Representing Priorities of Defaults in Possibility Theory

Churn-Jung Liau
Institute of Information Science
Academia Sinica, Taipei, Taiwan, ROC
e-mail: liaucj@iis.sinica.edu.tw

Abstract

It is suggested by Yager[5] that default knowledge can be represented in the framework of possibility theory. In this note, we show a potential application of the proposal. That is, the priorities of defaults can be encoded into the formalism naturally and the results produced by such encoded default theories reflect the strength of the conclusions.

1 Introduction

Possibility theory[7] and default logic[4] deal with vague and incomplete information respectively. Yager[5] suggestes that default knowledge can be represented in the framework of possibility theory. This results in an integrated formalism to represent vague and incomplete information. On the other hand, to eliminate unintended extensions in default theories, Brewka[1] proposes the prioritized default logic to partition a default theory into levels according to the priorities of defaults. In this note, we will show that a prioritized default theory can be transalted into Yager's formalism naturally. This in turn shows a potential complementarity between symbolic and numeric reasoning schemes.

In what follows, we first introduce Reiter's default logic and the prioritized one. Then, Yager's proposal is given. The formalism is called possibilistic default logic. In Sec. 4, we present a translation of prioritized default logic into the possibilistic one. Finally, a brief conclusion follows.

2 Default Logic

We restrict our attention to propositional default logic. Let L denote the set of all well-formed formulas (wffs) in a propositional language. For any $S \subseteq L$, define $Th_L(S) = \{\gamma | \gamma \in L \text{ and } S \vdash \gamma\}$, where $S \vdash \gamma$ means that γ is classically provable from S.

A *default* is expressed in the following form,

$$\frac{\alpha : \beta_1, \beta_2, \ldots, \beta_n}{\gamma}, \tag{1}$$

where α, β_i's and γ are wffs of L. α is called the prerequisite of the default, β_i's are its justifications and γ is its consequent. A default theory(DT) is a pair $\langle W, D\rangle$ where $W \subseteq L$ and D is a set of defaults.

The intended meaning of a default is as follows.

if α is known, and none of $\beta_1, \ldots, \beta_n$ are known to be false, then γ can be assumed by default.

The intuitive meaning motivates the following definition:

Definition 1 *([4]) Let $\Delta = \langle W, D\rangle$ be a DT based on a propositional language L.*

(1) For any $S \subseteq L$ let $\Gamma(S)$ be the smallest set satisfying the following three properties:

D1. $W \subseteq \Gamma(S)$

D2. $Th_L(\Gamma(S)) = \Gamma(S)$

D3. If $\frac{\alpha:\beta_1,\beta_2,\ldots,\beta_n}{\gamma} \in D$ *and* $\alpha \in \Gamma(S)$, *and* $\neg\beta_1, \ldots, \neg\beta_n \notin S$ *then* $\gamma \in \Gamma(S)$.

(2) A set $E \subseteq L$ is an extension of Δ iff $\Gamma(E) = E$.

When defaults in D have different priorities, we have the following definition.

Definition 2 *[1]*

1. *Let W be a set of wffs in L and D_i be a set of classical defaults for $1 \le i \le k$, then $\Delta = \langle W, D_1, \ldots, D_k\rangle$ is called a prioritized default theory(PDT).*

2. *E is a PDT-extension of Δ iff there exist sets of wffs $E_1, \ldots, E_k$ such that E_i is an extension of $\langle E_{i-1}, D_i\rangle$ for $1 \le i \le k$, where for convenience, $E_0 = W$.*

3 Possibilistic Default Logic

In what follows, let U be a universe of discourse and X be a variable (that may be attributes of a given object) taking its values on U. Also, let A, B, and C(possibly with indices) denote fuzzy subsets of U. A piece of vague or fuzzy information about X can usually be considered as some elastic constraints on X[8]. A basic statement to express these constraints is of the form:

$$X \text{ is } A.$$

The intent of this statement is to indicate that the possible values of X are constrained by the fuzzy set A.

Two indices that measure the degree of consistency and containment between two fuzzy sets are used extensively in approximate reasoning. These two indices, called possibility and certainty, are defined as

$$Poss[B|A] = \sup_{u \in U} \min[A(u), B(u)],$$

$$Cert[B|A] = 1 - Poss[\overline{B}|A].$$

The former measures the degree of A intersecting with B, and the latter measures the degree of A being a subset of B. When A and B are crisp sets, $Poss[B|A] = 1$ iff $A \cap B \neq \emptyset$ and $Cert[B|A] = 1$ iff $A \subseteq B$. Also note $Poss[B|A] = Poss[A|B]$. Based on the possibility index, a *level 2 fuzzy set* [3] can be constructed from an ordinary one. Let $\tilde{\wp}(U)$ denote the class of all fuzzy subsets of U. A level 2 fuzzy set is a fuzzy subset of $\tilde{\wp}(U)$. We will use $\tilde{A}$ to denote a general level 2 fuzzy set. For each A, a special level 2 fuzzy set A^+ can be defined with membership function $A^+(G) = Poss[A|G]$ for all $G \in \tilde{\wp}(U)$. The fuzzy subsets of U are also called level 1 fuzzy sets.

Yager suggests that a default can be represented in the framework of possibility theory as follows:

$$\frac{A : B_1, B_2, \ldots, B_n}{C} \tag{2}$$

A, B_i's, and C are the prerequisite, justifications, and consequent of the default respectively. A possibilistic default theory (ΠDT) is a pair $\Delta_\pi = \langle F, D_\pi \rangle$, where F is a fuzzy subset of U and D_π is a set of possibilistic defaults in the form of (2).

Assume $\Delta_\pi = \langle F, \{d_i\}_{1 \leq i \leq m} \rangle$ is a ΠDT with finite defaults where each default

$$d_i = \frac{A_i : B_{i1}, B_{i2}, \ldots, B_{in_i}}{C_i}. \tag{3}$$

Acccording to Yager[6], we can define

$$\delta_i = \{(\bar{A}_i)^+, \overline{B^+_{i1}}, \ldots, \overline{B^+_{in_i}}, C_i\},$$

for each d_i. Let $V \subset U$ and $\chi = \chi_1 \cup \chi_2$ be a class of level 1 or 2 fuzzy sets so that all elements in χ_1 are level 1 and those in χ_2 are level 2, we use $\chi[V]$ to denote the fuzzy subset of U such that for each $u \in U$

$$\chi[V](u) = \max(\{\tilde{A}(V) | \tilde{A} \in \chi_2\} \cup \{B(u) | B \in \chi_1\}).$$

We can further abbreviate $V \cap \chi[V]$ as $V \cap \chi$. Then for each permutation σ on $I_m = \{1, 2, \ldots, m\}$, we can define $V_0^\sigma = F$ and define

$$V_i^\sigma = V_{i-1}^\sigma \cap \delta_{\sigma(i)}$$

for $1 \leq i \leq m$. Finally, let $V^\sigma = V_m^\sigma$.

Definition 3 *Let V be a fuzzy subset of U, then V is said to be sequentially generated from Δ_π iff there exists a permutation σ on the index set of the defaults such that $V = V^\sigma$.*

Intuitively, for the given ΠDT, the initial knowledge is denoted by the fuzzy set F. That is, the possible values of X are constrained by F. Then, the defaults are applied to F in some fixed order so that the possible values of X are constrained further. Finally, a more specific fuzzy set V^σ is obtained that indicates the combination of the initial knowledge and defaults.

Analogous to default logic, we can give a fixed point definition for those fuzzy sets that are considered as reasonable restrictions of X in accordance with a ΠDT.

Definition 4 *Let $\Delta_\pi = \langle F, D_\pi \rangle$ be a ΠDT, then for each $V \in \tilde{\wp}(U)$, define $\Gamma_\pi(V)$ as the largest (least specific) fuzzy subset of F such that for each default in D_π of form (2) and $u \in U$,*

$$\begin{aligned} \Gamma_\pi(V)(u) \quad \leq \quad & \max(Poss[\overline{A}|\Gamma_\pi(V)], Cert[\overline{B_1}|V], \\ & \ldots, Cert[\overline{B_n}|V], C(u)). \end{aligned}$$

A fuzzy set V is called a possibility restriction(π-restriction) of Δ_π iff $\Gamma_\pi(V) = V$.

4 Translation of PDT into ΠDT

We will show that a PDT can be translated into a ΠDT naturally, and the results produced by the ΠDT refelect the certainty of the default conclusions in an explicit way. To represent the translation, we need to use the notion of the fuzzy default value defined by Dubois and Prade[2]. According to the notion, the default knowledge "Usually, X is A" is translated into the statement "X is A^λ", where $\lambda \in [0,1]$ and the membership function of A^λ is defined as

$$A^\lambda(u) = \max(A(u), 1-\lambda).$$

This means that though in general X takes its value inside A, it may take its value outside A with the possibility $1-\lambda$. That is, λ estimates the certainty that the value of X is not an exception. When $\lambda = 1$, A^λ is reduced to A, i.e., the knowledge is certain without exceptions.

In the translation, we let the universe U be the set of all truth assignments on L and let $Mod(\varphi)$ denote $\{u | u \models \varphi, u \in U\}$ for any $\varphi \in L$.

For each default $d = \frac{\alpha:\beta_1,\ldots,\beta_n}{\gamma}$, define the translation of d with certainty λ as

$$t(d,\lambda) = \frac{A : B_1, \ldots, B_n}{C^\lambda}$$

and let

$$Mod(d) = \frac{A : B_1, \ldots, B_n}{C}$$

where $A = Mod(\alpha)$, $B_i = Mod(\beta_i)$ for $1 \le i \le n$ and $C = Mod(\gamma)$. Then given a PDT $\Delta = \langle W, D_1, \ldots, D_k\rangle$ and an arbitrary strictly descreasing function $f : I_k \to (\frac{1}{2}, 1]$, we can define a ΠDT $Tr_f(\Delta) = \langle F, D_\pi\rangle$, where $F = Mod(W)$ and

$$D_\pi = \{t(d, f(i)) \mid d \in D_i, 1 \le i \le k\}.$$

The function f is called certainty assignment function. Though different certainty assignment functions may result in different translations for a PDT, the difference can be ignored when only the ordering is relevent. Thus, we will assume a fixed f from now on(e.g. $f(i) = \frac{i+2}{2(i+1)}$).

To investigate the relationship between prioritized default logic and the possibilistic one, let us consider a given finite PDT $\Delta = \langle W, D_1, \ldots, D_k\rangle$ such that $D_1 \cup \ldots \cup D_k = \{d_i | 1 \le i \le m\}$, where each

$$d_i = \frac{\alpha_i : \beta_{i1}, \beta_{i2}, \ldots, \beta_{in_i}}{\gamma_i}.$$

Let us assume that the index set is partitioned so that $d_i \in D_j$ iff $i \in I^j$, where $I^j = \{i | m_{j-1}+1 \le i \le m_j\}$ for $1 \le j \le k$ and $0 = m_0 < m_1 < \ldots < m_k = m$. Then for each d_i, we have

$$\delta_i = \{(\bar{A}_i)^+, \overline{B^+_{i1}}, \ldots, \overline{B^+_{in_i}}, C_i^{\lambda_i}\},$$

where $A_i = Mod(\alpha_i)$, $B_{ij} = Mod(\beta_{ij})(1 \le j \le n_i)$, $C_i = Mod(\gamma_i)$, and $\lambda_i = f(j)$ iff $i \in I^j$, for $1 \le i \le m$. For convinience, we also assume $\lambda_0 = 1$.

Let us first consider the fuzzy sets sequentially generated from $Tr_f(\Delta)$. A permutation σ on I_m is said to be compatible with $Tr_f(\Delta)$ iff $\lambda_{\sigma(i)} \ge \lambda_{\sigma(i+1)}$ for $1 \le i \le m-1$. We will only consider those fuzzy sets sequentially generated from $Tr_f(\Delta)$ by a permutation compatible with it. Without loss of generality, we may assume σ is the identity permutation and drop the superscript σ. Then we have the following definitions.

$$S_0 = W,$$

$$S_i = \begin{cases} S_{i-1} \cup \{\gamma_i\} & \text{if } S_{i-1} \vdash \alpha_i \text{ and} \\ & \forall 1 \le j \le n_i, S_{i-1} \not\vdash \neg\beta_{ij} \\ S_{i-1} & \text{otherwise,} \end{cases}$$

$$V_0 = F = Mod(W),$$

and

$$V_i = V_{i-1} \cap \delta_i,$$

for $1 \le i \le m$.

If V is a fuzzy subset of U, we define the binary approximation of V, denoted by $\widehat{V}$ as the crisp subset such that $u \in \widehat{V}$ iff $V(u) > \frac{1}{2}$. Then we can prove the following lemmas.

Lemma 1 *For $0 \le i \le m$, the following results hold*

(1) $\widehat{V_i} = Mod(S_i)$,

(2) for each $u \in U$, $V_i(u) \in \Lambda_i =_{df} \{\lambda_j, 1-\lambda_j | 0 \le j \le i\}$, and

(3) if $V_{i-1}(u) < \frac{1}{2}$, then $V_i(u) = V_{i-1}(u)$ for $1 \le i \le m$.

Define the threshold normalization of V_m as a fuzzy subset V so that

$$V(u) = \begin{cases} V_m(u) & \text{if } V_m(u) \le \frac{1}{2} \\ 1 & \text{otherwise.} \end{cases}$$

Lemma 2 *Let $\varphi \in L$ and V be the threshold normalization of V_m. Define $N(\varphi) = 1 - \sup_{u \notin Mod(\varphi)} V(u)$. Then*

1. *if $\varphi \in Th_L(S_m)$ and i is the smallest index so that $S_i \vdash \varphi$, then $N(\varphi) = \lambda_i$,*
2. *if $\varphi \notin Th_L(S_m)$, then $N(\varphi) = 0$.*

Note that because we adopt threshold normalization, V is truly normalized only when $height(V_m) > \frac{1}{2}$ in the lemma. Thus, if $height(V_m) \leq \frac{1}{2}$, then $N(\varphi) > 0$ for all $\varphi \in L$, i.e., S_m is inconsistent.

Let us define $E_i = Th_L(S_{m_i})$ for $0 \leq i \leq k$ and call a wff $\varphi \in E_k$ *i-rooted* iff i is the smallest index so that $\varphi \in E_i$. We assume $\lambda_1 < 1$ in the following theorem.

Theorem 1 *If φ_1 and φ_2 are i-rooted and j-rooted respectively, and $i > j$, then $N(\varphi_1) < N(\varphi_2)$.*

This theorem shows that the fuzzy sets sequentially generated by $Tr_f(\Delta)$ provides useful information about the derivation process of default conclusions for Δ when E_k is a PDT-extension of Δ. However, it can also easily seen that not all fuzzy sets sequentially generated from compatible permutation correspond to PDT-extensions. This motivates the following definition.

Definition 5 *A fuzzy subset V_m sequentially generated from $Tr_f(\Delta)$ by a peumutation compatible with it is called a prioritized π-restriction of $Tr_f(\Delta)$ iff for all $1 \leq j \leq k$, $\widehat{V_{m_j}}$ is a π-restriction of $\langle \widehat{V_{m_{j-1}}}, Mod(D_j)\rangle$, where $Mod(D_j) = \{Mod(d) | d \in D_j\}$.*

Theorem 2 *V is a prioritized π-extension of $Tr_f(\Delta)$ iff there is a PDT-extension E of Δ so that $\widehat{V} = Mod(E)$.*

Finally, let us use an example to illustrate the abovementioned results.

Example 1 Consider the (extended) high school dropout example. The example has the following three defaults

d_1 = *Typically high school dropouts are adults,*

d_2 = *Typically adults are employed,*

d_3 = *Typically high school dropout are not employed.*

Assume further that d_1 and d_3 has higher priority over d_2. Thus $D_1 = \{d_1, d_3\}$ and $D_2 = \{d_2\}$. Let h, a, and e denote "Jack is a high school dropout", "Jack is an adult", and "Jack is employed" respectively.

The eight truth assignments of the three propositional symbols are defined in the following table

	u_0	u_1	u_2	u_3	u_4	u_5	u_6	u_7
h	0	0	0	0	1	1	1	1
a	0	0	1	1	0	0	1	1
e	0	1	0	1	0	1	0	1

For $I \subseteq \{0, 1, 2, \ldots, 7\}$, let $A_I = \{u_i | i \in I\}$. Then

$$\delta_1^\lambda = \{A^+_{0123}, \overline{A^+_{2367}}, A^\lambda_{2367}\},$$

$$\delta_2^{\lambda'} = \{A^+_{0145}, \overline{A^+_{1357}}, A^{\lambda'}_{1357}\},$$

$$\delta_3^\lambda = \{A^+_{0123}, \overline{A^+_{0246}}, A^\lambda_{0246}\},$$

and assume $1 > \lambda > \lambda' > 0.5$ according to the priorities. Moreover, assume the knowledge base $F = A_{4567}$, i.e., the only known fact is that Jack is a high school dropout. Then $E_0 = Th_L(\{h\})$ and $E_1 = E_2 = Th_L(\{h, a, \neg e\})$. The fuzzy subset sequentially generated by the permutation $\sigma(1) = 1, \sigma(2) = 3, \sigma(3) = 2$, is $V^\sigma = (0, 0, 0, 0, 1-\lambda, 1-\lambda, \lambda, 1-\lambda)$ where the ith element of the tuple denotes $V^\sigma(u_i)$. By normalization, we get $V = (0, 0, 0, 0, 1-\lambda, 1-\lambda, 1, 1-\lambda)$. Thus $N(h) = 1, N(a) = N(\neg e) = \lambda$, and $N(e) = 0$. This means that h is known certainly, a and $\neg e$ are drawn as default conclusions and e is not derivable.

On the other hand, if F is equal to A_{2367} instead of A_{4567}, i.e., only the fact that Jack is an adult is known, then the same σ will produce $V^\sigma = (0, 0, 1-\lambda', 1, 0, 0, 1-\lambda', 1)$ that has been normalized. Thus $N(a) = 1, N(h) = N(\neg h) = 0$, and $N(e) = \lambda'$. This corresponds to the fact that $E_0 = E_1 = Th_L(\{a\})$ and $E_2 = Th_L(\{a, e\})$.

5 Conclusions

It has been shown that prioritized default theories have natural translations in ΠDTs. The priorities can be encoded into the consequents of defaults and the strength of a conclusion is reflected in its necessity measure according to the fuzzy sets sequentially generated from the resultant ΠDT.

References

[1] G. Brewka. *Nonmonotonic Reasoning : Logical Foundation of Commonsense.* Number 12 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1991.

[2] D. Dubois and H. Prade. "Default reasoning and possibility theory". *Artificial Intelligence*, 35:243–257, 1988.

[3] G. J. Klir and T. A. Folger. *Fuzzy Sets, Uncertainty, and Information.* Prentice-Hall International Inc., 1988.

[4] R. Reiter. "A Logic for default reasoning". *Artificial Intelligence*, 13:81–132, 1980.

[5] R. R. Yager. "Using approximate reasoning to represent default knowledge ". *Artificial Intelligence*, 31:99–112, 1987.

[6] R. R. Yager. "A generalized view of nonmonotonic knowledge: A set-theoretic perspective". *International Journal of General Systems*, 14:251–265, 1988.

[7] L. A. Zadeh. "Fuzzy sets as a basis for a theory of possibility ". *Fuzzy Sets and Systems*, 1:3–28, 1978.

[8] L. A. Zadeh. "PRUF - a meaning representation language for natural languages". *Int. J. Man-Machine Studies*, 10:395–460, 1978.