities used in Bayesian reasoning from high-quality data that they have gathered by collecting information on thousands of patients (Shortliffe, 2001). The system applied the sensitivity, specificity, and prevalence data for various signs, symptoms, and lab tests to calculate the conditional probabilities of seven possible diagnostics of abdominal pain, namely, *appendicitis*, *diverticulitis*, *perforated ulcer*, *cholecystitis*, *small bowel obstruction*, *pancreatitis*, and *nonspecific abdominal pain*. In one famous test the software performed diagnosis of abdominal pain in a sample of 304 patients, scoring 91.8% of accuracy, against 65% to 80% of the clinicians' diagnosis (de Dombal *et al.*, 1972). In addition, in six of the seven disease categories, the software was more likely to assign patients to the correct disease category than the senior clinician in charge of the case.

Some years after the development of de Dombal's system, the Stanford group, leaded by Shortliffe presented MYCIN (Shortliffe, 1976), a consultation system that applies the above mentioned heuristic certainty factor algebra. The system was designed to explain the advice it offered, to justify its performance using simple English sentences, to learn new information through interactions with experts, to encode knowledge in a modular format, and to have prompts, answers, and volunteered information that matched its users' needs. The developers evaluated MYCIN's performance on therapy selection for patients with bloodborne bacterial infections, and for those with meningitis. MYCIN is a landmark on developing medical aid systems, and is best viewed as an early exploration of methods for capturing and applying ill-structured expert knowledge to solve important medical problems.

Most decision-support systems have assumed a passive role in giving advice to clinicians. In contrast, some systems have been developed that play a more active role and do not wait for physicians specifically to ask for assistance (Shortliffe, 2001). This kind of active systems are, however, more turned to other medical applications than diagnosis. The latter are normally classified into two basic styles of interaction (Shortliffe, 2001): the *consulting model* or the *critiquing model*. In the first model, the system serves as an advisor, accepting patient-specific data, asking questions, and generating advise for the user about diagnosis or management. Examples are the MYCIN, the general medical diagnostic program DXplain (Barnett *et al.*, 1987), and Internist-1/QMR (Miller *et al.*, 1986). Critiquing model systems assume a preconceived notion of a diagnosis. The program then acts as an advisor for the physician, expressing agreement or

suggesting reasoned alternatives. Examples of this kind of model is ATTENDING, a program designed for anesthesia (Miller, 1983), and ONCOCIN, a system for clinical oncology (Shortliffe, 1986).

The softwares CASNET and ABEL apply causal networks, described above, and are examples of systems that deal very well with uncertainty in medical reasoning. ABEL, for instance, represented the state-of-the-art in clinical reasoning for its time and still remains unsurpassed in its hierarchical integration of causal reasoning across multiple level of details (Luger & Stubblefield, 1998).

Another interesting class of computer models applied to medical diagnosis is the so called *connectionist* systems, also known as *parallel distributed processing* systems. They hold that intelligence arises in systems of simple, interacting components through a process of learning or adaptation by which the connections between the components are adjusted. The process is distributed across layers of artificial neurons (these systems are also called *neural networks*). Neural networks mimic the human brain's methods of problem solving by constructing artificial "neurons" in the process of analyzing the data. Data can be entered in the form of numbers, patterns, or even sounds and images. Neurons are created by the program and connected in a straightforward, one-to-one manner as well as in complex arrays. The user is rarely aware of the number and relative importance (weighting) of these units. During analysis of a problem, a neural network can spontaneously grow (add new neurons) and re-weight itself.

In normal operation a user provides a data set (called inputs), and usually a single output, to "train" the network. The neural network "digest" the data, reweights the inputs, and add layers of neurons as necessary to accurately predict the output. This prediction is then compared with the known output of an actual experiment. In older programs of this type, a user could then readjust weightings and algorithms to make predicted results move closer to observed results. Newer systems, however, do that on automatic mode, although a software can be supplied to allow more user control.

Problem solving is parallel in the sense that all the neurons within the layers process their inputs simultaneously and independently. These systems have the enormous advantage of being able to learn when properly trained by one of a set of algorithms. On the other hand, they work like black boxes in the sense that when they err it is almost impossible to discover why they erred. In medical diagnosis, connectionist systems are very suitable for classification (deciding the category to which an input value belongs) and pattern recognition (identifying structure in sometimes noisy data).

Connectionist models for medical diagnosis have been rather popular, both as a subject of research and as application tools. The medical publication database *MEDLINE* presents 820 papers on neural networks and diagnosis published within the last two years (2006-2007). Illustrative samples are the articles by Falk *et al.* (1998) who use neural networks as an aid in the determination of disease status; by Rudzki *et al.* (1997), who published a paper on focal liver disease, using a neural network-aided diagnosis based on clinical and laboratory data; the paper by Tourassi *et al.* (1998), who present a cost-effectiveness

analysis of the effect of artificial neural networks on patient care of acute pulmonary embolism; the application of artificial neural networks by Shiomi *et al.* (1997) on the diagnosis of chronic liver disease; by Levin *et al.* (1996), who published a paper on neural network differentiation of optic neuritis and anterior ischemic optic neuropathy; by Venta *et al.* (1998), who demonstrates how diagnosis of breast implant rupture can be improved with sonographic findings and artificial neural networks; by Pesonen *et al.* (1998), who evaluate the diagnosis of acute appendicitis with neural networks; by Reategui *et al.* (1997), who combined neural network with case-based reasoning in a diagnostic system; by Holst *et al.* (1998), who applied an intelligent computer system reporting lack of confidence, a proposal of a confidence measure for decision support systems; by Downs *et al.* (1996), who present an application of a neural network model to medical pattern classification tasks; and more recently, by Mueller *et al.* (2004), who developed an expert system to predict the extubation outcome in preterm newborns and compared the neural networks approach with clinical expertise and statistical modeling; by Kahya and collaborators (2006), who classified respiratory sounds with different feature sets using neural network structures; by Marcos *et al.* (2007), who applied neural network classifiers in the diagnosis of the obstructive sleep apnea syndrome; in the same way, Emoto *et al.* (2007), developed a neural network system to extract the features of snore soundjust, which is the earliest and the most common symptom of obstructive sleep apnea; just to mention a few. A good review on neural networks can be found in Rocha (1992).

13.3 Fuzzy Diagnostic Systems

In the previous sections we showed how mathematical and computer models can aid to deal with uncertainty and lack of complete information. In this section we present fuzzy logic tools which, in addition of their great capacity to dealing with uncertainty and incomplete information, are also able to deal with the vagueness related with medical diagnosis.

The process of classifying different sets of symptoms, signs and laboratory tests under a single name is getting increasingly difficult. Several factors are contributing to this fact like the enormous amount of medical information available for clinicians and, most important, the great variety of uncertainties, vagueness and ambiguities involved in the diagnostic process. Therefore, alternative methods are desperately needed for the design of diagnostic systems. Fuzzy logic is one of the best current candidate for this role.

We have already mentioned in previous sections that the medical diagnosis process can be viewed as a process by which a clinician assigns a label to a patient. Therefore, at the heart of the diagnostic process is the idea of categorization. However, the terms diagnosis, symptoms and signs, are applied in a very loose way to refer to the labels and information available. Except in a few cases, the line dividing health and disease is irreducibly fuzzy. Furthermore, symptoms are often subjectively described by the patient and the signs collected by the

physicians, except when she/he has a measurement instrument like a thermometer or a sphygmomanometer, which measure crisp variables, are very frequently surrounded by uncertainties and ambiguities. In addition, the descriptions of symptoms from the patients and of disease entities by the physicians often use linguistic terms that are irreducibly vague (see discussions on chapters 1 and 5). Even complementary tests, frequently considered as crisp parametrization of medical practice, are, in some cases, extremely subjective, either in their interpretation (what is the meaning of a glycemia of 110 mg/dl, a normal test or an indication of diabetes ?) or in its very presentation (can you imagine anything more fuzzy than an ultrasonography image?).

Let us begin by the assumption that the aim of the diagnostic process is to assign a label to patients who, in addition to their complaints (often presented in a fuzzy way), present a cluster of clinical signs and laboratory (complementary) tests alterations. The clinical purpose then, is to differentiate a "normal" from a "sick" individual. As briefly discussed in chapter 2, a classical view is to consider a crisp divide between a healthy and a non-healthy individuals. So, we need to start by defining what we understand by "health". The World Health Organization defines health as *the complete absence of physical, mental or social wellbeing*. Have you ever heard a fuzziest definition?

Even when an objective, crisp laboratory test or any other auxiliary/complementary diagnostic test is available, its results' interpretation is often open to discussion, that is, vague. One way or another, the test report will wind up calling some results "normal" and others "abnormal". Sacket *et al.* (1997) recognize six definitions of "normal" in common use, listed in table 13.2.

In the *Gaussian* definition it is assumed a normal distribution and that all "abornomalities" have same frequency; the *Percentile* definition has the same basic defect as the *Gaussian* definition; the *Culturally desirable* definition confuses the role of medicine; the *Risk factor* definition labels the outliers, which may not be helped; the *Diagnostic* definition is the focus of this discussion; and the *Therapeutic* definition means that you have to keep up with advances in therapy.

The authors propose the use of definition 5 and comment that the others are practically useless. Even definition 5, however, is irreducibly vague and leave it clear that no sharp boundary between "normal" and "abnormal" results is possible.

In a classical, standard approach, the diagnostic process should be, ideally a classificatory process able to determine the crisp divide between healthy and non-healthy individual. Let us imagine, like Bellamy (1997), that a reliable measure of health is available and that a normal threshold, below which an individual is classified as non-healthy exists. A good clinician should then be the one able to classify individuals below (diseased) or above (healthy) such a threshold. Figure 2.12 in chapter 2 illustrates this situation in terms of the classical sets theory.

As mentioned in chapter 2, the most fundamental aspect of fuzzy set theory is the idea of graded membership. Classifying individuals as healthy and not

Table 13.2. Six definitions of "normal" (from Sacket *et al.*, 1997)

Number	Name	Definition
1	**Gaussian**	the mean +/− 2 standard deviations
2	**Percentile**	within the range, say, 5 - 95%
3	**Culturally desirable**	preferred by society
4	**Risk factor**	carrying no additional risk of disease
5	**Diagnostic**	range of results beyond which disease is probable
6	**Therapeutic**	range of results beyond which treatment does more good than harm

healthy by using fuzzy sets offers several advantages over the use of crisp sets. The two sets "healthy" and "not healthy" can be represented by triangular fuzzy sets, as illustrated by figure 2.13 in chapter 2. So, if we consider a measure of health, x, in the fuzzy sets approach an individual can presents some degree of membership to the set of healthy individuals, $healthy(x)$, and another degree of membership to the set of non-healthy individuals, $non - healthy(x)$. Furthermore, the membership degrees $healthy(x)$ and $non - healthy(x)$ do not need to sum up 1, as usually is required in the probability theory. Indeed, considering the diagnostic scenario described above, it is much more functional!

The fuzzy logic approach of modeling the diagnostic process is useful in all the diagnostic approaches described in section 13.1: the *pattern recognition* approach; the *multiple-branching* method; the *exhaustion* method; and the *hypothetic-deductive* approach. The pattern recognition approach, due to its intrinsic subjectiveness is probably the most prone to be aided by fuzzy logic modeling. However, all the other approaches have several uncertainties and vagueness associated with the classificatory process of diagnosis.

13.3.1 Fuzzy Relations Diagnostic Models

The fuzzy logic framework has been utilized in several different approaches to modeling the diagnostic process (Klir & Yuan, 1995). The most simple fuzzy structure applied in a diagnostic system is the *fuzzy relations*, defined in full detail in section 2.5, in chapter 2. Some examples of fuzzy relations applied in diagnostic process can be found in Sanchez (1979), and more recently in the works of Reis *et al.* (2004 and 2005) and Lopes *et al.* (2006).

Let us now see an example of the application of the type of linguistic model as a fuzzy relation between *exposure levels*, e, and *disease severity*, d. The example provided in this section is a simplified version of the system CADIAG-2, originally described in a set of seminal (and now historical) articles by Adlassnig and Kolars (1982), Adlassnig *et al.* (1984), Adlassnig *et al.* (1985), Kolarz and Adlassnig (1986), and Adlassnig (1986) system CADIAG-2. The improvement of this system, the CADIAG-II/RHEUMA, is presented by Leitich *et al.* (2000), and consists in a semiautomatic knowledge acquisition system for rheumatic diseases and relates symptoms with specific diagnosis.

This simplified version of CADIAG system here presented is due to Klir and Yuan (1995). As mentioned above, we need to assign grades of membership values to the linguistic quantifiers, which describe the fuzzy relationship between exposure and disease, like proposed in table 13.3:

Table 13.3. Membership grades values to the linguistic quantifiers (Klir & Yuan, 1995)

Fuzzy quantifier	**Grade of membership** (μ)
always	1.00
often	0.75
unspecified	0.50
seldom	0.25
never	0.00
very	μ^2

The knowledge about the occurrence of exposure to a certain environmental factor and a given disease may be described in such a relational model as a string of statements of the kind:

- Exposure level e_1 *very seldom* causes disease severity d_1;
- Exposure level e_1 *often* causes disease severity d_2;
- Exposure level e_2 *always* causes disease severity d_1;
- Exposure level e_3 *very often* causes disease severity d_2; and
- Exposure level e_3 *seldom* causes disease severity d_1.

The causality relation, R_c is then given by the matrix:

$$R_c = \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \end{array} \begin{array}{c} \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{bmatrix} 0.06 & 0.75 \\ 1.00 & 0.00 \\ 0.25 & 0.56 \end{bmatrix} \end{array}.$$

Now, suppose three distinct populations, p_1, p_2 and p_3, subject to the above three exposure levels, e_1, e_2 and e_3. A fuzzy relation, R_e, specifying the degree of exposure for those three populations is given by the matrix:

$$R_e = \begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \end{array} \begin{array}{c} \begin{array}{ccc} e_1 & e_2 & e_2 \end{array} \\ \begin{bmatrix} 0.4 & 0.8 & 0.7 \\ 0.6 & 0.9 & 0.0 \\ 0.9 & 0.0 & 1.0 \end{bmatrix} \end{array}.$$

The fuzzy compositional rule of inference for this situation is given by:

$$R = R_e \circ R_c, \tag{13.8}$$

whose grade of membership function is given by the $max - min$ composition of fuzzy relations (see chapter 2):

$$\mu_B(d) = \max_{e \in E}[\min(\mu_A(e), \mu_R(d)], \tag{13.9}$$

which provides:

$$R = \begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \end{array} \begin{array}{c} \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{bmatrix} 0.80 & 0.56 \\ 0.90 & 0.60 \\ 0.25 & 0.75 \end{bmatrix} \end{array}.$$

Therefore, population p_1 has possibility of 0.8 of developing disease severity d_1, and so for.

CADIAG-2 system incorporates relations between symptoms and diseases and also between diseases themselves, between symptoms themselves, and between combinations of symptoms and diseases. It demonstrated an accuracy of 94.5% in achieving correct diagnosis in rheumatological diseases (Adlassnig, 1986).

13.3.2 Fuzzy Cluster Analysis Models

Another example provided by Klir and Yuan (1995) is a set of models which utilize a technique proposed by Fordon and Bezdek (1979) and Esogbue and Elder (1979, 1980 and 1983). Models of diagnosis using fuzzy cluster analysis examine the similarity of the presence and severity of symptoms patterns, which can be designated with degrees of memberships in fuzzy sets representing each symptom category.

The patient is clustered to varying degrees with the prototypical patients whose symptoms are most similar. The most likely diagnostic candidates are those disease clusters in which the patient's degree of membership is greatest. The specific patient x presents itself displaying a set of symptoms s_i (again by symptoms we mean in addition to symptoms, signs and laboratory tests) at levels of severity given by a fuzzy set A_x, with $A_x(s_i)$ denoting the grade of membership of fuzzy set characterizing the patient and defined on the set of symptoms S, which indicates the severity level of each symptom presented by the patient.

Each of the likely diseases is described by a matrix B_l giving the upper and lower bounds of the normal range of severity of each of the symptoms that should be expected in a patient with the disease. We further define a fuzzy relation R on the set of symptoms and diseases that specifies the pertinence or importance of each symptom s_i in the diagnosis of the matrix of each likely disease d_j.

The clustering is performed by computing a similarity measure between the patient's symptoms and those typical of each disease d_j. The example provided

by Klir and Yuan (1995) uses a distance measure based on the Minkowski distance given by:

$$\begin{aligned} D_p(d_j, x) \quad = \quad & \left[\sum_{i \in I_l} |R(s_i, d_j)(B_{jl}(s_i) - A_x(s_i))|^p \right. \\ & \left. + \sum_{i \in I_u} |R(s_i, d_j)(B_{ju}(s_i) - A_x(s_i))|^p \right]^{1/p}, \end{aligned} \tag{13.10}$$

where $R(s_i, d_j)$ is the fuzzy relation between the symptom s_i and the disease d_j; $B_{jl}(s_i)$ is the matrix that describe the upper and lower bounds of the normal range of severity of the symptom s_i related to the disease j; $A_x(s_i)$ is the membership degree of the symptom s_i in the fuzzy set severity level related to the patient x, A_x; and the sets I_l and I_u is given by:

$$\begin{aligned} I_l &= \{i \in N_m | A_x(s_i) < B_{jl}(s_i)\} \\ I_u &= \{i \in N_m | A_x(s_i) < B_{ju}(s_i)\}, \end{aligned} \tag{13.11}$$

in which m denotes the total number of symptoms. The most likely disease candidate is the one for which the similarity measure attains the minimum value.

13.3.3 Smets' Model for Fuzzy Diagnosis

An interesting model was that proposed by Smets (1981) and analyzed with detail in Bandemer and Gottwald (1995). Smets proposed modeling the relationships between diseases and clinical findings (symptoms, signs and/or lab tests) by specifying *credibility degrees*. This author applied Shafer's (1976) concept, according to which it is possible to construct fuzzy measures starting from incomplete specifications. So, for a finite universe X, a random variable $x \in X$, and a set $P(X)$ it is possible to specify a mapping $p : P(X) \rightarrow [0, 1]$, which is called *basic probability assignment* (Bandemer & Gottwald, 1995).

In addition, due to the normalization to 1, we have for a subset B of $P(X)$:

$$\sum_{B \in P(X)} p(B) = 1. \tag{13.12}$$

So, according to Shafer's degree of credibility (or belief) we have for B:

$$Cr(B) = \sum_{A \subseteq p(B)} p(A), \tag{13.13}$$

in which $Cr(B)$ represents the degree of confidence concentrated in B from the events $A's$ that support B.

The Smets' model assumes that a clinical finding, x is a random variables in the clinical findings universe, $x \in X$, and d is a disease in the diseases universe, $d \in D$. The model assumes focal sets and two basic probability assignment: $p_X(\cdot, d)$, which is the probability of the disease d in relation to the clinical findings universe; and $p_D(\cdot, x)$, which is the probability of the clinical finding x in relation to the diseases universe. So from the equation (13.13) we can write:

$$Cr_X(A, d) = \sum_{A_h \subseteq A} p_X(A_h, d), \tag{13.14}$$

where $Cr_X(A,d)$ is the *credibility degree* for $A \in P(X)$, if the disease $d \in D$ is present, where $P(X)$ is a power set of X. The sets $A_h's$ are the focal sets, in the Shafer sense, in which $A_h \in \sup[p_X(\cdot,d)]$, i.e., the evidences that support A.

Assuming that the prior credibility degree over the situations expresses total ignorance, that is $p_D(D) = 1$ and $p_D(B) = 0$ for all $B \in P(D)$ with $D \neq B$, then the posterior credibility degree for $B \in P(D)$, if a clinical finding x from $A \in P(X)$ was observed, is (Smets, 1993; Bandemer & Gottwald, 1995):

$$Cr_{D|X:0}(B,A) = \frac{\sum_{d \in B^c} Cr_X\,(A^c;d) - a}{1-a}, \tag{13.15}$$

where

$$a = \prod_{d \in D} Cr_X\,(A^c;d), \tag{13.16}$$

where the index $\{D|X:0\}$ indicates that total ignorance with respect to D was assumed.

When an informative prior credibility degree Cr_D, different from total ignorance is assumed, this degree can be connected with $Cr_{D|X:0}$ by the Dempster's rule of combination. This connection, represented by $p_1 \cap p_2$, brings together two basic probability assignments over the same power set, where different direct or indirect assignment, called *conflicts*, for certain subsets are reconciled (Bandemer & Gottwald, 1995). When this conflicts are very hard, the application of Dempster's rule is problematic.

The operation product $p_1 \cdot p_2$ over a set A is commutative and associative and its basic probability assignments assume values in [0,1]. However, the sum over all sets A in the power set can be less than 1, since the probability of the empty set can be positive, i.e., $(p_1 \cdot p_2)(\emptyset) > 0$. In this case we say that there are conflicts between the assignments p_1 and p_2. If the case of total conflict $(p_1 \cdot p_2)(\emptyset) = 1$ is excluded, then we can renormalize $p_1 \cdot p_2$ and generate a conflict reconciling basic probability assignment for all $A \neq \emptyset$, similarly as been done in the Dempster's rule combination (Dempster, 1967; Bandemer & Gottwald, 1995), by:

$$(p_1 \cap p_2)(A) = \frac{(p_1 \cdot p_2)(A)}{1-(p_1 \cdot p_2)(\emptyset)}. \tag{13.17}$$

From this rule, it is possible to obtain the posterior probability assignment for the informative case by:

$$p_{D|X}(B,A) = \sum \frac{p_{D|X:0}(G;A)p_D(C)}{1-\sum_{G \cap C=\emptyset} p_{D|X:0}(G;A)p_D(C)}, \tag{13.18}$$

where p_D is the basic probability assignment for Cr_D and $p_{D|X:0}$ is the assignment for $Cr_{D|X:0}$ given by (13.15). In addition, the summation in (13.18) is conditioned to the following bonds: $G \in \sup[p_{D|X:0}]$, $C \in \sup[p_D]$ and $C \cap B$. The summation term in the denominator of the equation (13.18) can be interpreted as an expression of the inconsistency between the prior and the "ignorant" posterior credibility degree. Although the credibility degree was deduced for the

finite universe, it is also possible to deduce it for infinite universe sets (Smets, 1981).

Finally, the credibility of a *fuzzy* diagnosis B with the membership function μ_B over D, if $x \in A$ is observed, can be obtained through:

$$Cr_{D|X}(B|A) = E_*(\mu_B|A), \tag{13.19}$$

where E_* is the expected value of the conditional $(\mu_B|A)$ (for more detail see Smets, 1981 and 1993).

13.3.4 Bellamy's State-Space Approach

One of the most interesting article on fuzzy medical diagnostic systems is by Bellamy (1997), who proposes modeling diagnosis as a mapping between subsets of property measurements and subsets of diagnostic categories, described within multidimensional diagnostic spaces. This multidimensional space, combined with a fuzzy sets representation of the variables and states of the patient, leads to a fuzzy model of the diagnostic process.

As a matter of fact, the state-space approach for modeling the diagnostic process has already been proposed in a probabilistic setting by Miller and collaborators (1981). This model utilizes an m-dimensional symptom space in which each of the m axes represents a different clinical or laboratory finding. Within this m-dimensional space, the findings of patients with established diagnosis fall into definable, but frequently overlapping cluster. In figure 13.2 we show a two-dimensional space, illustrating this concept.

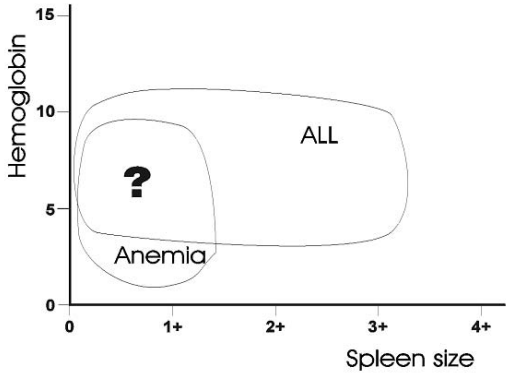

Fig. 13.2. Two-dimensional state-space illustration

When plotter against each other, the hemoglobin concentration and spleen size of patients with diagnosed acute lymphblastic leukemia (ALL) fall into the ALL area. The findings of patients with iron deficiency anemia (FeDef) fall in the Anemia area. If an undiagnosed patient (x) has hemoglobin of 6 gm/dl and a

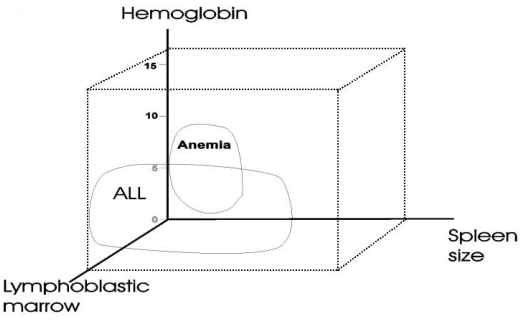

Fig. 13.3. Presence or absence of lymphoblasts in the bone marrow

enlarged spleen of 1+, findings that fall in the region of overlap, the differential diagnosis consists of those diseases having symptom spaces overlapping at point x, in this case ALL and FeDef. A probability distribution function over these possible disease alternatives can be estimated by:

$$p(ALL) = \frac{number\ of\ ALL\ with\ findings\ atx}{total\ number\ of\ patients\ with\ findings\ at\ x} \tag{13.20}$$

and

$$p(FeDef) = \frac{number\ of\ FeDef\ with\ findings\ at\ x}{total\ number\ of\ patients\ with\ findings\ at\ x}. \tag{13.21}$$

If the p value of the most probable diagnosis is too low, new tests are needed. Each new test adds an additional dimension to the diagnostic hyperspace. If the new test has high specificity, its results clarify the diagnostic problem. In figure 13.3, the presence or absence of lymphoblasts in the bone marrow has been added as an additional axis in the diagnostic space. Since the presence of lymphoblasts is highly specific for ALL, mutually exclusive diagnostic planes free of overlap are formed. Patient's x marrow is full of lymphoblasts. Therefore, this patient falls on the ALL diagnostic phase and the diagnosis becomes certain (Miller *et al.*, 1981).

Bellamy's model also assumes a multidimensional symptom space and plots the specific patient as a point in such a space. In addition, he adds a dynamic, representing how the system changes in time. As time passes, the point in the multidimensional space moves around and the trajectory of the point describes the behavior of the system. In figure 13.4 we show a five dimension diagnostic space with patient x plotted as a point in this space:

The behavior of a patient over time is similarly modeled as a trajectory in a multidimensional space, and the prognosis for a patient can be estimated from the direction and rate of change of the trajectory (see figure 13.5). For example, if the trajectory of many individuals patients with tuberculosis are followed during the course of their recovery to the "healthy region", a *prognostic corridor* can be defined that covers the range of trajectories that indicate a good prognosis.

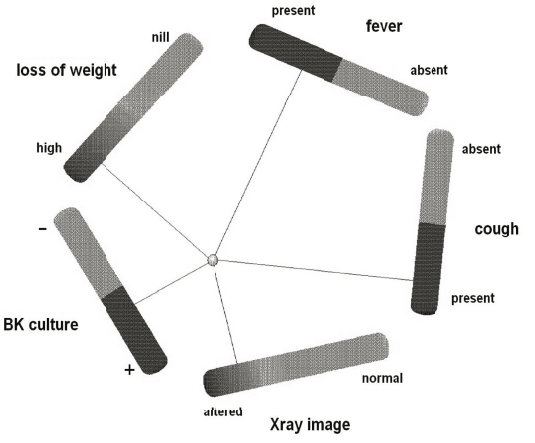

Fig. 13.4. Illustrations of Bellamy's model with five dimension diagnostic space

In his fuzzy diagnostic system Bellamy represents each dimension (input variable) of the diagnostic space as a series of overlapping fuzzy subsets, like normal, high, low, etc. (see figure 13.5):

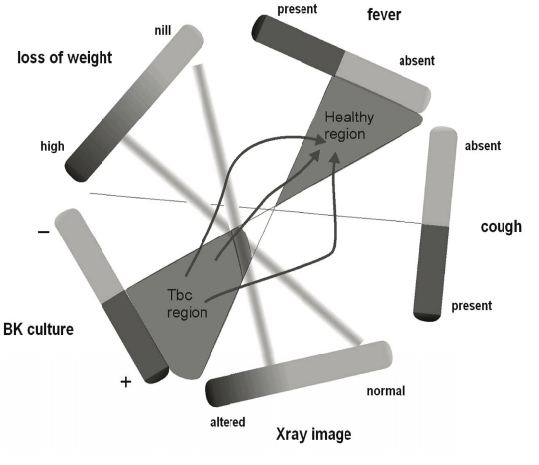

Fig. 13.5. Illustration of Bellamy' space with behavior of patients over time trajectories

Certain combinations of fuzzy subsets correspond to particular diagnostic categories. For example, if the blood neutrophils counting is high, fever is moderate, and low-back pain is intense, then the regions overlap in the pyelonephritis region, and so on. In addition, each bounded region representing a specific diagnostic category is equivalent to one of the fuzzy linguistic rules used to relate

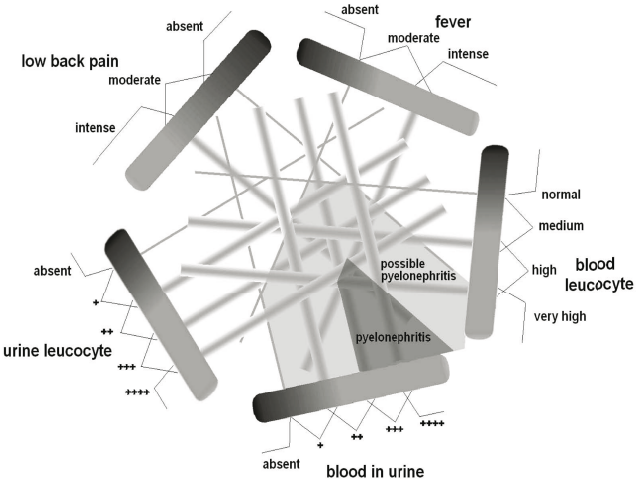

Fig. 13.6. Illustration of Bellamy' space for the diagnostic region for pyelonephritis

inputs to outputs, that is, the diagnostic region for pyelonephritis is equivalent to the fuzzy rule described in that region of figure 13.6.

13.3.5 Other Fuzzy Diagnostic Systems

As mentioned above, a substantial number of articles dealing with fuzzy diagnostic systems is available in the specialized literature. In the previous sections we presented some of the most interesting diagnosis models considering their mathematical structures to deal with uncertainties. In this section we complement this review on fuzzy medical diagnosis by briefly refer to some specific aspects of the fuzzy approach to diagnosis systems.

In Bartolin *et al.* (1982), the theory of fuzzy logic is approached as a diagnostic aid. A few theoretical considerations are followed by practical applications on the diagnosis of hyperlipoproteinemias and on the classification of the four most prevalent anemias encountered in internal medicine. The main conclusion of the authors is that the various parameters involved in the analysis should be grouped on a hierarchical basis.

Esogbue and Elder (1983) describe a study in which fuzzy diagnostic models were computerized, validated and compared with a physician hypothesis as well as existing mathematical models. The authors present a critic of classical mathematical models for medical diagnosis, which are known to perform very poorly when compared to diagnosis made by real doctors. Among the factors which contribute to this poor performance they include the omission by these models of important information on the patient such as symptoms of past undiagnosed diseases which can only be vaguely recalled by the patient. In addition, classical models fail to model the stage of development of the disease, and other intrinsically fuzzy aspects of the information necessary for a medical hypothesis.

The fuzzy models presented by these authors are exemplified by applications concerning valvular heart disease. They conclude that the fuzzy logic approach is not only practical but results in models of greater validity than those based on classical set theoretic approaches.

Fuzzy numbers were applied in a computerized electrocardiography system by Degani and Bortolan (1987). Those fuzzy numbers arise in the automatic processing of electrocardiographic signals. The authors present a computerized system for the diagnostic classification of the standard 12-lead electrocardiogram and the results from this work exemplifies the usefulness of the fuzzy set approach to electrocardiographic diagnosis. Recently, the fuzzy numbers and their arithmetic were deeply discussed and widely applied, to the understanding how the brain computes and how the cognitive process is supported, by Rocha *et al.* (2004).

A diagnostic method using fuzzy discrimination and connectivity analysis is presented by Norris *et al.* (1987). This method described by these authors constructs a numerical tabular knowledge base from historical cases, and derives inferences from particular case histories using discrimination and connectivity analysis which are based on a theory of fuzzy relations. The authors claim that the method can handle incomplete information, partial inconsistency and fuzzy descriptions of data in a natural way. The discriminating analysis ranks medical symptoms for their relative ability to distinguish between a well defined set of diseases. The connectivity analysis, in turn, establishes which sets of symptoms are representative of each of the diseases. The systems was tested against senior clinicians and performed favorably in diagnosing acute abdomen.

A route-choosing medical diagnostic technique is presented by Anderson *et al.*, (1987), in an article describing the central component of an expert system for medical diagnosis. The method consists of an inference technique with special reference to the use of fuzzy logic, a route-choosing heuristic method to reduce the cost of reaching a diagnosis, and the tree-constructing of the domain which follows clinicians' division into syndromes.

Kuncheva (1990) proposed a fuzzy patterns recognition model to handle problems with non-crisp and multi-class membership of the objects. The model was oriented to medical diagnostics, where patients suffer from more than one disease in different degrees. The author designed a multi-level fuzzy decision scheme in order to derive high performance, taking into account expert logic and human experience. The paper discuss two main topics, namely the criterion for evaluation of classification accuracy and the training rule. In addition, the implementation of the fuzzy multi-level classifier is illustrated with real clinical data, showing a good diagnostic accuracy.

Another quite interesting approach to fuzzy modeling is due to Torasso (1991), who designed a diagnostic expert system with heuristic learning capability. The role of the supervisor is analyzed and a set of strategies is defined which allows the system to implement different policies. The organization of the system has been strongly influenced by the results obtained so far in investigating the properties of neural nets and human learning. The learning system is able to revise

the knowledge bases used by the consultation system by taking into account the experience gained in solving cases as well as the confirmation (disconfirmation) of diagnosis provided by the external world. In particular the learning system revises the membership functions between findings and diagnostic hypotheses and the membership relations defined among diagnostic hypotheses. In this article the author describes the behavior of the learning system in the conservative approach and discusses some alternative solutions for memory organization.

The article by Marín and Mira (1991) discusses the role of knowledge in the problem of classification and presents a knowledge-oriented fuzzy classification system suitable for use in fields in which classification criteria, though numerically imprecise, can be formulated in natural language, and in which it is important to retain the expert's conceptual descriptions. This knowledge-oriented fuzzy classifier extends previous fuzzy nearest neighbor techniques in that it generalizes the concept of a design set, in order to allow both reference sets and their labeling to be defined in fuzzy terms by an expert, expressing items of his knowledge in production role format. The article exemplify the application of the knowledge-oriented fuzzy classifier with a study in fetal medicine.

A very interesting, albeit somewhat odd paper was presented by Demling (1992) dealing with chaos theory, fractals and fuzzy logic. The author argues that chaos researchers are attempting, in a non-linear world, to understand mathematically a dynamic, apparently unordered system. In this connection, the fractal dimension also appears, which can be employed in the area of diagnosis to define tumor contours.

Diagnostic imaging is the subject of a paper by Stroke (1993), who demonstrated that the single most important elements in the current development of imagenology is the feasibility of diagnostic imaging that results from the use of mathematical methods, implemented with digital computers. The author mentions computerized tomography, magnetic resonance imaging and other moderns techniques that use three-dimensional image recording and reconstruction, as promising fields of investigation for the use of fuzzy logic. He also considers the role of fuzzy logic as an important component of the dramatic change from empirical to scientific technology in radiology and general medicine.

A fuzzy expert computer-assisted diagnosis system for osteoporosis is presented by Binaghi *et al.* (1993). The article shows how the diagnosis system can be employed to build a fuzzy medical expert system in the domain of postmenopausal osteoporosis. The aims of the expert system are to standardize knowledge and support physicians in the early detection of postmenopausal osteoporosis. A wide range of diagnostic situations has been considered for both categories of the disease, with judgments that range from disease is excluded to disease is definite. The salient aspects of the approach are the use of fuzzy logic as an analytic language for the representation and manipulation of knowledge and strategies and the integration of structured interview techniques and learning-by-example to address the knowledge acquisition task.

The diagnosis of iron deficiency by a fuzzy computerized system is the subject of a paper by Causer *et al.* (1994). The aim of this study was to develop an expert

system that could reproduce a pathologist's diagnosis of iron deficiency from the data obtained from blood test. The diagnostic system used a combination of fuzzy set and cut-off points from 14 parameters to arrive at one of 5 diagnostic categories graded from *iron deficient* to *no evidence of iron deficiency*. The authors found an overall agreement between pathologist and expert system of 71%.

The diagnosis of acute abdominal pain is modeled by a fuzzy expert system proposed by Fathi-Torbagham and Meyer (1994). The authors considered that knowledge in acute abdominal pain is characterized by uncertainty, imprecision and vagueness, and therefore, it is rather amenable to the application of fuzzy logic techniques. The representation and application of uncertain and imprecise knowledge is carried out by fuzzy sets and fuzzy relations. The hybrid concept of the system enables the integration of rule-based, heuristic and case-based reasoning on the basis of imprecise information. The central idea of the integration is to use case-based reasoning for normal cases. The heuristic principle is ideally suited for making uncertain, hypothetical inferences on the basis of fuzzy data and fuzzy relations.

Neonatal assessment by the Apgar scoring system using fuzzy expert system is proposed by Shimomura *et al.* (1994). Three Apgar fuzzy expert systems were determined separately by each one of three physicians groups (four inexperienced obstetricians, four experienced obstetricians and four expert neonatologists) in order to demonstrate that the fuzzy system reflected the examiner's expertise situation. Two-hundred and sixty-seven neonates were assessed 1 minute after birth by an experienced obstetrician using the classical Apgar scoring system and the three Apgar fuzzy expert systems. Statistical analysis showed that the fuzzy Apgar system determined by four expert neonatologists had the highest sensitivity and were significantly different ($p < 0.05$) from the classical Apgar scoring system. Recently, as discussed in chapter 2, Reis *et al.* (2004 and 2005) proposed an expert system to predict the risk of perinatal asphyxia and the needing of the resuscitation maneuvers, and in chapter 7 we presented an expert system proposed by Nascimento and Ortega (2002) to estimate the risk of neonatal death.

Fuzzy reasoning was applied by Shiomi *et al.* (1995) for diagnosing chronic liver disease. The method was applied to standardize diagnosis of liver disease based on scintigraphic results and compared the result of the fuzzy system with those obtained when scintiscan were scored conventionally. Fuzzy logic was used to evaluate five items: the ratio of the sizes of the left and right lobes, splenomegaly, radioactivity in the bone marrow, deformity of the liver and distribution of radioactivity in the liver. The degree of conformity to each of the three liver diseases being investigated was substituted into the membership function for the conclusion. Distinctions between chronic persistent hepatitis and chronic aggressive hepatitis were difficult to assess with fuzzy reasoning and conventional scoring. The diagnostic accuracy was 95% for patients with cirrhosis and 88% for patients with chronic hepatitis with fuzzy reasoning. With conventional scoring the accuracy was 86% for patients with cirrhosis and 75% for patients

with chronic hepatitis. The method was considered simple and could be used routinely in clinical settings.

An interesting study by Phelps and Hutson (1995) estimated diagnostic test accuracy using fuzzy gold standards. The study used Monte Carlo simulations methods to analyze the consequences of having a criterion standard (the so-called *gold standard*) that contains some error when analyzing the accuracy of a diagnostic test using receiver operating curves. The authors mention that when diagnostic test errors are statistically independent from inaccurate fuzzy gold standard errors, estimated test accuracy declines. Also, when the test and the fuzzy gold standard have statistically dependent errors, test accuracy can become overstated. The article proposes two methods to eliminate the first of those errors, exploring the risk of exacerbating the second, one of them, called *two-truth* method, selectively eliminates those cases where the gold standard is most ambiguous. Fuzzy receiver operating characteristic curves (ROC) are developed (Campbell *et al.*, 1991) and its application in medical studies were recently verified by Castanho *et al* (2007), who applied the fuzzy ROC curves to evaluate diagnosis tests for prostate cancer.

An automatically-generated differential diagnosis system based upon a patient's recent history was proposed by Cordova and Goldman (1995). In order to make this diagnosis clinically reliable, however, the system must be sensitive enough to discriminate between physiologic events that could share nearly identical trends. In order to refine their discriminatory technique, the authors created a software monitor that used fuzzy logic to analyze physiologic signals to make a clear distinction between an arrhythmic cardiac arrest and the onset of cardiopulmonary bypass. At the time of the study the authors expected that the differentiation capabilities of that monitor could form the foundation for a comprehensive automated diagnostic system.

Computer-assisted radiologic diagnosis system of rheumatologic diseases applying fuzzy logic is presented by Boegl *et al.* (1995). The authors' approach make use of pre-existed sources of information to build an expert system that minimizes the interaction between radiologists and the computer. Given data of a specific case, a deductive inference procedure combines the observed radiological signs, establishes confirmed and excluded diagnosis as well as diagnostic hypotheses, and provides explanations for these conclusions. In addition, proposal for confirmation or exclusion of diagnostic hypotheses are offered. The system was tested on radiological disorders of the hip joint related to rheumatological diseases, reaching a diagnostic accuracy of bout 80%.

Another radiologic system applying fuzzy logic was proposed by Phillips *et al.* (1995), who enhanced the information available from magnetic resonance imaging with a computer-assisted diagnostic system. Image pixels were classified into tissue classes based on feature vectors using unsupervised fuzzy clustering techniques as the patterns recognition method. Correlation of fuzzy segmentations and gross and histopathology were successfully performed. Based on the results of neuropathological correlation, the application of fuzzy magnetic resonance image segmentation to a patient with a brain tumor and extensive edema

represents a viable technique for automatically displaying clinically important tissue differentiation. With this pattern recognition technique, it was possible to generate automatic segmentation images that displayed diagnostically relevant neuroanatomical and neuropathological tissue contrast information from raw magnetic resonance data for use in three-dimensional volume reconstructions.

An expert system for diagnostic decision support system in psychiatry is proposed by Kovács and Juranovics (1995). The system uses the methods of fuzzy logic and backward chaining. The diagnostic course is biphased as we can differ symptoms and criteria (duration of the illness, ethological factors). The authors managed to extend the traditional applications using yes-no logic with three factors that make the system more sensitive and flexible: "scaling", "sorting by importance" and "reliability-validity" results. The diagnostic expert system is a shell that can be filled up optionally with psychiatric traditional diagnostic systems, like DSM-IV, ICD-X, or other diagnostic system.

Applications of fuzzy classification systems to electrodiagnosis of peripheral polyneuropathy is the object of an article by Duckstein *et al.* (1995). The method accounts for uncertainty or imprecision in experimental observations and both normal and pathology definitions are developed on the basis of a distance measure between fuzzy numbers. The distance measure, called normalized fuzzy pathology index evaluates the difference of distance between observed experimental values for a given patient and normal on the one hand and pathology on the other hand. This normalized fuzzy pathology index characterizes patient status as a continuous index and categories if values are defined, to conform to medical usage. Each of these categories corresponds to a linguistic variable. The application presented is the electrodiagnosis of peripheral polyneuropathy in diabetic patients. Four linguistic categories are defined by a doctor: normal state, borderline state, clear-cut, and severe pathology. The index is calculated in three cases that provide a sensitivity analysis on measurement of fuzziness and distance function weighting. The model was calibrated with 203 cases and validated with 291 different cases. The authors state that the results corresponds very closely to the physician's diagnosis.

A fuzzy logic diagnosis system for classification of pharyngeal dysphagia is proposed by Surynarayanan *et al.* (1995). The system's purpose was to develop a fuzzy logic classification of the patient into four categories of risk for aspiration. Acceleration and swallow pressure measurements were obtained and five parameters were extracted from these measurements. A set of membership functions were defined for each parameter. The measured parameter values were fuzzified and fed to a rule base which provided a set of output membership values corresponding to each of the categories. The set of output values were then defuzzified, in order to obtain a continuous measure of classification. The fuzzy system was evaluated using data obtained from 22 subjects. There was a complete agreement between the fuzzy system classification and the clinician's classification in 18 of the 22 patients (82% sensitivity).

The Glucose Tolerance Test is generally used for diagnosis of diabetes mellitus and is one of the most interesting example of the arbitrariness of crisp

thresholds between normal and diseased patients. Patients are offered a 75g glucose loading dose by oral intake and the value of blood glucose is measured some time afterward. Individuals with blood glucose of 201 mg/dk is diagnosed as diabetes mellitus and other whose value is 199 mg/dl are considered as impaired glucose tolerance. Arita *et al.* (1996) proposed an alternative approach in which they analyzed the dynamical response of glucose tolerance tests and proposed a new diagnostic system for diabetes using a fuzzy inference.

Truth-qualification and fuzzy relations in natural languages as applied to medical diagnosis is the subject of a paper by Sanchez (1996), one the pioneer in fuzzy logic and medical problems. In this articles Sanchez addresses the problem of given two fuzzy propositions, how to truth-qualify one of them to induce the other in a semantical equivalence. Two fuzzy subsets of the unit interval (τ_0 and τ_1, representing linguistic truth values) are introduced that provide best lower and upper approximations when no exact solution can be found: best semantic entailments of propositions are thus derived. The problem is reformulated in a new way, in terms of fuzzy relation equations, from which results are retrieved and extended. Also, a truth-possibility index, defined from τ_0 to τ_1, is introduced, that serves pattern-matching purposes, in addition to the usual possibility and necessity measures. The example provided, in which medical knowledge is expressed in a rule form, with fuzzy propositions in the antecedent, illustrates the aggregation of these measures, for medical diagnosis assistance.

Klein *et al.* (1996) proposed a patterns recognition system for focal liver lesions using crisp and fuzzy classifiers. Their aim was to determine the diagnostic performance of an artificial intelligence system for classification of focal liver lesions, in comparison to human observers. The pattern recognition was performed in two steps with initial extraction of textural features: training of a classifier and classification of the lesions. The system accuracy when compared with the classification achieved by human observers was 90.2%.

Fuzzy logic concepts were applied by Mir *et al.* (1996) to enhance diagnostic features of computed tomography images. The authors considered the vague nature (fuzziness) of functional characteristics in organ pathologies and argued that classical image enhancements techniques cannot adapt to the characteristics of that nature. The fuzzy method transforms the image of interest into a fuzzy plane using fuzzifiers which can changed to select a crossover point. At the early stages of a disease, when the contrast of the pathological tissues is very low, the visibility of the disease could be considerably improved using those techniques.

A framework for the design of diagnostic monitors called DIAMON-1 was presented by Steimann (1996). The system considers that the methods of artificial intelligence to clinical monitoring requires some kind of signal-to-symbol conversion as a prior set. Subsequent processing of the derived symbolic information must also be sensitive to history and development, as the failure to address temporal relationships between findings leads to poor results. The method proposed provided two methods for the interpretation of time-varying data, one for the detection of trends based on classes of courses, and one for the tracking of disease histories modeled through deterministic automata. Both methods use fuzzy set

theory taking into account of the elasticity of medical categories and allowing discrete disease models to mirror the patient's continuous progression through the stages of illness.

Fuzzy reasoning in an expert system for ultrasonography is the subject of the paper by Tanaka *et al.* (1997), who also evaluated the clinical utility of the proposed fuzzy system as a diagnostic aid for the unskilled clinician. The diagnostic system was designed to differentiate metastatic from inflammatory lymph nodes. Three fuzzy production rules were set up according to the diagnostic criteria for lymphadenopathy. The system was tested with clinicians who were one to three years after graduation and inexperienced in ultrasonography. The average increase in accuracy was 8.5% and the sensitivity and specificity 10.7% and 6.4% respectively, which were statistically significant. The authors concluded that the application of fuzzy reasoning in an expert system for ultrasonography improves the diagnostic performance of inexperienced clinicians.

Another group presented a computer-aided diagnosis system using fuzzy inference for breast ultrasonography in the same year (Koyama *et al.*, 1997). The ultrasonographic features of a breast mass were used as input data and included shape, border, halo, internal echoes, posterior echoes, and edge shadows. The probability of malignancy was described by an actual number ranging from 0 to 1. The fuzzy inference method demonstrated a sensitivity of 94.5% and specificity of 76% for cancer diagnosis.

An expert laboratory system using fuzzy sets and pattern recognition as its inference mechanism was proposed by Innis (1997). The program coupled the fuzzy inference mechanism with a data base comprised of hematological and biochemical responses to diseases collected over a period of 10 years in a teaching hospital. The author found that the system often presented diagnosis not thought of by the clinician and concluded that the computer, programmed to recognize a disease by the patterns of its response to routine hematological and biochemical investigations, could contribute significantly to diagnosis.

Electromyograms in erectile dysfunctions were interpreted by a system designed by Gorek and collaborators (1997). The system extracted signal patterns of higher activity, form stored data and described those patterns in mathematical terms of the features obtained from the electromyogram of the corpora cavernosa. Using fuzzy logic, the features were used to effect pattern evaluation. A correspondence of some diagnostic classes of 70% was found. In addition, the accuracy achieved in each of the individual classes was better than 50%. Finally, discrimination between normal and abnormal evaluation, which was of particular interest in the diagnostic test, reached 80%.

In a paper of 1998, Sanchez discusses the application of fuzzy logic to inflammatory protein variations. The model is of special interest in the processing of borderline cases, allowing a graded assignment of diagnosis to patients. Relationships between signs and diagnosis are interpreted as labels of fuzzy sets and it is shown how diagnosis can be derived from soft machine processing. When pattern matching is achieved, the final ranking of inflammatory syndromes assigned to a given patient might change to better fit the actual classification.

In 2000 Adlassnig published in a book a compilation of diagnostic systems based on fuzzy logic, from which we selected the ones presented below.

Shtovba and Chernovolik (2000) presented a system which provides support for a pathology anatomist in decision making about instant nontrauma death occurrence time determination. Their system is based on some linguistic expert rules formalized in the form of fuzzy knowledge bases.

Georgopoulos *et al.* (2003) discuss the use of fuzzy cognitive maps as a diagnostic model for specific language impairment, a disorder of spoken language ability where a variety of problems in many aspects of language exists. Since then, the application of the fuzzy cognitive maps in medicine is an approach in developing and several works are available (Papageorgiou *et al.*, 2003, 2006a, 2006b and 2008; Glies *et al.*, 2007; Georgopoulos & Stylios, 2008).

Some interesting examples of fuzzy image processing are presented in the paper by Axer *et al.* (2000). Fuzzy methods are used to analyze image from confocal laser scanning microscopy, polarized light microscopy, and magnetic resonance. Those fuzzy methods were applied from low level image to high level image processing, and included the use of linguistic variables. The authors argued that the use of linguistic variables may lead to a higher acceptance in medicine.

Multicriteria decision problems is the focus of the paper by Dujet (2000) in which two types of fuzzy methodologies are applied in medical decision aid: data fusion and theory of possibilities combined with mathematical morphology. Those methods are illustrated via a problem of classification of patients. The methods demonstrated to be accurate and are very generic and easily transferable in many other fields.

The difficulties in defining fuzzy membership functions for medical implementation is addressed by Straszecka (2000), who gives interesting suggestions for solving the problems with the use of the classical definition of a fuzzy set as well as by introducing the Dempster-Shaffer theory of evidence. In this paper, fuzzy sets are proposed for a similarity interpretation while a basic probability assignment is suggested for an estimation of a diagnosis quality.

As illustrated in figure 1.1 and exemplified in the review above, hundred of papers have been published applying fuzzy logic and fuzzy sets theory in several areas of medicine and, particularly, in the diagnosis systems. In fact, only in the last two years (2006-2007) the *MEDLINE* database presented 196 articles concerning to fuzzy logic and diagnosis systems. Of these we highlight the following systems: by Shieh *et al.* (2002 and 2007), who developed a pain model based on fuzzy logic to control analgesia in patients; by Ju and collaborators (2005), who to designed a robot system for assisting in the rehabilitation of patients with neuromuscular disorders by performing various facilitation movements, using a fuzzy controller; by Duarte and collaborators (2006), who proposed the selection of patients for myocardial perfusion scintigraphy based on fuzzy sets theory applied to clinical-epidemiological data and treadmill test results; by Sousa et al. (2006) who applied fuzzy logic and logistic regression in the decision making for parathyroid scintingraphy study; by Campos-Delgado *et al.* (2006), who elaborated a fuzzy-based controller for glucose subcutaneous regulation in

type-1 diabetic patients; Saraoglu and Sanli (2007), who elaborated a fuzzy decision support system on anesthetic depth control for helping anesthetists in surgeries; and by Boissy *et al.* (2007), who applied the fuzzy logic approach in telemedicine. In the medical images segmentation we point out the recent work by Tanaka *et al.* (2008), who developed a fuzzy rule-based system to treat electrical impedance tomography images for lung and heart segmentation, through ventilation and perfusion pulmonary functions.

13.3.6 Hybrid Diagnostic Systems

As mentioned above the association of fuzzy logic with neural networks and distributed processing, also called *soft computing*, is one of the most advanced areas of research in artificial intelligence as applied to medicine.

Research in fuzzy neural networks, which started from application oriented fuzzy system tuning, then moving to the automatic generation of fuzzy systems from data, is reaching a more mature stage, especially after the proof of functional equivalence of certain fuzzy models and neural networks (Halgamuge & Glesner, 1994). Non-linear models, such as given by neural networks and fuzzy logic, have established a good reputation for medical data analysis as computational and logical counterparts to statistical methods. Whereas multilayer perceptrons perform well with large datasets, a combination of neural learning with fuzzy logical network interpretation provides a network reduction well suited for smaller datasets (Eklund & Forsström, 1995). In this section we briefly review some articles dealing with hybrid models, particularly fuzzy logic and neural network, as applied to medical diagnosis.

A really hybrid model was presented by Molnar *et al.* (1993), combining multivariate mathematics, fuzzy logic and neural networks for the diagnosis of cytological smears. The method was applied in the area of quantitative cytology and was compared with the traditional classifiers. The discriminant analysis classified correctly the 95.6% of malignant cases, 86.7% of the dysplasias, and 80.7% of normal cases in a sample of gastric imprint smears. The fuzzy logic module of the system made the diagnostic borders fine tunable and reliable, and the back propagation neural network classified all the diagnostic groups above 95% correctly. The authors stated that the application of nonlinear computational methods made the diagnostic system more reliable.

Heuristic combinations of pattern recognition and artificial intelligence tools is the subject of the hybrid diagnostic model by Kuncheva (1993). This model is comprised by a trainable fuzzy neuron which, according to the author, resembles some elements from the physician's decision process. The model was applied to a real case from aviation medicine and demonstrated the enhanced performance of the system.

Fuzzy neural networks were applied on a model for reduction of false-positive detections in digital chest radiographs by Lin *et al.* (1993). The mode's architecture was based on fuzzy set theory and convulsional neural network and was tested in an automatic lung nodule detection system. The neural network was trained by a supervised back-propagation algorithm based on fuzzy membership

functions for lung nodule areas. A linguistic label was assigned to each nodule candidate of the training set and the label was then converted to a membership value. The trained network's output was defuzzified and the system was evaluated throughout a receiver operating characteristic analysis and showed an average Az (the performance index) of 0.84 which is equivalent to 0.80 true-positive detection (sensitivity) with an average 2-3 false positive detection per chest image.

A tool for building hybrid connectionist expert systems for medical diagnosis was presented by Leão and collaborators (1994). The system, called HYCONES II offers, according to the authors, to the knowledge engineer a hybrid knowledge base that integrates frames with three different neural networks, namely, the combinatorial neural models, the fuzzy ARTMAP and the semantic ART-SMART models. These models have their performance to solve diagnostic problems compared. In addition, the system's knowledge representation features, built in the symbolic component of its hybrid knowledge-base to deal and represent fuzzy medical variables is presented.

Real-time hemodynamic diagnostics was the subject of the work by Goldman and Cordova (1994). These authors developed a real-time system to diagnose cardiopulmonary emergencies. The system was designed to utilize routinely-monitoring physiological data in order to automatically diagnose potentially fatal events. The diagnostic engine was based on a hybrid fuzzy logic/neural network and was applied to analyze physiological data during a simulated arrhythmic cardiac arrest in order to assess the validity of the diagnostic methodology. The system used data from capnogram, electrocardiogram and arterial blood pressure. The system had a good performance and the diagnostic engine effectively diagnosed the likelihood of arrhythmic cardiac arrest from the subtle hemodynamic trends which precede the complete arrest. As the clinical picture worsened the system accurately indicated the change in patient condition. The end of the simulation was rapidly detected by the diagnostic engine.

Ichimura *et al.* (1995) extracted fuzzy rules using neural networks with structure level adaptation and applied them to diagnosis of hepatobiliary disorders. Their method proposed a procedure to derive a neuron generation/annihilation automatically and the authors applied the procedure to the learning system. They next applied those procedures to the learning system in which the experimental data related to hepatobiliary data, containing ten biochemical terms test for four pathologies. After the learning phase the proposed system converged to a diagnostic with an accuracy of 70%. The fuzzy rules applied were related in meaning to the input of the weight vector. In addition, the authors used the extracted fuzzy rules for all databases to implement the feed-forward calculations.

A new classification strategy for magnetic resonance data, called computerized consensus diagnosis, was proposed by Somorjai *et al.* (1995). The strategy involved the cross-validated training of several classifiers of diverse conceptual and methodological origin on the same data, and combined their outcomes. The method was tested on proton magnetic resonance spectra of human thyroid biopsies. The authors used linear discriminant analysis, a neural network method

and genetic programming as independent classifiers on two spectral regions, and chose the median of the six classification (normal/malignant) outcomes as the consensus. The procedure showed a 100% specificity and 98% sensitivity on samples of known malignancy in the test sets. The authors discussed the importance of fuzziness and undecidability in robust classification methods.

Feature vectors were applied by Sànchez *et al.* (1995) for automatic detection of auditory brainstem. Features are quantitative descriptors of different aspects of the response commonly taken into consideration by expert to assess auditory brainstem responses. The authors applied discriminant analysis and neural networks, which were a modified version of the fuzzy ARTMap model. The accuracy of the classification into normal and abnormal was assessed with methods from signal detection theory. The methods proposed showed that the approaches based on feature vectors had a performance more efficient than the artificial networks with raw data, or the individual features.

A fuzzy-net with single layer perceptron to analyze laboratory data was presented by Forsström *et al.* (1995). The authors worked with databases from computerized patient records, which included much clinical knowledge that could be useful for clinicians if properly retrieved. The system proposed built a smart link between patient databases and clinicians. It utilized neural network-based machine learning techniques and could produce decision support which met the special needs of clinicians. In this paper the authors used two small datasets to show how this scheme worked in the diagnosis of acute appendicitis and in the diagnosis of myocardial infarction. The performance of the neuro-fuzzy tool, as compared with logistic regression or backpropagation neural networks was slightly better, although not statistically significant.

Alzheimer's disease diagnosis was the subject of the paper by Pizzi *et al.* (1995), who applied a neural network classification of infrared spectra of histopathological material. The authors applied principal component analysis as a preprocessing technique for some of those neural networks while others were trained using the original spectra, one of which applied a variation of the back-propagation algorithm using fuzzy encoding. The neural nets using the principal components consistently outperformed their linear discriminant counterparts whereas only one of the original spectra produced results comparable to the best corresponding principal component cases.

The work by Holzmann *et al.* (1995) presents an expert system based on fuzzy analog ganglionar lattices. The system's reasoning scheme is designed analogously to the expert's mental organization and it is realized on an analog operator called the ganglionar lattice. This connectionist system used the medical knowledge to define its architecture. In addition, it used non-approximate reasoning with multiple antecedents of different relative importance and limited uncertainty. The system produced numerical results which could be translated into restricted natural language. The paper presents a simple example of that technology and the method's potentials are discussed for future applications.

Guez and Nevo (1996) present an analysis of the computational features of neural networks and fuzzy logic architectures which attempts to explain their

recent popularity as well as their drawbacks. The authors describe a customized neural network architecture as a non-linear adaptive signal processor for integrating monitoring, employed in the adaptive real-time anesthesiologist associate system. In this application the neural network realizes a non-linear scalar map from the set of physiological signals to a vital function status indicator.

The paper by Downs *et al.* (1996) presents an interesting application of the fuzzy ARTMAP neural network model to medical pattern classification tasks. The authors considered a number of diagnostic and prognostic domains, each one demonstrating a particular aspect of the model's usefulness. The model's strategy involved pooled decision-making in coronary care prognosis, using a number of networks. In addition, the application to breast cancer diagnosis demonstrated the model's symbolic rule extraction capabilities which support the validation and explanation of he network's prediction. Finally, the diagnosis of acute myocardial infarction demonstrated a novel category pruning technique allowing performance of a trained network to be altered so as to favor predictions of one class over another.

A neuro-fuzzy algorithm for diagnosis of coronary artery stenosis was proposed by Sztandera *et al.* (1996). The method used a neural network approach for the diagnosis of stenosis in the three main coronary arteries. The fuzzy network is trained with data from scintigram and the images are preprocessed and the uncertainties treated by fuzzy logic techniques. The model performed very well when compared with traditional diagnostic alternatives.

Cluster analysis, an often used technique to determine the number and characteristics of patterns present in vectors of biomedical response parameters, was treated with a hybrid system by Thayer (1996). This combined supervised and unsupervised learning algorithm. The author illustrated those procedure using growth curves of indices of family functioning in adaptation to pediatric chronic illness. Those clustering procedures was based upon neural network approaches to supervised (discriminant analysis) and unsupervised (cluster analysis) learning and was similar to fuzzy set algorithm developed to assess the degree of relatedness among a number of discrete units.

Another interesting application of fuzzy logic in medical diagnosis is the system for classification of cardiac arrhythmias proposed by Ham and Han (1996). In this article, the authors investigate the QRS complex, extracted from electrocardiogram (ECG) data, using fuzzy adaptive resonance theory mapping (ARTMAP) to classify cardiac arrhythmias. Two different conditions have been analyzed: normal and abnormal premature ventricular contraction. Based on standardized database annotations, cardiac beats for normal and abnormal QRS complexes were extracted from this database, scaled, and Hamming windowed, after bandpass filtering, to yield a sequence of 100 samples for each QRS segment. From each of these sequences, two linear predictive coding coefficient were generated using Burg's maximum entropy method. The two coefficient, along with the mean-square value of the QRS complex segment, were utilized as features for each condition to train and test a fuzzy ARTMAP neural network for classification of normal and abnormal premature ventricular conditions. The

test results show that the fuzzy ARTMAP neural network can classify cardiac arrhythmias with greater than 99% specificity and 97% sensitivity.

Innocent and collaborators (1997) applied the fuzzy ARTMAP and MINMAXMAP neural network to radiographic image classification analysis. The work was concerned with the classification analysis of exercise-induced lower leg pain by applying competitive neural network clustering and mapping techniques to fuzzy descriptions of bone scan images of the tibia. The clusters were described and compared with each other and with the expert known classes that would be expected from medical findings. The discovery clusters provided training sets for supervised and unsupervised learning by the ARTMAP and similar neural network. The authors concluded that the use of the neural clustering method improves the classification process of the shin images despite the paucity of the data and its inherent uncertainty.

Undoubtedly, the hybrid systems is one of the areas with the highest growing rate in the expert systems field. Concerning diagnostic systems we can find more than 2.200 articles in *MEDLINE* database, of which almost 400 only in 2006-2007 interval (Peña-Reyes & Sipper, 2000; Kannathal *et al.*, 2005; Zarkogianni *et al.*, 2007; Bosl, 2007; Maraziotis *et al.*, 2007; Tu & Toga, 2007; Xu *et al.*, 2007; Grossi *et al.*, 2007; Chen *et al.*, 2007; Bommanna *et al.*, 2008; Ertas *et al.*, 2008; Maglogiannis *et al.*, 2008; Jekova *et al.*, 2008).

Finally, we want to finish this brief (and acknowledgedly incomplete) review highlighting that the main objective here was to illustrate how different and creative have been the applications of fuzzy logic in biomedicine, both from the theoretical point of view and in the diversity of medical areas involved. In fact, the word "fuzzy", and its sets theory meaning, has becoming common among the medical communities. The progress of the fuzzy logic and all AI tools, and their valiant contribution, in the health sciences makes us to believe that the insertion of these techniques in medicine is a road without turn. In fact, the development of expert systems, as to support decision in the diagnosis as for the elaboration of modern laboratory and hospital equipments, together with the progress of the automation, telemedicine and electronic patient record, have, little by little, transformed the way to do medicine in the world.

14 Final Reflexions

In this book we pursued the challenge posed by the need to tame uncertainty in the context of medical and epidemiological problems. The fact that epidemiology deals with population of individuals leads to uncertainties of different natures. In fact, much ink has been spent on the speculations about the competitive or complementary role of probability and fuzzy logic in dealing with uncertainty. We are fully convinced that both disciplines address complementary dimensions of uncertainty. The success of applied probability theory in describing sampling variations and model specification in epidemiological problems is undeniable. Yet, fuzzy logic can be an undispensible tool to address vagueness related to intended meaning or linguistic uncertainty. As thoroughly discussed along this book, we believe that they can be seen as subsets of a more general logic. Paraconsistent logic, which in some way generalizes the non-binary properties of fuzzy logic, has been proposed as an appropriate candidate (Sylvan & Abe, 1996; Batens *et al.*, 2000; Abe, 2004; Abe *et al.*, 2005), particularly when combined with neural network structures. More recently, the Neutrosophic logic is appearing as a promise of a new powerful approach with great applicability, due to its flexibility and its connection with fuzzy sets theory (Wang *et al.*, 2005; Kandasamy & Smarandache, 2005).

The challenges in recasting fuzzy logic and probability theory under a common framework seems the natural subsequent step to be pursued. The seminal work by the school founded by Zadeh point to possible promising directions. The specific applications of hybrid models combining both probability theory and fuzzy reasoning proposed in this book testify the richness and usefulness of this complementary approach.

The historical development of probability theory, since the work by the founding fathers, Pascal and Fermat among others, until its axiomatic foundation provided by Kolmogorov, spanned for at least three centuries. Fuzzy logic was first proposed in 1965 being therefore a much younger field. A long road still lies ahead before one could expect to see an axiomatic formulation that could accommodate the complementary principles that inspired their original founders. The epidemiological examples illustrated here make a strong case for the need of a unifying theory that overcomes the current hurdles. Under this unifying

E. Massad et al.: Fuzzy Log. in Act.: Appl. in Epidem. & Bey., STUDFUZZ 232, pp. 311–312, 2008.
springerlink.com

axiomatic theory, a formal definition of uncertainty, imprecision, and vagueness would emerge naturally from first principles.

Biosciences have decisively inspired the theoretical development of probability theory, since the work by Fisher to the current post-genomic era. The predominant role of statistics in modern epidemiology is no exception and has also triggered key theoretical developments in applied probability. It is, therefore, natural to expect that medicine and epidemiology will continue to play a key role in the theoretical development of fuzzy logic. This common ground is one additional factor leading to a common theoretical formalism to the underlying treatment of uncertainty provided by both schools of thought. The current popularity of fuzzy logic applications in the field of engineering and control may add a refreshing perspective to the traditional influence offered by medicine and epidemiology so far. The time seems ripe for the extended influence on fuzzy sets theory originating from disciplines that have to deal equally with uncertainty, such as social sciences, economics and demography.

Modern developments in numerical procedures, such as the bootstrap and *Markov Chain Monte Carlo* (MCMC) simulations, have sharply broadened the statistical horizons. These procedures have successfully overcome the constraints imposed by the analytical limitations unavoidable to the application of formal approaches to complex real problems. Similarly, the inherent flexibility of fuzzy logic tools could greatly benefit from modern numerical techniques. Linguistic models confer enormous flexibility to the fuzzy logic approach, however, as mentioned elsewhere in the book, they are plagued by the multiplicity of rules that may explode to several orders of magnitude. This is indeed an area in which numerical procedures could show their immense power and usefulness.

The strength of fuzzy logic as well as statistics emanates from their proved abilities of solving real problems. In medicine and epidemiology, this equates to reducing human suffering. We hope that the medical and epidemiological applications of fuzzy logic described in this work might work catalytically to encourage theoreticians as well as applied researchers to further develop the foundations laid down here. In spite of all the **uncertainty** one might think of, this will **certainly** contribute to a better world.

References

1. Abbey, H.: An examination of the Reed-Frost theory of epidemics. Hum. Biol. 24, 201–233 (1952)
2. Abe, J.M.: Paraconsistent artificial neural networks: an introduction. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3214, pp. 942–948. Springer, Heidelberg (2004)
3. Abe, J.M., Ortega, N.R.S., Mario, M.C., Del Santo, M.: Paraconsistent artificial neural network: an application in cephalometric analysis. In: Khosla, R., Howlett, R.J., Jain, L.C. (eds.) KES 2005. LNCS (LNAI), vol. 3682, pp. 716–723. Springer, Heidelberg (2005)
4. Abrams, B., Newman, V.: Small for gestational age birth: maternal predictors and comparison with risk factors of spontaneous preterm delivery in the same cohort. Am. J. Obstet. Gynecol. 164, 785–790 (1991)
5. Adlassnig, K.P., Kolarz, G.: CADIAG-2: Computer-assisted medical diagnosis using fuzzy subsets. In: Gupta, M.M., Sanchez, E. (eds.) Approximate Reasoning in Decision Analysis, pp. 219–247. North-Holland, New York (1982)
6. Adlassnig, K.P., Scheithauer, W., Grabner, G.: Computer-Assisted Diagnosis and its Application in Pancreatic Diseases. Acta. Med. Austriaca. 11(3-4), 125–134 (1984)
7. Adlassnig, K.P., Kolarz, G., Scheithauer, W., Effenberger, H., Grabner, G.: CADIAG: Approaches to Computer-Assisted Medical Diagnosis. Comput. Biol. Med. 15(5), 315–335 (1985)
8. Adlassnig, K.P.: A survey on medical diagnosis and fuzzy subsets. In: Gupta, M.M., Sanchez, E. (eds.) Approximate Reasoning in Decision Analysis, pp. 203–217. North-Holland, New York (1986)
9. Adlassnig, K.P.: Fuzzy diagnostic and therapeutic decision support. Österreichische Computer Gesellschaft, Wien (2000)
10. Almond, R.G.: Graphical belief modeling. Chapman & Hall, London (1995)
11. Amaku, M., Coutinho, F.A.B., Azevedo-Neto, R.S., Burattini, M.N., Lopez, L.F., Massad, E.: Vaccination against rubella: Analysis of the temporal evolution of the age-dependent force of infection and the effects of different contact patterns. Phys. Rev. E. 67(051907), 1–11 (2003)
12. Anderson, R.M.: Population dynamics of infectious diseases: theory and applications. Chapman & Hall, London (1982)
13. Anderson, J., Bandler, W., Kohout, L.J., Trayner, C.: A route-choosing medical diagnostic technique. Fuzzy Sets Syst. 23, 89–96 (1987)

14. Anderson, R.M.: Epidemiological models and predictions. Tropical and Geographical Medicine 40(3), S30–S39 (1988)
15. Anderson, R.M.: The epidemiology of HIV infection: variable incubation plus infections periods and heterogeneity infection behavior. Journal of Royal Statistical Society A151, 66–93 (1988)
16. Anderson, R.M., May, R.M.: Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford (1991)
17. Arita, S., Yoneda, M., Hori, Y.: Quantitative analysis of glucose tolerance tests, responses using fuzzy inference. Nippon Rinsho. 54(10), 2807–2811 (1996)
18. Aubin, J.P., Cellina, A.: Differential inclusions-set-valued maps and viability theory. Springer, Heidelberg (1984)
19. Axer, H., Berks, G., Graf, V.: Fuzzy approaches to medical image processing. In: Adlassnig, K.P. (ed.) Fuzzy Diagnostic and Therapeutic Decision Support, pp. 89–97. Österreichische Computer Gesellschaft, Wien (2000)
20. Babbott Jr., F.L., Gordon, J.E.: Modern measles. Am. J. Med. Sci. 228, 334–361 (1954)
21. Bailey, N.T.J.: The mathematical theory of infectious diseases and its applications. Charles Griffin & Co., London (1975)
22. Bandemer, H., Gottwald, S.: Fuzzy sets, fuzzy logic, fuzzy methods. John Wiley & Sons, Chichester (1995)
23. Barillot, C., Lemoine, D., Le Briquer, L., Lachmann, F., Gibaud, B.: Data fusion in medical imaging: merging multimodal and multipatient images, identification of structures and 3D display aspects. Eur. J. Radiol. 17(1), 22–27 (1993)
24. Barnett, G.O., Cimino, J.J., Hupp, J.A., Hoffer, E.P.: DXPlain. An evolving diagnostic decision-support system. JAMA 258(1), 67–74 (1987)
25. Barros, L.C.: Deterministic models with subjective parameters, Master Thesis, University of Campinas, São Paulo, Brazil (in Portuguese) (1992)
26. Barros, L.C., Tonelli, P.A., Bassanezi, L.C.: Fuzzy modeling in population dynamics. Ecological Modeling 128, 27–33 (2000)
27. Barros, L.C., Oliveira, R.Z.G., Bassanezi, R.C., Leite, M.B.F.: A disease evolution with uncertain parameters. In: Proceedings of the Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, Canada, pp. 1626–1630 (2001)
28. Barros, L.C., Oliveira, R.Z.G., Bassanezi, R.C., Leite, M.B.F.: Type SIS epidemiological model with subjective parameters. In: Proceedings of the First Brazilian Symposium on Mathematical and Computational Biology, Rio de Janeiro, Brazil, pp. 149–164 (2002)
29. Barros, L.C., Bassanezi, R.C., Leite, M.B.F.: The epidemiological models SI with a fuzzy transmission. Computer and Mathematics with Applications 45, 1619–1628 (2003)
30. Barros, L.C., Bassanezi, R.C., Oliveira, R.Z.G.: Fuzzy differential inclusion: an application to epidemiology. In: Lopes-Diaz, M., Gil, M.A., Grzegorzewski, P., Hyrniewicz, O., Lawry, J. (eds.) Soft Methodology and Random Information Systems, vol. I, pp. 631–637. Springer, Warsaw (2004)
31. Barros, L.C., Bassanezi, R.C.: Topics of fuzzy logic and biomathematics. Didactic Books, vol. 5, University of Campinas, São Paulo, Brazil (2006) (in Portuguese)
32. Barros, L.C., Oliveira, R.Z.G., Leite, M.B.F., Bassanezi, R.C.: Models to directly transmitted diseases based on fuzzy rules. In: Hyun, M.Y. (ed.) 2nd International Conference on Computational and Mathematical Population Dynamics, Campinas-São Paulo, Brazil, p. 112 (2007)

33. Bartle, R.G.: The elements of integration of Lebesgue measure. Wiley, USA (1995)
34. Bartolin, R., Bouvenot, G., Soula, G., Sanchez, E.: The fuzzy set theory as a biomedical diagnostic aid. Sem. Hop. 58(22), 1361–1365 (1982)
35. Bassanezi, R.C., Barros, L.C.: A simple model of life expectancy with subjective parameters. Kybernetes: The International Journal of Systems and Cybernetics 4(7), 91–98 (1995)
36. Bassanezi, R.B., Bergamin Filho, A., Amorim, L., Gimenes-Fernandes, N., Gottwald, T.R.: Spatial and temporal analysis of Citrus Sudden Death as a tool to generate hypothesis its etiology. Phytopathology 93, 502–512 (2003)
37. Bastian, A.: Identifying fuzzy models utilizing genetic programming. Fuzzy Sets Syst. 113, 333–350 (2000)
38. Batens, D., Mortensen, C., Priest, G., Van Bendegem, J.-P.: Frontiers of paraconsistent logic. Research Studies Press, Beldock, Hertfordshire (2000)
39. Bailey, N.T.: The mathematical theory of infectious diseases. John Wiley & Sons, London (1975)
40. Becker, N.G.: Analysis of infectious disease data. Chapman & Hall, London (1979)
41. Bellamy, J.E.: Medical diagnosis, diagnostic spaces, and fuzzy systems. J. Am. Vet. Med. Assoc. 210(3), 390–396 (1997)
42. Bellman, R., Zadeh, L.A.: Decision-making in a fuzzy environment. Management Science 17, 141–154 (1970)
43. Bergner, M.: Measurement of health status. Med. Care 23, 696–704 (1985)
44. Bettiol, H., Rona, R.J., Chinn, S., Goldani, M., Barbieri, M.A.: Factors associated with preterm births in Southeast Brazil: a comparison of two birth cohorts born 15 years apart. Pediatric & Perinatal Epidemiol. 14, 30–38 (2000)
45. Binaghi, E., De Giorgi, O., Maggi, G., Motta, T., Rampini, A.: Computer-assisted diagnosis of postmenopausal osteoporosis using a fuzzy expert system shell. Comput. Biomed. Res. 26(6), 498–516 (1993)
46. Boegl, K., Kainberger, F., Adlassnig, K.P., Kolousek, G., Leitich, H., Kolarz, G., Imhof, H.: New approaches to computer-assisted diagnosis of rheumatologic diseases. Radiologe 35(9), 604–610 (1995)
47. Boissy, P., Choquette, S., Hamel, M., Noury, N.: User-based motion sensing and fuzzy logic for automated fall detection in older adults. Telemed. J. E. Health 13(6), 683–693 (2007)
48. Bolotin, A.: Replacing indicator variables by fuzzy membership functions in statistical regression models: examples of epidemiological studies. In: Barreiro, J.M., Martín-Sánchez, F., Maojo, V., Sanz, F. (eds.) ISBMDA 2004. LNCS, vol. 3337, pp. 251–258. Springer, Heidelberg (2004)
49. Bommanna Raja, K., Madheswaran, M., Thyagarajah, K.: A hybrid fuzzy-neural system for computer-aided diagnosis of ultrasound kidney images using prominent features. J. Med. Syst. 32(1), 65–83 (2008)
50. Bosl, W.J.: Systems biology by the rules: hybrid intelligent systems for pathway modeling and discovery. BMC Syst. Biol. 15(1), 1–13 (2007)
51. Boxler, P.: Lyapunov Exponents, Inst. für Dynamische Systeme, Report Nr. 197, Bremen, Germany (1988)
52. Boylan, R.D.: A note on epidemics in heterogeneous populations. Mathematical Biosciences 105, 133–137 (1991)
53. Bravo, R., Soriano, V., Martin, R., Delromero, J., Dronda, F., Gutierrez, M., Martines, P., Valencia, E., Moreno, V., Laguna, F.: The quantification of human immunodeficiency viremia in patients with rapid and slow progression of the disease. Med. Clin. (Barc) 104(14), 530–534 (1995)

54. Buchanan, B.G., Shortliffe, E.H.: Rule-based expert systems: the MYCIN experiments of the Stanford heuristic programming project, Addison-Wesley series in artificial intelligence, USA (1984)
55. Budd, W.: Malignant Cholera: Its mode of propagation, and its prevention. John Churchill, London (1849)
56. Burattini, M.N., Massad, E., Coutinho, F.A.B.: Dynamics of malarial transmission based on serological data. Epidemiol Infect. 111(3), 503–523 (1993)
57. Burattini, M.N., Massad, E., Coutinho, F.A.B., Azevedo-Neto, R.S., Menezes, R.X., Lopes, L.F.: A mathematical model of the impact of crack-cocaine use on the prevalence of HIV/AIDS among drug users. Mathematical and Computer Modeling 28(3), 21–29 (1998)
58. Burattini, M.N., Coutinho, F.A.B., Lopez, L.F., Massad, E.: Modeling the dynamics of Leishmaniasis considering human, animal host and vector populations. Journal of Biological Systems 6(4), 337–356 (1998)
59. Caetano, M., Yoneyama, T.: A comparative evaluation of open loop and closed loop drug administration strategies in the treatment of aids. Brazilian Academy of Sciences 71, 589–597 (1999)
60. Campbell, G., Levy, D., Bailey, J.J.: Bootstrap comparison of fuzzy ROC curves for ECG-LVH algorithms using data from the Framingham-Heart-Study. J. Electrocardiol. 23(S), 132–137 (1991)
61. Campisi, G., Giovannelli, L., Calvino, F., Matranga, D., Colella, G., Di Liberto, C., Capra, G., Leao, J.C., Lo Muzio, L., Capogreco, M., D'Angelo, M.: HPV infection in relation to OSCC histological grading and TNM stage. Evaluation by traditional statistics and fuzzy logic model. Oral Oncol. 42(6), 638–645 (2006)
62. Campos-Delgado, D.U., Hernández-Ordoñez, M., Femat, R., Gordillo-Moscoso, A.: Fuzzy-based controller for glucose regulation in type-1 diabetic patients by subcutaneous route. IEEE Trans. Biomed. Eng. 53(11), 2201–2210 (2006)
63. Casals, M.R.: Bayesian testing of fuzzy parametric hypotheses from fuzzy information. Oper. Res. 27, 189–199 (1993)
64. Castanho, M.J.P., Barros, L.C., Yamakami, A., Vendite, L.L.: Fuzzy receiver operating characteristic curve: An option to evaluate diagnostic tests. IEEE T. Inf. Technol. Biomed. 11(3), 244–250 (2007)
65. Castanho, M.J.P.: Elaboration and evaluation of a mathematical model to predict the evolution of prostate cancer and to describe its growing using fuzzy sets theory, Ph.D. Thesis, University of Campinas, São Paulo, Brazil (in Portuguese) (2005)
66. Caswell, H.: Matrix population models. Sinauer Associates, Inc., Sunderland (2000)
67. Causer, M.B., Findlay, G.A., Hawes, C.R., Boswell, D.R.: Assessment of a computerized system fom the diagnosis of iron deficiency. Pathology 26(1), 37–39 (1994)
68. CDC: CDC - Advances in global measles control and elimination. In: Proceedings of the 1997 International Meeting Cosponsored by the World Health Organization, the Pan American Health Organization, CDC, and the United Nations Children's Fund (UNICEF) 47 (RR-11), (MMWR), pp. 1–24 (1998)
69. CDC: CDC - Severe acute respiratory syndrome: What Everyone Should Know About SARS, Centers for Disease Control and Prevention (2003), `http://www.cdc.gov/ncidod/sars/basics.htm` (last visit in April 2008)
70. Chen, C.A., Li, Y.C., Lin, Y.F., Yu, F.C., Huang, W.H., Chiu, J.S.: Neuro-fuzzy technology as a predictor of parathyroid hormone level in hemodialysis patients. Tohoku J. Exp. Med. 211(1), 81–87 (2007)

71. Childs, R.: Genetic medicine: a logic of medicine. The Johns Hopkins University Press, Baltimore (1999)
72. Civanlar, M.R., Trussel, H.J.: Constructing membership function using statistical data. Fuzzy Sets and System 18, 1–13 (1986)
73. Civanlar, M.R., Trussel, H.J.: Constructing membership function and medical decisions with uncertainty. In: Bilal, M., Ayyub, A.A. (eds.) Proceedings of First International Symposium on Uncertainty Modeling Analysis, pp. 694–699. IEEE Computer Society Press, USA (1990)
74. Clayton, D., Hills, M.: Statistical models in epidemiology. Oxford Science Publications, New York (1993)
75. Col, N.F., Eckman, M.H., Karas, R.H., Pauker, S.G., Goldberg, R.J., Ross, E.M., Orr, R.K., Wong, J.B.: Patient-specific decisions about hormone replacement therapy in post menopausal women. JAMA 277(14), 1140–1147 (1997)
76. Coletti, G., Scozzafava, R.: Conditional probability, fuzzy sets, and possibility: a unifying view. Fuzzy Sets Syst. 144(1), 227–249 (2004)
77. Conte, S.D., Boor, C.: Elementary numerical analysis. McGraw Hill, USA (1981)
78. Cordova, M.J., Goldman, J.M.: Advanced clinical monitoring: similar scenario discrimination. Medinfo. 8(2), 880–884 (1995)
79. Costa, A.J.L., Massad, E., Ortega, N.R.S., Araújo, A.Q.C.: Estimation of the degree of functional disability through a fuzzy model. International Journal of Medical Informatics 73, 647–656 (2004)
80. Costa, C.E.C., Gotlieb, S.L.D.: Epidemiologic study of birth weight from birth certificates. Rev. Saude. Publica. 32, 328–334 (1998)
81. Coutinho, F.A.B., Massad, E., Lopez, L.F., Burattini, M.N., Struchiner, C.J., Azevedo-Neto, R.S.: Modelling heterogeneities in individual frailties in epidemic models. Mathematical and Computer modeling 30, 97–115 (1999)
82. Coutinho, F.A.B., Lopez, L.F., Burattini, M.N., Massad, E.: Modeling the natural history of HIV in individuals and its epidemiological implications. Bull. Math. Biol. 63, 1041–1062 (2001)
83. Cox, D.R.: Planning of experiments. Wiley, New York (1958)
84. de Dombal, F.T., Leaper, D.J., Staniland, J.R., McCann, A.P., Horrocks, J.C.: Computer-aided diagnosis of acute abdominal pain. Br. Med. J. 2(5804), 238–245 (1972)
85. Degani, R., Bortolan, G.: Fuzzy numbers in computerized electrocardiography. Fuzzy Sets Syst. 24, 345–362 (1987)
86. Degani, R., Bortolan, G.: Fuzzy decision-making in electrocardiography. Artif. Intell. Med. 1(2), 87–91 (1989)
87. Delgado, M., Verdegay, J.L., Vila, M.A.: Testing fuzzy-hypotheses, a Bayesian approach. In: Gupta, M.M., Ragade, R.K., Yager, R.R. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 307–316. North-Holland, Amsterdam (1979)
88. Demling, L.: Chaos reach, fractal, fuzzy logic. From stereotypes to reality-consequences for medicine. Fortschr. Med. 110(6), 77–79 (1992)
89. Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Stat. 38, 325–329 (1967)
90. de Oliveira, L.H., Struchiner, C.J.: Vaccine-associated paralytic poliomyelitis: a retrospective cohort study of acute flaccid paralysis in Brazil. Int. J. Epidemiol. 29(4), 757–763 (2000)
91. Department of Health (UK): Saving lives: our healthier nation, The Stationery Office, London, UK (1999)
92. Diamond, P.: Time-dependent differential inclusions, cocycle attractors and fuzzy differential equations. IEEE Trans. Fuzzy Systems 7, 734–740 (1999)

93. Diekmann, O., Heesterbeek, J.A.P., Metz, J.A.J.: On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. J. Math. Biol. 28, 365–382 (1990)
94. Diekmann, O., Heesterbeek, J.A.P.: Mathematical epidemiology of infectious diseases. John Wiley & Son Ltd., Chichester (2000)
95. Dietz, K.: The estimation of the basic reproduction number for infectious diseases. Stat. Meth. Med. 2, 23–41 (1993)
96. Downs, J., Harrison, R.F., Kennedy, R.L., Cross, S.S.: Application of the fuzzy ARTMAP neural network model to medical pattern classification tasks. Artif. Intell. Med. 8(4), 403–428 (1996)
97. Drumond, E.F., Machado, C.J., França, E.: Early neonatal mortality: An analysis of multiple causes of death by the Grade of Membership method. Cad Saúde Pública 23(1), 157–166 (2007) (in Portuguese)
98. Duarte, P.S., Mastrocolla, L.E., Farsky, P.S., Sampaio, C.R.E.P.S., Tonelli, P.A., Barros, L.C., Ortega, N.R., Pereira, J.C.R.: Selection of patients for myocardial perfusion scintigraphy based on fuzzy sets theory applied to clinical-epidemiological data and treadmill test results. Braz. J. Med. Biol. Res. 39(1), 9–18 (2006)
99. Dubois, D., Prade, H.: Fuzzy sets and systems: theory and applications. Academic Press, USA (1980)
100. Dubois, D., Prade, H.: Towards fuzzy differential calculus - Part 1: Integration of fuzzy mappings. Fuzzy Sets Syst. 8, 1–17 (1982a)
101. Dubois, D., Prade, H.: Towards fuzzy differential calculus - Part 2: Integration of fuzzy intervals. Fuzzy Sets Syst. 8, 105–116 (1982b)
102. Dubois, D., Prade, H.: Towards fuzzy differential calculus - Part3: Differentiation. Fuzzy Sets Syst. 8, 225–233 (1982c)
103. Dubois, D., Prade, H., Sandri, S.: On possibility/probability transformations. In: Lowen, R., Roubens, M.R. (eds.) Fuzzy Logic: State of Art, pp. 103–112. Kluwer Academic Publishers, Dordrecht (1993)
104. Dubois, D., Prade, H., Grabish, M.: Gradual rules and the approximation of control laws. In: Nguyen, H.T., Sugeno, M., Tong, R., Yager, R. (eds.) Theoretical Aspects of Fuzzy Control, pp. 147–181. Wiley, New York (1995)
105. Duckstein, L., Blinowska, A., Verroust, J.: Fuzzy classification of patient state with application to electrodiagnosis of peripheral polyneuropathy. IEEE Trans. Biomed. Eng. 42(8), 786–792 (1995)
106. Dujet, C.: Fuzzy methodologies in decision aid. In: Adlassnig, K.P. (ed.) Fuzzy Diagnostic and Therapeutic Decision Support, Österreichische Computer Gesellschaft, Wien, pp. 103–110 (2000)
107. Edelstein-Keshet, L.: Mathematical models in biology. McGraw-Hill, USA (1987)
108. Eklund, P., Forsström, J.J.: Computational intelligence for laboratory information systems. Scan J. Clin. Lab. Invest. Suppl. 222, 21–30 (1995)
109. Emoto, T., Abeyratne, U.R., Akutagawa, M., Nagashino, H., Kinouchi, Y.: Feature extraction for snore sound via neural network processing. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 5477–5480 (2007)
110. Ertas, G., Gulcur, H.O., Tunaci, M., Osman, O., Ucan, O.N.: A preliminary study on computerized lesion localization in MR mammography using 3D nMITR maps, multilayer cellular neural networks, and fuzzy c-partitioning. Med. Phys. 35(1), 195–205 (2008)
111. Esogbue, A.O., Elder, R.C.: Fuzzy sets and the modeling of physician decision processes: Part I: The initial interview-information gathering process. Fuzzy Sets Syst. 2, 279–291 (1979)

112. Esogbue, A.O., Elder, R.C.: Fuzzy sets and the modeling of physician decision processes: Part II: Fuzzy diagnosis decision models. Fuzzy Sets Syst. 3(1), 1–9 (1980)
113. Esogbue, A.O., Elder, R.C.: Measurement and valuation of a fuzzy mathematical model for medical diagnosis. Fuzzy Sets Syst. 10, 223–242 (1983)
114. Estrada-Peña, A.: Epidemiological surveillance of tick populations: a model to predict the colonization success of Ixodes ricinus (Acari: Ixodidae). Eur. J. Epidemiol. 13, 571–580 (1997)
115. Falk, C.T., Gilchrist, J.M., Perical, V.M.A., Speer, M.C.: Using neural networks as an aid in the determination of disease status: comparison of clinical diagnosis to neural-network predictions in a pedigree with autosomal dominant limb-girdle muscular dystrophy. Am. J. Hum. Genet. 62(4), 941–949 (1998)
116. Fanshel, S., Bush, J.W.: A health-status index and its application to health-services outcomes. Oper. Res. 18, 1021–1065 (1970)
117. Fathi-Torbaghan, M., Meyer, D.: MEDUSA: A fuzzy expert system for medical diagnosis of acute abdominal pain. Methods Inf. Med. 33(5), 522–529 (1994)
118. Fava, V.L.: Urbanization and Poverty in Brazil, Ph.D. Thesis, University of São Paulo, São Paulo, Brazil (in Portuguese) (1983)
119. Feller, W.: An introduction to probability theory and its applications. Wiley, New York (1968)
120. Fine, P.E.: Herd immunity: history, theory, practice. Epidemiol. Rev. 15(2), 265–302 (1993)
121. Filter, R.A., Xia, X., Gray, C.M.: Dynamic HIV/AIDS parameter estimation with application to a vaccine readiness study in Southern Africa. IEEE Trans. on Biomedical Engineering 52(5), 784–791 (2005)
122. Fordon, W.A., Bezdek, J.C.: The application of fuzzy set theory to medical diagnosis. In: Gupta, M.M., Ragade, R.K., Yager, R.R. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 445–461. North-Holland, Amsterdam (1979)
123. Forsström, J.J., Irjala, K., Seln, G., Nystrm, M., Eklund, P.: Using data preprocessing and single layer perceptron to analyse laboratory data. Scand. J. Clin. Lab. Invest. Suppl. 222, 75–81 (1995)
124. Frangakis, C.E., Rubin, D.B.: Principal stratification in causal inference. Biometrics 58(1), 21–29 (2002)
125. Gaulter, H.: The origin and progress of the malignant cholera in Manchester, London, UK (1833)
126. Georgopoulos, V.C., Malandraki, G.A., Stylios, C.D.: A fuzzy cognitive map approach to differential diagnosis of specific language impairment. Artif. Intell. Med. 29(3), 261–278 (2003)
127. Georgopoulos, V.C., Stylios, C.D.: Complementary case-based reasoning and competitive fuzzy cognitive maps for advanced medical decisions. Soft Computing 12(2), 191–199 (2008)
128. Gil, M.A., Corral, N., Gil, P.: The fuzzy decision problem: an approach to the point estimation problem with fuzzy information. European J. Oper. Res. 22, 26–34 (1985)
129. Gilbert, P.B., Bosch, R.J., Hudgens, M.: Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. Biometrics 59(3), 531–541 (2003)
130. Glies, B.G., Findlay, C.S., Haas, G., LaFrance, B., Laughing, W., Pembleton, S.: Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. Soc. Sci. Med. 64(3), 562–576 (2007)

131. Goldman, J.M., Cordova, M.J.: Advanced clinical monitoring: considerations for real-time hemodynamic diagnostics. In: Proc. Annu. Symp. Comput. Appl. Med. Care, pp. 752–755 (1994)
132. Gorek, M., Hartung, C., Stief, C.G.: Electromyograms in erectile dysfunction and computer-assisted interpretation. Biomed. Tech (Berl) 42(3), 48–54 (1997)
133. Greene, P.H.: On looking for neural networks and "cell assemblies" that underlie behavior I: a mathematical model. Bull. Math. Biophys. 24, 247–275 (1962)
134. Gravena, S.: O controle biológico na cultura algodoeira. Informe Agropecuário 9, 3–15 (2003) (in Portuguese)
135. Greenhalgh, D.: Vaccination campaigns for common childhood diseases. Math. Biosci. 100, 201–240 (1990)
136. Greenland, S., Robins, J.M.: Identifiability, exchangeability, and epidemiologic confounding. Int. J. Epidemiol. 15(3), 413–419 (1986)
137. Greenland, S.: Interpretation and choice of effect measures in epidemiologic analysis. Am. J. Epidemiol. 125(5), 761–768 (1987)
138. Greenwood, M., Yule, U.G.: The statistics of anti-typhoid anti-cholera inoculations, and the interpretation of such statistics in general. Proc. R Soc. Med. 8(part 2), 113–194 (1915)
139. Guez, A., Nevo, I.: Neural networks and fuzzy logic in clinical laboratory computing with application to integrated monitoring. Clin. Chim. Acta. 248(1), 73–90 (1996)
140. Grossi, E., Mancini, A., Buscema, M.: International experience on the use of artificial neural networks in gastroenterology. Dig Liver Dis. 39(3), 278–285 (2007)
141. Guarini, G.: New horizons in medicine. The application of "Fuzzy Logic" in Clinical and Experimental Medicine. Recenti. Prog. Med. 85(6), 335–339 (1994)
142. Halgamuge, S.K., Glesner, M.: Neural networks in designing fuzzy systems for real world applications. Fuzzy Sets Syst. 65(1), 1–12 (1994)
143. Halloran, M.E., Haber, M., Longini Jr., I.M., Struchiner, C.J.: Direct and indirect effects in vaccine efficacy and effectiveness. Am. J. Epidemiol. 133(4), 323–331 (1991)
144. Halloran, M.E., Struchiner, C.J.: Study designs for dependent happenings. Epidemiology 2(5), 331–338 (1991)
145. Halloran, M.E., Haber, M., Longini Jr., I.M.: Interpretation and estimation of vaccine efficacy under heterogeneity. Am. J. Epidemiol. 136(3), 328–343 (1992)
146. Halloran, M.E., Struchiner, C.J.: Causal inference in infectious-diseases. Epidemiology 6(2), 142–151 (1995)
147. Halloran, M.E.: Concepts of infectious disease epidemiology. In: Rothman, K.J., Greenland, S. (eds.) Modern epidemiology, Lippincott-Raven, Philadelphia, USA, pp. 529–554 (1998)
148. Halloran, M.E., Longini Jr., I.M., Struchiner, C.J.: Design and interpretation of vaccine field studies. Epidemiol. Rev. 21(1), 73–88 (1999)
149. Ham, F.M., Han, S.: Classification of cardiac arrhythmias using fuzzy ARTMAP. IEEE Trans. Biomed. Eng. 43(4), 425–430 (1996)
150. Hammad, T.A., Abdel-Wahab, M.F., DeClaris, N., El-Sahly, A., El-Kady, N., Strickland, G.T.: Comparative evaluation of the use of artificial neural networks for modeling the epidemiology of schistosomiasis mansoni. Trans. R Soc. Trop. Med. Hyg. 90(4), 372–376 (1996)
151. Health Development Agency (UK): Introducing health impact assessment (HIA): informing the decision making process. Health Development Agency, London, UK (2002)

152. Hethcote, H.W., Van Ark, J.W.: Epidemiological models for heterogeneous populations: Proportionate mixing, parameter estimation, and immunization programs. Math. Biosci. 84, 85–118 (1987)
153. Hethcote, H.W.: Optimal ages for vaccination for measles. Math. Biosc. 89, 29–52 (1988)
154. Hethcote, H.W.: The mathematics of infectious diseases. SIAM Review 42(4), 599–653 (2000)
155. Hiramatsu, K., Kabasawa, K., Kaihara, S.: Applications of fuzzy logic to medical diagnosis. Iyodenshi To Seitai Kogaku 12(3), 148–155 (1974) (in Japanese)
156. Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M.: Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature 373, 123–126 (1995)
157. Hodek, I.: Biology of Coccinellidae. Academic of Science. Prague, Czech Republic (1973)
158. Holland, P.W.: Statistics and causal inference. J. Am. Stat. Assoc. 81(396), 945–960 (1986)
159. Holling, C.S.: Some characteristics of simple types of predation and parasitism. The Canadian Entomologist 91(7), 385–398 (1959)
160. Holst, H., Ohlsson, M., Peterson, C., Edenbrandt, L.: Intelligent computer reporting 'lack of experience': a confidence measure for decision support systems. Clin. Physiol. 18(2), 139–147 (1998)
161. Holzmann, C.A., Ehijo, A., Perez, C.A.: Methodology for a medical expert system on fuzzy analog ganglionar lattices. Non-approximate reasoning with multiple antecedents of different relative importance and limited uncertainty. Med. Prog. Technol. 21(3), 147–158 (1995)
162. Hsin, C., Yang, T.: Two-sex life table and predation rate of Propylaea japonica Thunberg (Coleoptera: Coccinellidae) Fed on Myzus persicae (Sulzer) (Homoptera: Aphididae). Environmental Entomology 32(2), 327–333 (2003)
163. Hudgens, M.G., Hoering, A., Self, S.G.: On the analysis of viral load endpoints in HIV vaccine trials. Stat. Med. 22(14), 2281–2298 (2003)
164. Hüllermeier, E.: An approach to modeling and simulation of uncertain dynamical systems. Intern. J. Uncertainty Fuzziness Knowledge-Based Systems 5, 117–137 (1997)
165. Hutchins, S.S., Markowitz, L.E., Mead, P., Mixon, D., Sheline, J., Greenberg, N., Preblud, S.R., Orenstein, W.A., Hull, H.F.: A school-based measles outbreak: the effect of a selective re-vaccination policy and risk factors for vaccine failure. Am. J. Epidemiol. 132, 157–168 (1990)
166. Ichimura, T., Tazaki, E., Yoshida, K.: Extraction of fuzzy rules using neural networks with structure level adaptation - verification to the diagnosis of hepatobiliary disorders. Int. Biomed. Comput. 40(2), 139–146 (1995)
167. Innis, M.D.: Clinical problem solving-the role of expert laboratory systems. Med. Inf. (Lond). 22(3), 251–266 (1997)
168. Innocent, P.R., Barnes, M., John, R.: Application of the fuzzy ART/MAP and MINMAX/MAP neural network models to radiographic image classification. Artif. Intell. Med. 11(3), 241–263 (1997)
169. Jafelice, R., Barros, L.C., Bassanezi, R.C., Gomide, F.: Fuzzy modeling in symptomatic HIV virus infected population. Bull. Math. Biol. 66, 1597–1620 (2004a)

170. Jafelice, R., Barros, L.C., Bassanezi, R.C., Gomide, F.: Fuzzy set-based model to compute the life expectancy of HIV infected populations. In: Dick, S., Kurgan, L., Musilek, P., Pedrycz, W., Reformat, M. (eds.) Proceedings of Annual Meeting of the North American Fuzzy Information Processing, NAFIPS, Alberta, Canada, pp. 314–318 (2004b)
171. Jang, J.R.: ANFIS: adaptative-network-based fuzzy inference system. IEEE Trans. Syst. Man. Cybern. 23(3), 665–685 (1993)
172. Jekova, I., Bortolan, G., Christov, I.: Assessment and comparison of different methods for heartbeat classification. Med. Eng. Phys. 30(2), 248–257 (2008)
173. Joffe, M.M., Small, D., Hsu, C.Y.: Defining and estimating intervention effects for groups that will develop an auxiliary outcome. Statistical Science 22(1), 74–97 (2007)
174. Joffe, M.M., Ten Have, T.R., Feldman, H.I., Kimmel, S.E.: Model selection, confounder control, and marginal structural models: Review and new applications. American Statistician 58(4), 272–279 (2004)
175. Ju, M.S., Lin, C.C., Lin, D.H., Hwang, I.S., Chen, S.M.: A rehabilitation robot with force-position hybrid fuzzy controller: hybrid fuzzy control of rehabilitation robot. IEEE Trans. Neural Syst. Rehabil. Eng. 13(3), 349–358 (2005)
176. Kahya, Y.P., Yeginer, M., Bilgic, B.: Classifying respiratory sounds with different feature sets. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 1, pp. 2856–2859 (2006)
177. Kaleva, O.: Fuzzy differential equations. Fuzzy Sets Syst. 24, 301–317 (1987)
178. Kandasamy, W.B.V., Smarandache, F.: Fuzzy and neutrosophic analysis of women with HIV/AIDS (With specific reference to rural Tamil Nadu in India), Hexis, USA (2005)
179. Kandel, A.: Fuzzy mathematical techniques with applications. Addison-Wesley Publishing Company, USA (1986)
180. Kannathal, N., Puthusserypady, S., Choo Min, L., Rajendra Acharya, U., Laxminarayan, S.: Cardiac state diagnosis using adaptive neuro-fuzzy technique. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 4, pp. 3864–3867 (2005)
181. Kaplan, R.M., Bush, J.W., Berry, C.C.: Health status: types of validity and the index of Well-Being. Health Serv. Res. 11, 478–507 (1976)
182. Klein, H.M., Eisele, T., Klose, K.C., Stauss, I., Brenner, M., Ameling, W., Gnther, R.W.: Pattern recognition system for focal liver lesions using "crisp" and "fuzzy" classifiers. Invest. Radiol. 31(1), 6–10 (1996)
183. Keyfitz, N.: On future population. J. Am. Stat. Assoc. 67, 347–363 (1972)
184. Kleinbaum, D.G., Kupper, L.L., Morgenstern, H.: Epidemiologic research: principles and quantitative methods. Lifetime Learning Publications, Belmont (1982)
185. Klir, G., Yuan, B.: Fuzzy sets and fuzzy logic: theory and applications. Prentice Hall, USA (1995)
186. Kolarz, G., Adlassning, K.P.: Problems in establishing the medical expert systems CADIAG-2 in rheumatology. J. Med. Syst. 10(4), 395–405 (1986)
187. Kolmogorov, A.: Grundbegriffe der Wahrscheinlichkeitsrechnung. Springer, Berlin (1933); English translation Foundations of the theory of probability, Chelsea, New York (1950)
188. Kosko, B.: Fuzzy Engineering. Prentice-Hall, USA (1997)
189. Kovács, M., Juranovics, J.: "Auctoritas" Psychiatric Expert System Shell. Medinfo. 8 pt 2, 997 (1995)
190. Koyama, S., Obata, Y., Shimamoto, K., Ishigaki, T., Ishii, N., Isomoto, Y., Yoshine, K.: Breast ultrasonography computer-aided diagnosis using fuzzy inference. J. Ultrasound Med. 16(10), 665–672 (1997)

191. Krivan, V., Colombo, G.: A non-stochastic approach for modeling uncertainty in population dynamics. Bull. Math. Biol. 60(4), 721–751 (1998)
192. Kuncheva, L.: Fuzzy multi-level classifier for medical applications. Comput. Biol. Med. 20(6), 421–431 (1990)
193. Kuncheva, L.: An aggregation of pro and con evidence for medical decision support systems. Comput. Biol. Med. 23(6), 417–424 (1993)
194. Kuncheva, L., Andreeva, K.: DREAM: a shell-like software system for medical data analysis and decision support. Comput. methods Programs Biomed. 40(2), 73–81 (1993)
195. Kwok, H.F., Linkens, D.A., Mahfouf, M., Mills, G.H.: Rule-base derivation for intensive care ventilator control using ANFIS. Artif. Intell. Med. 29(3), 185–201 (2003)
196. Kwok, H., Linkens, D., Mahfouf, M., Mills, G.: Adaptive ventilator FiO2 advisor: use of non-invasive estimations of shunt. Artif. Intell. Med. 32(3), 157–169 (2004)
197. Landigran, P.J., Witte, J.J.: Neurologic disorders following live measles-virus vaccination. JAMA 223, 1459–1462 (1973)
198. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics 33, 159–174 (1977)
199. Leão, B.F., Guazzelli, A., Mendonça, E.A.: HYCONES II: A tool to build hybrid connectionist expert systems. In: Proc. Annu. Symp. Comput. Appl. Med. Care, pp. 747–751 (1994)
200. Lee, N., Hui, D., Wu, A., Chan, P., Cameron, P., Joynt, G.M., Ahuja, A., Yung, M.Y., Leung, C.B., To, K.F., Lui, S.F., Szeto, C.C., Chung, S., Sung, J.Y.: A major outbreak of severe acute respiratory syndrome in Hong Kong. N. Engl. J. Med. 348(20), 1986–1994 (2003)
201. Lefèvre, C., Picard, P.: A non-standard family of polynomials and the final size distribution of Reed-Frost epidemic processes. Advances in Applied Probability 22, 25–48 (1990)
202. Leite, M.B.F., Bassanezi, R.C., Yang, H.M.: The basic reproduction ratio for a model of directly transmitted infections considering the virus charge and the immunological response. IMA-Journal of Mathematics Applied in Medicine and Biology 17, 15–31 (2000)
203. Leitich, H., Adlassnig, K.P., Kolarz, G.: Strengths and limitations of automatic knowledge acquisition for the medical consultation system CADIAG-II/RHEUMA. In: Adlassnig, K.P. (ed.) Fuzzy Diagnostic and Therapeutic Decision Support, Österreichische Computer Gesellschaft, Wien, pp. 152–155 (2000)
204. Levin, L.A., Rizzo III, J.F., Lessell, S.: Neural network differentiation of optic neuritis and anterior ischaemic optic neuropath. Br. J. Ophthalmol. 80(9), 835–839 (1996)
205. Levy, J.A.: The transmission of AIDS - The case of the infected cell. JAMA 259(20), 3037–3038 (1988)
206. Lilienfeld, D.E.: The greening of epidemiology: sanitary physicians and the London Epidemiological Society (1830-1870). Bull. Hist. Med. 52, 503–528 (1979)
207. Lopez, L.F., Coutinho, F.A.B., Buratinni, M.N., Massad, E.: Threshold conditions for infection persistence in complex host-vectors interactions. Comptes Rendus Biologies Académie des Sciences Paris 325, 1073–1084 (2002)
208. Lopes, M.H.B.M., Marin, H.F., Ortega, N.R.S., Massad, E.: Fuzzy logic model based on the differential nursing diagnosis of alterations in urinary elimination. Stud. Health Technol. Inform. 122, 117–120 (2006)

209. Lin, J.S., Ligomenides, P.A., Freedman, M.T., Mun, S.K.: Application of artificial neural networks for reduction of fase-positive detections in digital chest radiographs. In: Proc. Annu. Symp. Comput. App. Med. Care, pp. 434–438 (1993)
210. Luger, G.F., Stubblefield, W.A.: Artificial Intelligence. Addison Wesley Longman Inc., Harlow, England (1998)
211. Macartney, F.: Logic in medicine. BMJ 295, 1325–1332 (1987)
212. Macdonald, G.: The analysis of equilibrium in malaria. Trop. Dis. Bull. 49, 813–828 (1952)
213. Maglogiannis, I., Sarimveis, H., Kiranoudis, C.T., Chatziioannou, A.A., Oikonomou, N., Aidinis, V.: Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images. IEEE Trans. Inf. Technol. Biomed. 12(1), 42–54 (2008)
214. Maia, J.O.C.: Some mathematical developments on the epidemic theory formulated by Reed and Frost. Hum. Biol. 24, 167–200 (1952)
215. Mahfouf, M., Abbod, M.F., Linkens, D.A.: A survey of fuzzy logic monitoring and control utilization in medicine. Artif. Intell. Med. 21, 27–42 (2001)
216. Mamdani, E.H., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies 7(1), 1–13 (1975)
217. Maraziotis, I.A., Dragomir, A., Bezerianos, A.: Gene networks reconstruction and time-series prediction from microarray data using recurrent neural fuzzy networks. IET Syst. Biol. 1(1), 41–50 (2007)
218. Marcos, J.V., Hornero, R., Alvarez, D., Del Campo, F., López, M.: Applying neural network classifiers in the diagnosis of the obstructive sleep apnea syndrome from nocturnal pulse oximetric recordings. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 5174–5177 (2007)
219. Marín, R., Mira, J.: On knowledge-based fuzzy classifiers: a medical case study. Fuzzy Sets Syst. 44(3), 421–430 (1991)
220. Markowitz, L.E., Katz, S.L.: Measles vaccine. In: Plotkin, S.A., Mortmer, E.A. (eds.) Vaccines, pp. 229–276. W.B.Saunders, Philadelphia (1994)
221. Mason, J.J., Owens, D.K., Harris, R.A., Cooke, J.P., Hlatky, M.A.: The role of coronary angiography and coronary revascularization before noncardiac vascular surgery. JAMA 273(24), 1919–1924 (1995)
222. Massad, E., Coutinho, F.A.B., Yang, H.M., De Carvalho, H.B., Mesquita, F., Burattini, M.N.: The basic reproduction ratio of HIV among intravenous-drug-users. Math. Biosci. 123(2), 227–247 (1994a)
223. Massad, E., Burattini, M.N., Azevedo Neto, R.S., Yang, H.M., Coutinho, F.A.B., Zanetta, D.M.T.: A model-based design of a vaccination strategy against rubella in a non-immunized community of São Paulo State, Brazil. Epidemol. Infect. 112(3), 579–594 (1994b)
224. Massad, E., Azevedo-Neto, R.S., Burattini, M.N., Zanetta, D.M.T., Coutinho, F.A.B., Yang, H.M., Moraes, J.C., Panutti, C.S., Souza, V.A.U.F., Silveira, A.S.B., Struchiner, C.J., Oselka, G.W., Camargo, M.C.C., Omoto, T.M., Passos, S.D.: Assessing the efficacy of a mixed vaccination strategy against rubella in São Paulo, Brazil. Int. J. Epidemiol. 24(4), 842–850 (1995)
225. Massad, E.: Mathematical and Computer Models Epidemiology. MD Computing 13(2), 147–154 (1996)
226. Massad, E., Burattini, M.N., Reale, D., Carvalho, H.B., Zanetta, D.M.: Fuzzy logic in the analysis of vulnerability to HIV/AIDS infection in sexual partners of injecting drug users. In: 8th International Conference on the Reduction of Drug Related Harm, Maison de la Mutualité, Paris, France (1997a)

227. Massad, E., Struchiner, C.J., Coutinho, F.A.B., Burattini, M.N., Azevedo-Neto, R.S., Lopez, L.F.: Fuzzy dynamic systems in epidemic modeling. In: 11th International Conference on Mathematical and Computer Modeling and Scientific Computing, p. 139. International Science and Technology Publishers, Washington (1997b)
228. Massad, E., Burattini, M.N., Coutinho, F.A.B., Struchiner, C.J.: A fuzzy model in HIV/AIDS epidemiology: Predicting the clinical course basing on HIV viral load. In: Pearson, D.W. (ed.) Proceedings of SOCO 1997 - Second International Computer Science Conventions Symposium on Soft Computing, pp. 58–63. Academic Press, Nimes (1997c)
229. Massad, E., Burattini, M.N., Ortega, N.R.S.: Fuzzy logic and measles vaccination: designing a control strategy. Int. J. Epidemiol. 28, 550–557 (1999)
230. Massad, E., Ortega, N.R.S., Struchiner, C.J., Burattini, M.N.: Fuzzy Epidemiology. Artif. Intell. Med. 29, 241–259 (2003)
231. Massad, E., Coutinho, F.A.B., Burattini, M.N., Lopez, L.F., Struchiner, C.J.: Yellow fever vaccination: How much is enough? Vaccine 23(30), 3908–3914 (2005a)
232. Massad, E., Burattini, M.N., Lopez, L.F., Coutinho, F.A.B.: Forecasting versus projection models in epidemiology: The case of the SARS epidemics. Med. Hypotheses 65, 17–22 (2005b)
233. May, R.M.: Stability and complexity in model ecosystems. Princeton University, USA (1974)
234. Mayr, E.: The growth of biological thought. Harvard University Press, Cambridge (1982)
235. McDowell, I., Newell, C.: Measuring health: a guide to rating scales and questionnaires. Oxford University Press, Oxford (1987)
236. Mellors, J.W., Rinaldo Jr., C.R., Gupta, P., White, R.M., Todd, J.A., Kingsley, L.A.: Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science 272, 1167–1170 (1996)
237. Mendel, J.M.: Rule-based fuzzy logic systems: introduction and new directions. Prentice-Hall, USA (2001)
238. Menezes, R.X., Ortega, N.R.S., Massad, E.: A Reed-Frost model taking into account uncertainties in the diagnostic of the infection. Bull. Math. Biol. 66, 689–706 (2004)
239. Merilan, J.E., Roe, D.J.: Application of fuzzy numbers to assess opinion in epidemiologic studies. Am. J. Epidemiol. 138(8), 663–664 (1993)
240. Miller, P.: Critiquing anesthetic management: The "ATTENDING" computer system. Anesthesiology 53, 362–369 (1983)
241. Miller, M.C., Westphal Jr., M.C., Reigart, J.R.: Mathematical models in medical diagnosis. Praeger Publishers, New York (1981)
242. Miller, R., McNeil, M., Challinor, S.: The INTERNIST-1/Quick medical reference project: status report. Western Journal of Medicine 145, 816–822 (1986)
243. Mir, A.H., Hanmandlu, M., Tandon, S.N.: Computed tomography image enhancement using the concept of fuzzy logic. Front Med. Biol. Eng. 7(2), 93–109 (1996)
244. Molnar, B., Szentirmay, Z., Bodo, M., Sugar, J., Feher, J.: Application of multivariate, fuzzy set and neural network analysis in quantitative cytological examinations. Anal. Cell. Pathol. 5(3), 161–175 (1993)
245. Morales, J., Buranr Jr., V.: Interactions between Cycloneda sanhuine and the Brown Citrus Aphid: adult feeding and larval mortality. Environmental Entomology 14(4), 520–522 (1985)
246. Morgenstern, H., Kleinbaum, D.G., Kupper, L.L.: Measures of disease of disease incidence used in epidemiologic research. Int. J. Epidemiol. 9(1), 97–104 (1980)

247. Mueller, M., Wagner, C.L., Annibale, D.J., Hulsey, T.C., Knapp, R.G., Almeida, J.S.: Predicting extubation outcome in preterm newborns: a comparison of neural networks with clinical expertise and statistical modeling. Pediatr. Res. 56(1), 11–18 (2004)
248. Murray, J.: Mathematical Biology. Springer, Berlin (1990)
249. Murray, C.J.L., Lopez, A.D.: The global burden of disease and injury series. Global Health Statistics, vol. II. Harvard University Press, USA (1996)
250. Murray, C.J.L.: Rethinking DALYs. In: Murray, C.J.L., Lopez, A.D. (eds.) The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020, pp. 1–89. Harvard University Press, Boston (1996)
251. Nascimento, L.F.C., Ortega, N.R.S.: Fuzzy Linguistic Model to Evaluation of the Risk of Neonatal Death. Rev. Saude. Publica. 36, 686–692 (2002)
252. Negoita, C.V., Ralescu, D.A.: Application of fuzzy sets to system analysis. Birkhäuser Verlag, Basel und Stuttgart, Germany (1975)
253. Nguyen, H.T., Walker, E.A.: A first course in fuzzy logic. Chapman & Hall, USA (2000)
254. Nokes, D.J., Swinton, J.: Vaccination in pulses: a strategy for global eradication of measles and polio? Trends Microbiol. 5, 14–19 (1997)
255. Norris, D., Pilsworth, B.W., Baldwin, J.F.: Medical diagnosis from patient records - a method using fuzzy discrimination and connectivity analyses. Fuzzy Sets Syst. 23, 73–87 (1987)
256. Nowak, M., Bangham, C.: Population dynamics if immune responses to persistent viruses. Science 272, 74–79 (1996)
257. Nowak, M.A.: The mathematical biology of human infections. Conservation Ecology 3(2), 12 (1999), `http://www.consecol.org/vol3/iss2/art12` (last visit in April 2008)
258. Ohayon, M.M.: Improving decision making processes with the fuzzy logic approach in the epidemiology of sleep disorders. J. Psychosom. Res. 47(4), 297–311 (1999)
259. Onho-Machado, L., Lacson, R., Massad, E.: Decision trees and fuzzy logic: a comparison of models for selection of measles vaccination strategies in Brazil. Journal of American Medical Informatics Society, S625–S629 (2000)
260. Okuda, T., Tanaka, H., Asai, L.: A formulation of fuzzy decision problems with fuzzy information using probability measures of fuzzy events. Inform. Control 38, 135–147 (1978)
261. Ortega, N.R.S., Sallum, P.C., Massad, E.: Fuzzy dynamical systems in epidemic modeling. Kybernetes: The International Journal of Systems and Cybernetics 29(1-2), 201–218 (2000)
262. Ortega, N.R.S., Barros, L.C., Massad, E.: Fuzzy gradual rules in epidemiology. Kybernetes: The International Journal of Systems and Cybernetics 32(4), 460–477 (2003)
263. Ortega, N.R.S., Santos, F.S., Zanetta, D.T., Massad, E.: A fuzzy Reed-Frost model for epidemic spreading. Bull. Math. Biol. (in press, 2008a)
264. Ortega, N.R.S., Massad, E., Struchiner, J.C.: A Bayesian approach to fuzzy hypotheses testing for the estimation of optimal age for vaccination against measles. Mathematics and Computers in Simulation (in press, 2008b)
265. Ouattara, D.A., Mhawej, M.J., Moog, C.H.: Clinical tests of therapeutical failures based on mathematical modeling of the HIV infection. Special Issue on Systems Biology, 230–241 (2008)

266. Papageorgiou, E.I., Stylios, C.D., Groumpos, P.P.: An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps. IEEE Trans. Biomed. Eng. 50(12), 1326–1339 (2003)
267. Papageorgiou, E.I., Spyridonos, P.P., Stylios, C.D., Ravazoula, P., Groumpos, P.P., Nikiforidis, G.N.: Advanced soft computing diagnosis method for tumour grading. Artif. Intell. Med. 36(1), 59–70 (2006a)
268. Papageorgiou, E.I., Stylios, C., Groumpos, P.: A combined Fuzzy Cognitive Map and decision trees model for medical decision making. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., vol. 1, pp. 6117–6120 (2006b)
269. Papageorgiou, E.I., Spyridonos, P.P., Glotsos, D.T., Stylios, C.D., Ravazoula, P., Nikiforidis, G.N., Groumpos, P.P.: Brain tumor characterization using the soft computing technique of fuzzy cognitive maps. Applied Soft Computing 8(1), 820–828 (2008)
270. Partin, A., Carter, H.B., Chan, D.W., Epstein, J.I., Oesterling, J.E., Rock, R.C., Weber, J.P., Walsh, P.C.: Prostate specific antigen in the staging of localized prostate cancer: influence of tumor differentiation, tumor volume and benign hyperplasia. J. Urol. 143(4), 747–752 (1990)
271. Partin, A., Kattan, M., Subong, E., Walsh, P., Wojno, K., Oesterling, J., Scardino, P., Pearson, J.: Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. JAMA 277(18), 1445–1451 (1997)
272. Partin, A., Mangold, L.A., Lamm, D.M., Walsh, P., Epstein, J., Pearson, J.: Contemporary update of prostate cancer staging nomograms Partin Tables for the new millennium. Urology 58(6), 843–848 (2001)
273. Patil, R., Szolovis, P., Schawartz, W.: Causal understanding of patient illness in medical diagnosis. In: Hayes, P.J. (ed.) Proceedings of the International Joint Conference on Artificial Intelligence. William Kaufmann, Vancouver, Canada (1981)
274. Patrick, D.L., Bush, J.W., Chen, M.M.: Toward an operational definition of health. J. Health Soc. Behav. 14, 6–23 (1973)
275. Patrick, D.L., Erickson, P.: Health status and heath policy: quality of life in health care evaluation and resource allocation. Oxford University Press, UK (1993)
276. Pearl, J.: Causality: models, reasoning, and inference. Cambridge University Press, New York (2000)
277. Pearson, D.W.: A property of linear fuzzy differential equations. Appl. Math. Lett. 10(3), 99–103 (1997)
278. Pearson, D.W., Dray, G., Peton, N.: On linear fuzzy dynamical systems. In: Pearson, D.W. (ed.) Proceedings of SOCO 1997 - Second International Computer Science Conventions Symposium on Soft Computing, pp. 203–209. Academic Press, Nimes, France (1997)
279. Pedrycz, W., Gomide, F.: An introduction to fuzzy sets: Analysis and design. MIT Press, Cambridge (1998)
280. Pedrycz, W., Gomide, F.: Fuzzy systems engineering: Toward human-centric computation. John Wiley & Sons Publications, USA (2007)
281. Peña-Reyes, C.A., Sipper, M.: A fuzzy-genetic approach to breast cancer diagnosis. Artif. Intell. Med. 17(2), 131–155 (1999)
282. Peña-Reyes, C.A., Sipper, M.: Evolutionary computation in medicine: an overview. Artif. Intell. Med. 19(1), 1–23 (2000)
283. Peixoto, M.S., Barros, L.C., Bassanezi, R.C.: Predator-Prey Fuzzy Model. Ecological Modeling (in press, 2008a)

284. Peixoto, M.S., Barros, L.C., Bassanezi, R.C.: A model of cellular automata for the spatial and temporal analysis of Citrus Sudden Death with the fuzzy parameter. Ecological Modeling (in press, 2008b)
285. Pereira, C.C., Machado, C.J., Rodrigues, R.N.: Profiles of multiple causes of death related to HIV/AIDS in the cities of São Paulo and Santos, Brazil. Cad Saúde Pública 23(3), 645–655 (2001) (in Portuguese)
286. Perelson, A., Nelson, P.W.: Mathematical analysis of HIV-1 dynamics in vivo. SIAM Review 41, 3–44 (1999)
287. Pesonen, E., Ohmann, C., Eskelinen, M., Juhola, M.: Diagnosis of acute appendicitis in two databases. Evaluation of different neighborhoods with an LVQ neural network. Methods Inf. Med. 37(1), 59–63 (1998)
288. Peterman, T., Curran, J.W.: Epidemiology of the acquired immunodeficiency syndrome AIDS. Epidemiology Reviews 7, 7–21 (1985)
289. Phelps, C.E., Hutson, A.: Estimating diagnostic test accuracy using a Fuzzy Gold Standard. Med. Decis. Making 15(1), 44–57 (1995)
290. Phillips, W.E., Phuphanich, S., Velthuizen, R.P., Silbiger, M.L.: Automatic magnetic resonance tissue characterization for three-dimensional magnetic resonance imaging of the brain. J. Neuroimaging 5(3), 171–177 (1995)
291. Picard, P., Lefèvre, C.: The dimension of Reed-Frost epidemic models with randomized susceptibility levels. Math. Biosci. 107, 225–233 (1991)
292. Pizzi, N., Choo, L.P., Mansfield, J., Jackson, M., Halliday, W.C., Somorjai, R.L.: Neural network classification of infrared spectra of control and Alzheimer's diseased tissue. Artif. Intell. Med. 7(1), 67–79 (1995)
293. Plaisier, A.P., Van Oortmarssen, G.J., Habbema, J.D.F., Remme, J., Alley, E.S.: ONCHOSIM: a model and computer simulation program for the transmission and control of onchocerciasis. Comput. Methods Programs Biomed. 31, 43–56 (1990)
294. Porter, R.: Cambridge illustrated history of medicine. Cambridge University Press, UK (1996)
295. Proudfoot, T.: Account of the epidemic cholera of Kendal. Edinburgh Medical and Surgical Journal 39, 70–103 (1833)
296. Puri, M.L., Ralescu, D.A.: Differentials of fuzzy functions. J. Math. Analysis and Applications 91, 552–558 (1983)
297. Quigley, R.J., Taylor, L.C.: Evaluating health impact assessment. Public Health 118(8), 544–552 (2004)
298. Ralescu, D.A., Adams, G.: The fuzzy integral. Journal of Mathematical Analysis and Applications 75, 562–570 (1980)
299. Reategui, E.B., Campbell, J.A., Leão, B.F.: Combining a neural network with case-based reasoning in a diagnostic system. Artif. Intell. Med. 9(1), 5–27 (1997)
300. Reed, S.C., Markowitz, L.E., Katz, S.L.: Measles vaccine. In: Plotkin, A.S., Orenstein, W.A. (eds.) Vaccines, pp. 222–266. Saunders, Philadelphia (1999)
301. Reis, M.A.M., Ortega, N.R.S., Silveira, P.S.P.: Fuzzy expert system in the prediction of neonatal resuscitation. Braz. J. Med. Biol. Res. 37, 755–764 (2004)
302. Reis, M.A.M., Ortega, N.R.S., Silveira, P.S.P.: A fuzzy sets theory application in the prediction of newborn's life-threatening situation in the delivery room. In: Liu, Y., Chen, G., Ying, M. (eds.) Fuzzy Logic, Soft Computing and Computational Intelligence - Proceedings of the Eleventh International Fuzzy Systems Association World Congress (IFSA 2005), pp. 1191–1194. Tsinghua University Press, Springer, Beijing, China (2005)
303. Reznik, L.: Fuzzy controllers. Newnes, Great Britain (1997)
304. Rocha, A.F.: Neural nets: a theory for brains and machines. Springer, USA (1992)

305. Rocha, A.F., Massad, E., Pereira Jr., A.: The brain: fuzzy arithmetic to quantum computing. Springer, USA (2004)
306. Rosenbaum, P.R., Rubin, D.B.: The central role of propensity score in observational studies for causal effects. Biometrika 70(1), 41–55 (1983)
307. Rothman, K.J., Greenland, S.: Modern epidemiology. Lippincott-Raven, Philadelphia (1998)
308. Rothman, K.J., Greenland, S., Lash, T.L.: Modern epidemiology. Lippincott Williams & Wilkins, Philadelphia (2008)
309. Rubin, D.B.: Comment: Neyman (1923) and causal inference in experiments and observational studies. Statistical Science 5(4), 472–480 (1990)
310. Rubin, D.B.: Randomization analysis of experimental data - The Fisher randomization test comment. J. Am. Stat. Assoc. 75(371), 591–593 (1980)
311. Rubin, D.B.: Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology 66(5), 688–701 (1974)
312. Rudzki, K., Hartleb, M., Sadowski, T., Rudzka, J.: Focal liver disease: neural network-aided diagnosis based on clinical and laboratory data. Gastroenterol Clin. Biol. 21(2), 98–102 (1997)
313. Sackett, D.L., Haynes, R.B., Guyatt, G.H., Tugwell, P.: Clinical epidemiology a basic science for clinical medicine. Little, Brown and Company, Boston (1991)
314. Sackett, D.L., Richardson, W.S., Rosenberg, W., Haynes, R.B.: Evidence-based medicine. How to practice and teach EBM. Churchill Livingstone, New York (1997)
315. Sadegh-Zadeh, K.: Fundamentals of clinical methodology: 1. Differential indication. Artif. Intell. Med. 6, 83–102 (1994)
316. Sadegh-Zadeh, K.: Fundamentals of clinical methodology: 2. Etiology. Artif. Intell. Med. 12, 227–270 (1998)
317. Sadegh-Zadeh, K.: Fundamentals of clinical methodology: 3. Nosology. Artif. Intell. Med. 17, 87–108 (1999)
318. Sadegh-Zadeh, K.: Fuzzy Genomes. Artif. Intell. Med. 18, 1–28 (2000)
319. Sallum, P.C., Almeida, M.F., Massad, E.: Rabies seroprevalence of street dogs from São Paulo City, Brazil. Prev. Vet. Med. 2000 44(3-4), 131–139 (2000)
320. Sanchez, E.: Inverse of fuzzy relations. Application to possibility distributions and medical diagnosis. Fuzzy sets and System 2, 75–86 (1979)
321. Sanchez, E.: Truth-qualification and fuzzy relations in natural languages, application to medical diagnosis. Fuzzy Sets Syst. 84, 155–167 (1996)
322. Sanchez, E.: Fuzzy logic and inflammatory protein variations. Clin. Chim. Acta. 271(1), 31–42 (1998)
323. Sánchez, R., Riquenes, A., Pérez-Abalo, M.: Automatic detection of auditory brainstem responses using feature vectors. Int. J. Biomed. Comput. 39(3), 287–297 (1995)
324. Saraoglu, H.M., Sanli, S.: A fuzzy logic-based decision support system on anesthetic depth control for helping anesthetists in surgeries. J. Med. Syst. 31(6), 511–519 (2007)
325. Sattenspiel, L., Simon, C.P.: The spread and persistence of infectious diseases in structured populations. Math. Biosci. 90, 341–366 (1988)
326. Scalia-Tomba, G.: Asymptotic final size distribution of the multitype Reed-Frost process. Journal of Applied Probability 23, 563–584 (1986)
327. Seikkala, S.: On the fuzzy initial value problem. Fuzzy Sets Syst. 24, 319–330 (1987)
328. Shafer, G.: A mathematical theory of evidence. Princeton University Press, Princeton (1976)

329. Shafer, G.: Belief functions and possibility measures. In: Bezdek, J.C. (ed.) Analysis of fuzzy information mathematics and logic, USA, pp. 51–84 (1987)
330. Shieh, J.S., Chang, L.W., Wang, M.S., Sun, W.Z., Wang, Y.P., Yang, Y.P.: Pain model and fuzzy logic patient-controlled analgesia in shock-wave lithotripsy. Med. Biol. Eng. Comput. 40(1), 128–136 (2002)
331. Shieh, J.S., Dai, C.Y., Wen, Y.R., Sun, W.Z.: A novel fuzzy pain demand index derived from patient-controlled analgesia for postoperative pain. IEEE Trans. Biomed. Eng. 54(12), 2123–2132 (2007)
332. Shimomura, K., Shono, H., Kohara, M., Uchiyama, A., Ito, Y., Sugimori, H.: Neonatal assessment using the Apgar fuzzy expert system. Comput. Biol. Med. 24(3), 171–178 (1994)
333. Shiomi, S., Kuroki, T., Jomura, H., Ueda, T., Ikeoka, N., Kobayashi, K., Ikeda, H., Ochi, H.: Diagnosis of chronic disease from liver scintiscans by fuzzy reasoning. J. Nucl. Med. 36(4), 593–598 (1995)
334. Shiomi, S., Kuroki, T., Kuriyama, M., Morikawa, H., Masaki, K., Ikeoka, N., Tanaka, T., Ikeda, H., Ochi, H.: Diagnosis of chronic liver disease from liver scintiscans by artificial neural networks. Ann. Nucl. Med. 11(2), 75–80 (1997)
335. Shirahama, N., Yanaru, T., Nagamatsu, M.: An emotion-processing system based on fuzzy inference and subjective observations. In: Teodorescu, H.N., Kandel, A., Jain, L.C. (eds.) Soft computing in human-related sciences, pp. 293–334. CRC Press, USA (1999)
336. Shortliffe, E.H.: Computer based medical consultations: mycin. Elsevier-North Holland, New York (1976)
337. Shortliffe, E.H.: Update on ONCOCIN: a chemotherapy advisor for clinical oncology. Med. Inform. (Lond.) 11(1), 19–21 (1986)
338. Shortliffe, E.H.: Clinical decision-support systems. In: Shortliffe, E.H., Perreault, L.E. (eds.) Medical Informatics: Computer Applications in Health Care, pp. 466–502. Addison-Wesley, Reading (2001)
339. Shtovba, S., Chernovolik, G.: Fuzzy rule-based system for decision making support of pahtology anatomist. In: Adlassnig, K.P. (ed.) Fuzzy Diagnostic and Therapeutic Decision Support, pp. 45–50. Österreichische Computer Gesellschaft, Wien (2000)
340. Silveira, P.S.P., Yang, H.M., Massad, E.: Computer-Based environment for the study of ecological system. Environment Management and Health 6(4), 19–28 (1995)
341. Silveira, P.S.P., Yang, H.M., Azevedo-Neto, R.S., Massad, E.: Mathematical and computer models for infective contact rates. Mathematical Modeling Scientific Computing 6, 793–798 (1996)
342. Smets, P.: Medical diagnosis: fuzzy sets and degrees of belief. Fuzzy Sets Syst. 5, 259–266 (1981)
343. Smets, P.: Belief functions: the disjunctive rules of combination and generalized Bayesian Theorem. Internet J. Approx. Reason. 9, 1–35 (1993)
344. Smith, G.D., Ebrahim, S.: Epidemiology - is it time to call it a day? Int. J. Epidemiol. 30, 1–11 (2001)
345. Snow, J.: On the mode of communication of cholera. John Churchill, London (1855)
346. Somorjai, R.L., Nikulin, A.E., Pizzi, N., Jackson, D., Scarth, G., Dolenko, B., Gordon, H., Russell, P., Lean, C.L., Delbridge, L.: Computerized consensus diagnosis: a classification strategy for the robust analysis of MR spectra. Application to 1H spectra of thyroid neoplasms. Magn. Reason. Med. 33(2), 257–263 (1995)

347. Sousa, C.A., Duarte, P.S., Pereira, J.C.R.: Fuzzy logic and logistic regression in the decision making for parathyroid scintingraphy study. Rev Saude Publica 40(5), 898–906 (2006)
348. Spiegel, M.R.: Theory and problems of probability and statistics. McGraw-Hill, New York (1992)
349. Splawa-Neyman, J.: On the Application of Probability Theory to Agricultural Experiments. Essay on Principles, Section 9, English translation (1990). In: Dabrowska, D.M., Speed, T.P. (eds.) Statistical Science, vol. 5(4), pp. 465–472 (1923)
350. Steimann, F.: The interpretation of time-varying data with DIAMON-1. Artif. Intell. Med. 8(4), 343–357 (1996)
351. Stewart, J.: Calculus - early transcendentals. Brooks/Cole Publishing Company, USA (1999)
352. Straszecka, E.: Membership functions and a basic probability assignment as certainty measures for a medical diagnosis support. In: Adlassnig, K.P. (ed.) Fuzzy Diagnostic and Therapeutic Decision Support, pp. 138–143. Österreichische Computer Gesellschaft, Wien (2000)
353. Stroke, G.W.: The dramatic change from empirical to scientific technology in medicine: computerized imaging. Keio. J. Med. 42(2), 70–77 (1993)
354. Struchiner, C.J., Halloran, M.E., Brunet, R.C., Ribeiro, J.M., Massad, E.: Malaria vaccines: lessons from field trials. Cad Saude Publica 10(2), 310–326 (1994)
355. Struchiner, C.J., Ortega, N.R.S., Massad, E., Burattini, M.N.: Fuzzy risk estimators in epidemiology. In: Pearson, D.W. (ed.) Proceedings of SOCO 1997 - Second International Computer Science Conventions Symposium on Soft Computing, pp. 64–68. Academic Press, Nimes, France (1997)
356. Struchiner, C.J., Luz, P.M., Dourado, I., Sato, H.K., Aguiar, S.G., Ribeiro, J.G.L., Soares, R.C.R., Codeco, T.C.: Risk of fatal adverse events associated with 17DD yellow fever vaccine. Epidemiol. Infect. 132(5), 939–946 (2004)
357. Struchiner, C.J., Halloran, M.E.: Randomization and baseline transmission in vaccine field trials. Epidemiol. Infect. 135(2), 181–194 (2007)
358. Sudcamp, T.: On probability-possibility transformations. Fuzzy Sets Syst. 51, 73–81 (1992)
359. Sugeno, M.: Theory of fuzzy integrals and applications, Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, Japan (1974)
360. Sugeno, M., Takagi, T.: A new approach to design of fuzzy controller. In: Wang, P.P. (ed.) Advances in fuzzy sets, possibility theory and applications, pp. 325–334. Plenum Press, New York (1983)
361. Sugeno, M., Kang, G.T.: Structure identification of fuzzy model. Fuzzy Sets Syst. 28(1), 15–33 (1988)
362. Surynarayanan, R., Reddy, N.P., Canilang, E.P.: A fuzzy logic diagnosis system for classification of pharyngeal dysphagia. Int. J. Biomed. Comput. 38(3), 207–215 (1995)
363. Sutherland, J.: Appendix (A) to report of the general board of health on the epidemic cholera of 1848-1849. W Clowes & Sons, Stamford Street, for Her Majesty's Stationery Office (HMSO), London (1850)
364. Svirezhev, Y.M., Logofet, D.O.: Stability of Biological Communities. Mir publishers, Moscow (1983)
365. Sylvan, R., Abe, J.M.: On general annotated logics, with an introduction to full accounting logics. Bulletin of Symbolic Logic 2, 118–119 (1996)
366. Szczepaniak, P.S., Lisboa, P.J.G., Kacprzyk, J.: Fuzzy systems in Medicine. Physica-Verlag, Springer, Heidelberg, Germany (2000)

367. Sztandera, L.M., Goodenday, L.S., Cios, K.J.: A neuro-fuzzy algorithm for diagnosis of coronary artery stenosis. Comput. Biol. Med. 26(2), 97–111 (1996)
368. Taheri, S.M., Behboodian, J.: A Bayesian approach to fuzzy hypotheses testing. Fuzzy Sets Syst. 123, 39–48 (2001)
369. Tanaka, H., Ortega, N.R.S., Galizia, M., Borges, J.B., Amato, M.B.P.: Fuzzy modeling of electrical impedance tomography image of the lungs. Clinics (in press in 63(3), 2008)
370. Tanaka, H., Okuda, T., Asai, K.: Fuzzy information and decision in a statistical model. In: Gupta, M.M., Ragade, R.K., Yager, R.R. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 303–320. North-Holland, Amsterdam (1979)
371. Tanaka, T., Miwa, K., Kanda, S.: Application of fuzzy reasoning in an expert system for ultrasonography. Dentomaxillofac Radiol. 26(2), 125–131 (1997)
372. Tanner, J.T.: The stability and the intrinsic growth rates of prey and predator populations. Ecology 56, 855–867 (1975)
373. Teodorescu, H.-N., Kandel, A., Jain, L.C.: Soft computing in human-related sciences. CRC Press LLC, USA (1999a)
374. Teodorescu, H.-N., Kandel, A., Jain, L.C.: Fuzzy and neuro-fuzzy systems in medicine. CRC Press LLC, USA (1999b)
375. Thayer, J.F.: A hybrid approach to clustering biomedical data. Biomed. Sci. Instrum. 32, 39–46 (1996)
376. Torasso, P.: Supervising the heuristic learning in a diagnostic expert system. Fuzzy Sets Syst. 44, 357–372 (1991)
377. Tourassi, G.D., Floyd, C.E., Coleman, R.E.: Acute pulmonary embolism: cost-effectiveness analysis of the effect of artificial neural networks on patient care. Radiology 206(1), 81–88 (1998)
378. Tu, Z., Toga, A.W.: Towards whole brain segmentation by a hybrid model. Med. Image Comput. Comput. Assist. Interv. 10(2), 169–177 (2007)
379. Turelli, M.: Stochastic community theory: a partially guided tour. Biomathematics 17, 321–339 (1986)
380. Uemura, Y.: A decision rule on fuzzy events. Japanese J. Fuzzy Theory System 3, 291–300 (1991)
381. Venta, L.A., Salchenverger, L.M., Venta, E.R.: Improved diagnosis of breast implant rupture with sonographic findings and artificial neural networks. Acad. Radiol. 5(4), 238–244 (1998)
382. Viertl, R.: Statistical methods for non-precise data. CRC Press, Boca Raton (1996)
383. Xu, L.S., Meng, M.Q., Wang, K.Q.: Pulse image recognition using fuzzy neural network. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 3148–3151 (2007)
384. Wang, L., Mendel, J.M.: Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man. Cybern. 22(6), 1411–1427 (1992)
385. Wang, H., Smarandache, F., Zhang, Y.-Q., Sunderraman, R.: Interval neutrosophic sets and logic: Theory and applications in computing, Hexis, USA (2005)
386. Weeks, R.M., Barenzi, J.F., Wayira, J.R.: A low-cost, community-based measles outbreak investigation with follow-up action. Bull WHO 70, 317–321 (1992)
387. Weiss, S.M., Kulikowski, C.A., Amarel, S., Safir, A.: A model-based method for computer-aided medical decision-making. Artificial Intelligence 11(1-2), 145–172 (1978a)
388. Weiss, S.M., Kulikowski, C.A., Safir, A.: Glaucoma consultation by computer. Comput. Biol. Med. 8(1), 25–40 (1978b)
389. WER: Severe acute respiratory syndrome (SARS). Wkly. Epidemiol. Rec. 78(12), 81–88 (2003)

390. WHO: Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS), World Health Organization (2003), http://www.who.int/csr/sars/country/en/ (last visit in April 2008)
391. Winslow, C.-E.A., Smillie, W.G., Doull, J.A., Gordon, J.E., Top, F.H.: The history of american epidemiology. Mosby CV Company, USA (1952)
392. Yager, R.R.: A characterization of the extension principle. Fuzzy Sets Syst. 18(3), 205–217 (1986)
393. Yager, R.R., Filev, D.P.: Essentials of fuzzy modeling and control. Wiley-Interscience, USA (1994)
394. Yen, J., Langari, R.: Fuzzy logic: intelligence, control, and information. Prentice Hall, USA (1999)
395. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)
396. Zadeh, L.A.: Biological applications of the theory of fuzzy sets and systems. In: Proctor, L.D. (ed.) Biocybernetics of the Central Nervous System, pp. 199–212. Little Brown & Co., Boston (1969)
397. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Syst. Man. Cybern. SIC-3, 28–44 (1973)
398. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning I. Information Science 8, 199–251 (1975a)
399. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning II. Information Science 8, 301–357 (1975b)
400. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning III. Information Science 9, 43–80 (1975c)
401. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets Syst. 1(1), 3–28 (1978)
402. Zanetta, R.A.C., Amaku, M., Azevedo-Neto, R.S., Zanetta, D.M.T., Burattini, M.N., Massad, E.: Optimal age for vaccination against measles in the State of São Paulo, Brazil, taking into account the mother's serostatus. Vaccine 20, 226–234 (2002)
403. Zarkogianni, K., Mougiakakou, S.G., Prountzou, A., Vazeou, A., Bartsocas, C.S., Nikita, K.S.: An insulin infusion advisory system for type 1 diabetes patients based on non-linear model predictive control methods. In: Conf. Proc. IEEE Eng. Med. Biol. Soc., pp. 5972–5975 (2007)
404. Zimmermann, H.J.: Operators in models decision making. In: Dubois, D., Prade, H., Yafer, R.R. (eds.) Fuzzy Information Engineering, pp. 471–496. Wiley, USA (1996)

Index