On the continuous extension of a four valued logic for preference modelling

Patrice Perny
LIP6, Université Pierre et Marie Curie
Case 169, 4 place Jussieu
75252 Paris Cedex 05, France
Patrice.Perny@lip6.fr

Alexis Tsoukiàs
LAMSADE, Université Paris Dauphine
Place du Maréchal de Lattre de Tassigny
75775 Paris Cedex 16, France
tsoukias@lamsade.dauphine.fr

Abstract

This paper presents a fuzzy logic useful for preference modelling from possibly incomplete or conflicting sources of information. This logic is based on the standard predicate language with a semantic admitting an infinite number of truth values defined as convex combinations of reference values $\{true, false, contradictory, unknown\}$. This logic is introduced as a fuzzy extension of Belnap's four valued logic and its specific interest for multidimensional preference modelling is discussed.

Introduction

You do not know who Anaxagoras is. But someone told you that he is intelligent. Therefore if they ask you if Anaxagoras is intelligent you may answer "yes" (being confident to the one who told you that ...). On the contrary someone else told you that Anaxagoras is not intelligent. Your answer now should be "no" (same confidence of course). But suppose you have been told both that Anaxagoras is intelligent and that he is not intelligent (two equally reliable information sources). Then your answer will be "yes" and "no", "perhaps" or something like this. Finally, if no one tolds you anything about Anaxagoras, your answer should be "I do not know".

This small example represents a well known problem used by Belnap in order to introduce his four-valued paraconsistent logic (see [2]). In this logic, the four possible interpretations of an atomic formula are "true", "false", "contradictory", "unknown", corresponding to the four situations listed in the previous example. However, there are some problems that may be present and which we want to discuss here. Coming back to the initial example, the following questions can be raised:

1. How confident are the two potentially conflicting sources of information? When it is possible to establish a degree of "confidence" in the two sources of information, how can we "measure" the credibility of the sentence "Anaxagoras is (not) intelligent"?
2. What does the statement"Anaxagoras is (not) intelligent" really mean? Intelligence is not a clear and well established attribute. It is easier to state that someone is more or less intelligent or that he is intelligent to a certain degree. Under this perspective "Anaxagoras is (not) intelligent" is not a crisp sentence since the predicate $intelligent(x)$ is ill-defined. When it is possible to establish a degree of "intensity" or "importance" of the positive and negative arguments related to this statement, how can we "measure" the intensity to which "Anaxagoras is (not) intelligent"?

In the first case the problem of attaching an uncertainty distribution to the sentence "Anaxagoras is (not) intelligent" is addressed. In the second case a continuous modulation of the predicate $intelligent(x)$ has to be established. In both the cases it seems necessary to introduce a continuous extension of the four crisp cases introduced by Belnap. In this paper, we mainly focus on the second case, trying to define the intensity of truth, contradiction, unknown and falsity in each logical formula.

This type of question is of particular interest for preference modelling in real decision problems where partial information and conflicting opinions are always present. This statement has led many authors to extend the standard predicate logic used for preference modelling.

On the one hand, the usefulness of four valued logics for preference modelling in the context of multicriteria decision making is well illustrated in [13] and [14]. In this context, the natural state of the overall preference aPb may be "unknown" (lack of information) or "contradictory" (conflicting criteria) as well as "true" or "false".

On the other hand, the interest of considering a continuous extension of the standard binary predicate logic for preference modelling has been illustrated by many authors (see *e.g.* [7], [8], [12], [6], [3], [11]).

Fuzzy preference relations allow either the representation of uncertainty about the preferences between alternatives whose consequences are ill-known, or the representation of intensity of preferences between alternatives whose consequence are perfectly known. In the first case, fuzziness of preference reflects the uncertainty in a statement that would be crisp with a complete information. In the second case, fuzziness reflects the graduality of preference and allow continuous transitions from non-preference to preference situations.

In this paper, we investigate the possibilities provided by the merging of these two parallel extensions. The aim of the paper is to define a fuzzy logic for preference modelling where truth values are fuzzy subsets of $\{true, false, unknown, contradictory\}$. In the first part of this paper, we recall the semantic of Belnap's four-valued logic (§ 1), introduce the fuzzy extension of this semantic (§ 2) for atomic formulae, and discuss the extension to more complex formulae (§ 3).

Then we provide illustrations of the use of such extended logics for preference modelling in decision problems with multiple criteria or multiple judges (§ 4).

1 Semantic of the four-valued logic

The basic idea in the paraconsistent approach is to distinguish between negation and complement in order to avoid inconsistency of the sentence $\alpha \wedge \neg\alpha$. From a set theory point of view that means the extension of a predicate (in a specific domain) and the extension of its negation do not form a partition (of the specific domain), but have a non empty intersection and their union is a proper subset of the domain.

More formally, suppose that $\mathcal{L}$ is a first order language. A similarity type ρ is a finite set of predicate constants S where each S has an arity n_s. Relative to a given similarity type ρ, $S(x_1, ..., x_m)$ is an atomic formula iff $x_1, ..., x_m$ are individual variables, $S \in \rho$ and $n_s = m$. Formula will be denoted by letters $\alpha, \beta, \gamma, \delta, ...$. A structure, for similarity type ρ, consists of a non empty domain X of tuples, and for each predicate symbol $S \in \rho$ an ordered pair $\langle S^+, S^- \rangle$ of sets of n_s-tuples from X. Intuitively S^+ represents the models of the predicate S and S^- the models of $\neg S$, the negation of S. Within X, the complement of a subset S of X is denoted $\sim S = X \setminus S$. Hence, the ordered pairs $\langle (\neg S)^+, (\neg S)^- \rangle$ and $\langle (\sim S)^+, (\sim S)^- \rangle$ are defined by:

$$(\neg S)^+ = S^-, \qquad (\neg S)^- = S^+ \qquad (1)$$
$$(\sim S)^+ = \sim (S^+), \qquad (\sim S)^- = \sim (S^-) \qquad (2)$$

Following the paraconsistent approach we do not impose that S^+ and S^- form a partition of X. In the general case indeed, $S^- \neq \sim (S^+)$. Given a m-tuple $x = (x_1, ..., x_m)$ a substitution σ is a mapping from the set of variables to the set X. We denote the resulting instance as $\sigma(x) = (\sigma(x_1), ..., \sigma(x_m))$. We are able now to introduce two membership functions:

$$\mu_{S^+}: \begin{array}{ll} X & \to [0,1] \\ a & \mapsto \mu_{S^+}(a) \end{array} \qquad \mu_{S^-}: \begin{array}{ll} X & \to [0,1] \\ a & \mapsto \mu_{S^-}(a) \end{array}$$

where $a = \sigma(x)$ is an instance of the x tuple. These functions evaluate the degree of membership of the instance a to the sets S^+ and S^- respectively. Intuitively, quantities $\mu_{S^+}(a)$ and $\mu_{S^-}(a)$ reflect the degree to which we believe in $S(a)$ and *not* $S(a)$ respectively. For this reason, we use also the following notation:

$$\mu_{S^+}(a) = B(S(a)) \qquad \mu_{S^-}(a) = B(\neg S(a))$$

where $\neg S(a)$ represents the negation of $S(a)$.

In Belnap's logic, S^+ and S^- are crisp subsets of X and $B(S(a))$ equals 0 or 1, as well as $B(\neg S(a))$. In this case the domain X is partitioned in four sets (with respect to the predicate S) which are:

$$S^t = S^+ \cap \sim (S^-) \qquad (3)$$
$$S^k = S^+ \cap S^- \qquad (4)$$
$$S^u = \sim (S^+) \cap \sim (S^-) \qquad (5)$$
$$S^f = \sim (S^+) \cap S^- \qquad (6)$$

Hence we get the following relations:

$$S^t \cup S^k = S^+ \qquad (7)$$
$$S^f \cup S^k = S^- \qquad (8)$$
$$S^t \cup S^u = \sim (S^-) \qquad (9)$$
$$S^f \cup S^u = \sim (S^+) \qquad (10)$$
$$S^t \cap S^k = \emptyset \qquad (11)$$
$$S^t \cap S^u = \emptyset \qquad (12)$$
$$S^t \cap S^f = \emptyset \qquad (13)$$
$$S^f \cap S^k = \emptyset \qquad (14)$$
$$S^f \cap S^u = \emptyset \qquad (15)$$
$$S^k \cap S^u = \emptyset \qquad (16)$$
$$S^t \cup S^k \cup S^u \cup S^f = X \qquad (17)$$

and from equations (1–6) we get:

$$S^t = (\neg S)^f = (\sim S)^f \qquad (18)$$

$$S^k = (\neg S)^k = (\sim S)^u \quad (19)$$
$$S^u = (\neg S)^u = (\sim S)^k \quad (20)$$
$$S^f = (\neg S)^t = (\sim S)^t \quad (21)$$

2 Continuous extension of the four-valued logic

In our fuzzy extension, we want to consider S^+ and S^- as fuzzy subsets of X. Hence we have to define the fuzzy subsets S^t, S^k, S^u, S^f of X so as to extend equations (3–17). This implies to extend intersection, union, and complementation to fuzzy subsets of X. A standard definition of these operations is provided by considering a De Morgan triple (N, T, V) where N is a strict negation on $[0,1]$, T a t-norm and V the conorm of T for negation N. In this case we get:

$$\mu_{\sim S}(a) = N(\mu_S(a)) \quad (22)$$
$$\mu_{S\cap S'}(a) = T(\mu_S(a), \mu_{S'}(a)) \quad (23)$$
$$\mu_{S\cup S'}(a) = V(\mu_S(a), \mu_{S'}(a)) \quad (24)$$

If we try to combine equations (22-24) with (3), (4) and (7) denoting $u = \mu_{S^t}(a)$, $v = \mu_{S^k}(a)$, $x = \mu_{S^+}(a)$, $y = \mu_S{}^-(a)$, we have:

$$\begin{aligned} u &= T(x, N(y)) \\ v &= T(x, y) \\ x &= V(u, v) \end{aligned}$$

Hence, we get the following functional equation:

$$\forall x, y \in [0,1], \quad x = V(T(x, N(y)), T(x, y))$$

Unfortunately, it is easy to prove that this equation has no solution (see [1]). Therefore, we must investigate partial solutions relaxing some constraints of the problem. Similar investigations have been considered in the context of fuzzy preference modelling. In this context, many studies consider the fuzzy partition of a cartesian product into three fuzzy binary relations (preference, indifference and incomparability) and provide solutions that could be imported in our context (see [9], [10], [12], [6], [3], [11]). As an example, we introduce now a useful result which directly follows from those obtained in [9], [10] and [6].

Denoting $\alpha = S(a)$, $t(\alpha) = \mu_{S^t}(a)$, $k(\alpha) = \mu_{S^k}(a)$ $u(\alpha) = \mu_{S^u}(a)$, $f(\alpha) = \mu_{S^f}(a)$, our aim is to define quantities $t(\alpha)$, $k(\alpha)$, $u(\alpha)$, $f(\alpha)$ as functions of quantities $B(a)$ and $B(\neg\alpha)$ and viceversa. Considering the negative result presented before, a more general hypothesis to investigate possibilities of deriving a continuous extension of (3–24) is the following:

$$t(\alpha) = T_1(B(\alpha), N(B(\neg\alpha))) \quad (25)$$
$$k(\alpha) = T_2(B(\alpha), B(\neg\alpha)) \quad (26)$$
$$u(\alpha) = T_3(N(B(\alpha)), N(B(\neg\alpha))) \quad (27)$$
$$f(\alpha) = T_4(N(B(\alpha)), B(\neg\alpha)) \quad (28)$$

where T_1, T_2, T_3, T_4 are continuous t-norms and N is a strict negation.

As a consequence of (18–21) we want:

$$t(\alpha) = f(\neg\alpha) = f(\sim\alpha) \quad (29)$$
$$k(\alpha) = k(\neg\alpha) = u(\sim\alpha) \quad (30)$$
$$u(\alpha) = u(\neg\alpha) = k(\sim\alpha) \quad (31)$$
$$f(\alpha) = t(\neg\alpha) = t(\sim\alpha) \quad (32)$$

Moreover, the translation of equations (7–10) using (22–24) gives:

$$B(\alpha) = V(t(\alpha), k(\alpha)) \quad (33)$$
$$B(\neg\alpha) = V(f(\alpha), k(\alpha)) \quad (34)$$
$$N(B(\alpha)) = V(t(\alpha), u(\alpha)) \quad (35)$$
$$N(B(\neg\alpha)) = V(f(\alpha), u(\alpha)) \quad (36)$$

Finally, because it makes no sense to consider situations admitting simultaneously "unknown" (lack of information) and "contradiction" (excess of information) for a same formula, we translate (16) to the strong condition:

$$\forall\alpha, \quad \min(u(\alpha), k(\alpha)) = 0 \quad (37)$$

Proposition 2.1 *$< T_1, T_2, T_3, T_4, T, V, N >$ is solution of equations (22–37) if and only the following conditions hold:*

$$\begin{aligned} N &= N_\phi \\ T &= T_2 = T_3 = LT_\phi \\ V &= LV_\phi \\ T_1 &= T_4 = \min \end{aligned}$$

where $(N_\phi, LT_\phi, LV_\phi)$ is a Lukasiewicz triple, i.e.*:*

$$\begin{aligned} LN_\phi(x) &= \phi^{-1}(1 - \phi(x)) \\ LT_\phi(x, y) &= \phi^{-1}(\max(\phi(x) + \phi(y) - 1, 0)) \\ LV_\phi(x, y) &= \phi^{-1}(\min(\phi(x) + \phi(y), 1)) \end{aligned}$$

where ϕ is an automorphism of $[0,1]$.

For the sake of simplicity, we will only consider here $\phi(x) = x$ for all x in [0, 1] but the following results can easily be extended for any automorphism ϕ of the unit interval. From (25–28) we get:

$$t(\alpha) = \min(B(\alpha), 1 - B(\neg\alpha)) \quad (38)$$
$$k(\alpha) = \max(B(\alpha) + B(\neg\alpha) - 1, 0) \quad (39)$$
$$u(\alpha) = \max(1 - B(\alpha) - B(\neg\alpha), 0) \quad (40)$$
$$f(\alpha) = \min(1 - B(\alpha), B(\neg\alpha)) \quad (41)$$

Remark 2.1 *If $B(\alpha)$, which may reflect uncertainty in α, is seen as the probability $P(\alpha)$, we get $B(\alpha) + B(\neg\alpha) = 1$ and therefore:*

$$\begin{aligned} t(\alpha) &= P(\alpha) \\ k(\alpha) &= 0 \\ u(\alpha) &= 0 \\ f(\alpha) &= 1 - P(\alpha) \end{aligned}$$

Remark 2.2 *If $B(\alpha)$, which may reflect uncertainty in α, is seen as a standard necessity, $N(\alpha)$ then $1 - B(\neg\alpha)$ is the possibility $\Pi(\alpha)$ and $B(\alpha) + B(\neg\alpha) \geq 1$; therefore we get:*

$$\begin{aligned} t(\alpha) &= N(\alpha) \\ k(\alpha) &= 0 \\ u(\alpha) &= \max(\Pi(\alpha) - N(\alpha), 0) \\ f(\alpha) &= N(\neg\alpha) \end{aligned}$$

From equations (38–41) we get the following properties that can be seen as the multivalued counterpart of (7–17):

$$\begin{aligned} t(\alpha) + k(\alpha) &= B(\alpha) & (42) \\ f(\alpha) + k(\alpha) &= B(\neg\alpha) & (43) \\ t(\alpha) + u(\alpha) &= 1 - B(\neg\alpha) & (44) \\ f(\alpha) + u(\alpha) &= 1 - B(\alpha) & (45) \\ \min(k(\alpha), u(\alpha)) &= 0 & (46) \\ t(\alpha) + k(\alpha) + u(\alpha) + f(\alpha) &= 1 & (47) \end{aligned}$$

3 Interpreting non-atomic formulae

Let $\mathcal{M}$ be the set of 2×2 matrices of reals whose elements m_{ij} verify:

$$\begin{aligned} \forall i, j \in \{1, 2\}, \quad m_{ij} &\geq 0 \\ m_{11} + m_{12} + m_{21} + m_{22} &= 1 \\ \min\{m_{12}, m_{21}\} &= 0 \end{aligned}$$

Given a first order language $\mathcal{L}$ each atomic formula α in the language can be evaluated using the interpretation $v : \mathcal{L} \mapsto \mathcal{M}$ such that for each $\alpha \in \mathcal{L}$

$$v(\alpha) = \begin{bmatrix} f(\alpha) & k(\alpha) \\ u(\alpha) & t(\alpha) \end{bmatrix}$$

A a simple example, consider the truth values "true", "contradictory", "unknown" and "false" defined by matrices T, K, U, F respectively:

$$T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad K = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$U = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad F = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

Moreover, to each atomic formula α corresponds the ordered pair $(B(\alpha), B(\neg\alpha))$. Equations (42–43) define a one-to-one correspondence linking the two representations:

$$\begin{aligned} \psi : \quad \mathcal{M} &\rightarrow [0,1] \times [0,1] \\ v(\alpha) &\mapsto (B(\alpha), B(\neg\alpha)) \end{aligned}$$

Conversely, ψ^{-1} is defined by equations (38–41).

Consider (x, y) and (z, w) two ordered pairs of $[0,1]^2$. We define their lower and upper bound, and the pseudo-complementation by:

$$\begin{aligned} \bot_1[(x,y),(z,w)] &= (\min\{x,z\}, \max\{y,w\}) & (48) \\ \top_1[(x,y),(z,w)] &= (\max\{x,z\}, \min\{y,w\}) & (49) \\ N_1(x,y) &= (y,x) & (50) \end{aligned}$$

defining a pseudo-complemented lattice structure L_1 on $[0,1]^2$.

In the same way, considering the Lukasiewicz triple (LN, LT, LV), we could define a complemented lattice structure L_2 on $[0,1]^2$ as follows:

$$\begin{aligned} \bot_2[(x,y),(z,w)] &= (LV(x,z), LT(y,w)) \\ \top_2[(x,y),(z,w)] &= (LT(x,z), LV(y,w)) \\ N_2(x,y) &= (LN(x), LN(y)) \end{aligned}$$

By ψ we can derive from each lattice $L_j, j = 1, 2$ a lattice structure $L_j^{\mathcal{M}}$ on $\mathcal{M}$ by setting, for all X, Y in $\mathcal{M}$:

$$\begin{aligned} \bot_j^{\mathcal{M}}(X,Y) &= \psi^{-1}(\bot_j[\psi(X), \psi(Y)]) \\ \top_j^{\mathcal{M}}(X,Y) &= \psi^{-1}(\top_j[\psi(X), \psi(Y)]) \\ N_j^{\mathcal{M}}(X) &= \psi^{-1}(N_j[\psi(X)]) \end{aligned}$$

Knowing the interpretation of atomic formulae, we are now able to extend the interpretation function v to any complex formula in $\mathcal{L}$ involving connectives $\neg$ ("not"), $\wedge$ ("and"), $\vee$ ("or") by setting:

$$\begin{aligned} v(\neg\alpha) &= N_1^{\mathcal{M}}(v(\alpha)) & (51) \\ v(\alpha \wedge \beta) &= \bot_1^{\mathcal{M}}(v(\alpha), v(\beta)) & (52) \\ v(\alpha \vee \beta) &= \top_1^{\mathcal{M}}(v(\alpha), v(\beta)) & (53) \end{aligned}$$

Consider now the "truth" partial order $\succeq$ on $[0,1]^2$ defined by:

$$(x,y) \succeq (z,w) \quad \Leftrightarrow \quad (x \geq z \text{ and } y \leq w) \qquad (54)$$

Proposition 3.1 *The relation $\succeq^{\mathcal{M}}$ defined on $\mathcal{M}$ by:*

$$X \succeq^{\mathcal{M}} Y \;\Leftrightarrow\; \psi(X) \succeq \psi(Y)$$

is a partial order.

Proposition 3.2 *For all X, Y in $\mathcal{M}$, we have:*

$$X \succeq Y \quad \Leftrightarrow \quad [\bot_1^{\mathcal{M}}(X,Y) = Y \text{ and } \top_1^{\mathcal{M}}(X,Y)) = X]$$

Therefore, relation $\succeq^{\mathcal{M}}$ partially orders formulae according to their truth value. In order to make implication (denoted $\supset$) and equivalence (denoted $\equiv$) compatible with the partial order $\succeq^{\mathcal{M}}$ we want:

$$v(\alpha) \preceq^{\mathcal{M}} v(\beta) \quad \Rightarrow \quad v(\alpha \supset \beta) = T \tag{55}$$

$$v(\alpha) = v(\beta) \quad \Rightarrow \quad v(\alpha \equiv \beta) = T \tag{56}$$

Remark that the second condition immediately derives from the first one if we interpret $A \equiv B$ as $(A \supset B) \wedge (B \supset A)$ Unfortunately, the first condition does not hold if we set:

$$v(\alpha \supset \beta) = \top_1^{\mathcal{M}}[N_1^{\mathcal{M}}(v(\alpha)), v(\beta)] \tag{57}$$

There is a twofold technical reason for that:

- In Belnap's logic, since complementation differs from negation: $(\neg\alpha \vee \alpha)$ is not a tautology. Therefore, assuming $(\alpha \supset \alpha)$ can be rewritten as $(\neg\alpha \vee \alpha)$ it contradicts condition (55).
- In multivalued logics whose semantic is based on the interpretation domain $[0,1]$, it is not possible to define an idempotent disjunction such that $(\neg\alpha \vee \alpha)$ is a tautology. More precisely, there exists no co-norm V verifying simultaneously:

$$\forall x \in [0,1], \qquad V(x,x) = x \tag{58}$$

$$\forall x \in [0,1], \quad V(1-x,x) = 1 \tag{59}$$

In order to overcome the first difficulty, we will substitute negation by complementation as suggested by [4] and [5]. This amounts to interpreting $(\alpha \supset \alpha)$ by $(\sim \alpha \vee \alpha)$. In order to overcome the second difficulty, we will substitute in equation (55) $\top_1^{\mathcal{M}}$ by $\top_2^{\mathcal{M}}$ while keeping disjunction unchanged in (53). This amounts to use Lukasiewicz co-norm in (59) while keeping the *max* co-norm in (58). This solution has been adopted by many authors in the context of preference modelling (see *e.g.* [7], [12], [6]).

Hence, we obtain the following result:

Proposition 3.3 *Equations (55–56) hold when implication and equivalence are defined by:*

$$\begin{aligned} v(\alpha \supset \beta) &= \top_2^{\mathcal{M}}[N_2^{\mathcal{M}}(v(\alpha)), v(\beta)] \\ v(\alpha \equiv \beta) &= v((\alpha \supset \beta) \ \wedge \ v(\beta \subset \alpha)) \end{aligned}$$

Hence we get the following formulae:

$$v(\neg\alpha) = \begin{bmatrix} t(\alpha) & k(\alpha) \\ u(\alpha) & f(\alpha) \end{bmatrix}$$

$$\begin{aligned} t(\alpha \wedge \beta) &= \min\{t(\alpha), t(\beta)\} \\ k(\alpha \wedge \beta) &= \max\{\min\{t(\alpha)+k(\alpha), t(\beta)+k(\beta)\} \\ &\quad -\min\{t(\alpha)+u(\alpha), t(\beta)+u(\beta)\}, 0\} \\ u(\alpha \wedge \beta) &= \max\{\min\{t(\alpha)+u(\alpha), t(\beta)+u(\beta)\}, 0\} \\ &\quad -\min\{t(\alpha)+k(\alpha), t(\beta)+k(\beta)\} \\ f(\alpha \wedge \beta) &= \min\{\max\{f(\alpha)+u(\alpha), f(\beta)+u(\beta)\}, \\ &\quad \max\{f(\alpha)+k(\alpha), f(\beta)+k(\beta)\} \end{aligned}$$

$$\begin{aligned} t(\alpha \vee \beta) &= \min\{\max\{t(\alpha)+u(\alpha), t(\beta)+u(\beta)\}, \\ &\quad \max\{t(\alpha)+k(\alpha), t(\beta)+k(\beta)\} \\ k(\alpha \vee \beta) &= \max\{\min\{f(\alpha)+k(\alpha), f(\beta)+k(\beta)\} \\ &\quad -\min\{f(\alpha)+u(\alpha), f(\beta)+u(\beta)\}, 0\} \\ u(\alpha \vee \beta) &= \max\{\min\{f(\alpha)+u(\alpha), f(\beta)+u(\beta)\}, 0\} \\ &\quad -\min\{f(\alpha)+k(\alpha), f(\beta)+k(\beta)\} \\ f(\alpha \vee \beta) &= \min\{f(\alpha), f(\beta)\} \end{aligned}$$

$$\begin{aligned} t(\alpha \supset \beta) &= \min\{f(\alpha)+u(\alpha)+t(\beta)+k(\beta), \\ &\quad f(\alpha)+k(\alpha), t(\beta)+u(\beta), 1\} \\ k(\alpha \supset \beta) &= \max\{\min\{f(\alpha)+u(\alpha)+t(\beta)+k(\beta), 1\} \\ &\quad -\min\{f(\alpha)+k(\alpha)+t(\beta)+u(\beta), 1\}\} \\ u(\alpha \supset \beta) &= \max\{\min\{f(\alpha)+k(\alpha), t(\beta)+u(\beta)\}, 1\} \\ &\quad -\min\{f(\alpha)+u(\alpha)+t(\beta)+k(\beta), 1\}\} \\ f(\alpha \supset \beta) &= \min\{\max\{t(\alpha)+k(\alpha)-t(\beta)-k(\beta), 0\}, \\ &\quad \max\{f(\beta)+k(\beta)-f(\alpha)-k(\alpha), 0\}\} \end{aligned}$$

Remark 3.1 *In order to give an example of how the previous formulae are obtained we present extensively the calculation of $f(\alpha \wedge \beta)$.*
From (41), $f(\alpha \wedge \beta) = \min(1 - B(\alpha \wedge \beta), B(\neg(\alpha \wedge \beta))$.
From (48) and (52) we get: $1 - B(\alpha \wedge \beta) = 1 - \min(B(\alpha), B(\beta)) = \max(1 - B(\alpha), 1 - B(\beta)) = \max(f(\alpha)+u(\alpha), f(\beta)+u(\beta))$
and also: $B(\neg(\alpha \wedge \beta)) = \max(B(\neg\alpha), B(\neg\beta)) = \max(f(\alpha)+k(\alpha), f(\beta)+k(\beta))$.

Finally, we define the interpretation of a quantified formulae for a predicate S whose domain is A, as follows:

$$v(\forall x S(x)) = \bot_1^{\mathcal{M}}\{v(S(a)), a \in A\} \tag{60}$$

$$v(\exists x S(x)) = \top_1^{\mathcal{M}}\{v(S(a)), a \in A\} \tag{61}$$

This completes the definition of the interpretation of first order logical formulae. For any formula α, $v(\alpha)$ is a matrix of $\mathcal{M}$ defined as a continuous function of the truth values of atomic components of α.

4 Application to Preference Modelling

We illustrate now the use of our logic to model preferences in the context of group decision making.

Consider a set of alternatives $A = \{a, b, c, d\}$ and a set of 100 voters whose preferences can be synthesized as follows:

46 voters have the preference $a \succ b \succ c \succ d$
17 voters have the preference $b \succ c \succ d \succ a$
18 voters have the preference $c \succ d \succ a \succ b$
19 voters have the preference $d \succ a \succ b \succ c$

The problem is now to elaborate an overall preference model reflecting the collective opinion of the population of voters.

For all (x, y) in $A \times A$, a voter is supposed to support the preference of x over y (denoted $P(x, y)$) if he ranks x strictly before y. The total number of voters supporting $P(x, y)$ is denoted γ_{xy}. Conversely, a voter is supposed to be *opposed* to the preference $P(x, y)$ if he ranks y strictly before x. In this problem we suppose that, for each voter j, the strength of this opposition is proportional to the rank difference between x and y and can be represented, for any $j \in [1, 100]$, by the normalized difference:

$$\delta_{xy}(j) = \max\{r_j(x) - r_j(y), 0\}/3$$

Therefore, the overall opposition to $P(x, y)$ within the population of voters is given by:

$$\delta_{xy} = \sum_{j=1}^{100} \delta_{xy}(j)$$

Remark 4.1 $\forall x, y \in A,\ \gamma_{xy} + \delta_{xy} \leq 100$

For any ordered pair (x, y) we believe that x is preferred to y by the population when γ_{xy} is "large enough". Independently, we believe that x is not preferred to y when δ_{xy} is "significant". This implicitly implies to consider thresholds separating "large" and "not large" proportions of voters on the one hand, "significant" and not "significant" oppositions on the other hand. Since such thresholds cannot be determined precisely, we will represent these notions with fuzzy subsets of $[0, 100]$. We suppose here that a proportion of voter below 50% is not large enough to support the preference $P(x, y)$ and becomes more and more credible when it raises to 70%. If the proportion exceeds 70% we totally believe in $P(x, y)$. Moreover, we will consider here that an opposition of strength $\delta_{xy} \geq 30$ vetoes the preference $P(x, y)$ whereas an opposition $\delta_{xy} \leq 15$ is not significant. Thus, we set:

$$B(P(x, y)) = \begin{cases} 0 & if \quad \gamma_{xy} \leq 50 \\ \frac{\gamma_{xy} - 50}{20} & if \quad 50 \leq \gamma_{xy} \leq 70 \\ 1 & if \quad \gamma_{xy} \geq 70 \end{cases}$$

$$B(\neg P(x, y)) = \begin{cases} 0 & if \quad \delta_{xy} \leq 15 \\ \frac{\delta_{xy} - 15}{15} & if \quad 15 \leq \delta_{xy} \leq 30 \\ 1 & if \quad \delta_{xy} \geq 30 \end{cases}$$

Hence we get the following results (see table 1).

Table 1: Truth values for preference

α	$B(\alpha)$	$B(\neg\alpha)$	$t(\alpha)$	$k(\alpha)$	$u(\alpha)$	$f(\alpha)$
P(a, b)	1.00	0.13	0.87	0.13	0.00	0.00
P(a, c)	0.75	0.56	0.44	0.31	0.00	0.25
P(a, d)	0.00	0.20	0.00	0.00	0.80	0.20
P(b, a)	0.00	0.84	0.00	0.00	0.16	0.84
P(b, c)	1.00	0.20	0.80	0.20	0.00	0.00
P(b, d)	0.65	0.64	0.36	0.29	0.00	0.35
P(c, a)	0.00	1.00	0.00	0.00	0.00	1.00
P(c, b)	0.00	0.82	0.00	0.00	0.18	0.82
P(c, d)	1.00	0.27	0.73	0.27	0.00	0.00
P(d, a)	0.20	1.00	0.00	0.20	0.00	0.80
P(d, b)	0.00	1.00	0.00	0.00	0.00	1.00
P(d, c)	0.00	0.80	0.00	0.00	0.20	0.80

Concerning the pair (a, b) we can observe that preference is well established since $t(P(a, b)) = 0.87$. This seems consistent with the rules defined previously since the large majority of voters support this preference and the opposition is weak ($\delta_{ab} = 17$).

Concerning the pair (a, c), the preference is not so well established. $k(P(a, c)) = 0.31$ reflects a conflict of arguments in the evaluation of preference. This seems consistent with the initial information since a majority of criteria support this preference (65%) but a significant discordant coalition ($\delta_{ac} = 27.33$) is conflicting with this preference. Moreover, since the majority is weak and the opposition is significant, it seems also natural to observe that $f(P(a, c)) = 0.25$ whereas $t(P(a, c)) = 0.44$.

Concerning the preference $P(a, d)$, only a minority of criteria support this assertion but the opposition is weak. The above rules do not allow to establish that a is better than d but they provide very few arguments against this preference. Therefore it seems consistent to observe $u(P(a, d)) = 0.80$ and $f(P(a, d)) = 0.20$.

Suppose now that we have to decide which is the best alternative on the basis of this information. One can try to identify a *Condorcet Winner i.e* an alternative which is preferred to all other alternatives. In our language, the extend to which an alternative is a Condorcet Winner can be evaluated in $\mathcal{M}$ by the truth value:

$$CW(a) = v(\forall x \in A \setminus \{a\}, P(a, x))$$

From equation (60) and table 1 we get the following results (see table 2):

Table 2: Search of a Condorcet Winner

α	$t(\alpha)$	$k(\alpha)$	$u(\alpha)$	$f(\alpha)$
CW(a)	0.00	0.00	0.44	0.56
CW(b)	0.00	0.00	0.16	0.84
CW(c)	0.00	0.00	0.00	1.00
CW(d)	0.00	0.00	0.00	1.00

Since there is no obvious Condorcet winner, we can extend Orlovski's choice function (see [7]) in this context by defining in our logic the set ND of non-dominated alternatives as follows:

$$ND = \{a \in A, \forall x \in A \setminus \{a\}, \neg P(x, a)\}$$

Thus, for each element $a \in A$, membership in ND is measured by the truth value $ND(a) \in \mathcal{M}$ defined by:

$$ND(a) = v(\forall x \in A \setminus \{a\}, \neg P(x, a))$$

From table 1 propositions of type $\neg P(x, y)$ must be interpreted as follows (see table 3):

Table 3: Truth values for non preference

α	$t(\alpha)$	$k(\alpha)$	$u(\alpha)$	$f(\alpha)$
¬P(b, a)	0.84	0.00	0.16	0.00
¬P(c, a)	1.00	0.00	0.00	0.00
¬P(d, a)	0.80	0.20	0.00	0.00
¬P(a, b)	0.00	0.13	0.00	0.87
¬P(c, b)	0.82	0.00	0.18	0.00
¬P(d, b)	1.00	0.00	0.00	0.00
¬P(a, c)	0.25	0.31	0.00	0.44
¬P(b, c)	0.00	0.20	0.00	0.80
¬P(d, c)	0.80	0.00	0.20	0.00
¬P(a, d)	0.20	0.00	0.80	0.00
¬P(b, d)	0.35	0.29	0.00	0.36
¬P(c, d)	0.00	0.27	0.00	0.73

From equation (60) and table 3 we get the following results (see table 4):

Table 4: Non-dominated alternatives

α	$t(\alpha)$	$k(\alpha)$	$u(\alpha)$	$f(\alpha)$
ND(a)	0.80	0.04	0.00	0.16
ND(b)	0.00	0.13	0.00	0.87
ND(c)	0.00	0.20	0.00	0.80
ND(d)	0.00	0.20	0.00	0.80

Clearly a is the best non-dominated alternative.

Conclusions

A fuzzy extension of Belnap's logic is presented in the paper mainly in order to model a continuous "intensity" of truthness and knowledge in a sentence. Such a logic has been conceived mainly for preference modeling purposes, since it is a field where a continuous "modulation" of the concept of preference is very useful.

As shown by some little examples, our fuzzy logic makes explicit the truth, falsity, unknown and contradictory parts in any first order logical formula and thus provides useful and synthetic information about the real nature of arguments supporting this formula. Moreover the calculus of a truth value involves only continuous functions and make it impossible that small variations in the initial information lead to drastic changes in the output. This seems well fitted to the context of preference modelling.

Several problems are of course open, including:

- the experimentation of different families of De Morgan triples, basically in the case where "uncertainty" instead of "intensity" has to be modeled;
- a complete truth calculus for logics conceived as fuzzy extensions of four valued paraconsistent logics;
- a more thorough investigation of valued sets and valued relations (when the valuation domain is $\mathcal{M}$) and their potential use in the context of preference modelling.

References

[1] Alsina C., "On a family of connectives for fuzzy sets", *Fuzzy sets and systems*, 16, 231–235, 1985.

[2] Belnap N.D., (1977), "A useful four-valued logic", in G. Epstein, J. Dumm, (eds.), *Modern uses of multiple valued logics*, D. Reidel, Dordrecht, 8-37.

[3] De Baets B., Van de Walle B. and Kerre E., "Fuzzy preference structures and their characterization", *The Journal of Fuzzy Mathematics*, 3, 373–381,1995.

[4] Doherty P., Driankov D., Tsoukiàs A., "Partiality, Para-consistency and Preference Modeling", *IDA research report Lith-IDA-R-92-18*, Linköping University, 1992.

[5] Fages F., Ruet P., "Combining explicit negation and negation by failure via Belnap's logic", LIENS research report 94-15, Ecole Normale Supérieure, 1994, (to appear in Theoretical Computer Science).

[6] Fodor J. and Roubens M., *Fuzzy preference modelling and multicriteria decision support*, Kluwer Academic Publishers, Dordrecht, 1994.

[7] Orlovski S., "Decision Making with a fuzzy preference relation", *Fuzzy Sets and Systems*, 1, 155–167, 1978.

[8] Ovchinnikov S., "Structure of fuzzy binary relations", *Fuzzy Sets and Systems*, 1, 155–167, 1981.

[9] Ovchinnikov S. and Roubens M., "On strict preference relations", *Fuzzy Sets and Systems*, 43, 319–326, 1991.

[10] Ovchinnikov S. and Roubens M., "On fuzzy strict preference, indifference and incomparability relations", *Fuzzy Sets and Systems*, 47, 313–318, 1992.

[11] Perny P. and Roubens M., "Fuzzy Preference Modelling", in *Handbook of fuzzy sets and possibility theory*, vol. IV, R. Slowinski (Ed.), Kluwer Academic Publishers, to appear.

[12] Perny P. and Roy B., "The use of fuzzy outranking relations in preference modelling", *Fuzzy sets and Systems*, 49, 33–53, 1992.

[13] Tsoukiàs A., Vincke Ph., "A new axiomatic foundation of partial comparability", *Theory and Decision*, vol. 39, 79 - 114, 1995.

[14] Tsoukiàs A., Vincke Ph., "Extended preference structures in MCDA", in J. Clímaco (ed.), *Multicriteria Analysis*, Springer Verlag, Berlin, 37 - 50, 1997.