Book review of

Uncertainty Models for Knowledge-Based Systems

I.R. Goodman - H.T. Nguyen

North-Holland, Amsterdam, 1985

The problem of handling uncertainty in knowledge-based systems is one of the most appropriate grounds for the meeting of several distinct traditions in mathematics, especially measure theory and logic. The book by Goodman and Nguyen is very ambitious. Its aim is to tightly tie these two traditions by relating the notions of random set and fuzzy set membership function (also called "dispersion" in the book). It is based on the idea that random set theory is a good setting for unifying various uncertainty theories including probability theory, Shafer's belief functions and Zadeh's fuzzy set and possibility theories. This idea is indeed very important for any one who is interested in clarifying many lively (and sometimes sterile) debates between probability and fuzzy sets. Informally, the main results of the book are as follows : a fuzzy set is equivalent to a class of random sets, including a nested random set obtained as a convex combination of levels-cuts ; a probability measure is equivalent to a random singleton ; Shafer's belief functions are equivalent to random sets in a finite setting, fuzzy set theoretic operations (at least some of them) can be interpreted in terms of random set combinations. Dempster rule of combination is an intersection of statistically independent random sets. These few sentences roughly summarize the main messages of that thick book.

While the intended purpose of the book is definitely justified, and worthwhile, the title of the book may be misleading for an application-oriented reader. Although random set theory (at least an elementary version of it) is a useful background for any scholar involved in uncertain knowledge-based systems research, this book is not exactly about knowledge-based systems, and not primarily meant for knowledge enginers. This is a monograph for mathematicians, and seemingly good mathematicians, having background in category theory, logic, measure theory at a rather high level. The density of mathematical symbols and its variety is astonishing in many pages. It prevents any useful reading for the non-mathematician, most of the time. A regular artificial intelligence reader (with a limited background in mathematics) may get a taste of the intended purpose of the book by considering only the non-mathematically formalized parts of the text. Anyway this review does not adopt the point of view of the mathematician.

Chapter one is a brief introduction where the reader can indeed capture some of the authors motivation for this book. Chapter two is especially lengthy and abstract in its presentation. It proposes a categorial setting for a formal language powerful enough to grasp the random set theory, and to encompass most many valued logics. It introduces dispersions, as membership functions of generalized sets, and a generalized calculus of dispersions, i.e. definitions for usual connectives such as conjunction, disjunction and negation, which rely on the triangular-norm setting of Schweizer and Sklar. It also deals in a very abstract way with generalized versions of conditioning, projections, Bayes theorem, quantifiers. Comparisons of the categorial setting with previous ones due to Higgs and Goguen are made. The end of this chapter indicates that what is obtained is a general tool for natural language representation and approximate reasoning encompassing Zadeh's meaning representation language PRUF. The reader can try to get a flavor of the usefulness of the proposed approach on the 43 examples of natural language sentences dealt with at the end of the chapter.

Chapter 3 devoted to uncertainty measures contrasts by its brevity with the preceding one. It gives a quick overview of several topics in uncertainty modelling such as Kampé de Fériet and Forte's generalized information theory, and possibility theory where the minimum operation is optionally changed into a product. The treatment of the links between these theories is only touched upon and considered at a deeper level further on. Chapter 4 provides a high-level introduction to random set theory which consists of 11 theorems, supported by an appendix on topology. The topics considered in that section are the links with Dempster upper and lower probabilities, the approach to uncertainty measures as random set coverage functions (e.g. the degree of uncertainty g(C) of event C viewed as the probability that C contains or intersects a random set S), random variables as singleton-valued random sets, the links with Choquet capacities and with the integration theory for set- valued mappings as studied by Debreu and Aumann. Chapter 5 is devoted to the relationship between membership functions and random sets. It reports previous works of the authors starting from the fact that a fuzzy set F on X is equivalent to a class of random sets S which satisfy the following identity

$$\forall x \in X \ \mu_F(x) = \text{Prob}(x \in S)$$

$$= \text{Plausibility}(\{x\}) \text{ in Shafer terminology}$$

The random set equivalent to F with maximum random set entropy is also characterized, in the finite setting.

Chapter 6 is devoted to homomorphisms between fuzzy set theoretic operators and random set counterparts. Things behave nicely only for a subclass of fuzzy set operations (those based on the minimum and the product, for intersection). Chapter 7 criticizes several alternative theories of uncertainty which were originally motivated by a desire to argue against fuzzy set theory. The authors assess such critical comments under the light of the random set setting. The following works are reviewed : Manes' fuzzy theories, Watanabe's non-distributive logic, Schefe's agreement probabilities, and Gaines's uncertainty logic. Chapter 8 on inference is the closest to knowledge-based systems concerns ; it is rather concise and scatters over several issues which are not directly related to each other : a decision theory based on dispersions, the combination of evidence in random set theory, counterpart of asymptotic results in probability theory, and a brief discussion on the subjective view of probability theory, pushing forward Giles' ideas on the normative foundations of possibility theory. Only the second of these points is of direct relevance to knowledge-based systems : the authors review the three well-know approaches to the combination of evidence (MYCIN rule, Dempster rule, and the Bayesian rule based on conditional independence assumption) and promote the idea of combining the dispersions of random sets (thus encompassing the set theoretic combination of fuzzy sets). However no attempt is made to compare the relevance of these approaches, and difficulties encountered by these approaches as methodologies for controlling uncertainty propagation are not evoked. Chapter 9 is an attempt to illustrate the random set approach on a case study in data association. Chapter 10 contains a technical summary and it also proposes a listing of open problems and research issues, for the working mathematician mainly.

On the whole it should be clear from this review that the strong features of the book lie at the theoretical level, and that its direct relevance for knowledge-based system practitioners is rather limited.

Didier DUBOIS - Henri PRADE