Qatar Univ. Sci. J. (2004), 24 : 57 - 67

On Symbolic Reasoning and Direct Inference Principle

Jihad Mohamad Jaam
Department of Computer Science and Engineering
Qatar University, P.O. Box 2731, Doha, Qatar

حول منطق التحليل الرمزي ومبدأ الاستنتاج المباشر

د. جهاد محمد الجعم
كلية الهندسة - قسم هندسة وعلوم الحاسب
جامعة قطر - قطر

هذا البحث يناقش طريقة نوعية تعتمد على المنطق الرمزي المحدود ومبادىء الاستنتاج المنطقية الثلاثية والمباشرة (الاستنتاج من المعلومات الإحصائية للحصول على الخلاصات النهائية والنتائج) لمعالجة البيانات الإحصائية المقيمة بطريقة نوعية ، علماً بأن مقاييس التدرج المنطقية المستخدمة ممثلة بدرجات حقائق متنوعة تعتمد بصورة حقيقية ومباشرة على المعلومات المتوفرة من قاعدة البيانات المتاحة، بدلاً من أن تكون عشوائية.

Keywords: *Statistical Information, Syllogism, Direct Inference.*

ABSTRACT

This paper discusses an approach using the M-valued symbolic logic with syllogistic reasoning and direct inference principle (reasoning from statistical information to conclusions about individuals), to manipulate statistical knowledge evaluated in a qualitative way. The graduation scale of M symbolic quantifiers is expressed in terms of truth degrees, which should not be arbitrary, but rather should be based on the available information from the knowledge base.

Introduction

Consider an agent with some knowledge base KB, who has to make specific decisions about his actions in the world. For example, a doctor may need to decide on a treatment for a particular patient, P. The doctor's knowledge base should contain information of different types, including: statistical information, "*75% of patients with jaundice have hepatitis*"; first-order information, "*all patients with hepatitis have jaundice*"; and some information about the particular patient at hand, e.g., "*X has jaundice*". In most cases, the knowledge base will not contain complete information about a particular individual. For example, the doctor may be uncertain about the exact disease that P may have. Since the efficacy of a treatment will almost likely depend on the disease itself, it is important for the doctor to be able to quantify the relative likelihood of various possibilities. To apply standard tools for decision making an agent must assign some degrees of belief, to various events. For example, the doctor may wish to assign a degree of belief to an event such as "P has hepatitis". Thus, the representation of statistical information is made by quantified statements like "*all*", "*almost all*", "*most*", "*few*", "*a bit*", "*a little*", etc. A given model is important, if it has an inference process like the *syllogistic reasoning* of Zadeh [12], or the "*direct inference principle*" of Bacchus [2], which can allow deducing new assertions (i.e., reasoning from statistical information to conclusions about individuals). For example, knowing that "*most teachers are old*" and "*almost all teachers are married*", we can deduce that "*most teachers are old or married*" or knowing that "*most birds fly*" and "*Tweety is a bird*", so we can deduce that "*it is very probable that Tweety flies*". We interpret a statement such as "*Birds typically fly*" as expressing the statistical assertion that "*almost all birds fly*". Many approaches of these problems are generally based on probability theory [1,2,3], or fuzzy set theory [4,10,11]. However, in this work we propose an intuitive approach to manipulate efficiently some knowledge based on statistical information and evaluated in a qualitative way, using the M-valued symbolic logic introduced by Pacholczyk [7,8] with syllogistic reasoning and direct inference principle presented in [5,6]. We consider a graduation scale with seven adverbial expressions. The first scale degrees of truth, denoted by L_7, enables us to express the graduation of vagueness: "*Frank is very smart*" is equivalent to say that "*Frank*" satisfies the predicate "*smart*" with the degree "*very*". A second scale of "*degrees of statistical probability*" denoted by Q_7, that permits to express the *graduation of proportion*: given the basic space Ω, "Q_α Ω'are A's" means that an absolute proportion Q_α of individuals of Ω are in A with respect to the uniform probability distribution on Ω. A third scale of "*degrees of certainty*" denoted by U_7 is used to express the "*graduation of certainty*": "*it is very probable that Tweety flies*" means that "*very-probable*" is the certainty degree of the assertion "*Tweety flies*". The graduations scales that we use are the following:

(1) $L_7 = \{\tau_\alpha \mid \alpha = 1, 7\}$ = {*Not-at-all-true, Very-little-true, Little-true, Moderately-true, Very-true, Almost-true, Totally-true*},
(2) $Q_7 = \{Q_\alpha \mid \alpha = 1, 7\}$ = {*None, Very-few, Few, About-half, Most, Almost-all, All*}, and
(3) $U_7 = \{\upsilon_\alpha \mid \alpha = 1, 7\}$ = {*Not-at-all-probable, Very-little-probable, Little-probable, Moderately-probable, Very-probable, Almost-certain, Certain*}.

The M-Valued Logic and Satisfaction Formulas

Let $M \geq 2$ be an integer and let I be the integer interval $[1, M]$ ordered by the relation $\leq$, and let n be the mapping function defined by the following formula: $n(\alpha) = M + 1 - \alpha$. Then, $\{I, \vee, \wedge, n\}$ is a De Morgan lattice with: $\alpha \vee \beta = max(\alpha, \beta)$ and $\alpha \wedge \beta = min(\alpha, \beta)$. Let $L_M = \{\tau_\alpha, \alpha \in I\}$ be a set of M elements totally ordered by the relation $\leq$: $\tau_\alpha \leq \tau_{\alpha'} \Leftrightarrow \alpha \leq \beta$. Thus $\{L_M, \leq\}$ is a chain in which the least element is τ_1 and the greatest element is τ_M.

We define then in L_M the following operators:

- $\tau_\alpha \vee \tau_\beta = \tau_{max(\alpha,\beta)}$,
- $\tau_\alpha \wedge \tau_\beta = \tau_{min(\alpha,\beta)}$,
- $\sim \tau_\alpha = \tau_{n(\alpha)}$.

We interpret L_M as a set of linguistic truth degrees dealing with vague predicates. With $M = 7$, we can introduce the previous set L_7. We call an *interpretation structure* ϑ of the M-valued predicate language L, a pair $< D, I >$, where D is the domain of ϑ and I the interpretation function. We denote by R_n the multiset associated with the predicate P_n. We call a valuation of variables, a sequence denoted by $s = <s_0, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots>$ with $s_i \in D$. The valuation s (i/a) is defined by the following:

$$s\,(i/a) = <s_0, \ldots, s_{i-1}, a, s_{i+1}, \ldots>.$$

Definition 1: The premise is given by $\Gamma \cup \varphi \Rightarrow \psi$, and the conclusion is given by $\Gamma \Rightarrow \varphi_1 \wedge \varphi_2 \wedge \ldots \wedge \varphi_n$. The relation "s τ_α-satisfies ϕ in-ϑ", denoted by $\vartheta \models^s_\alpha \phi$, is defined as follows:

(*1*) $\vartheta \models^s_\alpha P_n(z_{i1}, \ldots, z_{ik}) \Leftrightarrow <s_{i1}, \ldots, s_{ik}> \in_\alpha R_n$,
(*2*) $\vartheta \models^s_\alpha \neg \phi \Leftrightarrow \models^s_\beta \phi$ with $\tau_\alpha = \sim \tau_{(\beta)}$,
(*3*) $\vartheta \models^s_\alpha \phi \cap \psi \Leftrightarrow \{\vartheta \models^s_\beta \phi$ and $\vartheta \models^s_\gamma \Psi$ with $\tau_\alpha = \tau_{(\beta)} \wedge \tau_{(\gamma)}\}$,
(*4*) $\vartheta \models^s_\alpha \phi \cup \psi \Leftrightarrow \{\vartheta \models^s_\beta \phi$ and $\vartheta \models^s_\gamma \psi$ with $\tau_\alpha = \tau_{(\beta)} \vee \tau_{(\gamma)}\}$,
(*5*) $\vartheta \models^s_\alpha \phi \supset \psi \Leftrightarrow \{\vartheta \models^s_\beta \phi$ and $\vartheta \models^s_\gamma \psi$ with $\tau_\alpha = \tau_{(\beta)} \rightarrow \tau_{(\gamma)}\}$,
(*6*) $\vartheta \models^s_\alpha \exists z_n \psi \Leftrightarrow \tau_\alpha = max\ \{\tau_{(\gamma)} \mid \vartheta \models^{s(n/a)}_\gamma \psi, a \in D\}$,
(*7*) $\vartheta \models^s_\alpha \forall z_n \psi \Leftrightarrow \tau_\alpha = min\ \{\tau_{(\gamma)} \mid \vartheta \models^{s(n/a)}_\gamma \psi, a \in D\}$,

We generalize the definition of conditional statistical probability in a symbolic context, by using a new predicate with a "*symbolic probabilistic division*"' operator, denoted by **C**, or equivalently a "*symbolic probabilistic division*" operator, denoted by **I**. Note that, the operator **C** is deduced from **I** by a unique way as follows:

$$Q_\mu \in \mathbf{C}(Q_\alpha, Q_\lambda) \Leftrightarrow Q_\lambda \Leftrightarrow \mathbf{C}(Q_\alpha, Q_\mu)$$

To define the probability space, we have to choose an appropriate set of possible worlds. Given some domain of individuals, we stipulate that the set of worlds is simply the set of all first-order models over this domain. That is, a possible world corresponds to a particular way of interpreting the symbols in the agent's vocabulary over the domain. We can assume that the "*true world*" has a finite domain of size *N*. In fact, without loss of generality, we assume that the domain is $\{1..N\}$. Having defined the probability space, we can construct a probability distribution over this set. We assume that all the possible worlds are equally likely (that is, each world has the same probability). This can be viewed as an application of the principle of indifference.

Definition 2: Consider the following space Ω, thus:

(1) Any quantified assertion "Q_α Ω's are *A*'s" means that an *absolute proportion* Q_α of individuals of Ω are in *A* with respect to the uniform probability distribution on Ω.

(2) Any quantified assertion "Q_μ *A*'s are *B*'s" means that among the elements of the basic space Ω which belong to *A*, *a relative proportion* Q_μ of these elements belong to *B*, and this with respect to the uniform probability distribution on Ω. It is defined as follows:

if {Q_α Ω's are *A*'s and Q_λ Ω's are A∩B's}, then "Q_μ *A*'s are *B*'s" with $Q_\mu \in \mathbf{C}(Q_\alpha, Q_\lambda)$.

Axiom 1: $A \cap B \neq A$, "Q_α Ω's are *A*'s" and "Q_α Ω's are (A∩B)'s" and $Q_\alpha \in [Q_3, Q_7] \Rightarrow$ "Almost all *A*'s are *B*'s.

Axiom 2: "Q_α Ω's are *A*'s", $Q_\alpha \in [Q_2, Q_7]$ and "Almost-All *A*'s are *B*'s" ⇒ "Q_α Ω's are (A∩B)'s".

Axiom 3: "Q_α Ω's are *A*'s" ⇔ "$Q_{n(\alpha)} \in$ Ω's are $\bar{A}$ with $n(\alpha) = M + 1 - \alpha$".

Axiom 4: "Q_α Ω's are *A*'s", "Q_β Ω's are *B*'s", $A \cup B \neq \Omega$ and $A \cap B = \varnothing \neq \Rightarrow$ "Q_r Ω's are (A∪B)'s" with $Q_r \in S(Q_\alpha, Q_b)$ where *S* is a symbolic addition.

If *A* and A∪B represent the same significant absolute proportion, that means "*almost all A*'s *are B*'s" (Axiom 1). Secondly (Axiom 2), knowing that "*almost-all A*'s are *B*'s", then A∩B has the absolute proportion of *A*. Thirdly, *A* and $\bar{A}$ have symmetrical absolute proportions (Axiom 3). Finally, the absolute proportion of A∪B results from the ones of A, B and A∩B (Axiom 4). The "*symbolic sum*" is denoted by **S** and the "*symbolic difference*" is denoted by **D** which can be deduced from **S**.

Proposition 1: Let *A* and *B* be subsets of Ω. If "Q_α Ω's are *A*'s" and A⊂B then "Q_β Ω's are *B*'s" with $Q_\alpha \leq Q_\beta$. If Q_α Ω's are *A*'s", "Q_λ Ω's are $\gamma^{(A \setminus B)}$'s" with $Q_\gamma \in \mathbf{D}(Q_\alpha, Q_\lambda)$.

Definition 3: A *syllogism* is an inference rule that consists of deducing a new quantified statement from quantified statements.

Let *R* be the set of available quantified assertions, then we can deduce from *R*, by using syllogisms, a set R^* containing *R* and new quantified assertions. We then adapt the following syllogisms:

Relative Duality: If R contains "$Q_{\mu 1}$ A's are B's" then R^* contains "$Q_{\mu 2}$ A's are B's" with $Q_{\mu 2} = Q_{n(\mu 1)}$ if $Q_{\mu 1} \neq Q_{n(\mu 1)}$ and $Q_{n(\mu 2)} \in [Q_{n(\mu 1)}, Q_{n(\mu 1)+1}]$.

Mixed Transitivity: If *R* contains $Q_{\mu 1}$ A's are B's" then R^* contains "$Q_{\mu 2}$ A's are B's" with $Q_{\mu 2} = Q_{n(\mu 1)}$ if $Q_{\mu 1} \neq Q_{n(\mu 1)}$ and

Intersection (Product Syllogism): If *R* contains "$Q_{\mu 1}$ 1 A's are B's" and "$Q_{\mu 2}$ (A∩B)'s are C's" then R^* contains "Q_{μ} A's are (B∩C)'s", with $Q_{\mu} = I(Q_{\mu 1}, Q_{\mu 2})$.

Intersection (Quotient Syllogism): If "$Q_{\mu 1}$ A's are B's", "$Q_{\mu 2}$ A's are C's" and "$Q_{\mu 3}$ (A∩B)'s are C's" then Q_{μ} (A∩C)'s are B's", with $Q_{\mu} \in C(Q_{\mu 2}, I(Q_{\mu 1}, Q_{\mu 3}))$.

Contraction: If *R* contains "*Almost-all A's are B's*" and "*Almost-all (A∩B)'s are C's*" then R^* contains "*Almost-all A's are C's*".

Cumulativity: If R contains "*Almost-all A's are B's*" and "*Almost-all A's are C's*" then R^* contains Q_{μ} (A∩B)'s are C's", with $Q_{\mu} \in$ [Most, Almost-all].

Union Left: If R^* contains "*Almost-all A's are C's*" and "*Almost-all B's are C's*" then R^* contains "Q_{μ} (A∪B)'s are C's", with $Q_{\mu} \in$ [Most, Almost-all].

Subjective Probability

An intelligent agent will often be uncertain about various properties of its environment, and when acting in that environment it will frequently need to quantify its uncertainty. For example, if the agent wishes to employ the expected-utility paradigm of decision theory to guide its actions, he will need to assign degrees of belief (i.e., subjective probabilities) to various assertions. Once we describe the language in which our knowledge base is expressed, we may need to decide how to assign degrees of belief given a knowledge base. Perhaps the most widely used framework for assigning degrees of belief (which are essentially subjective probabilities) is the Bayesian paradigm. There, one assumes a space of possibilities and a probability distribution over this space (the prior distribution), calculates posterior probabilities by conditioning on what is known (in our case, the knowledge base). To use this approach, we must specify the space of possibilities and the distribution over it. In Bayesian reasoning, relatively little is said about how this should be done. Indeed, the usual philosophy is that these decisions are subjective. More precisely, knowing that a particular individual "a" belongs to *A* (or "a is *A*"), we wish to deduce from "Q_{μ} A's are B's" and available knowledge, a symbolic certainty degree to which the particular individual *a* belongs to *B* (or "a is B"). We introduce then a certainty function *F* which is applied to Boolean formulas. Thus, the statement "A(a) is υ_{α} (or v_{α}-probable)" is translated into $F(A(a)) = \upsilon_{\alpha}$.

Definition 4: $F(\mathrm{A(a)}) = \upsilon_\alpha$ is equivalent to say that "A(a) is v_α-probable" is totally true. The function F satisfies the following axiomatics:

C_1: $\mathrm{A(a)} \equiv \mathrm{B(a)} \Rightarrow F(\mathrm{A(a)}) = \mathrm{F(B(a))}$,
C_2: A(a) is true $\Rightarrow F(\mathrm{B(a)}) = \upsilon_7$,
C_3: A(a) is false $\Rightarrow F(\mathrm{B(a)}) = \upsilon_1$,
C_4: $F(\mathrm{A(a)}) = \upsilon_\alpha \Rightarrow F(\neg\, \mathrm{A(a)}) = \upsilon_{8-\alpha}$,
C_5: $\{F(\mathrm{A(a)}) = \upsilon_\alpha, F(\mathrm{B(a)}) = \upsilon_\beta, F(\mathrm{A(a)} \cap \mathrm{B(a)}) = \upsilon_1\} \rfloor F(\mathrm{A(a)} \lceil \mathrm{B(a)}) = \upsilon_1$ with $\upsilon_\gamma = \mathrm{S}\,(\upsilon_\alpha, \upsilon_\beta)$.

This certainty concept has to satisfy a number of postulates, each of them being justified at a Meta-logical level. First of all (Axioms C_2 and C_3), if a statement is *true* (respectively, *false*), its certainty degree is certain (respectively, impossible). If two statements are equivalent, they receive the same certainty degree (Axiom C_1). The certainty degree of the negation is the symmetrical value in the graduation scale of the one of the affirmation (Axiom C_4). Finally, if the intersection of two statements is false, then the certainty associated with their union is the "symbolic sum" (Axiom C_5) of their uncertainty.

Direct Inference Principle

The quantified assertion "*α% of individuals of the domain verify a property*" can be interpreted as "*the probability that a randomly selected domain individual satisfies the property is equal to α*". This interpretation can be seen as a way of justifying the deduction of uncertain conclusions about particular individuals (i.e., subjective probabilities) from statistical knowledge (i.e., statistical probabilities) via the direct *inference* [2,9]. Indeed, the principle of direct inference is based on the idea that a particular individual in the domain is considered as a member randomly selected from a population, if no particular information distinguishes it from other members of this population. For example, if all we know about *Tweety* is that it is a bird, then *Tweety* can be viewed as a randomly selected member of the population of birds since we do not have any other information that distinguishes it from other birds. Thus, knowing that *Tweety* is a bird, the (subjective) probability that *Tweety* flies is equal to the (statistical) probability that a bird randomly selected from the set of birds flies, i.e. the proportion of flying birds among the birds.

We have proposed a symbolic generalization of the direct *inference principle* allowing us to infer a *symbolic subjective probability degree from a symbolic statistical probability degree.*

Definition 5: The available knowledge base can be formally represented in the basic domain by the couple $KB = (\chi,\mathrm{R})$ where χ is the conjunction of formulae representing the available knowledge about the particular individuals of the basic domain, and R is the set of quantified assertions.

(1) By using syllogisms, we can deduce from R a set R^* containing R and new quantified assertions.
(2) $\chi(a)$ will be the conjunction of formulas appearing in χ mentioning *a*.
(3) $\chi(a|z)$ is the formula obtained when textually substituting each occurrence of a in the formula $\chi(a)$ by the free variable *z*.

(4) $\neg\chi$ (a|z) and $\neg$ B (a|z) denote respectively the sets associated with the formulas χ(a|z) and B(a|z).

(5) Given *a* as an individual of the basic domain, a *reference class* of a given *knowledge base KB* for a formula *B*(*a*) (in which we want to generate a certainty degree) is a subset of the basic domain to which belongs the individual *a*.

Intuitively, the substitution χ(a|z) denotes a form related to the process of "*random selection*". The constant a is considered as a "*random member*" by replacing it in χ(a) by the free variable *z*. This leads to suppose that the individual denoted by a is randomly chosen among the individuals sharing all its properties, i.e. the individuals satisfying χ(a|z). Let us now present the basic notions leading to our basic definition of *symbolic direct inference*. Given a knowledge base *KB*, we suppose that *a* denotes an *individual constant* of the basic domain. Given *a*, we search its certainty degree u_α-probable resulting from the available knowledge base *KB* = (χ,R). It is defined in the following way:

The Direct Inference Principle

Let us suppose that *a* denotes an individual constant of the domain, *z* a variable, and *KB* = (χ,R) the available knowledge base. We say that *F*(B(a)) = υ_α results from direct inference principle, if the quantified assertion {Q_α χ(a|z)'s are B(a|z)'s} belongs to R^*.

The Reference Class

The only necessary relationship between objective knowledge about frequencies and proportions on the one hand and degrees of belief on the other hand is the simple mathematical fact that they both obey the axioms of probability. However, practically, we usually hope for a deeper connection: the latter should be based on the former in some "specific" way. Definitely, the random-worlds approach is precisely a theory of how this connection can be made. Most of the previous work is based on the idea of finding a suitable reference class. In this section, we review some of this work and we discuss the suitable reference class that we have selected.

The Basic Approach

The first sophisticated attempt at clarifying the connection between objective statistical knowledge and degrees of belief, and the basis for most subsequent proposals, is due to Reichenbach [9], using the following idea:

> If we are asked to find the probability holding for an individual future event, we must first incorporate the case in a suitable reference class. An individual thing or event may be incorporated in many reference classes. We then proceed by considering the narrowest (smallest) reference class for which suitable statistics can be compiled.

Reichenbach's approach was to equate the degree of belief in the individual event with the statistics from the chosen reference class. As an example, suppose that we want to determine a probability, or a degree of belief, that Frank, a particular patient with jaundice, has the disease hepatitis. The particular individual Frank is a member of the class of all patients with jaundice. Hence, following Reichenbach, we can use the class of all such patients as a reference class, and assign a degree of belief equal to our statistics concerning the frequency of hepatitis among this class. If we know that this frequency is 90%, then we would assign a degree of belief of 0:9 to the assertion that Franc has hepatitis. This approach consists of

(1) the postulate that we use the statistics from a particular reference class to infer a degree of belief with the same numerical value, and

(2) some guidance as to how to choose this reference class from a number of competing reference classes. In general, a reference class is simply a set of domain individuals that contains the particular individual about whom we wish to reason and for which we have "*suitable statistics*". We may take the set of individuals satisfying a formula $\psi(x)$ to be a reference class. The requirement that the particular individual *c* we wish to reason about belongs to the class is then represented by the logical assertion $\psi(c)$. However, what does the phrase "*suitable statistics*" mean? Assume for now we take a "suitable statistic" to be a closed interval that is nontrivial (i.e., that is not [0; 1]), in which the proportion or frequency lies.

Competing Reference Classes

Even if the problem of defining the set of "legitimate" reference classes can be resolved, the reference-class approach must still address the problem of choosing the "right" class out of the set of legitimate ones. The solution to this problem has typically been to posit a collection of rules indicating when one reference class should be preferred over another. The basic criterion is the one we already mentioned: choose the most specific class. However, even in the cases to which this specificity rule applies, it is not always appropriate. Assume, for example, that we know that between 70% and 80% of birds chirp and that between 0% and 99% of magpies chirp. If *Tweety* is a magpie, the specificity rule would tell us to use the more specific reference class, and conclude that Pr (*Chirps*(*Tweety*)) [0; 0:99]. Although the interval [0;0:99] is certainly not trivial, it is not very meaningful. Had the 0:99 been a 1, the interval would have been trivial, and we could have then ignored this class and used the more detailed statistics of [0:7; 0:8] derived from the class of birds. The knowledge base above might be appropriate for someone who knows little about magpies, and so feels less confidence in his statistics for magpies than in his statistics for the class of birds as a whole. But since [0:7; 0:8] [0; 0:99], we know nothing that indicates that magpies are actually different from birds in general with respect to chirping. There is an alternative intuition that says that if the statistics for the less specific reference class (the class of birds) are more precise, and they do not contradict the statistics for the more specific class (magpies), then we should use them. That is, we should conclude that Pr (*Chirps*(*Tweety*)) [0:7; 0:8]. This intuition is captured and generalized in Kyburg's strength rule. Unfortunately, neither the specificity rule nor its extension by Kyburg's strength rule are adequate in most cases. In typical examples, the agent generally has several incomparable classes relevant to the problem, so that neither rule applies. Reference-class systems such as Kyburg's and Pollock's simply give no useful answer in these cases. For example, suppose we know that Fred has high cholesterol and is a heavy smoker, and that 15% of people with high cholesterol get heart disease. If this is the only suitable reference class, then (according to all the systems) Pr (*Heart-disease*(*Fred*)) = 0:15. On the other hand, suppose we then acquire the additional information that 9% of heavy smokers develop heart disease (but still have no nontrivial statistical information about the class of people with both attributes). In this case, neither class is the single right reference class, so approaches that rely on finding a single reference class generate a trivial degree of belief that Fred will contract heart disease in this case. For example, Kyburg's system will generate the interval [0; 1] as the degree of belief. Giving up completely in the face of conflicting evidence seems to us to be inappropriate. The entire enterprise of generating degrees of belief is geared to providing the agent with some guidance for its actions (in the form of degrees of belief) when deduction is insufficient to provide a definite answer. That is, the aim is to generate plausible inferences. The presence of conflicting information does not mean that the agent no longer needs guidance. When we have several competing reference classes, none of which dominates the others according to specificity or any other rule that has been proposed, then the degree of belief should most reasonably be some combination of the corresponding statistical values. In general, we can find three conflict types as follows:

(1) Conflict between less and more specific classes,
(2) Conflict between classes associated with less and more precise information, and
(3) Conflict between incomparable classes.

To solve the first and the second conflict types, we are going to modify the basic definition of the direct inference by a symbolic formalization of the *specificity rule* of Reichenbach and the *strength rule* of Kyburg. For the third type, we are going to propose a combination function of symbolic degrees associated with incomparable reference classes. The specificity rule Reichenbach consists of choosing among reference classes, the smallest (specific) class for which we have meaningful information. We propose a symbolic formalization of the specificity rule allowing us to infer the certainty symbolic degree in *B* (*a*) from KB, by choosing information associated with the smallest reference class designed by $\chi'(a|z)$.

Definition 6: Let us suppose that: $KB = (\chi,R)$. The *specificity rule* allows us to infer "$C(B(a)) = \upsilon_\alpha$" if the three following conditions are satisfied:

(1) We have $\{Q_{[1,M]}\ \chi(a|z)$'s are $B(a|z)$'s$\}$ (i.e. from Q_1 to Q_M: total ignorance),
(2) $\exists\ \chi'(a|z)$ such that R^* contains $\{Q_7;\ \chi(a|z)$'s are $\chi'(a|z)\}$ and $\{Q_\alpha\ \chi'(a|z)$'s are $B(a|z)$'s$\}$,
(3) $\neg\ \exists\ \chi''(a|z)$ such that R^* contains $\{Q_7;\ \chi''(a|z)$'s are $\chi'(a|z)\}$ and $\{Q_{\beta\chi}''(a|z)$'s are $B(a|z)$'s$\}$.

Intuitively, the three conditions above express respectively the following:
(*1*) We don't have any meaningful information for the smallest reference class $\chi(a|z)$. Otherwise, the corresponding definition will be used.
(*2*) The existence of a reference class $\chi'(a|z)$ for which we possess a meaningful information.
(*3*) There is no smaller reference class $\chi''(a|z)$ that $\chi'(a|z)$ for which we possess a meaningful information.

Definition 7: Let $KB = (\chi,R)$. The strength rule allows us to derive $F(B(a)) = \upsilon_\alpha$ with $\upsilon\alpha \llcorner [\upsilon c,\upsilon d]$, if the following conditions are satisfied:
(1) R^* contains $\{Q_{\alpha 1}\ \chi(a|z)$'s are $B(a|z)$'s with $Q_{\alpha 1} \in [Q_a,Q_b]\}$,
(2) P^* contains $\{Q_{\alpha 2}\ \chi'(a|z)$'s are $B(a|z)$'s with $Q_{\alpha 2} \in [Q_c,Q_d]\}$, and $[Q_c,Q_d] \subset [Q_c,Q_d]\}$.

We define a combination function denoted by T which is an application of U_M^2 into U_M possessing the following properties:

P_1: $\forall\alpha,\beta \in [2..M]$, $T(\upsilon_\alpha,\upsilon_\beta) = T(\upsilon_\beta,\upsilon_\alpha)$,
P_2: $\forall\alpha,\beta \in [2..M]$, $T(\upsilon_\alpha,\upsilon_\beta) \in [\upsilon_{min(\alpha,\beta)},\upsilon_{max(\alpha,\beta)}]$,
P_3: $\forall\alpha \in [2..M]$, $T(\upsilon_\alpha,\upsilon_1) = \upsilon_1$, where υ_1 is an absorbent element for any $\alpha \in [2..M]$,
P_4: $\forall\alpha \in [1..M\text{-}1]$, $T(\upsilon_\alpha,\upsilon_M) = \upsilon_M$, where υ_M is an absorbent element for any $\alpha \in [1..M\text{-}1]$,
P_5: $\forall\alpha \in [2..M\text{-}1]$, $T(\upsilon_\alpha,\upsilon_{n(\alpha)}) = \upsilon_4$,
P_6: $\forall\alpha,\beta,\gamma \in [1..M]$, $T(T((\upsilon_\alpha,\upsilon_\beta),\upsilon_\gamma) = T(\upsilon_\alpha,T(\upsilon_\beta,\upsilon_\gamma))$,
P_7: $\forall\alpha,\beta,\gamma \in [2..M\text{-}1]$, $T((\upsilon_\alpha,\upsilon_\beta) = \upsilon_\gamma) \Rightarrow T(\upsilon_\alpha,\upsilon_{\beta+1}) \llcorner [\upsilon_\gamma,\upsilon_{\gamma+1}]$.

We can choose the function T as follows: $\forall\alpha,\beta \in [2..M\text{-}1]$:

$$T(\upsilon_\alpha,\upsilon_\beta) = \upsilon_{[(\alpha+\beta)/2]} \text{ if } \alpha+\beta \leq M$$
$$T(\upsilon_\alpha,\upsilon_\beta) = \upsilon_{\alpha+\beta)/2} \text{ if } \alpha+\beta > M$$

The table of the function T is given as follows:

T	υ_1	υ_2	υ_3	υ_4	υ_5	υ_6	υ_7
υ_1	υ_1	υ_1	υ_1	υ_1	υ_1	υ_1	
υ_2	υ_1	υ_2	υ_2	υ_3	υ_3	υ_4	υ_7
υ_3	υ_1	υ_2	υ_3	υ_3	υ_4	υ_5	υ_7
υ_4	υ_1	υ_3	υ_3	υ_4	υ_5	υ_5	υ_7
υ_5	υ_1	υ_3	υ_4	υ_5	υ_5	υ_6	υ_7
υ_6	υ_1	υ_4	υ_5	υ_5	υ_6	υ_6	υ_7
υ_7		υ_7	υ_7	υ_7	υ_7	υ_7	υ_7

Definition 8: Given $KB = (\chi,R)$ for which we have the following:

(1) $\chi(a) \ldots A_1(a)\cap\ldots\cap A_n(a)$, with $n \geq 2$,
(2) R^* contains $Q_{\alpha 1}$ $A_1(a|z)$'s where $\neg\exists\, \alpha_i \in [1..n], \neg\exists\, \alpha_j \in [1..n]$, such that $\alpha_i = 1$ and $\alpha_M = M$.
(3) The classes referred by $A_i(a|z)$ are incomparable.

The certainty degree $\upsilon\alpha$ results from a combination (using the function T) of the certainty degree $\upsilon_{\alpha i}$ associated with these classes, i.e. $T(\upsilon_{\alpha 1}, T(\upsilon_{\alpha 2}, \ldots)) = \upsilon_\alpha$

Bacchus [2] was interested mainly in representing the quantifier "most" e.g., denoting the majority. In our approach, this quantifier can be represented either by the quantifiers "most", "almost-all", or "all" which correspond to "at least most" and we get similar results like as Bacchus. For instance, "*Most native speakers of German are not born in America*", "*All native speakers of Pennsylvanian Dutch are native speakers of German*", "*Most native speakers of Pennsylvanian Dutch are born in Pennsylvania*", "*All people who are born in Pennsylvania are born in America*" and "*Hermann is a native speaker of Pennsylvanian Dutch*". In [2], Bacchus deduced that the probability that "Hermann is born in America is > 0.5", that is to say "*it is probable that Hermann is American*". In our framework, we have a similar statement which is "*it is very probable that Hermann is born in America*". Thus, the results found in our framework are in accordance with those found in Bacchus' direct inference.

Conclusion

In this paper we have proposed a symbolic logical approach using the M-valued logic with syllogistic reasoning and direct inference principle, to manipulate some statistical and qualitative knowledge. This approach allows to reason qualitatively on quantified assertions, since we provide inference rules based upon statements involving linguistic quantifiers. Thus our approach is consistent with the common sense reasoning and similar to Bacchus approaches [2,3]. It can be used in different areas of sciences like Artificial Intelligence and Linguistics for an explicit treatment of uncertainty and fuzzy environment.

REFERENCES

[1] **Akdag, H., M. De Glas, and D. Pacholczyk, 1992,** A Qualitative Theory of Uncertainty. *Fundamenta Informatica*, 17(4): 333-362.

[2] **Bacchus, F., 1990.** Representation and Reasoning with Probabilistic Knowledge. *MIT Press, Cambridge*, MA.

[3] **Bacchus, F., A.J. Grove, J.Y. Halpern, and D. Koller, 1997.** From Statistical Knowledge Bases to Degrees of Belief. *Artificial Intelligence*, 87, 75-143.

[4] **Dubois, D., H. Prade, L. Godo, and R. Mantaras, 1992,** A Symbolic Approach to Reasoning with Linguistic Quantifiers. *Uncertainty in Artificial Intelligence*, Stanford, 74-82.

[5] **Garcia, L., Y Khayata, and D. Pacholczyk, 2001,** A Symbolic Approach to Syllogistic Reasoning and Direct Inference Principle. In Proc. of the International Conference on Artificial Intelligence and Soft Computing, *ASC*'2001, 73-78.

[6] **Garcia, L., Y. Khayata, and D. Pacholczyk, 2001,** A Symbolic Formalisation of Direct Inference Principle, In Proc. of International Conference on Artificial Intelligence IC-AI'2001, Las Vegas, Nevada, USA, 205 - 209.

[7] **Pacholczyk, D. 1994,** A Logico-Symbolic Probability Theory for the Management of Uncertainty. *Artificial Intelligence*, 11(4): 417-484.

[8] **Pacholczyk, D., 1999,** A New Approach to the Linguistic Negation in the Context of Fuzzy Set Theory Based on the Reference Frame and the Scope of a Negation Operator, In Proc. of the Int. Congress on Computational Intelligence: *Methods and Applications, Academic Press*, USA, 314-322.

[9] **Reichenbach, H., 1949,** Theory of Probability. *University of California Press*, Berkeley, USA.

[10] **Yager, R.R. 1985,** Reasoning with Quantified Statements. *Part I, Kybernetes*, 14, 233-240.

[11] **Yager, R.R., 1986,** Reasoning with Quantified Statements. *Part II, Kybernetes*, 14, 111-120.

[12] **Zadeh, L.A., 1985,** Syllogistic Reasoning in Fuzzy Logic and its Application to Usuality and Reasoning with Dispositions. *In IEEE Transaction on Systems, Man and Cybernetics*, 15(6): 754-763.