Approximate Reasoning Systems: A Personal Perspective

Piero P. Bonissone

Artificial Intelligence Laboratory
General Electric Corporate R&D
Schenectady, New York 12301
Bonissone@crd.ge.com

Introduction

In this presentation we provide a personal perspective of the progress made in the field of approximate reasoning systems. Because of time and space limitations, we will limit the scope of our discussion to cover the most notable trends and efforts in reasoning with uncertainty and vagueness. The existing approaches to representing this type of information can be subdivided in two basic categories according to their qualitative or quantitative characterizations of uncertainty.

Models based on qualitative approaches are usually designed to handle the aspect of uncertainty derived from the incompleteness of the information, such as Reasoned Assumptions (Doyle, 1983), and Default Reasoning (Reiter, 1980). With a few exceptions, they are generally inadequate to handle the case of imprecise information, as they lack any measure to quantify confidence levels (Doyle, 1983). A few approaches in this group have addressed the representation of uncertainty, using either a formal representation, such as Knowledge and Belief (Halpern and Moses, 1986), or a heuristic representation, such as the Theory of Endorsements (Cohen, 1985).

We will further limit our presentation by focusing on the development the quantitative approaches. Over the past few years, quantitative uncertainty management has received a vast amount of attention from the researchers in the field, (Shachter *et al.*, 1990, Henrion *et al.*, 1990, Bonissone *et al.*, 1991, D'Ambrosio *et al.*, 1991), leading to the establishment of two well-defined approaches, that differ in the semantics of their numerical representation. One is the *probabilistic* reasoning approach, based on probability theory. The other one is the *possibilistic* reasoning approach, based on the semantics of fuzzy sets and many-valued logics. In this paper we will illustrate and compare these approaches and we will conclude with a review of efforts aimed at improving their real-time performance.

Approximate Reasoning Systems

Reasoning systems must attach a truth value to statements about the state or the behavior of a real world system. When this hypothesis evaluation is not possible due to the lack of complete and certain information, approximate reasoning techniques are used to determine a set of possibilities (possible worlds) that are logically consistent with the available information. These possible worlds are characterized by a set of propositional variables and their associated values. As it is generally impractical to describe these possible worlds to an acceptable level of detail, approximate reasoning techniques seek to determine some properties of the set of possible solutions or some constraints on the values of such properties. The above possible-world interpretation for approximate reasoning systems was originally proposed by Ruspini (Ruspini, 1987, Ruspini, 1990, Ruspini, 1991).

A large number of approximate reasoning techniques have been developed over the past decade to provide these solutions. The reader is referred to references (Bonissone, 1987a, Pearl, 1988a) for a survey. We will now analyze the two most common approximate reasoning techniques: probabilistic and possibilistic reasoning.

Probabilistic Reasoning. Probability-based, or probabilistic reasoning seeks to describe the constraints on the variables that characterize the possible worlds by identifying their conditional probability distributions given the evidence in hand. Its supporting formalisms are based on the concept of set-measures, additive real functions defined over certain subsets of some space. Probabilistic methods seldom make categorical assertions about the actual state of the system being investigated. Rather, they indicate that there is an experimentally determined (or believed) tendency or propensity for the system to be in some specified state. Thus, they are oriented primarily toward decisions that are optimal in the long run, describing the tendency or propensity of truth of a proposition without assuring its actual validity. Depending on the nature of the information, probabilistic reasoning estimates the frequency of the truth of a hypothesis as determined by prior observation (objectivist interpretation) or a degree of gamble based on the actual truth of the hypothesis (subjectivist interpretation).

The basic inferential mechanism used in probabilistic reasoning is the conditioning operation.

Possibilistic Reasoning. Conversely, possibilistic reasoning, which is rooted in fuzzy set theory (Zadeh,

1965) and many-valued logics, seeks to describe the constraints on the values of the variables of the possible worlds in terms of their similarity to other sets of possible worlds. The supporting formalisms are based on the mathematical concept of set distances or metrics instead of set measures. These methods focus on single situations and cases. Rather than measuring the tendency of the given proposition to be valid, they seek to find another related, similar proposition that is valid. This proposition is usually less specific and resembles (according to some measure of similarity) the original hypothesis of interest.

The basic inferential mechanism used in possibilistic reasoning is the *generalized modus-ponens* (Zadeh, 1979), which makes use of inferential chains and the properties of a similarity function to relate the state of affairs in the two worlds that are at the extremes of an inferential chain.

It should be noted that there are other interpretations for possibilistic reasoning which are not based on possible-world semantics. Of special interest is Zadeh's interpretation of a possibility distribution "as a fuzzy restriction which acts as an elastic constraint on the values that may be assigned to a variable" (Zadeh, 1978). This interpretation, in conjunction with the linguistic variable approach (Zadeh, 1979), is the basis for the development of numerous fuzzy control applications (Sugeno, 1985).

As a final comment regarding the proposed typology for approximate reasoning systems, we want to consider those hybrid situations which require the simultaneous representation of probability and possibility. Such cases have been analyzed by Zadeh in the definition of the probability measure of a fuzzy events (Zadeh, 1968) and by Smets in the extension of belief functions to fuzzy sets (Smets, 1988).

Given the duality of purpose and characteristics between probabilistic and possibilistic methods, we can conclude that these technologies ought to be regarded as being complementary rather than competitive.

Probabilistic Approaches

Some of the earliest techniques found among the approaches derived from probability are based on single-valued representations. These techniques started from approximate methods, such as the modified Bayesian rule (Duda *et al.*, 1976) and confirmation theory (Shortliffe and Buchanan, 1975), and evolved into formal methods for propagating probability values over Bayesian Belief Networks (Pearl, 1982, Pearl, 1988a),

Another trend among the probabilistic approaches is represented by interval-valued representations such as Dempster-Shafer theory (Dempster, 1967, Shafer, 1976, Lowrance *et al.*, 1986).

Bayesian Belief Networks.

Over the last five years, considerable efforts have been devoted to improve the computational efficiency of Bayesian belief networks for trees and small poly-trees, and for directed acyclic graphs (influence diagrams) (Howard and Matheson, 1984), (Schachter, 1986).

An efficient propagation of belief on Bayesian Networks has been originally proposed by J. Pearl (Pearl, 1982). In his work, Pearl describes an efficient updating scheme for trees and, to a lesser extent, for poly-trees (Kim and Pearl, 1983, Pearl, 1988b). However, as the complexity of the graph increases from trees to poly-trees to general graphs, so does the computational complexity. The complexity for trees is $O(n^2)$, where n is the number of values per node in the tree. The complexity for poly-trees is $O(K^m)$, where K is the number of values per parent node and m is the number of parents per child. This number is the size of the table attached to each node. Since the table must be constructed manually, (and updated automatically), it is reasonable to assume it to be small. The complexity for multi-connected graphs is $O(K^n)$, where K is the number of values per node and n is the size of the largest non-decomposable subgraph. To handle such complexity, techniques such as moralization and propagation in a tree of cliques (Lauritzen and Spiegelhalter, 1988) and loop cutset conditioning (Suermondt *et al.*, 1991, Stillman, 1991) are typically used to decompose the original problem (graph) into a set of smaller problems (subgraphs).

When this problem decomposition process is not possible, exact methods must be abandoned in favor of approximate methods. Among these methods the most common are clustering, bounding conditioning (Horvitz *et al.*, 1989), and simulation techniques, such as logic samplings and Markov simulations (Henrion, 1989). Figure 1 (provided by Max Henrion) illustrates a taxonomy of these Bayesian inference mechanisms.

Dempster-Shafer Theory (Belief Functions)

Belief functions have been introduced in an axiomatic manner by Shafer (Shafer, 1976). Their original purpose was to compute the degree of belief of statements made by different sources (or witnesses) from a subjective probability of the sources' reliability.

Many other interpretations of belief functions have also been presented, ranging from functions induced from some probability measure and multivalued mappings (Dempster, 1967) or compatibility relations (Lowrance *et al.*, 1986), to probability of provability (Pearl, 1988b), to inner measures (Ruspini, 1987, Fagin and Halpern, 1989), to a non-probabilistic model of transferable belief (Smets, 1991).

All these interpretation share the same static component of the theory: the Möbius Transform that defines a mapping from basic probability assignments (masses assigned to subsets of the frame of discernment) to the computation of the lower bound (belief) of a proposition (a region defined in the same frame of discernment). An inverse Möbius transform can be used to recover the masses from the belief. All these interpretations also share the same definition of the upper bound (usually referred to as plausibility).

More specifically, this formalism defines a function

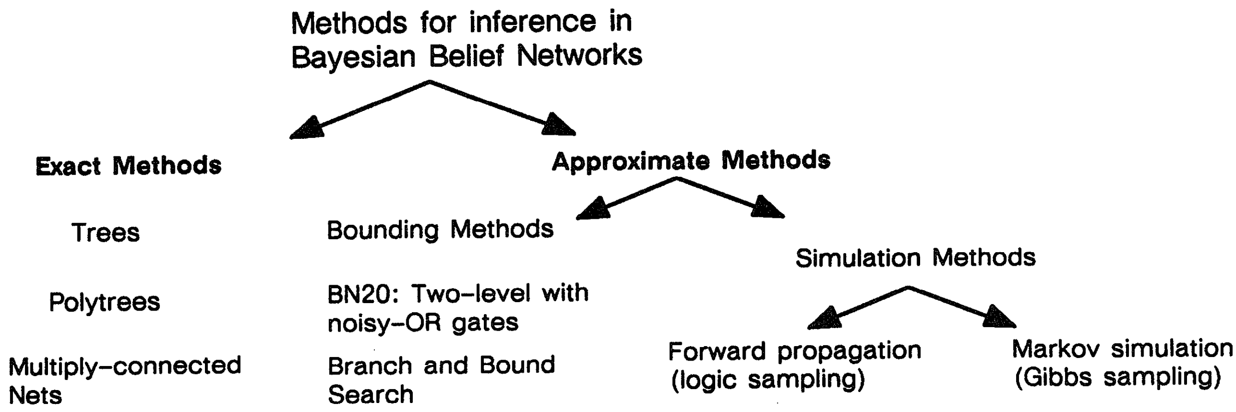

Figure 1: Taxonomy of Inference Mechanisms for Bayesian Belief Networks

that maps subsets of a space of propositions Θ on the [0,1] scale. The sets of partial beliefs are represented by mass distributions of a unit of belief across the propositions in Θ. This distribution is called basic probability assignment (bpa). The total certainty over the space is 1. A non-zero *bpa* can be given to the entire space Θ to represent the degree of ignorance. Given a space of propositions Θ, referred to as frame of discernment, a function $m : 2^{\Theta} \rightarrow [0,1]$ is called a basic probability assignment if it satisfies the following three conditions:

$$m(\phi) = 0 \quad \text{where } \phi \text{ is the empty set} \tag{1}$$

$$0 < m(A) < 1 \tag{2}$$

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{3}$$

The certainty of any proposition A is then represented by the interval $[Bel(A), P^*(A)]$, where $Bel(A)$ and $P^*(A)$ are defined as:

$$Bel(A) = \sum_{x \subseteq A} m(x) \tag{4}$$

$$P^*(A) = \sum_{x \cap A \neq \phi} m(x) \tag{5}$$

From the above definitions the following relation can be derived:

$$Bel(A) = 1 - P^*(\neg A) \tag{6}$$

Equations 4 and 5 represent the *static* component of the theory, which is common to all interpretations, However, these interpretations do not share the same *dynamic* component of the theory: the process of updating (i.e., conditioning or evidence combination). This issue has been recently addressed by various researchers (Halpern and Fagin, 1990, Smets, 1991).

Given the beliefs (or masses) for two propositions A and B, Dempster's rule of combination can be used (under assumptions of independence) to derive their combined belief (or mass).

If m_1 and m_2 are two *bpas* induced from two independent sources, a third *bpa*, $m(C)$, expressing the pooling of the evidence from the two sources, can be computed by using Dempster's rule of combination:

$$m(C) = \frac{\sum_{A_i \cap B_j = C} m_1(A_i) \cdot m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i) \cdot m_2(B_j)} \tag{7}$$

If proposition B is true (i.e, event B has occurred), then $Bel(B) = 1$ and from Demspster rule of combination, we can derive a formula for conditioning A given B, $Bel(A \mid B)$:

$$Bel(A \mid B) = \frac{Bel(A \cup \neg B) - Bel(\neg B)}{1 - Bel(\neg B)} \tag{8}$$

This expression is compatible with the interpretation of Belief as evidence, and as inner measure. However, this expression is not compatible with the interpretation of belief as the lower envelope of a family of probability distributions. Under such interpretation, the correct expression for conditioning is:

$$Bel(A \parallel B) = \frac{Bel(A \cap B)}{Bel(A \cap B) + Pl(\neg A \cap B)} \tag{9}$$

The interested reader is referred to reference (Shafer, 1990) for a very clear explanation and an updated bibliography on belief functions.

As for the case of belief networks, a variety of exact and approximate methods have been proposed to perform inferences using belief functions. Typically, the exact methods require additional constraints on the structure of the evidence. Figure 2 illustrates a taxonomy of Dempster-Shafer inference mechanisms.

D-S Exact Methods In the general case, the evaluation of the degrees of belief requires time exponential in $|\Theta|$, the cardinality of the frame of discernment. This is caused by the need of possibly enumerating all the subset and superset of a given set. Barnett (Barnett, 1981) showed that, when the frame of discernment is discrete and each piece of evidence supports only a singleton proposition or its negation, the

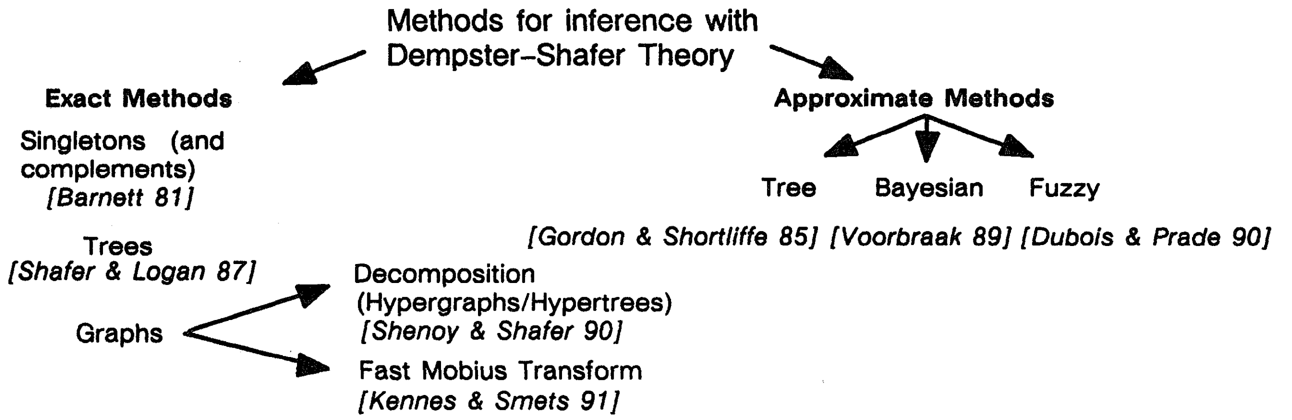

Figure 2: Taxonomy of Inference Mechanisms for Dempster-Shafer

computational-time complexity can be reduced from exponential to linear by combining the belief functions in a simplifying order.

Strat (Strat, 1984) proved that, when the frame of discernment is continuous and the evidence supports only contiguous intervals along the number line, the complexity can be reduced to $O(n^2)$, where n is the number of atomic propositions, i.e., intervals of unit length.

By imposing a tree structure restriction on the evidence, Shafer and Logan (Shafer and Logan, 1987) implemented a system which derives exact Belief functions. Hierarchical evidence enables a partitioning of the frame of discernment Θ and a great resulting efficiency over the unrestricted domain of size $|\Theta|$. While the unrestricted D-S calculus is NP-complete, the Shafer and Logan algorithm is order $O(nf)$, where $n = |\Theta|$ and f is the branching factor of the tree structure.

Shenoy and Shafer have addressed the issue of computing belief functions over an unrestricted graph, by creating a hypergraph, decomposing it into smaller subgraphs (via a covering hypertree), and finally using the Markov tree associated with the hypertree to perform local propagations (Shenoy and Shafer, 1990).

Another useful algorithm aimed at decreasing the computational complexity of belief propagation is the Fast Möbius Transform (Smets and Kennes, 1991), which is the Fast Fourier transform equivalent for the Möbius transform described by equation 4.

D-S Approximate Methods There are many *approximate* methods to efficiently compute belief functions. Because of space constraints we will only cite a few. Using a tree structure restriction, Gordon and Shortliffe (Gordon and Shortliffe, 1985) propose an algorithm to compute approximate Belief functions when different pieces of evidence are relevant to different levels of abstraction in a hierarchy of diseases. A Bayesian approximation in which the probability mass is distributed uniformly among the elements of the subset of Θ has been proposed by Voorbraak (Voorbraak, 1989). Finally, Dubois and Prade (Dubois and Prade, 1990) propose a consonant approximation of Belief functions which is mathematically equivalent to a fuzzy set.

Decisions with Belief Functions Belief theory is a relatively new theory and, unlike probability, does not have yet a fully developed decision theory. Initial work in this direction has been proposed by Jaffray (Jaffray, 1989) and Strat (Strat, 1990).

Possibilistic Approaches

Among the possibilistic approaches, the most notable ones are based on a fuzzy-valued representation of uncertainty. These include the Possibility Theory and Linguistic Variable Approach (Zadeh, 1978, Zadeh, 1979), and the Triangular-norm based approach (Bonissone, 1987b, Bonissone and Decker, 1986, Bonissone *et al.*, 1987).

The approach proposed by Zadeh is based on the notion that fuzziness is fundamentally different from randomness and, as such, cannot be modeled by traditional probabilistic techniques. The distinction between randomness and fuzziness is based on the different types of uncertainty captured by each concept. In *randomness*, the uncertainty is derived by the nondeterministic membership of a point (in the sample space) to a well-defined region in that space. The sample space represents the set of possible values for the random variable. The well-defined region represents an event. The region's characteristic function creates a dichotomy in the universe of discourse of the possible values for the random variable: when the point falls within the boundary of the region, it fully belongs to such region. Otherwise the point's membership in the region is zero. Probability measures the tendency (frequency) with which such random variable takes values inside the region.

In *fuzziness*, the uncertainty is derived from the partial membership of a point (in the universe of discourse) to an imprecisely defined region in that space. The region represents a fuzzy set. The characteristic function of the fuzzy set does not create a dichotomy in the universe of discourse: it maps the universe of discourse into the interval [0,1] instead of the set {0,1}. The partial membership of the point does not represent any frequency. It defines the degree to which that particular value of the universe of discourse satisfies the property which characterizes the fuzzy set.

Our space constraints prevent us from citing the numerous fuzzy sets literature. However, we would like to refer the reader to a compilation of selected papers by Zadeh, which provides a chronological description of the evolution of this theory (Yager *et al.*, 1987).

We will conclude the analysis of possibilistic reasoning by briefly discussing an approach based on many-valued logic operators (Triangular norms) and the generalized *modus ponens.*

Triangular-Norm Based Reasoning Systems

Triangular norms (T-norms) and their dual T-conorms are two-place functions from [0,1]x[0,1] to [0,1] that are monotonic, commutative and associative (Schweizer and Sklar, 1963, Schweizer and Sklar, 1983, Bonissone and Decker, 1986, Bonissone, 1987b). They are the most general families of binary functions which satisfy the requirements of the conjunction and disjunction operators respectively. Their corresponding boundary conditions satisfy the truth tables of the Boolean AND and OR operators.

Any triangular norm $T(A,B)$ falls in the interval $T_w(A,B) \leq T(A,B) \leq Min(A,B)$, where

$$T_w(A,B) = \begin{cases} \min(A,B) & \text{if } \max(A,B) = 1, \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

The corresponding DeMorgan dual T-conorm, denoted by $S(A,B)$, is defined as:

$$S(A,B) = 1 - T(1-A, 1-B) \quad (11)$$

In fuzzy logic the conjunction and disjunction operators are the minimum and maximum, respectively (upper and lower bounds of the T-norm and T-conorm ranges, respectively). These operators play a key role in the definition of the generalized modus ponens, which is the basis of possibilistic reasoning.

These possibilistic reasoning techniques have been implemented in RUM, a Reasoning with Uncertainty Module described in reference (Bonissone *et al.*, 1987).

Uncertainty in RUM is represented in both facts and rules. Facts represent the assignment of values to a set of propositional variables. Rules are acyclic quantitative Horn clauses in which a conjunct of antecedents implies (to a certain degree of belief) the rule consequent.

Facts are qualified by an uncertainty interval. The interval's lower bound represents the minimal degree of confirmation for the value assignment. The upper bound represents the degree to which the evidence failed to refute the value assignment. The interval's width represents the amount of ignorance attached to the value assignment. The uncertainty intervals are propagated and aggregated by T-norm based calculi.

Rules are discounted by a degree of sufficiency, indicating the strength with which the antecedent implies the consequent and a degree of necessity, indicating the degree to which a failed antecedent implies a negated consequent.

Real-Time Approximate Reasoning Systems

We conclude this discussion with a few remarks on the applicability of approximate reasoning systems to real-world problems requiring real-time performance. To achieve real-time performance levels, probabilistic reasoning systems need an efficient updating algorithm. The main problem consists in conditioning the existing information with respect to the new evidence: the computation of the new posterior probabilities in general belief networks is NP-hard (Cooper, 1989). A variety of solutions have been proposed, ranging from compilation techniques to shift the burden from run-time to compile-time, to the determination of bounds of the posterior probabilities. In particular, Heckerman et al. (Heckerman *et al.*, 1989) have proposed a decision-theoretic based analysis of computation versus compilation to determine the conditions under which run-time computation is preferable to lookup tables (generated at compile time). Horvitz et al. (Horvitz *et al.*, 1989) have proposed a method to approximate the posterior probabilities of the variables in each subgraph of a belief network. This method, called bounded conditioning, defines the upper and lower bounds of these probabilities and, if given enough resources, converges on the final point probabilities.

Due to its different underlying theory, possibilistic reasoning does not exhibit the same complexity problems as probabilistic reasoning. Most of the efforts aimed at achieving real-time performance from possibilistic reasoning systems have been based on translation/compilation techniques (Bonissone and Halverson, 1990) or hardware solutions (Corder, 1989, Watanabe and Dettloff, 1987).

Among the compilation techniques, a notable effort is RUMrunner, RUM's run-time system. RUMrunner is a software tool that transforms the customized knowledge base generated during the development phase into a fast and efficient real-time application. This goal is achieved by a combination of efforts: the translation of RUM's complex data structure into simpler, more efficient ones (to reduce overhead); the compilation of the rule set into a network (to avoid run-time search); the load-time estimation of each rule's execution cost (to determine, at run-time, the execution cost of any given deductive path); and the planning mechanism for path selection (to determine the largest relevant rule subset which could be executed within a given time-budget).

Among the hardware solutions to the problem of real-time performance for possibilistic reasoning systems, the most notable are the fuzzy chips (Corder, 1989, Watanabe and Dettloff, 1987). These chips are used in the application of fuzzy logic to industrial control. Fuzzy controllers (Sugeno, 1985) represent one of the the earliest instances of simple yet effective knowledge-based systems successfully deployed in the field. Their main use has been the replacement of the human operator in the feedback control loop of industrial processes. Their applications range from the development of the controller of a subway train system (Yasunobu and Miyamoto, 1985), to the use of a predictive fuzzy controller for container crane operation (Yasunobu and Hasegawa, 1986), to their application in the control of a continuously variable automobile transmission (Kasai and Morimoto, 1988).

References

Barnett J. A. 1981. Computational methods for a mathematical theory of evidence. In *Proc. 7th. Intern. Joint Conf. on Artificial Intelligence*, 868–875, Vancouver, British Columbia, Canada.

Bonissone P. P., and Decker K. S. 1986. Selecting Uncertainty Calculi and Granularity: An Experiment in Trading-off Precision and Complexity. In L. N. Kanal and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence*, 217–247. North Holland.

Bonissone P. P., and Halverson P. C. 1990. Time-Constrained Reasoning Under Uncertainty. *The Journal of Real Time Systems*, 2(1/2):22–45.

Bonissone P. P.; Gans S.S.; and Decker, K.S. 1987. RUM: A layered architecture for reasoning with uncertainty. In *Proceedings 10th Intern. Joint Conf. Artificial Intelligence*, 891–898. AAAI.

Bonissone, P. P.; Henrion, M.; Kanal, L. N.; and Lemmer, J. F. eds. 1991. *Uncertainty in Artificial Intelligence 6.* North-Holland, Amsterdam.

Bonissone, P. P. 1987a. Plausible Reasoning: Coping with Uncertainty in Expert Systems. In Stuart Shapiro, editor, *Encyclopedia of Artificial Intelligence*, 854–863. John Wiley & Sons Co., NY.

Bonissone, P. P. 1987b Summarizing and Propagating Uncertain Information with Triangular Norms. *International Journal of Approximate Reasoning*, 1(1):71–101, January 1987.

Cohen, P. , 1985. *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach.* Pittman, Boston, Massachusetts.

Cooper, G. F. 1989. Probabilistic Inference Using Belief Networks is NP-Hard. *Artificial Intelligence.*

Corder R. J. 1989. A High Speed Fuzzy Processor. In *Proceedings of the Third International Fuzzy Systems Association*, 379–389. IFSA.

D'Ambrosio, B.; Smets, P.; and Bonissone, P. P. eds. 1991. *Proceedings of the 7th Conference on Uncertainty in AI.* Morgan Kaufmann, San Mateo, CA.

Dempster, A. P. 1967. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339.

Doyle, J. 1983. Methodological simplicity in expert system construction: The case of judgements and reasoned assumptions. *The AI Magazine*, 4(2):39–43.

Dubois, D., and Prade, H. ,1990. Consonant Approximations of Belief Functions. *International Journal of Approximate Reasoning*, 4(5/6):419–449.

Duda, R. O.; Hart, P. E.; and Nilsson, N. J. 1976. Subjective Bayesian methods for rule-based inference systems. In *Proc. AFIPS 45*, 1075–1082, New York, AFIPS Press.

Fagin, R., and Halpern, J. H. 1989. Uncertainty, Belief, and Probability. In *Proc. 11th Intern. Joint Conf. on Artificial Intelligence*, 1161–1167, Detroit, MI.

Gordon, J., and Shortliffe, E. H. 1985. A method for managing evidential reasoning in a hierarchical hypothesis space. *Artificial Intelligence*, 26:325–339.

Halpern, J. Y., and Fagin, 1990. Two views of belief: Belief as generalized probability and belief as evidence. In *Proc. Eight National Conference on Artificial Intelligence*, 112–119, Boston, MA.

Halpern, J. Y., and Moses, Y. 1986. A Guide to Modal Logics of Knowledge and Belief. In *Proceedings of the 5th National Conference on Artificial Intelligence*, 480–490. AAAI.

Heckerman, D. E.; Breese, J. S.; and Horvitz, E. J. 1989. The Compilation of Decision Models. In *Proceedings of the the 5th Workshop on Uncertainty in Artificial Intelligence*, 162–173.

Henrion, M.; Shachter, R.D.; Kanal, L.N.; and Lemmer, J.F. eds. 1990. *Uncertainty in Artificial Intelligence 5.* North-Holland, Amsterdam.

Henrion, M. 1989. Practical Issues in Constructing a Bayes' Belief Network. In L.N. Kanal, T.S. Levitt, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 3*, 161–173. North-Holland.

Horvitz, E. J.; Suermondt, H. J.; and Cooper G. F. 1989. Bounded Conditioning Flexible Inference for Decisions Under Scarce Resources. In *Proceedings of the 5th Workshop on Uncertainty in Artificial Intelligence*, 182–193.

Howard, R.A., and Matheson, J.E. 1984. Influence diagrams. In R.A. Howard and J.E. Matheson, eds., *The Principles and Applications of Decision Analysis*, volume 2, 719–762. Strategic Decisions Group, Menlo Park, California.

Jaffray, J. Y. 1989. Linear Utility Theory for Belief Functions. *Oper. Res. Lett.*, 8:107–112.

Kasai, Y., and Morimoto, Y. 1988. Electronically Controlled Continuously Variable Transmission. In *Proceedings of the 1988 International Congress on Transportation Electronics*, 33–42. IEEE.

Kim, J. H., and Pearl, J. 1983. A Computational Model for Causal and Diagnostic Reasoning in Inference Engines. In *Proc. 8th. Intern. Joint Conf. on Artificial Intelligence*, 190–193, Karlsruhe, Germany.

Lauritzen, S.L., and Spiegelhalter, D. 1988. Local computations with probabilities on graphical structures and their application to expert systems. *J. Roy. Stat. Soc. Ser. B*, 50.

Lowrance, Y; Garvey, T.D.; and Strat, T.M. 1986. A Framework for Evidential-Reasoning Systems. In *Proc. 5th National Conference on Artificial Intelligence*, 896–903, Menlo Park, CA., AAAI.

Pearl, J. 1982. Reverend Bayes on Inference Engines: a Distributed Hierarchical Approach. In *Proceedings Second National Conference on Artificial Intelligence*, 133–136. AAAI.

Pearl, J. 1988. Evidential Reasoning Under Uncertainty. In Howard E. Shrobe, editor, *Exploring Artificial Intelligence*, 381–418. Morgan Kaufmann, San Mateo, CA.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan-Kaufmann, San Mateo, California.

Reiter, R. 1980. A Logic for Default Reasoning. *Artificial Intelligence*, 13:81–132.

Ruspini, E. H. 1987. Epistemic logic, probability, and the calculus of evidence. In *Proc. Tenth Intern. Joint Conf. on Artificial Intelligence*, 924-931, AAAI.

Ruspini, E. H. 1990. The semantics of vague knowledge. *Revue de Systemique*.

Ruspini, E. H. 1991. On the Semantics of Fuzzy Logic. *International Journal of Approximate Reasoning*, 5(1):45–88.

Schachter, R.D. 1986. Evaluating influence diagrams. *Operations Research*, 34:871–882.

Schweizer, B., and Sklar, A. 1963. Associative Functions and Abstract Semi-Groups. *Publicationes Mathematicae Debrecen*, 10:69–81.

Schweizer, B., and Sklar, A. 1983. *Probabilistic Metric Spaces*. North Holland, New York.

Shachter, R.D.; Levitt, T.S.; Kanal, L.N.; and Lemmer, J.F. eds. 1990. *Uncertainty in Artificial Intelligence 4*. North-Holland, Amsterdam.

Shafer, G., and Logan, R. 1987. Implementing Dempster Rule for Hierarchical Evidence. *Artificial Intelligence*, 33:271–298.

Shafer,G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shafer,G. 1990. Perspectives on the Theory and Practice of Belief Functions. *International Journal of Approximate Reasoning*, 4(5/6):323–362.

Shenoy, P., and Shafer, G. 1990. Axioms for Probability and Belief-Function Propagation. In R.D. Shachter, T.S. Levitt, L.N. Kanal, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 4*, 169–198. North-Holland, Amsterdam.

Shortliffe, E.H., and Buchanan, B. 1975. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23:351–379.

Smets, P., and Kennes, R. 1991. Computational Aspects of the Möbius Transform. In P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 6*. North-Holland, Amsterdam. Forthcoming.

Smets P. 1988. Belief functions. In P. Smets, A. Mamdani, D. Dubois, and H. Prade, eds., *Non-Standard Logics for Automated Reasoning*. Academic Press, NY.

Smets P. 1991. The Transferable Belief Model and Other Interpretations of Dempster-Shafer's Model. In P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 6*. North-Holland, Amsterdam. Forthcoming.

Stillman, J. 1991. On Heuristics for Finding Loop Cutsets in Multiply-Connected Belief Networks. In P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 6*. North-Holland, Amsterdam. Forthcoming.

Strat, T.M. 1984 Continuous belief functions for evidential reasoning. In *Proc. National Conference on Artificial Intelligence*, pages 308–313, Austin, Texas.

Strat, T.M. 1990. Decision Analysis Using Belief Functions. *International Journal of Approximate Reasoning*, 4(5/6):391–417.

Suermondt, J; Cooper, G.; and Heckerman, D. 1991. A Combination of Cutset Conditioning with Clique-Tree Propagation in the Pathfinder System. In P. Bonissone, M. Henrion, L.N. Kanal, and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence 6*. North-Holland, Amsterdam. Forthcoming.

Sugeno, M. editor. 1985. *Industrial Applications of Fuzzy Control*. North Holland, Amsterdam.

Voorbraak, F. 1989. A Computationally Efficient Approximation of Dempster-Shafer Theory. *Intl. J. Man-Machine Studies*, 30:525–536.

Watanabe, H, and Dettloff, W. 1987. Fuzzy Logic Inference Processor for Real Time Control: A Second Generation Full Custom Design. In *Proceedings of the Twenty-first Asilomar Conference on Signal, Systems & Computers*, 729–735. IEEE.

Yager, R. R.; Ovchinnikov, S.; Tong, R. M.; and Nguyen, H. T. eds. 1987. *Fuzzy Sets And Applications: Selected Papers by L.A. Zadeh*. John Wiley.

Yasunobu, S., and Hasegawa, G. 1986. Evaluation of an Automatic Crane Operation System based on Predictive Fuzzy Control. *Control Theory and Advanced Technology*, 2:419–432.

Yasunobu, S., and Miyamoto, S. 1985 Automatic Train Operation by Predictive Fuzzy Control. In M. Sugeno, editor, *Industrial Applications of Fuzzy Control*, 1–8. North Holland, Amsterdam.

Zadeh, L. A. 1965. Fuzzy Sets. *Information and Control*, 8:338–353.

Zadeh, L. A. 1968. Probability Measures of Fuzzy Events. *J. Math. Analysis and Appl.*, 10:421–427.

Zadeh, L. A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28.

Zadeh, L. A. 1979. A Theory of Approximate Reasoning. In P. Hayes, D. Michie, and L.I. Mikulich, eds., *Machine Intelligence*, 149–194. Halstead Press, NY.