

New Directions For Uncertainty Reasoning In Deductive Databases

U. Guntzer¹

W. Kießling²

H. Thöne¹

¹Wilhelm-Schickard-Institut, Universität Tübingen, Sand 13, D-7400 Tübingen 1

²Institut für Informatik, Ludwig-Maximilians-Universität, Theresienstr. 39, D-8000 München 2

Abstract

This paper contributes a novel approach to non-monotonic uncertainty reasoning, which is ubiquitous in many real-life applications. Founded on the paradigm of conditional probabilities we develop a rule-based calculus and prove that it is sound, even in the presence of incomplete information. Thus the merits of doing consistent judgements in uncertain domains and the advantages of modularity and incrementality of rule-based application development come together. We also can offer mechanisms to trace down inconsistencies that may be hidden in very large collections of uncertain rules. As next-generation applications will have to handle vast amounts of uncertain data, an integration into databases is mandatory. We give a direct implementation of our calculus on top of a database system with a DATALOG-interface. In this way we extend current database technology towards providing new applications with new suitable primitives and with a database platform for coping with uncertainty.

1. Introduction

Databases have been very successful over the years in managing large amounts of *certain* data, multi-user access and high-level query languages to ease application development. Such languages like SQL, declarative rule-based languages like DATALOG ([Ull 89], [CGT 89]) or object-oriented query languages cannot deal with uncertain knowledge, though.

As stressed in [Lag 90] recently, next-generation applications, such as the identification of features from satellite images, require reasoning under uncertainty to deduce conclusions from several inter-related partial or alternative results. This points out the need for a fundamental theory that can deal with non-monotonic uncertainty. Medical diagnosis, fault analysis, genetics or robot vision are other important application domains falling into this category.

While the database community so far has brought limited attention to this problem (see e.g. [BMP 90]), it has been a major area of research within AI for many years. The variety of solution proposals differs tremendously in the basic foundations, like Dempster/Shافر evidence theory, default reasoning, fuzzy logic systems, only to name a few ([Ric 86], [Sha 76], [HeHo 87], [Zad 86]). Statistical expert systems are another main line to attack the problem. For the efficient solution of small, well-defined problems or for selected subproblems within a larger task a variety of techniques based on statistics and graph-theoretical methods have proven useful ([Hun 75],[LaSp 88]). Our emphasis here, however, is on large and dynamically growing applications requiring a general-purpose rule-based approach, where changes can be accommodated in small increments understandable by the human user.

The success of expert systems in uncertain domains has been very modest until now, with little impact on commercial data processing. The reasons for this are twofold. First, the inference mechanisms used often led to unpredictable, uninterpretable or counterintuitive conclusions or depend on unreasonable assumptions (independence of events, complete information available). Second, only main-memory implementations exist presently. However, there have been gained many fundamental insights into the nature of uncertainty, to mention in particular [Pea 88]. On the other hand, nowadays we know how to manage declarative, rule-based certain knowledge efficiently in a database environment (e.g. LDL [NaTs 89],[Zan 88] or commercial systems like Declare [KiGu 90]). Similar to the situation back

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1991 ACM 0-89791-425-2/91/0005/0178...\$1.50

in the mid 80's, where ideas from AI and databases merged to give birth to deductive database technology, a confluence of the virtues of AI- and DB-technology in the area of uncertainty reasoning is intended.

The rest of this paper is organized as follows: Section 2 faces the reader with some startling phenomena inherent in uncertainty reasoning, pointing out the difficulties. In section 3 we introduce our solution of uncertain rules with conditional probabilities. The presented novel calculus is proved to be sound. Several examples show that it treats the mentioned phenomena correctly, like rule-chaining or explaining-away. Section 4 then describes how to implement this calculus on top of a deductive DB-system with a DATALOG interface and discusses questions of optimization and consistency support. In conclusion, section 5 summarizes our contributions and points out areas of future work.

2. Reasoning Phenomena Under Uncertainty

Some introductory examples demonstrate a typical derivation strategy applied in human processing of imperfect information: the combination of two rules by chaining, which yields sometimes strange conclusions.

Example (Barking cats) For illustrative purpose, assume that 70% of all domestic animals are dogs and 90% of all dogs can bark:

$$\begin{array}{l} \text{domestic_animals} \xrightarrow{0.7} \text{dogs} \\ \text{dogs} \xrightarrow{0.9} \text{bark} \end{array}$$

Putting these assumptions together by multiplying certainties we may come to the conclusion that (at least) 63% of all domestic animals can bark. This conclusion seems to be true. But now, consider the following situation:

$$\begin{array}{l} \text{cats} \xrightarrow{1.0} \text{domestic_animals} \\ \text{domestic_animals} \xrightarrow{0.63} \text{bark} \end{array}$$

With the same naive chaining method we come to the conclusion that 63% of all cats can bark, what is obviously strange. Thus, there must be **hidden** parameters distinguishing between the above cases, influencing the certainty of the conclusion. Here, the hidden parameters are evident. We didn't represent the information which we implicitly assumed: *All* dogs are domestic animals, and cats and dogs are *different* domestic animals. Nevertheless, a reasoning system mustn't come to the wrong conclusion above.

In the next example ([Pea 88]), we like to combine causal and diagnostic information.

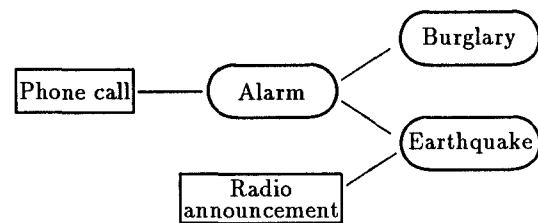
Example (Sprinkler)

$$\begin{array}{l} \text{Sprinkler_was_on} \xrightarrow{1.0} \text{Ground_is_wet} \quad (\text{causal}) \\ \text{It_rained_last_night} \xrightarrow{1.0} \text{Ground_is_wet} \quad (\text{causal}) \\ \text{Ground_is_wet} \xrightarrow{0.9} \text{It_rained_last_night} \quad (\text{diagnostic}) \\ \text{Ground_is_wet} \xrightarrow{0.1} \text{Sprinkler_was_on} \quad (\text{diagnostic}) \end{array}$$

When we combine the first causal and the first diagnostic rule in a naive way again, we get that "Sprinkler_was_on" implies "It_rained_last_night" with 90 percent. Here we do not combine the causal and diagnostic knowledge properly. By utilizing the given knowledge we would at least expect that the special cause "Sprinkler_was_on" doesn't support the general cause "It_rained_last_night".

The non-monotonic nature in uncertainty reasoning will be demonstrated in the following example ([Pea 88]).

Example (Burglary) Imagine, Mr. Holmes receives a phone call at work from his neighbor, who states that he hears an alarm from the direction of Mr. Holme's house. Mr. Holmes believes that a burglary is the cause for the alarm sound and decides to rush home. In his car he turns on his radio and listens to the news telling that an earthquake had occurred. Now Mr. Holmes reduces his initial belief in a burglary, since the alarm may be triggered by the earthquake.



The radio announcement reduces the likelihood of a burglary; i.e. the second event earthquake "explains away" burglary as cause for the alarm sound. This non-monotonic reasoning phenomenon is caused by considering different contexts. In the first situation, only the simple event alarm is relevant, whereas then two events alarm and earthquake must be considered.

While the simplicity of material implications has proved viable in predicate logic¹, it immediately breaks down in uncertain domains. In going from logic implications to uncertain implications, we lose the feature of modularity and incrementality ([Pea 88]), that

¹Quoting Tarski ([Tar 66]): " ... daß die auf diesem einfachen Implikationsbegriff begründete Logik sich als eine zufriedenstellende Basis für die kompliziertesten und subtilsten mathematischen Überlegungen erwiesen hat."

is the impact of new evidence can neither be imparted in stages nor separately. Naive approaches to reason with uncertainty fail because uncertainty is combined sloppily. But we must not give up the rule-based paradigm indispensably needed for the construction of large and dynamic applications.

3. A New Probabilistic Calculus

We investigate the problem of uncertainty reasoning between entire sets of events. This sort of reasoning seems to be prevailing in practice; uncertain reasoning about individuals seems to be the exception and will be taken care of in a companion paper. As we have to deal with uncertain correlations between sets, probabilistic reasoning is the prime paradigm to consider, because probability theory is highly developed and rests on a firm ground. Since we aim at a rule-based calculus with an intuitive meaning to everybody, we must make the right choice what is the primitive of our language. Conditional probabilities turn out to be an adequate candidate.

Definition 3.1 Let A, B be sets of events and let AB denote the intersection of A and B . The conditional probability of B given A is defined as

$$P(B|A) = \frac{P(AB)}{P(A)}, \text{ if } P(A) > 0.$$

Note that $P(B|A)$ is only partially defined. The equivalent rule-based interpretation is:

$$A \xrightarrow{P(B|A)} B$$

That is, given A , $P(B|A)$ among the events in A are also events in B . If we have both $A \xrightarrow{x} B$ and $B \xrightarrow{y} A$, we can also write $A \xleftrightarrow[x]{y} B$. We often want to combine rules. This can be modeled by a graph representation:

$$A \xrightarrow{x} B \xrightarrow{y} C$$

A strong clue that conditional probabilities are suitable to describe the previous non-monotonic phenomena adequately is due to the following: given $P(A|B)$, then $P(A|BC)$ can take any value between 0 and 1. In other words, the additional evidence C may diminish or amplify the confidence in A . Conditional probabilities are psychologically meaningful to most people and therefore provide an excellent guidance for decisions, based on the numerical probabilities. But since precise conditional probabilities are often hard to get or not available, working with intervals makes much more sense in practice.

Definition 3.2 Let C_1, C_2, \dots, C_k be the set of events. C_l and its complement $\bar{C}_l, 1 \leq l \leq k$, are called basic events. We consider conjunctive events $A = A_1 \dots A_n, B = B_1 \dots B_m$, where $n, m \geq 1$ and A_i, B_j are basic events, and $P(A) > 0$. An uncertain rule consists of an upper and a lower bound for a conditional probability:

$$A \xrightarrow{x_1, x_2} B \text{ iff } 0 \leq x_1 \leq P(B|A) \leq x_2 \leq 1.$$

If lower and upper bound coincide we simply write $A \xrightarrow{x} B$.

The axioms of our calculus are given by a collection \mathcal{R} of uncertain rules, respecting the laws of probability theory. Note that we do not impose an a priori causal or diagnostic interpretation on the implication arrow " $\xrightarrow{\quad}$ ". This is left to the application completely. Given a knowledge base for an application involving uncertainty in form of such axioms \mathcal{R} among a set of events C_1, \dots, C_k , we want to draw other uncertain conclusions, not explicitly listed in \mathcal{R} . This will be accomplished by a collection of inference rules, which achieve to combine few pieces of influence into new evidence.

Definition 3.3 Let \mathcal{R} be a set of uncertain rules, A and B be conjunctive events. $\mathcal{R} \vdash A \xrightarrow{x_1, x_2} B$ iff $A \xrightarrow{x_1, x_2} B$ can be generated, given \mathcal{R} , by the following inference rules in a finite number of steps.

Inference Rules

(A, B, C and D denote conjunctive events, F denotes a basic event.)

(1) Chaining (C):

$$\{ A \xrightarrow{x_1, x_2} FC, A \xrightarrow{y_1, y_2} \bar{F}C \} \vdash A \xrightarrow{z_1, z_2} C, \\ z_1 = x_1 + y_1, z_2 = \min(1, x_2 + y_2)$$

(2) Sharpening (S):

$$\{ A \xrightarrow{x_1, x_2} B, A \xrightarrow{y_1, y_2} B \} \vdash A \xrightarrow{z_1, z_2} B, \\ z_1 = \max(x_1, y_1), z_2 = \min(x_2, y_2)$$

(3) Conjunction left (CL):

$$\{ A \xrightarrow{x_1, x_2} B, x_1 > 0, A \xrightarrow{y_1, y_2} BC \} \vdash AB \xrightarrow{z_1, z_2} C, \\ z_1 = \frac{y_1}{x_2}, z_2 = \min(1, \frac{y_2}{x_1})$$

(4) Conjunction right (CR):

$$\{ A \xrightarrow{x_1, x_2} B, AB \xrightarrow{y_1, y_2} C \} \vdash A \xrightarrow{z_1, z_2} BC, \\ z_1 = x_1 \cdot y_1, z_2 = x_2 \cdot y_2$$

(5) Weak conjunction left (WCL):

$$\{ A \xrightarrow{x_1, x_2} B, x_1 > 0, B \xrightarrow{y_1, y_2} C \} \vdash AB \xrightarrow{z_1, z_2} C,$$

$$z_1 = \frac{1}{x_1} \cdot T_1(x_1, y_1), z_2 = \frac{1}{x_1} \cdot T_3(x_1, y_2), \text{ where}$$

$$T_1(a, b) = \max(0, a + b - 1), T_3(a, b) = \min(a, b)$$

(6) Weak conjunction right (WCR):

$$\{ A \xrightarrow{x_1, x_2} B, D \xrightarrow{y_1, y_2} C, D = B \text{ or } D = \bar{B}, A \neq C \}$$

$$\vdash A \xrightarrow{z_1, z_2} BC,$$

if $(y_1 = 1 \wedge D = B)$ then $z_1 = x_1$ else $z_1 = 0$,

if $(y_2 = 0 \wedge D = B)$ then $z_2 = 0$ else $z_2 = x_2$

(7) Negation (N):

$$\{ A \xrightarrow{x_1, x_2} F \} \vdash A \xrightarrow{z_1, z_2} \bar{F}, z_1 = 1 - x_2, z_2 = 1 - x_1$$

(8) Weak conjunction right with negation (WCRN):

$$\{ A \xrightarrow{u_1, u_2} F, v_1 > 0, F \xrightarrow{x_1, x_2} C, y_1 > 0, A \neq C \}$$

$$\vdash A \xrightarrow{z_1, z_2} \bar{F}C,$$

$$z_1 = 0, z_2 = \min(1, (1 - y_1) \cdot \frac{u_2 \cdot x_2}{v_1 \cdot y_1})$$

(9) Annulment (A):

$$\{ A \xrightarrow[0]{} B, A \xrightarrow{x_1, x_2} B \} \vdash A \xrightarrow[0]{} B$$

We denote all uncertain rules that can be generated out of a set of axioms \mathcal{R} by applying these inference rules as *deduced rules*. Already deduced rules can be used as premises for deducing further rules.

Definition 3.4 Let \mathcal{R} be a set of axioms, A and B be conjunctive events.

$\mathcal{R} \models A \xrightarrow{x_1, x_2} B$ iff $x_1 \leq P(B|A) \leq x_2$ follows from \mathcal{R} and the laws of probability.

As an indispensable property of any rule-based uncertainty calculus, we now show its soundness within a well-understood mathematical framework.

Theorem 3.5 (Soundness)

$$\text{If } \mathcal{R} \vdash A \xrightarrow{x_1, x_2} B \text{ then } \mathcal{R} \models A \xrightarrow{x_1, x_2} B$$

Proof: By induction on the number of steps in a deduction it suffices to show that each rule is valid within probability theory. For instance, take the chaining rule: We have $P(C|A) = P(CF|A) + P(C\bar{F}|A)$. Since we have $x_1 \leq P(CF|A) \leq x_2$ and $y_1 \leq P(C\bar{F}|A) \leq y_2$, we get $x_1 + y_1 \leq P(C|A) \leq \min(1, x_2 + y_2)$. For the remaining inference rules we leave the proof to the reader. ■

One of the most important uses of any rule-based calculus is rule chaining. Concretely, we want to answer the question: Given an uncertain relationship between A and B , and an uncertain relationship between B and C , what is the relationship between A and C ? This amounts to making *transitive* conclusions under uncertainty.

Theorem 3.6 (Rule chaining RC) Let \mathcal{R} be a set of axioms or deduced rules, A and C conjunctive events, B a basic event.

$\mathcal{R} = \{ A \xrightarrow{u_1, u_2} B, B \xrightarrow{x_1, x_2} C \}$. Then $\mathcal{R} \vdash A \xrightarrow{z_1, z_2} C$ with

$$z_1 = \begin{cases} \frac{u_1}{v_1} \cdot T_1(v_1, x_1) & \text{if } v_1 > 0 \\ u_1 & \text{if } v_1 = 0 \text{ and } x_1 = 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$z_2 = \begin{cases} \min(1, u_2 + \tau \cdot (1 - y_1), 1 - u_1 + \tau \cdot y_1, \tau) & \\ \quad \text{with } \tau = \frac{u_2 x_2}{v_1 y_1} & \text{if } v_1 > 0 \text{ and } y_1 > 0 \\ \min(1, 1 - u_1 + \frac{u_2 x_2}{v_1}) & \text{if } v_1 > 0 \text{ and } y_1 = 0 \\ 1 - u_1 & \text{if } v_1 = 0 \text{ and } x_2 = 0 \\ 1 & \text{otherwise.} \end{cases}$$

Note that the rule-chaining theorem can also be applied when $B \xrightarrow{y_1, y_2} C$ is not given. Only u_1 , v_1 and x_1 are needed to compute the lower bound z_1 ; the upper bound z_2 can be computed, for instance, with the formula $\min(1, 1 - u_1 + \frac{u_2 x_2}{v_1})$, if $v_1 > 0$.

Proof of theorem 3.6: (By deduction in the calculus.)

$$(1) \{ A \xrightarrow{u_1, u_2} B, v_1 > 0, B \xrightarrow{x_1, x_2} C \} \vdash_{\text{WCL}} AB \xrightarrow{a_1, a_2} C,$$

$$a_1 = \frac{1}{v_1} \cdot T_1(v_1, x_1), a_2 = \frac{1}{v_1} \cdot T_3(v_1, x_2)$$

$$(2) \{ A \xrightarrow{u_1, u_2} B, AB \xrightarrow{a_1, a_2} C \} \vdash_{\text{CR}} A \xrightarrow{a_3, a_4} BC,$$

$$a_3 = u_1 \cdot a_1, a_4 = u_2 \cdot a_2$$

$$(3) \mathcal{R} \vdash_{\text{WCRN}} A \xrightarrow{a_5, a_6} \bar{B}C, a_5 = 0,$$

$$a_6 = \min(1, (1 - y_1) \cdot \frac{u_2 \cdot x_2}{v_1 \cdot y_1}) \text{ if } v_1 > 0 \text{ and } y_1 > 0$$

$$(4) \{ A \xrightarrow{a_3, a_4} BC, A \xrightarrow{a_5, a_6} \bar{B}C \} \vdash_C A \xrightarrow{a_7, a_8} C,$$

$$a_7 = a_3 + a_5, a_8 = \min(1, a_4 + a_6)$$

$$(5) \{ A \xrightarrow{u_1, u_2} B \} \vdash_N A \xrightarrow{a_9, a_{10}} \bar{B},$$

$$a_9 = 1 - u_2, a_{10} = 1 - u_1$$

$$(6) \{ A \xrightarrow{a_9, a_{10}} \bar{B}, B \xrightarrow{x_1, x_2} C \} \vdash_{\text{WCR}} A \xrightarrow{a_{11}, a_{12}} \bar{B}C,$$

$$a_{11} = 0, a_{12} = a_{10}$$

$$(7) \{ A \xrightarrow{a_3, a_4} BC, A \xrightarrow{a_{11}, a_{12}} \bar{B}C \} \vdash_C A \xrightarrow{a_{13}, a_{14}} C,$$

$$a_{13} = a_3 + a_{11}, a_{14} = \min(1, a_4 + a_{12})$$

$$(8) \{ A \xrightarrow{u_1, u_2} B, B \xrightarrow{x_1, x_2} C \} \underset{wcr}{\vdash} A \xrightarrow{a_{15}, a_{16}} BC,$$

if $x_1 = 1$ then $a_{15} = u_1$ else $a_{15} = 0$,

if $x_2 = 0$ then $a_{16} = 0$ else $a_{16} = u_2$

$$(9) \{ A \xrightarrow{a_{15}, a_{16}} BC, A \xrightarrow{a_5, a_6} \overline{BC} \} \underset{c}{\vdash} A \xrightarrow{a_{17}, a_{18}} C,$$

$a_{17} = a_{15} + a_5, a_{18} = \min(1, a_{16} + a_6)$

$$(10) \{ A \xrightarrow{a_{15}, a_{16}} BC, A \xrightarrow{a_{11}, a_{12}} \overline{BC} \} \underset{c}{\vdash} A \xrightarrow{a_{19}, a_{20}} C,$$

$a_{19} = a_{15} + a_{11}, a_{20} = \min(1, a_{16} + a_{12})$

Applying the sharpening rule and basic arithmetic we finally get the result. ■

As an important observation, it is now clear that sound rule chaining necessarily requires *bi-directional inference* (" \longleftrightarrow "): also the "back-arrows" $A \xleftarrow{v_1, v_2} B$ and $B \xleftarrow{y_1, y_2} C$ are essential. Our introductory examples now behave as follows:

Example (Barking cats) The set of axioms, i. e. the base rules \mathcal{R} , are:

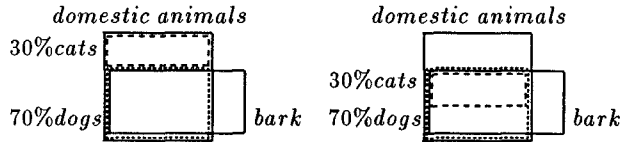
$$\{ domestic \xleftarrow{0.7} dogs, dogs \xrightarrow{0.9} bark, cats \xleftarrow{0.3} domestic \}$$

Applying the rule chaining theorem (RC), we get:

$$\{ domestic \xleftarrow{0.7} dogs, dogs \xrightarrow{0.9} bark \} \underset{RC}{\vdash} domestic \xrightarrow{0.63, 0.93} bark$$

$$\{ cats \xleftarrow{0.3} domestic, domestic \xrightarrow{0.63, 0.93} bark \} \underset{RC}{\vdash} cats \xrightarrow{0.0, 1.0} bark$$

We can give a plausible interpretation of these results using Venn diagrams:



Observe that the three sets can overlap as much as possible (*positive correlation*, see right picture), which yields $cats \subseteq dogs$ or overlap as little as possible (*negative correlation*, see left picture), yielding $cats \cap dogs = \emptyset$. The functions $T_1(a, b)$ and $T_3(a, b)$, given in the set of inference rules, precisely account for negative or positive correlation, resp. (T_1 and T_3 are known as T-norms [BoDe 86]).

As it turns out, the information specified in the base rules is not sufficient to conclude anything about the barking behavior of cats which is reflected in our calculus by the conclusion $cats \xrightarrow{0.0, 1.0} bark$. So imagine that you learn today that cats aren't dogs and that the dogs are the only barking animals for sure. Then you can *incrementally* update your knowledge base by adding two rules: $\mathcal{R} := \mathcal{R} \cup \{ cats \xrightarrow{0.0} dogs, dogs \xleftarrow{1.0} bark \}$

Now we get: $\mathcal{R} \underset{RC, A}{\vdash} cats \xleftarrow{0.0} bark$.

It comes to no surprise that our calculus does not produce erroneous results. Even better there are also *plausible explanations* in form of correlation behavior as shown above.

Example (Sprinkler)

$$\{ Sprinkler_was_on \xleftarrow{1.0} Ground_is_wet,$$

$$Ground_is_wet \xleftarrow{0.9} It_rained_last_night \}$$

$$\underset{RC}{\vdash} Sprinkler_was_on \xleftarrow{0.0, 1.0} It_rained_last_night$$

Again we avoid the erroneous conclusion that the activities of the sprinkler imply rain. Instead our calculus admits honestly that nothing can be concluded about the probability of rain given the sprinkler event. Everything else, computing preciser estimates than what is implied by the input rules, would amount to a deception of the user.

As direct corollary from the rule chaining theorem, we get the following important special cases for our calculus, which again proves its consistency in treating uncertain reasoning.

Corollary 3.7 (Certain chaining) Let \mathcal{R} be a set of axioms or deduced rules:

$$\mathcal{R} = \{ A \xleftarrow{1} B, B \xleftarrow{1} C \}. \text{ Then } \mathcal{R} \underset{RC}{\vdash} A \xrightarrow{1} C.$$

Proof: Apply the rule chaining theorem with $u_1 = u_2 = x_1 = x_2 = 1$. ■

Corollary 3.8 (Product rule) Let \mathcal{R} be a set of axioms or deduced rules:

$$\mathcal{R} = \{ A \xleftarrow{u_1, u_2} B, B \xleftarrow{x_1, x_2} C \}. \text{ Then } \mathcal{R} \underset{RC}{\vdash} A \xrightarrow{u_1 \cdot x_1, u_2 \cdot x_2} C.$$

Proof: Apply the rule chaining theorem with $v_1 = v_2 = y_1 = y_2 = 1$. ■

Coming back to our initial attempts with naive rule chaining we now see that this is sound in this special case.

Let us draw our attention to the question, how our calculus deals with non-monotonic behavior.

Theorem 3.9 (Multiple events) Let \mathcal{R} be the following set of axioms or deduced rules:

$$\mathcal{R} = \{ A \xleftarrow{u_1, u_2} B, u_1 > 0 \text{ or } v_1 > 0, A \xrightarrow{x_1, x_2} C, B \xrightarrow{y_1, y_2} C \}.$$

$$\text{Then } \mathcal{R} \vdash AB \xrightarrow{z_1, z_2} C,$$

$$z_1 = \max\left(\frac{1}{u_1} \cdot T_1(u_1, x_1), \frac{1}{v_1} \cdot T_1(v_1, y_1)\right),$$

$$z_2 = \min\left(\frac{1}{u_1} \cdot T_3(u_1, x_2), \frac{1}{v_1} \cdot T_3(v_1, y_2)\right).$$

Proof:

$$(1) \{ A \xrightarrow{u_1, u_2} B, u_1 > 0, A \xrightarrow{x_1, x_2} C \} \underset{\text{wCL}}{\vdash} AB \xrightarrow{a_1, a_2} C, \\ a_1 = \frac{1}{u_1} \cdot T_1(u_1, x_1), a_2 = \frac{1}{u_1} \cdot T_3(u_1, x_2)$$

$$(2) \{ B \xrightarrow{v_1, v_2} A, v_1 > 0, B \xrightarrow{y_1, y_2} C \} \underset{\text{wCL}}{\vdash} AB \xrightarrow{a_3, a_4} C, \\ a_3 = \frac{1}{v_1} \cdot T_1(v_1, y_1), a_4 = \frac{1}{v_1} \cdot T_3(v_1, y_2)$$

$$(3) \{ AB \xrightarrow{a_1, a_2} C, AB \xrightarrow{a_3, a_4} C \} \underset{s}{\vdash} AB \xrightarrow{z_1, z_2} C, \\ z_1 = \max(a_1, a_3), z_2 = \min(a_2, a_4)$$

Substituting we get the result. ■

The use of this formula is required in the following situations:

Definition 3.10 *A and B are mutually supportive for C iff $P(C|AB) > \max(P(C|A), P(C|B))$.*

C is a more probable explanation of A than B iff $P(C|A) > P(C|B)$.

B explains away C iff $P(C|AB) < P(C|A)$.

Example (Burglary) For A = alarm, B = earthquake, C = burglary assume

$$\mathcal{R} = \{ A \xleftarrow[0.7]{0.06} B, A \xrightarrow{0.95} C, B \xrightarrow{0.2} C \}.$$

Given A (alarm is reported), C is a more probable explanation than B . Now given also B (earthquake is reported after alarm), we get:

$$\mathcal{R} \vdash AB \xrightarrow{0.17, 0.29} C.$$

Since $[0.17, 0.29] < [0.95, 0.95]$ B explains away C ; i.e. , your expectation of burglary is sharply reduced and hence A and B are definitely not supporting events.

Example (Symptoms) Let A and B be symptoms for a disease C and suppose

$$\mathcal{R} = \{ A \xleftarrow[0.54]{0.64} B, A \xrightarrow{0.36} C, B \xrightarrow{0.40} C \}.$$

We get: $\mathcal{R} \vdash AB \xrightarrow{0.0, 0.56} C$.

Thus, in case of positive correlation we have $P(C|AB) = 0.56 > \max(0.36, 0.40)$, meaning A and B are mutually supportive of C , whereas in case of negative correlation they aren't. Continuing the example, assume we somehow get to know that - with everything else unchanged - $BC \xrightarrow{0.67} A$. This allows us to infer:

$$\mathcal{R} \cup \{ BC \xrightarrow{0.67} A \} \underset{\text{CR,CL}}{\vdash} AB \xrightarrow{0.50} C$$

Note that we have inferred a 1-point interval again. As now $0.50 > \max(0.36, 0.40)$, we safely know that A and B are mutually supportive of C .

As this example demonstrates, it is often desirable to have some sort of correlation information available. In the absence of concrete measurements, in many cases the assumption of conditional independence is helpful.

Definition 3.11 *A is independent of C under condition B, denoted by $I(A, B, C)$, iff $P(A|BC) = P(A|B)$.*

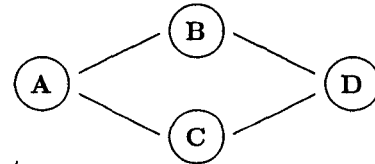
As shown in [Pea 88], there are sound inference rules, which are conjectured to be complete, to generate new conditional independences out of a given set. The combination of independence information and conditional probabilities is done by the following inference rules.

$$\text{Invariance I1: } \{ B \xrightarrow{x_1, x_2} C, I(A, B, C) \} \vdash AB \xrightarrow{x_1, x_2} C$$

$$\text{Invariance I2: } \{ AB \xrightarrow{x_1, x_2} C, I(A, B, C) \} \vdash B \xrightarrow{x_1, x_2} C$$

The two more rules are obviously sound because $I(A, B, C) \iff I(C, B, A)$. They generate more precise intervals than, e.g. , weak conjunction left.

Example (Generic trait) (cf.[AnHo 90]) In the following figure, we can imagine that node D represents the proposition that a patient has a certain symptom. Nodes B and C represent assertions that the patient has either of two diseases that may cause the symptom, which we may call disease B and disease C . Node A represents the proposition that the patient has a genetic trait that predisposes him to both diseases.



Assume that

$$\mathcal{R} = \{ I(C, A, B), I(\bar{C}, A, B), I(A, BC, D), I(A, \bar{B}\bar{C}, D), \\ I(A, B\bar{C}, D), I(A, \bar{B}C, D), BC \xrightarrow{0.95} D, \bar{B}C \xrightarrow{0.8} D, \\ B\bar{C} \xrightarrow{0.7} D, \bar{B}\bar{C} \xrightarrow{0.1} D, A \xrightarrow{0.4} B, A \xrightarrow{0.2} C \}.$$

Then we can generate the following rules:

$$\mathcal{R} \underset{\text{I1,N}}{\vdash} U_1 = \{ ABC \xrightarrow{0.95} D, A\bar{B}C \xrightarrow{0.8} D, AB\bar{C} \xrightarrow{0.7} D, \\ A\bar{B}\bar{C} \xrightarrow{0.1} D, AC \xrightarrow{0.4} B, A\bar{C} \xrightarrow{0.4} B, A \xrightarrow{0.6} \bar{B}, A \xrightarrow{0.8} \bar{C} \}$$

$$U_1 \underset{\text{CR,N}}{\vdash} U_2 = \{ AC \xrightarrow{0.38} BD, A\bar{C} \xrightarrow{0.28} BD, AC \xrightarrow{0.6} \bar{B}, A\bar{C} \xrightarrow{0.6} \bar{B} \}$$

$$U_1 \cup U_2 \underset{\text{CR}}{\vdash} U_3 = \{ AC \xrightarrow{0.48} \bar{B}D, A\bar{C} \xrightarrow{0.06} \bar{B}D \}$$

$$U_2 \cup U_3 \underset{c}{\vdash} U_4 = \{ AC \xrightarrow{0.86} D, A\bar{C} \xrightarrow{0.34} D \}$$

$$\mathcal{R} \cup U_1 \cup U_4 \underset{\text{CR}}{\vdash} U_5 = \{ A \xrightarrow{0.17} CD, A \xrightarrow{0.27} \bar{C}D \}$$

$$U_5 \underset{c}{\vdash} U_6 = \{ A \xrightarrow{0.44} D \}$$

Note that rule chaining via B and C has produced a sharp (1-point) interval again. At this point it's worth contemplating a bit about the way our calculus copes with transitive uncertainty. As we just demonstrated, with sufficient correlation information, precise intervals will be deduced. This clearly distinguishes our approach from other correlation-techniques (e.g. implemented in INFERNO [Qui 83]), where accumulated intervals have a tendency to diffuse towards $[0, 1]$.

4. Extending Deductive Database Technology

Unlike other approaches for uncertainty reasoning, our calculus, in addition to its semantical virtues, can be implemented directly with deductive database technology. To capture the semantics of an application domain, the knowledge engineer has to determine the following factual knowledge:

- (1) The relevant sets of events C_1, \dots, C_k : On a deductive DB these are database relations, on an object-oriented DB it might be classes or relations.
- (2) Basic conditional probabilities between those sets of events.
- (3) Conditional independence information.

The result of this knowledge acquisition process, which can be done in a modular and incremental manner, is a set of axioms \mathcal{R} as described previously. Now, \mathcal{R} plus the set of sound inference rules can be implemented directly on a deductive database system with a $DATALOG^{f_{un+set}}$ rule interface. We define three base relations as follows:

$$\begin{aligned} br(A, B, X1, X2) &: \iff A \xrightarrow{X1, X2} B \in \mathcal{R} \\ ir(A, B, C) &: \iff I(A, B, C) \in \mathcal{R} \\ bar(C, Cbar) &: \iff Cbar = \overline{C}, C \text{ is a basic event.} \end{aligned}$$

The relation bar is symmetric, too. New conditional probabilities, which can be deduced by our inference mechanism ("deduced rules") are implemented by a virtual relation

$$dr(A, B, X1, X2) : \iff \mathcal{R} \vdash A \xrightarrow{X1, X2} B$$

The attribute types for dr are chosen as

- (1) $X1, X2 \in [0, 1]$ (of type real or better type infinite precision if available).
- (2) A, B are of type set of strings: This allows us conveniently to remove "redundant" conjunctions. E.g. ,

if $A = C_1 C_2 C_1, B = C_3 C_3 C_1$ then $A \xrightarrow{0.3, 0.8} B$ is represented as $dr(\{C_1, C_2\}, \{C_1, C_3\}, 0.3, 0.8)$. This set implementation also accounts for associativity, commutativity and idempotence of conjunctions. The identity $C_i = \overline{\overline{C_i}}$ is controlled by means of the bar -relation. For instance, $D = C_1 \overline{\overline{C_3}} C_2$ is recognized as $\{C_1, C_2, C_3\}$ and $D = C_1 C_3 \overline{C_1}$ as $\{\}$. The evaluable predicate $union(A, B, S)$ is used to implement the *intersection* of conjunctive events. For the above example AB is mapped onto $union(A, B, S)$, where S becomes $\{C_1, C_2\} \cup \{C_1, C_3\} = \{C_1, C_2, C_3\}$.

Then dr can be defined by the following nonlinearly recursive $DATALOG^{f_{un+set}}$ program in LDL-notation ([NaTs 89]).²

$$\begin{aligned} dr(A, B, Z1, Z2) &\leftarrow br(A, B, Z1, Z2) \quad \% \text{ axioms} \\ dr(A, C, Z1, Z2) &\leftarrow \\ &\quad dr(A, BC, X1, X2), \quad union(B, C, BC), \quad bar(B, Bbar), \\ &\quad dr(A, BbarC, Y1, Y2), \quad union(Bbar, C, BbarC), C^{\sim}=\{\}, \\ &\quad Z1 = X1 + Y1, \quad Z2 = \min(1.0, X2 + Y2). \\ dr(A, B, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad dr(A, B, Y1, Y2), \\ &\quad Z1 = \max(X1, Y1), \quad Z2 = \min(X2, Y2). \\ dr(AB, C, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad union(A, B, AB), \quad X1 > 0.0, \\ &\quad dr(A, BC, Y1, Y2), \quad union(B, C, BC), \quad C^{\sim}=\{\}, \\ &\quad Z1 = \frac{Y1}{X2}, \quad Z2 = \min(1.0, \frac{Y2}{X1}). \\ dr(A, BC, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad union(A, B, AB), \\ &\quad dr(AB, C, Y1, Y2), \quad union(B, C, BC), \\ &\quad Z1 = X1 \cdot Y1, \quad Z2 = X2 \cdot Y2. \\ dr(AB, C, Z1, Z2) &\leftarrow \\ &\quad dr(B, A, X1, X2), \quad union(A, B, AB), \quad X1 > 0.0, \\ &\quad dr(B, C, Y1, Y2), \quad intersection(A, C, \{\}), \\ &\quad Z1 = \max(0.0, \frac{X1 + Y1 - 1}{X1}), \quad Z2 = \min(1.0, \frac{Y2}{X1}). \\ dr(A, BC, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad union(B, C, BC), \\ &\quad dr(B, C, Y1, Y2), \quad intersection(A, C, \{\}), \\ &\quad if(Y1 = 1.0 \text{ then } Z1 = X1 \text{ else } Z1 = 0.0), \\ &\quad if(Y2 = 0.0 \text{ then } Z2 = 0.0 \text{ else } Z2 = X2). \\ dr(A, BC, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad union(B, C, BC), \quad bar(B, Bbar), \\ &\quad dr(Bbar, C, Y1, Y2), \quad intersection(A, C, \{\}), \\ &\quad Z1 = 0.0, \quad Z2 = X2. \\ dr(A, Bbar, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, X1, X2), \quad bar(B, Bbar), \\ &\quad Z1 = 1.0 - X2, \quad Z2 = 1.0 - X1. \\ dr(A, BbarC, Z1, Z2) &\leftarrow \\ &\quad dr(A, B, \neg U2), \quad dr(B, A, V1, \neg), \quad bar(B, Bbar), \\ &\quad dr(B, C, \neg X2), \quad dr(C, B, Y1, \neg), \quad union(Bbar, C, BbarC), \\ &\quad intersection(A, C, \{\}), \quad V1 > 0.0, \quad Y1 > 0.0, \\ &\quad Z1 = 0.0, \quad Z2 = \min(1.0, (1.0 - Y1) \cdot \frac{U2 \cdot X2}{V1 \cdot Y1}). \end{aligned}$$

²Actually, some of the stated evaluable functions have to be defined as external predicates, using "C".

$$\begin{aligned}
dr(A, B, Z1, Z2) \leftarrow & \\
& dr(A, B, X1, X2), dr(B, A, 0.0, 0.0), Z1=0.0, Z2=0.0. \\
dr(AB, C, Z1, Z2) \leftarrow & \\
& dr(A, B, Z1, Z2), \text{union}(A, B, AB), \text{ir}(A, B, C). \\
dr(B, C, Z1, Z2) \leftarrow & \\
& dr(AB, C, Z1, Z2), \text{union}(A, B, AB), \text{ir}(A, B, C).
\end{aligned}$$

User queries:

Recall that our set of relevant events is C_1, \dots, C_k . Let $A_i (1 \leq i \leq n)$ and $B_j (1 \leq j \leq m)$ be basic events; i. e. any of these C_i or $\overline{C}_i (1 \leq i \leq k)$. Typically one then would like to ask questions like: " $P(B_1 \dots B_m | A_1 \dots A_n) = ?$ " ($m, n \geq 1$). Since our calculus deals with intervals, of course we would like to know the sharpest possible interval to the previous question. This means, a *standard user query* actually has the format:

"Give me the sharpest interval $[Z1, Z2]$ for $A_1 \dots A_n \xrightarrow{Z1, Z2} B_1 \dots B_m$ that can be deduced."

Some of these queries can be answered efficiently right away:

- (1) If $\{A_1, \dots, A_n\} \supseteq \{B_1, \dots, B_m\}$, then $[Z1, Z2] = [1, 1]$.
- (2) If $A_i = \overline{B}_j$ for some i, j , then $[Z1, Z2] = [0, 0]$.

In general, however, this query must be answered by generating deduced rules from the axioms \mathcal{R} . Since the given DATALOG-program for dr has a fixpoint semantics, we can realize this inference mechanism by a fixpoint iteration for dr . The query to be posed to dr then becomes an aggregate query:

$$\begin{aligned}
hq(A, B, \langle Z1 \rangle, \langle Z2 \rangle) \leftarrow & dr(A, B, Z1, Z2). \\
q(A, B, X, Y) \leftarrow & hq(A, B, \text{Set}Z1, \text{Set}Z2), \\
& \text{aggregate}(\text{max}, \text{Set}Z1, X), \\
& \text{aggregate}(\text{min}, \text{Set}Z2, Y).
\end{aligned}$$

Besides such standard queries with the binding pattern (b = bound, f = free) of $dr(b, b, f, f)$ one can imagine more fancy queries, directly implementable with the full power of DATALOG systems. For example, queries like "tell me the names of those sets of events, which are implied by A with more than 75%". This corresponds to the binding pattern $dr(b, f, b, b)$.

Example (Burglary) The set of axioms given earlier translates into

$$\begin{aligned}
br = & \{(\{'alarm'\}, \{'earthquake'\}, 0.06, 0.06), \\
& (\{'earthquake'\}, \{'alarm'\}, 0.7, 0.7), \\
& (\{'alarm'\}, \{'burglary'\}, 0.95, 0.95), \\
& (\{'earthquake'\}, \{'burglary'\}, 0.2, 0.2)\}
\end{aligned}$$

The user query " $alarm_and_earthquake \xrightarrow{?} burglary$ " is transformed into the query:

$$? q(\{'alarm'\}, \{'earthquake'\}, \{'burglary'\}, X, Y).$$

As always with recursive DATALOG-programs, massive use of sophisticated optimization techniques must be made to achieve safety of execution ([SaVa 89]) and sufficient efficiency for practical use. A complete discussion of this matter would take us beyond the scope of this paper, so we want to point out only some aspects here.

The Optimization-Challenge

- (1) The LDL-program in the form given above won't terminate because of unsafety of execution. Safety and dramatic efficiency gains can be accomplished by a variation of the optimization technique known as "push selection by certainty" ([SSGKAB 89]): If $dr(U, V, X1, X2)$ and $dr(U, V, Y1, Y2)$ are deduced during the fixpoint iteration, then these two tuples can be replaced by one, namely

$$dr(U, V, \max(X1, Y1), \min(X2, Y2)).$$

That is the aggregation should be applied as soon as possible *and* it should discard its two inputs. Actually, the optimizer has to push the aggregation on q into dr , violating stratification, but not local stratification ([ShNa 87]).³

- (2) The given DATALOG-program is nonlinearly recursive. Whether some special transitive closure operators (like closed semiring or generalized transitive closure), can be applied, needs to be examined. Since some of the functions occurring are not associative, this might be not the case. (This might also imply that no equivalent linear recursive solution can be found in general).
- (3) The use of heuristics for pruning the search space, like ([SKGB 89]) must be explored.
- (4) The cost of set unification is of concern, too. While LDL employs a costly general-purpose algorithm, special solutions using bit vector implementations are promising here.

This optimization challenge will be investigated more closely in a forthcoming paper.

Consistency Support

If the numbers in br are statistical measured conditional probabilities or come from scientific models for interaction, no inconsistencies can arise. But often numbers are subjective probabilities; e. g. , expert opinions or speculative estimates of the knowledge engineer. So any automatic support for maintaining the consistency of the

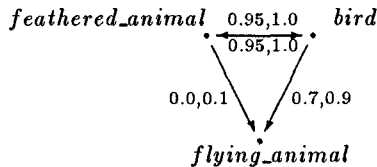
³This is once more a very practical case indicating that DATALOG-systems should be freed from rigorous stratification checks.

base rules is of big practical importance. In general, not every potential inconsistency of the base rules can be detected by a compile time analysis easily. But the implementation of our calculus gives a detection mechanism for inconsistencies at run-time almost for free.

$$\text{contradiction}(P,Q) \leftarrow \text{dr}(P,Q,X1,X2), \text{dr}(P,Q,Y1,Y2), \\ X2 < Y1 \text{ or } Y2 < X1.$$

Since our calculus is sound, this contradiction can only happen because of an inconsistency in the base rules. Now suppose that during the fixpoint iteration two tuples in dr with $(A, B, U1, U2)$, $(A, B, V1, V2)$ and $U2 < V1$ are deduced. Then application of sharpening during the next iteration round produces a tuple $(A, B, V1, U2)$ with $V1 > U2$. So a simple *trigger* during fixpoint iteration, watching for tuples $(P, Q, X1, X2)$ with $X1 > X2$, is enough for detection.

Example (Flying_animal) (cf.[Pea 88]) We know, that "birds can fly" with the exception of penguins, ostriches, etc. Therefore, we estimate that between 70% and 90% of all birds can fly. Typically, almost all "birds are feathered animals", and almost all "feathered animals are birds", at least more than 95% of each. And finally, we assume that "feathered animals do not fly", for instance, we estimate that less than 10% of all feathered animals can fly. Our set of axioms \mathcal{R} is depicted graphically below:



In the above statements, there is a contradiction, which can easily be detected with our rule chaining mechanism:

$$\{ \text{feathered_animal} \xrightarrow{0.95, 1.0} \text{bird}, \text{bird} \xrightarrow{0.7, 0.9} \text{flying_animal} \}$$

$$\frac{}{\text{RC}} \text{feathered_animal} \xrightarrow{0.65, 0.99} \text{flying_animal}.$$

In contrast to the assumption in the knowledge base that "feathered animals do not fly", we have deduced that many "feathered animals can fly", at least more than 65% of them.

Summarizing, we can detect consistencies even in the case of cycles as above. A powerful consistency maintenance component should offer a strategy to resolve inconsistencies by picking suitable victims for removal from \mathcal{R} . Last but not least, we can give support for keeping the knowledge base \mathcal{R} minimal by detecting outdated base rules.

$$\text{outdated}(P,Q,X1,X2) \leftarrow \text{br}(P,Q,X1,X2), \text{dr}(P,Q,Y1,Y2), \\ X1 < Y1 \text{ or } X2 > Y2.$$

This situation indicates that the information in a base rule is subsumed by deducible knowledge from other sources and hence might be a candidate for removal from \mathcal{R} , but only after all inconsistencies have been resolved before. Otherwise it might happen that $\text{dr}(P, Q, Y1, Y2)$ has been derived by means of some other base rule which has to be retracted.

System Architecture

We conclude this discussion by showing the gross architecture for implementing sound and non-monotonic uncertainty reasoning on a deductive database system (which may be part of an object-oriented DBMS as well).

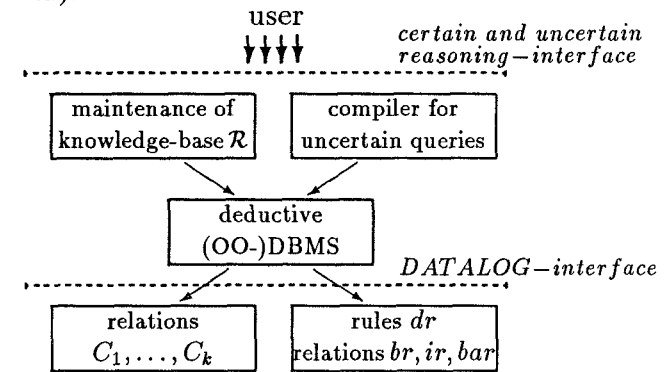


Figure 5.1: The DUCK-system architecture (Deduction with UnCertain Knowledge)

Depicted relations (classes) C_1, \dots, C_k are supposed to be those appearing in \mathcal{R} . If, as normally, these relations have their own tuples (objects), the question naturally comes up how to combine standard deductive queries on the instances of C_1, \dots, C_k with the uncertain queries on the C_1, \dots, C_k themselves. As it turns out, a combination of both is achievable by another sound inference rule, the *uncertain modus ponens*, which will be described in detail elsewhere. This, among other things, provides us with a method for dealing with exceptions.

5. Summary and Outlook

We have presented a novel calculus for non-monotonic uncertainty reasoning. This approach, relying on conditional probabilities, exhibits several salient features:

- (1) It provides a sound inference mechanism for rule-based uncertainty, while maintaining the virtues of modularity and incrementality of rule-based application development.
- (2) Sound inferences under uncertainty are feasible even in the presence of incomplete information.

Thus it can be employed for expert systems in diagnosis (like in medicine) or pattern recognition (like analysis of satellite data).

- (3) The consistency of the knowledge base can be monitored and acquisition of new base rules be optimized.
- (4) The calculus lives naturally in a database environment. The direct mapping onto a DATALOG-interface integrates uncertainty reasoning into databases and commercial data processing.

Having solved several problems with our calculus so far, naturally a range for further investigations is wide open, predominantly in the area of efficiency. Certainly a comfortable user language for manipulating the base rules and for posing queries must be designed. The language should be a natural extension of existing popular database languages.

References

- [AnHo 90] K.A. Anderson and J.N. Hooker: Bayesian Logic. Working Paper 1-90-91, Carnegie Mellon University, Pittsburgh, 1990.
- [BMP 90] D. Barbara, H. Garcia-Molina and D. Porter: The Management of Probabilistic Data. *Proc. EDBT*, Venice, 1990.
- [BoDe 86] P.P. Bonissone and K.S. Decker: Selecting Uncertainty Calculi and Granularity: An Experiment in Trading-off Precision and Complexity. In *Uncertainty in AI*, L.N. Kanal and J.F. Lemmer (ed.), Elsevier, North Holland, 1986, pp. 217-247.
- [CGT 89] S. Ceri, G. Gottlob and L. Tanca: Logic Programming and Databases. *Springer*, New York, Berlin, Heidelberg, 1989.
- [HeHo 87] D.E. Heckermann, E.J. Horvitz: On the Expressiveness of Rule-based Systems for Reasoning with Uncertainty. *Proc. 6th Nat. Conf. on AI*, Washington, 1987, pp. 121-126.
- [Hun 75] E. Hunt: Artificial Intelligence. *Academic Press*, New York, 1975.
- [KiGu 90] W. Kießling and U. Güntzer: Deduktive Datenbanken auf dem Weg zur Praxis. *Informatik Forschung u. Entwicklung*, Vol. 5, No. 4, *Springer*, 1990, pp. 177-187.
- [Lag 90] Lagunita Beach Report: Database Systems: Achievements and Opportunities. Report of the *NFS Invitational Workshop on Future Directions in DBMS Research*, Palo Alto, 1990.
- [LaSp 88] S.L. Lauritzen and D.J. Spiegelhalter: Local Computation with Probabilities on Graphical Structures and their Application to Expert Systems. *Journ. Royal Statistical Society B*, 50(2), 1988, pp. 157-224.
- [NaTs 89] S. Naqvi and S. Tsur: A Logical Language for Data and Knowledge Bases. *Computer Science Press*, New York, 1989.
- [Pea 88] J. Pearl: Probabilistic Reasoning in Intelligent Systems. *Morgan Kaufmann*, San Mateo, 1988.
- [Qui 83] J.R. Quinlan: INFERNO: A cautious approach to uncertain inference. *The Computer Journal* 26. 1983, pp. 255-269.
- [Ric 86] E. Rich: Artificial Intelligence. *McGraw-Hill*, 5th ed., 1986.
- [SaVa 89] Y. Sagiv and M.Y. Vardi: Safety of datalog queries over infinite databases. *Proc. ACM SIGACT-SIGART-SIGMOD Symp. on Principles of Database Systems*, Philadelphia, 1989, pp. 160-171.
- [Sha 76] G. Shafer: A mathematical theory of evidence. *Princeton University Press*, Princeton, 1976.
- [ShNa 87] O. Shmueli and S. Naqvi: Set Grouping and Layering in Horn Clause Programs. *Proc. 4th. Int. Conf. on Logic Programming*, 1987.
- [SKGB 89] H. Schmidt, W. Kießling, U. Güntzer and R. Bayer: DBA*: Solving Combinatorial Problems with Deductive Databases. *Proc. GI/SI-Conf. on "Datenbanksysteme in Büro, Technik und Wissenschaft"*, Zürich, 1989.
- [SSGKAB 89] H. Schmidt, N. Steger, U. Güntzer, W. Kießling, R. Azone and R. Bayer: Combining Deduction by Certainty with the Power of Magic. *Proc. 1st Int. Conf. on Deductive and Object-Oriented Databases*, Kyoto, 1989, pp. 205-224.
- [Tar 66] A. Tarski: Einführung in die mathematische Logik. *Vandenhoeck & Ruprecht*, Göttingen, 1966.
- [Ull 89] J. Ullman: Principles of Database and Knowledge-Base Systems. Vols 1 and 2, *Computer Science Press*, New York, 1989.
- [Zad 86] L.A. Zadeh: A Simple View of the Dempster-Shafer Theory of Evidence and its Implication for the Rule of Combination. *AI Magazine* 7, 2, 1986, pp. 85-90.
- [Zan 88] C. Zaniolo: Design and implementation of a logic based language for data intensive applications. *Proc. 5th Int. Symposium on Logic Programming*, 1988, pp. 1666-1687.