A FAST GENERALIZED MODUS PONENS

Roger MARTIN-CLOUAIRE*

Langages et Systèmes Informatiques
ENSEEIHT
2 rue Charles Camichel
31071 TOULOUSE CEDEX
FRANCE

ABSTRACT

The generalized modus ponens is a fuzzy logic pattern of reasoning that permits to deduce an imprecise conclusion from imprecise premises. Unfortunately its computation can be unacceptably slow if one simply relies on a direct implementation of the definition formula. This short paper presents an algorithm for performing an efficient deduction by means of a special case of the generalized modus ponens (i.e. based on the Brouwer-Gödel implication). The exhibited method is of particular interest for application in expert system technology.

I - INTRODUCTION

In the course of the extensive research that is done on knowledge processing in artificial intelligence, a specific problem concerns the ways of representing and treating imperfect informations. Among the most pertinent works that have arisen in response let us mention those :

- in the field of non-monotonic logics [1] that deal with what might be called 'inference from incomplete or insufficient evidence',
- developed in the frameworks of the MYCIN [9] and PROSPECTOR [3] expert systems for propagating uncertainty through reasoning chains,

* The author is partially supported by a grant of the Société Nationale Elf Aquitaine (Production) in the framework of a Ph.D.Thesis.

- pertaining to fuzzy logic [11], [12] which takes its ground in the possibility theory [10], [2] and encompasses treatment of both imprecision and uncertainty (that can be of several types e.g. probabilistic or possibilistic).

In fuzzy logic (or equivalently approximate reasoning [8]) a feature of particular importance to expert systems [5] [6] is the ability to draw an imprecise conclusion from a set of imprecise premises and a set of imprecise facts matched against them. The main pattern of reasoning to perform such a deduction is known as the generalized modus ponens that was introduced by Zadeh [11]. Actually several versions of the generalized modus ponens exist [7], each being defined with respect to a multivalued logic implication and a t-norm. As far as computation is concerned, the application of the generalized modus ponens is basically equivalent to the solution of a non-linear program [12]. Unfortunately, a naïve implementation relying directly on the formula of the definition may render it time-inefficient and even obsolete for inferences involving more than two premises. Indeed, generally in order to achieve a satisfactory level of validity for an inferred conclusion such an implementation would require to go through a thin discretisation of the cartesian product of possibility distributions involved in the premises and would therefore lead to numerous expensive iterations.

This short paper presents an algorithm for performing a fast generalized modus ponens defined with respect to the Brouwer-Gödel multivalued logic implication and the t-norm 'min'. The efficiency exhibited by this method stems from the fact that it does not require any discretisation of the possibility distributions involved in the premises. In addition this method produces an optimally valid conclusion since it is not subject to the thinness of a discretisation.

The next section provides some background on the generalized modus ponens and introduces the notation and hypotheses. The algorithm is presented in section III.

II - Generalized modus ponens

Stated in the form of a syllogism the generalized modus ponens looks like :

$$\frac{\begin{array}{l}\text{If X is A then Y is C}\\ \text{X is A'}\end{array}}{\text{Y is C'}} \qquad (1)$$

Basically this means that from a rule which associates a variable X specified by an elastic (or fuzzy) constraint A with a variable Y specified by an elastic constraint C and a fact "X is A'" expressing the value (eventually imprecise) of X one can infer the fact "Y is C'" where C' is the deduced elastic constraint on Y. X and Y are supposed to take their values in U and V respectively. The constraints A, C, A' and C' are respectively expressed by the possibility distributions μ_A, μ_C, $\mu_{A'}$ and $\mu_{C'}$ that represent the possible values which X and Y may take in the rule and in the facts. The possibility distribution $\mu_{C'}$ is computed from both $\mu_{A'}$ and a conditional possibility distributions $\mu_{\rightarrow}$ consistent with a multivalued logic implication derived from μ_A and μ_C. In this paper $\mu_{C'}$ is assumed to be given by :

$$\forall v \in V,\ \mu_{C'}(v) = \sup_{u \in U} \min(\mu_{A'}(u),\ \mu_{\rightarrow}(u, v)) \qquad (2)$$

$$\text{where } \mu_{\rightarrow}(u, v) = \begin{cases} 1 & \text{if } \mu_A(u) \leq \mu_C(v) \\ \mu_C(v) & \text{otherwise} \end{cases}$$

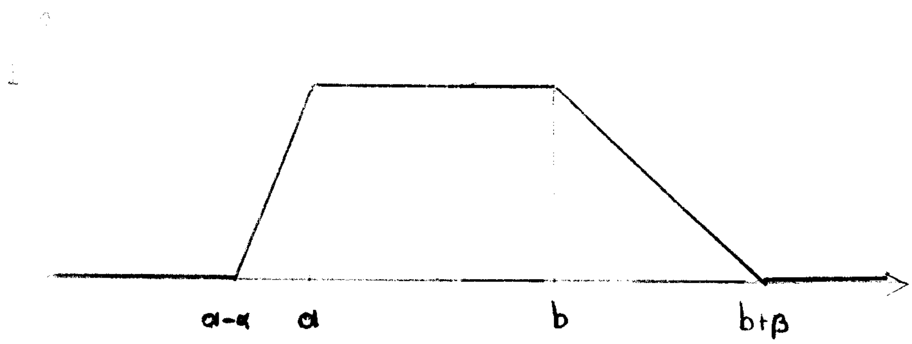
$\mu_{\rightarrow}$ is consistent with the Brouwer-Gödel implication. See [7][8] for other possibilities than (2) (they are obtained by making use of other multivalued logic implications and other t-norms than 'min'). One of the most interesting feature of (2) is that when A' is the same constraint than A (or more generally when $\forall u \in U$, $\mu_{A'}(u) \leq \mu_A(u)$) then the deduced constraint C' is exactly C (i.e. $\forall v \in V$, $\mu_{C'}(v) = \mu_C(v)$).

The rule involved in the modus ponens may be multidimentional. In effect, the variable X may actually stand for a finite collection of n variables X_i (i.e. $X = X_1, \ldots, X_n$)) that are assumed here to be non-interactive [10]. Each X_i can be regarded as the implicit or explicit constrained variable of the ith premise "X_i is A_i" involved in the rule. For i = 1, ..., n the constraint A_i is expressed by the possibility distribution μ_{A_i} on U_i. Under the assumption of non-interactivity the possibility distribution μ_A and $\mu_{A'}$ are computed as follows :

$$\mu_A(u) = \min_{1 \leq i \leq n} \mu_{A_i}(u_i) \qquad (3)$$

$$\mu_{A'}(u) = \min_{1 \le i \le n} \mu_{A'_i}(u_i) \qquad (4) \text{ in which } u = (u_1, \ldots, u_n)$$

In the sequel the possibility distributions μ_{A_i}, $i = 1, \ldots, n$ and μ_C are assumed to be unimodal, normalized and represented by four place parameterized functions symbolically written $(a_i\ b_i\ \alpha_i\ \beta_i)$, $i = 1, \ldots, n$ and $(c\ e\ \gamma\ \varepsilon)$ respectively. The meaning of the four parameters is shown in picture 1 in the simple (but not restricted to) case of a trapezoidal distribution $(a\ b\ \alpha\ \beta)$.



Picture 1

For brevity and clarity of the exposure, any distribution $\mu_{A'_i}$, $i = 1, \ldots, n$ is assumed to be continuous and represented by $(a'_i\ b'_i\ \alpha'_i\ \beta'_i)$. However the extention to discrete normalized distributions (necessary in case of chaining) does not present any theoretical problem.

III - Algorithm

The algorithm starts with an evaluation of the global level of indetermination that appears in the conclusion as soon as a significant part of $\mu_{A'}$ falls outside of μ_A ($\mu_{A'}$ and μ_A being considered as the fuzzy sets they are membership of) i.e. $\exists\, u \in U$ such that $\mu_{A'}(u) > \mu_A(u)$. It is easy to see on formula (2) that this global level of indetermination is given by

$$\zeta = \sup_{u \in \{u \in U_1 \times \ldots \times U_n \,/\, \mu_A(u) = 0\}} \mu_{A'}(u) \qquad (5)$$

Then, three different treatments have to be considered depending on the result obtained for ζ. The second and third treatments encompass both two different situations. For clarity the more complex situations are illustrated by pictures corresponding to single-premise rules.

(0)- Computation of the global level of indetermination.

$$\zeta = \max_{1 \leq i \leq n} \zeta_i \quad \text{where } \forall\, i \in \{1, \dots, n\} \;\; \zeta_i \text{ is obtained by}$$

$$\zeta_i = \begin{cases} \max(\mu_{A'_i}(a_i - \alpha_i),\ \mu_{A'_i}(b_i + \beta_i)) & \text{if } a'_i < b_i + \beta_i \text{ and } b'_i > a_i - \alpha_i \\ 1 & \text{otherwise.} \end{cases}$$

(1)- If $\zeta = 1$ then $\mu_{C'}(v) = 1 \quad \forall\, v \in V$

This is the case of complete indetermination.

(2)- If $\zeta = 0$ let us define for any $i \in \{1, \dots, n\}$ the set I_i by :

$I_i = \{u_i \in \,]a_i - \alpha_i, a_i[\, \cup \,]b_i, b_i + \beta_i[$ such that $\mu_{A_i}(u_i) = \mu_{A'_i}(u_i)$ (this intersection has to be considered exclusively on the parts where μ_{A_i} and $\mu_{A'_i}$ are both increasing of both decreasing without being merged*)}.

a) If $\forall\, i \in \{1, \dots, n\}$ $I_i = \emptyset$ then $\mu_{C'}(v) = \mu_C(v) \quad \forall\, v \in V$.

b) If $\exists\, i \in \{1, \dots, n\}$ such that $I_i \neq \emptyset$ let us define Ω and one** by :

$$\Omega = \min_{\substack{1 \leq i \leq n \\ I_i \neq \emptyset}} \; \min_{u_i \in I_i} \; \mu_{A_i}(u_i)$$

$$\text{one} = \min_{1 \leq i \leq n} \min(\mu_{A_i}(a'_i),\ \mu_{A_i}(b'_i)) \,.$$

Then :

- $\mu_{C'}(v) = \mu_C(v) \;\; \forall\, v \leq \inf \mu_C^{-1}(\Omega)$ and $\forall\, v \geq \sup \mu_C^{-1}(\Omega)$
  where $\mu_C^{-1}(\Omega) = \{v \in V \,/\, \mu_C(v) = \Omega\}$
- $\mu_{C'}(v) = 1 \;\; \forall\, v \in \,]\inf \mu_C^{-1}(\text{one}), \sup \mu_C^{-1}(\text{one})[$
- $\mu_{C'}(v) = \max\limits_{1 \leq i \leq n} \;\; \sup\limits_{u_i \in \mu_{A_i}^{-1}(\mu_C(v))} \;\; \mu_{A'_i}(u_i)$

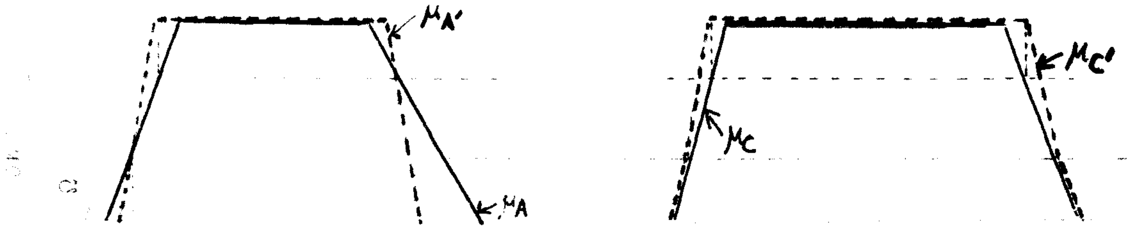
  $\forall\, v \in \,]\inf \mu_C^{-1}(\Omega), \inf \mu_C^{-1}(\text{one})] \cup [\sup \mu_C^{-1}(\text{one}), \sup \mu_C^{-1}(\Omega)[$.

Picture 2 illustrates, in an approximate manner, the situation 2b with $n = 1$.

---

* This means, in particular, that if $\mu_{A'_i} = \mu_{A_i}$ then $I_i = \emptyset$.

** More generally, 'one' is defined by : $\text{one} = \inf\limits_{u \in \{u \,/\, \mu_{A'}(u) = 1\}} \mu_A(u)$



Picture 2.

③ - If $\zeta \in]0,1[$ then two situations have to be considered :

a) If $\forall\ i \in \{1, \ldots, n\}$ $a_i \leq a'_i \leq b'_i \leq b_i$ then let us define :

- for any $i \in \{1, \ldots, n\}$ the set I_i as in treatment 2 (i.e. for $\zeta = 0$)
- Ω by $\Omega = \max\limits_{\substack{1 \leq i \leq n \\ I_i \neq \emptyset}} \max\limits_{u_i \in I_i} \mu_{A_i}(u_i)$

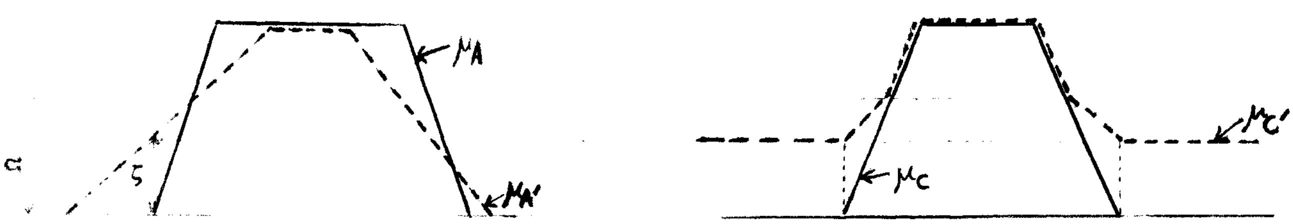
Then . $\mu_{C'}(v) = \zeta\ \forall\ v \leq c - \gamma$ and $\forall\ v \geq e + \varepsilon$

. $\mu_{C'}(v) = \mu_C(v)\ \forall\ v \in]\inf \mu_C^{-1}(\Omega), \sup \mu_C^{-1}(\Omega)[$

. $\mu_{C'}(v) = \max\limits_{1 \leq i \leq n} \sup\limits_{u_i \in \mu_{A_i}^{-1}(\mu_C(v))} \mu_{A'_i}(u_i)$

$\forall\ v \in]c - \gamma, \inf \mu_C^{-1}(\Omega)] \cup [\sup \mu_C^{-1}(\Omega), e + \varepsilon[$.

The situation 3a is roughly sketched in picture 3.



Picture 3.

b) If $\exists\ i \in \{1, \ldots, n\}$ such that $a'_i < a_i$ or $b'_i > b_i$ then let us define 'one' by :

$$\text{one} = \min_{1 \leq i \leq n} \min(\mu_{A_i}(a'_i), \mu_{A_i}(b'_i)).$$

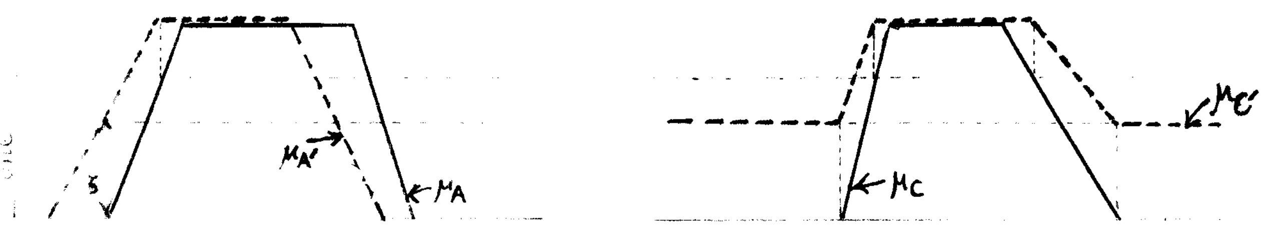
Then : . $\mu_{C'}(v) = \zeta \;\forall\; v \leq c - \gamma$ and $\forall\; v \geq e + \varepsilon$

. $\mu_{C'}(v) = 1 \quad \forall \quad v \in \,]\inf \mu_C^{-1}(\text{one}),\ \sup \mu_C^{-1}(\text{one})[$

. $\mu_{C'}(v) = \max\limits_{1 \leq i \leq n} \quad \sup\limits_{u_i \in \mu_{A_i}^{-1}(\mu_C(v))} \quad \mu_{A'_i}(u_i)$

$\forall\; v \in \,]\, c - \gamma,\ \inf \mu_C^{-1}(\text{one})] \cup [\sup \mu_C^{-1}(\text{one}),\ e + \varepsilon[.$

The situation 3b is roughly sketched in picture 4.



Picture 4.

IV - Concluding remarks

This short paper has described an efficient algorithm for performing a generalized modus ponens based on Brouwer-Gödel implication. This method drastically improves the performance of the naïve implementation and therefore permits to retain the generalized modus ponens its potentiality for serving as the basic pattern of reasoning with imprecise premises and conclusions. It has actually been developed for use in the inference engine ELFIN [4] that is designed for a class of petroleum geology expert systems. Among directions the current work can be extended in are the similar investigations based on other t-norms and other multivalued logic implications.

REFERENCES

[1] - Bobrow, D.G., (Ed.). Special issue on non-monotonic logic. Artificial Intelligence, 13, (1 and 2), (1980).

[2] - Dubois, D., Prade, H. Fuzzy sets and systems : theory and applications. Academic Press, New-York (1980).

[3] - Duda, R., Hart, P., Nilson, N. Subjective Bayesian method for rule-based inference systems. Tech. Note n° 124, SRI International, Menlo Park, CA. (1976).

[4] - Martin-Clouaire, R. Le moteur d'inférence ELFIN : conception et premières réalisations. Internal Rep., LSI (ENSEEIHT), Toulouse (1984).

[5] - Martin-Clouaire, R., Prade, H. On the problems of representation and propagation of uncertainty in expert systems. 2nd NAFIP Workshop, Schenectady, N.Y., June 29-30, July 1. (1983).

[6] - Martin-Clouaire, R., Prade, H. Managing uncertainty and imprecision in petroleum geology. Int. Colloquium "Informatique dans les Sciences de la Terre pour la Caractérisation des Ressources Naturelles". Nancy, France, April 9-13 (1984).

[7] - Prade, H. Modèle mathématiques de l'imprécis et de l'incertain en vue d'applications au raisonnement naturel. Thèse d'Etat, Univ. Paul Sabatier, Toulouse (1982).

[8] - Prade, H. A synthetic view of approximate reasoning techniques. Proc. 8th Int. Joint Conference on Artificial Intelligence, Vol. 1, 130-136, (1983).

[9] - Shortliffe, E.H., Buchanan, B.G. A model of inexact reasoning in medicine. Mathematical Bioscence, 23, 351-379, (1975).

[10] - Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3-28, (1978).

[11] - Zadeh, L.A. A theory of approximate reasoning. machine Intelligence, Vol. 9, (Hayes J., Michie D., Mikulich L. Eds.), Elsevier, 149-194, (1979).

[12] - Zadeh, L.A. The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems, 11, 199-227, (1984).