Syllogistic Reasoning as a Basis for Combination of Evidence in Expert Systems

Computer Science Division, University California, Berkeley, California, 94720

Abstract

In the presence of uncertainty, computation of the certainty factor of a hypothesis requires, in general, the availability of rules for combining evidence under chaining, disjunction and conjunction. The method described in this paper is based on the use of what may be viewed as a generalization of syllogistic reasoning in classical logic - - a generalization in which numerical or, more generally, fuzzy quantifiers assume the role of probabilities. For example, the proposition QA's are B's, in which Q is a numerical or fuzzy quantifier, may be interpreted as "the conditional probability of B given A is Q." In this sense, the knowledge base of an expert system may be assumed to consist of propositions of the general form "QA's are B's."

It is shown that six basic syllogisms are sufficient to provide a systematic framework for the computation of certainty factors. A comparison with the rules of combination of evidence in PROSPECTOR, MYCIN and other expert systems is presented and a connection between syllogistic reasoning and the Dempster-Shafer theory is established. The syllogistic method of reasoning lends itself to a computationally efficient implementation and thus provides an effective tool for the management of uncertainty in expert systems.

I. SYLLOGISTIC REASONING

In the existing expert systems, the computation of certainty factors of hypotheses and conclusions is carried out through the use of probability-based methods (Barr and Feigenbaum, 1982). More recently, the use of belief functions in the context of the Dempster-Shafer theory has attracted increasing attention (Wesley, Lowrance and Garvey, 1984).

In a different approach which is outlined in this paper, syllogistic reasoning is employed to provide a basis for the formulation of rules for combination of evidence. Classically, the major and minor premises in syllogistic reasoning are allowed to contain only the standard quantifiers *all* and *some*. For example, the paradigmatic syllogism *Barbara* is expressed by the rule

$$\frac{\begin{array}{c}\text{All A's are B's}\\ \text{All B's are C's}\end{array}}{\text{All A's are C's}}$$

where A, B and C are arbitrary predicates in two-valued logic. This syllogism expresses the transitivity of set containment and, as such, provides the basis for property inheritance in knowledge representation systems.

Research supported in part by NASA Grant NCC2-276, NSF Grant 1ST-8320410 and DARPA/NESC Contract N00039-84-C-0089

In a more general sense which is employed in this paper, a syllogism is an inference rule of the form

$$\frac{\begin{array}{c}Q_1\text{A's are B's}\\ Q_2\text{C's are D's}\end{array}}{Q_3\text{E's are F's}}$$

where Q_1, Q_2 and Q_3 are numerical or, more generally, fuzzy quantifiers (e.g., 80%, *most*, *many*, *few*), and A,B,C,D,E and F are crisp or fuzzy predicates (e.g., *positive*, >5, *small*, *tall*, *heavy*). The predicates A, B, ..., F are assumed to be related in a specified way, giving rise to different types of syllogisms. For example, the ***consequent conjunction syllogism*** corresponds to the relations A=C=E, F=B $\wedge$ D, where $\wedge$ stands for conjunction.

The following syllogisms play a basic role in the formulation of rules for combining evidence in expert systems.

Intersection/product syllogism: C=A B, E=A, F=B $\wedge$ D.
Consequent conjunction syllogism: A=C=E, F=B$\wedge$ D.
Consequent disjunction syllogism: A=C=E, F=B$\vee$ D. ($\vee \triangleq$ disjunction)
Antecedent conjunction syllogism: B=D=F, E=A$\wedge$ C.
Antecedent disjunction syllogism: B=D=F, E=A$\vee$ C.
Chaining syllogism: B=C, E=A, F=B.

(The chaining syllogism may be viewed as a special case of the intersection/product syllogism).

A quantifier in a premise of a syllogism plays a role analogous to that of conditional probability. More specifically, the proposition "QA's are B's," may be interpreted as "the conditional probability of B given A is Q," with the understanding that A and B may be fuzzy events such that the conditional probability of B given A is a fuzzy probability Q. Viewed in this perspective, the consequent conjunction syllogism, for example, may be viewed as a generalization of the combining rule for the conjunction of hypotheses in PROSPECTOR and MYCIN.

A combining rule expressed as a syllogism differs from standard combining rules in several important respects. First, when Q_1 and Q_2 are numerical quantifiers or, equivalently, numerical probabilities, the resultant quantifier Q_3 in the conclusion is, in general, interval-valued. For example, in the case of the consequent conjunction syllogism, Q_3 is given by

$$0 \vee (Q_1 + Q_2 - 1) \leq Q_3 \leq Q_1 \wedge Q_2,$$

where $\wedge$ and $\vee$ stand for min and max, respectively. The interval-valuedness of Q_3 reflects the incomplete state of knowledge regarding the dependence between the major and minor premises. In the conventional approaches, the underlying events are assumed to be conditionally independent, which accounts for a numerical value--rather than an interval value--for the certainty factor of the conclusion.

In the more general case where Q_1 and Q_2 are fuzzy quantifiers, Q_3 is also a fuzzy quantifier given by

$$0 \circledvee (Q_1 \oplus Q_2 \ominus 1) \leq Q_3 \leq Q_1 \circledwedge Q_2,$$

where $\circledvee$, $\oplus$, $\ominus$ and $\circledwedge$ are the operations of max, +, − and min in fuzzy arithmetic (Kaufmann and Gupta, 1985). In this sense, when Q_1 and Q_2 are fuzzy numbers, so is Q_3. As an illustration, an instance of the consequent conjunction syllogism is

Most students are young
Most students are healthy
Q students are young and healthy

where $0 \circledvee (2\ most \ominus 1) \leq Q \leq most$.

When B is contained in A, and Q_1 and Q_2 are monotonic (i.e., *at least* Q_1 is equivalent to Q_1, and likewise for Q_2) the intersection/product syllogism reduces to the chaining syllogism:

Q_1A's are B's
Q_2B's are C's
$(Q_1 \otimes Q_2)$A's are C's

where $\otimes$ denotes multiplication in fuzzy arithmetic. For example,

Most students are undergraduates
Most undergraduates are young
$Most^2$ students are young

where $most^2$ is a fuzzy quantifier which is the product of *most* with itself. In this way, the chaining syllogism provides a mechanism for dealing with chains of rules or facts whose constitutents (i.e., predicates and quantifiers (or probabilities)) are not sharply defined.

II. DEMPSTER-SHAFER THEORY

A special case of syllogistic reasoning which provides a link with the Dempster-Shafer theory of belief functions, relates to premises in which the quantifiers can take only three values: 1 (*all*), >0(*some)* and 0 *(none)*. Thus, all premises are of the form $Q_i A'_s$ are R, i=1, ..., n, where Q_i is 1, >0 or 0, and the A_i and R are subsets of the universe of discourse, with R representing a range query and A_i being the value of an attribute. For example, if the attribute is Age, the query may be, "What fraction of individuals in the database are in the age range [20, 25], with R being the interval [20, 25]. For an individual i, A_i would represent the possible values of the age of i, so that the proposition "All A_i are R," would signify that all possible values of the age of i are contained in the interval R and hence that it is certain that the age of i satisfies the condition of the query.

The answer to the query, then, would have two components. The first component, called *belief* in the Dempster-Shafer theory, is, in effect, the fraction of the Q_i which are equal to 1. The second component, *plausibility*, is the fraction of the Q_i which are greater than zero. These components may be interpreted, respectively, as the measures of certainty (or necessity) and possibility (Zadeh, 1981).

The advantage of viewing the Dempster-Shafer theory as a variant of syllogistic reasoning in which the quantifiers are three-valued, is that it suggests a more general approach in which the quantifiers may be quantized to any desired number of levels. In this way, the number of levels may be adjusted to fit the granularity of the information in the knowledge base.

III. USUALITY

An important application of syllogistic reasoning relates to what might be called the concept of *usuality*. In essence, usuality is concerned with the usual values of variables, e.g., the usual price of a cup of coffee, the usual duration of a particular event, etc. More specifically let X be a variable taking values in a universe of discourse U, and let F be a fuzzy subset of U characterized by a membership function μ_F which associates with each point $\mu \in U$ a number in the interval [0,1] which represents its grade of membership in F. As shown in (Zadeh, 1978), the proposition *X is F* - - which may be interpreted as the assignment of a fuzzy value F to X - - induces a possibility distribution in U such that the possibility that $X = u$ is given by $\mu_F(u)$. For example, the proposition *X is small* induces a possibility distribution on the space of real numbers such that the possibility that $X = u$, where u is a real number, is equal to $\mu_{SMALL}(u)$, (i.e., the degree to which u fits the definition of the predicate *small*, with *SMALL* being a fuzzy subset of real numbers which represents the denotation of *small*. Similarly, in the case of the proposition *Mary is young*, if X is taken to be the age of Mary and U is the interval [0,100], then the proposition in question induces a possibility distribution of X such that the possibility that *Age (Mary)* =26, say, equals the degree to which 26 fits the definition of the predicate *young* in the context in which the proposition *Mary is young* is asserted.

With this possibilistic interpretation of the proposition *X is F* as the point of departure, the meaning of the dispositional proposition *usually (X is F)* or, equivalently, *X is usually F*, may be defined as follows (Zadeh, 1984b):

$$usually\ (X\ is\ F) \rightarrow if\ Z\ is\ R\ then\ most\ X's\ are\ F,$$

where Z is the *conditioning variable* and R is a fuzzy value of Z which constrains Z away from exceptional values for which the proposition *most X's are F* does not hold. For example, in the case of the proposition *usually it takes about an hour to drive from Berkeley to Stanford*, Z may be the time at which the trip is started and R is the non-rush-hour period. If R is U, then *usually* is unconditioned and the definition of *usually (X is F)* reduces to

$$usually\ (X\ is\ F) \rightarrow most\ X's\ are\ F$$

which implies that *usually* plays the role of a fuzzy quantifier, that is, a fuzzy number which represents the relative count of elements in a fuzzy set (Zadeh, 1983). More specifically, if the observed values of X are $u_1, \ldots, u_n$ then the relative count of *X's* which are F is expressed by

$$r = \frac{1}{n}(\mu_F(u_1) + \ldots + \mu_F(u_n))$$

and the degree to which r fits the definition of *most* is given by

$$r = \mu_{MOST}(r)$$

where *MOST* is the fuzzy subset of the unit interval which represents the denotation of the fuzzy quantifier *moat.* In this way, the meaning of the propositipn *usually (X is F)* may be expressed in terms of the meanings of *F* and *most.* In a similar fashion, in the case of conditioned usuality, the meaning of *usually (X is F)* may be expressed as

$$\Sigma Count\ (F/R)\ is\ MOST$$

where $\Sigma Count(F/R)$ denotes the relative count of elements of *F* which are in *R.*

The concept of usuality makes it possible to extend the domain of applicability of syllogistic reasoning by allowing a premise in a syllogism to be qualified - - implicitly or explicitly - - by the adjunction of *usually,* with the understanding that *usually* may be interpreted as a conditioned or unconditioned fuzzy quantifier. As a simple illustration, a special case of the intersection/product syllogism is the *dispositional modus ponens*

X is F
if X is F then usually (Y is G)

---

usually (Y is G)

or, more generally

(usually)(X is F)
if X is F then (usually) (Y is G)

(usually ^{2}*)(Y is G)*

where *usually* ^{2} is the product of the fuzzy number *usually* with itself in fuzzy arithmetic. Syllogisms of this type are of particular relevance to commonsense reasoning.

In summary, syllogistic reasoning provides a systematic framework for inference from premises which are imprecise and/or not totally reliable. In this mode of reasoning, numerical quantifiers play a role analogous to that of probabilities in probabilistic reasoning (Nilsson, 1984). However, the facility provided by syllogistic reasoning for manipulation of fuzzy quantifiers as fuzzy numbers within the framework of fuzzy arithmetic goes beyond the capabilities of existing probability-based methods and makes it possible to deal computationally with rules and facts whose certainty factors are not known with sufficient precision to justify the use of numerical probabilities.

References

Barr, A. and Feigenbaum, E.A., eds., *The Handbook of Artificial Intelligence,* vols. 2 and 3, Kaufmann, Palo Alto, 1982.

Buchanan, B.B. and Shortliffe, E.H., *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project,* Addison-Wesley, Reading, MA, 1984.

Cohen, M.S., Schum, D.A., Freeling, A.N.S. and Chinnis, J.O., "On the art and science of hedging a conclusion: alternative theories of uncertainty in intelligence analysis," *Tech. Rep. 84-6,* Decision Science Consortium, Falls Church, VA, 1984.

Duda, R., Gaschnig, J. and Hart, P., "Model design in the PROSPECTOR consultant system for mineral exploration," Final Tech. Report, SRI International, Menlo Park, CA.1979.

Dubois, D., and Prade, H., "On several representations of an uncertain body of evidence," in *Fuzzy Information and Decision Processes,* Gupta, M.M. and Sanchez, E., (ed), North-Holland, Amsterdam, pp. 167-181, 1982.

Goguen, J.A., "The logic of inexact reasoning," *Synthese 19,* pp. 325-373, 1969.

Kaufmann, A. and Gupta, M.M., *Introduction to Fuzzy Arithmetic.* Van Nostrand Reinhold, New York, 1985.

Mamdani, E.H. and Gaines, B.R., *Fuzzy Reasoning and its Applications.* Academic Press, London, 1981.

McCarthy, J., "Applications of circumscription to formalizing commonsense knowledge." *Proc. of Non-monotonic Reasoning Workshop,* AAA1, pp. 295-324, 1984.

Nilsson, N., "Probabilistic logic," SRI Tech. Note 321, Menlo Park, CA, 1984.

Reiter, R. and Criscuolo, G., "Some representational issues in default reasoning," *Computers and Mathematics 9,* pp. 15-28, 1983.

Szolovits, P. and Pauker, S.G., "Categorical and probabilistic reasoning in medical diagnosis," *Artificial Intelligence 11,* pp. 115-144, 1978.

Wesley, L., Lowrance, J. and Garvey, T. "Reasoning about control: an evidential approach," SRI Tech. Rep. 327, Menlo Park, CA,1984.

Yager, R.R., "Quantifiers in the formulation of multiple objective decision functions." *Inf. Sciences 91,* pp. 107-139, 1983.

Zadeh, L.A., "PRUF - A meaning representation language for natural languages," *Inter. J. Man-Machine Studies 10,* pp. 395-460, 1978.

Zadeh, L.A., "Possibility theory and soft data analysis," in *Mathematical Frontiers of the Social and Policy Sciences,* Cobb, L., and Thrall, R.M., (eds.), Boulder: Westview Press, pp. 69- 129, 1981.

Zadeh, LA., "A computational approach to fuzzy quantifiers in natural languages," *Computers and Mathematics 9,* pp. 149-184, 1983.

Zadeh, LA., "Syllogistic reasoning in fuzzy logic and its application to reasoning with dispositions," in *Proceedings of the 1984 International Symposium on Multiple-Valued Logic,* pp. 148-153, Winnipeg, Canada, 1984a.

Zadeh, LA., "A simple view of the Dempster-Shafer theory of evidence," *Berkeley Cognitive Science Report 27, University of California, Berkeley. 1984b.*

Zemankova-Leech, M. and Kandel, A., *Fuzzy Relational Databases - A Key to Expert Systems.* Interdisciplinary Systems Research, Verlag TUV, Rhineland, Cologne, 1984.