APPLICATION SCHEMES OF POSSIBILITY THEORY-BASED DEFAULTS

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

Abstract

In this paper, we analyze the formalism proposed by Yager[6] to represent default knowledge in the framework of possibility theory. Three different application schemes are examined. The results produced by these application schemes are compared with those by Reiter's default logic. It is shown that for a restricted class of default theories, there is some kinds of correspondence between Yager's formalism and Reiter's default logic. It is also exemplified that there is much mismatch between them for general default theories. Finally, the fixed point mechanism is suggested as an auxiliary scheme to eliminate such mismatch.

1 Introduction

Vagueness and incompleteness of information are two main sources of uncertainty in knowledge based systems. Possibility theory and default logic are two main approaches for these two types of uncertainty respectively. Because the two types of uncertainty often occur simultaneously, it is desirable to have an integrated formalism to handle both of them in the same time. Yager[6] suggests to represent default knowledge in the framework of possibility theory and shows that the resultant formalism indeed produces reasonable results in many cases. However, when multiple defaults are represented, there is much mismatch between his proposal and default logic. In this paper, we formalize and analyze Yager's proposal in detail and prove some correspondence theorems between Reiter's default logic and it for some restricted class of default theories. We also show that for general default theories, there is much difference between the results produced by them. We show that by incorporating the fixed point characterization into Yager's formalism as an auxiliary, the mismatch can be eliminated.

In what follows, we first review Reiter's default logic and Yager's proposal to represent default knowledge in the framework of possibility theory. Second, the conjunctive application scheme, the sequential application scheme, and the parallel application scheme, which are mainly exemplified in [6, 7, 5], are presented and analyzed in turn. Then, an auxiliary scheme is proposed. Finally, a brief conclusion is given.

2 Basic Default Logic

Here, we restrict our attention to propositional default logic. Let L denote the set of all well-formed formulas (wffs) of a propositional language. For any $S \subseteq L$, define $Th_L(S) = \{\gamma | \gamma \in L, S \vdash \gamma\}$, 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 $\alpha, \beta_1, \beta_2, \ldots, \beta_n, \gamma$ are wffs of L. 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 $\Gamma(E) = E$.

3 Possibility Theory-Based Default Reasoning

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[9]. 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* [2] can be constructed from an ordinary one. Let $\widetilde{\wp}(U)$ denote the class of all fuzzy subsets of U. A level 2 fuzzy set is a fuzzy subset of $\widetilde{\wp}(U)$. We will use $\widetilde{A}$ to denote a general level 2 fuzzy set. Let $\widetilde{A}$ and B be level 2 and level 1 fuzzy sets of a same universe respectively, then the application of *effection procedure* to $\widetilde{A} \cap B$ results in a level 1 fuzzy set C such that $C(u) = \min(\widetilde{A}(B), B(u))$ for each $u \in U$.[8] For each A, a special level 2 fuzzy set A^+ can be defined with membership function $A^+(G) = Poss[A|G]$ for all $G \in \widetilde{\wp}(U)$.

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).

Given a DT $\Delta = \langle W, D \rangle$, we can find a corresponding representation in the framework of possibility theory. First, let U be the set of all truth assignments on the language L and define $Mod(\gamma)$ as the set of all models of γ for any $\gamma \in L$. Second, let $Mod(W) =_{df} \bigcap_{\gamma \in W} Mod(\gamma)$. Finally, define

$$Mod(D) = \{\frac{Mod(\alpha) : Mod(\beta_1), Mod(\beta_2), \ldots, Mod(\beta_n)}{Mod(\gamma)} \mid \frac{\alpha : \beta_1, \beta_2, \ldots, \beta_n}{\gamma} \in D\}.$$

Then the transalation of Δ into ΠDT is defined as $Tr(\Delta) = \langle Mod(W), Mod(D) \rangle$.

4 Application Schemes of Possibilistic Defaults

Throughout this section, let us assume a ΠDT $\Delta_\pi = \langle F, \{d_i\}_{1 \leq i \leq m} \rangle$ where each default

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

Given such a ΠDT, it is implicitly assumed that we are talking about a variable X taking its values on U. The fuzzy set F constrains the possible values of X and each default is intended to impose further constraints on X. F represents the known facts about X and the defaults reflect what the typical values of X should be. There are some different schemes according to which the defaults can be applied and the results produced are rather different. We discuss three kinds of schemes and their implications in what follows.

4.1 The conjunctive application scheme

This scheme is used to show that Yager's formalism derives reasonable results in the famous example of Nixon diamond[5]. First, let us define some notations. Assume $S_1, S_2, \ldots, S_k$ are k crisp sets, let $\otimes_{i=1}^{k} S_i = \{\{s_1, s_2, \ldots, s_k\} \mid s_i \in S_i, \forall 1 \leq i \leq k\}$. For each default d_i, define $\delta_i = \{(\bar{A}_i)^+, \overline{B_{i1}^+}, \ldots, \overline{B_{in_i}^+}, C_i\}$. Also let $\delta_0 = \{F\}$ for convenience. Suppose χ is a collection of level 1 or 2 fuzzy sets, then $eff(\chi)$ is the result of applying effection procedure to $\widetilde{X_2} \cap X_1$ where X_1 (resp. $\widetilde{X_2}$) is the intersection of all level 1 (resp. 2) fuzzy sets in χ. Combining these notations, we have the following definition.

Definition 2 *A fuzzy set $V_c \subseteq U$ is the conjunctive possibilistic restriction (cπ-restriction) of Δ_π iff*

$$V_c = \bigcup_{\chi \in \otimes_{i=0}^{m} \delta_i} eff(\chi).$$

Theorem 1 *Let $\Delta = \langle W, \{\frac{\alpha_i : \beta_i}{\beta_i} | 1 \leq i \leq m\}\rangle$ be a normal DT with $W \models \alpha_i$ for $1 \leq i \leq m$ and $E_j (1 \leq j \leq k)$ be all its extensions, then*

$$Mod(\bigcap_{j=1}^{k} E_j) = \bigcup_{j=1}^{k} Mod(E_j)$$

is equal to the cπ-restriction of $Tr(\Delta)$.

Both requirements of the theorem are important. If one of them is not satisfied, the result may not hold.

Example 1 Consider the following semi-normal DT $\Delta = \langle \emptyset, \{\frac{\top : p \wedge \neg q}{p}, \frac{\top : \neg p \wedge q}{q}, \frac{\top : \neg p \vee \neg q}{\neg p \vee \neg q}\}\rangle$, where p and q are propositional variables. Obviously, Δ has two extensions, i.e., $E_1 = Th_L(\{p, \neg q\})$ and $E_2 = Th_L(\{\neg p, q\})$. Let $U = \{u_i | 1 \leq i \leq 4\}$ be the set of all truth assignments on $\{p, q\}$ with the following truth table:

	u_1	u_2	u_3	u_4
p	1	1	0	1
q	1	0	1	0

By some tedious computation, the cπ-restriction of $Tr(\Delta)$ is equal to $\{u_1, u_2, u_3\}$. However, $Mod(E_1) \cup Mod(E_2) = \{u_2, u_3\}$.

Example 2 Consider the normal DT $\Delta = \langle \emptyset, \{\frac{p:q}{q}, \frac{\top:\neg q}{\neg q}\}\rangle$. The unique extension of Δ is $Th(\{\neg q\})$. Let U be defined as in the preceding example. Then the cπ-restriction of $Tr(\Delta)$ is $U \neq Mod(Th(\{\neg q\}))$.

4.2 The sequential application scheme

The scheme is exemplified in [7] by a default theory with two defaults that have no prerequisites. However, it is easily generalized to the present formalism. Let δ_i be defined as above and σ be a permutation on the index set of the defaults, i.e., $\sigma : \{1, 2, \ldots, m\} \rightarrow \{1, 2, \ldots, m\}$ is one-to-one and onto. If V is a fuzzy subset of U, define $V \cap \delta_i = \bigcup\{V \cap G | G \in \delta_i\}$, where we abuse $V \cap G$ to denote the result produced by the effection procedure to $V \cap G$ when G is a level 2 fuzzy set. Let $F_\sigma^0 = F$ and define

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

for $1 \leq i \leq m$, Finally, let $V_\sigma = F_\sigma^m$.

Definition 3 *For each permutation σ on the index set of the defaults, V_σ is called a sequential possibilistic restriction (sπ-restriction) of Δ_π.*

Unlike the cπ-restriction, which is unique for each ΠDT, there may be more than one sπ-restrictions for a general ΠDT, so each sπ-restriction of a ΠDT can be considered as an candidate of reasonable constraints on X. In fact, we have

Theorem 2 *If E is an extension of Δ, then $Mod(E)$ is an sπ-restriction of $Tr(\Delta)$.*

This is in fact a generalization of Theorem 2 of [3] where Δ is restricted to be a normal DT. However, the converse of the theorem does not hold in general.

Example 3 Consider a semi-normal DT $\Delta = \langle \emptyset, \{\frac{\top:p}{p}, \frac{\top:\neg p \wedge q}{q}\}\rangle$. Then the unique extension of Δ is $E = Th_L(\{p\})$. Let σ_1 be the identity permutation and σ_2 be the permutation such that $\sigma_2(1) = 2$ and $\sigma_2(2) = 1$ then $V_{\sigma_1} = Mod(E)$ but V_{σ_2} does not correspond to any extension of Δ.

Example 4 Let $\Delta = \langle \emptyset, \{\frac{\top:p}{p}, \frac{p:q}{q}\}\rangle$ be a normal DT. Then Δ has only one extension, i.e., $E = Th_L(\{p, q\})$. However, by using the same notations as above for σ_1 and σ_2, we have $V_{\sigma_1} = Mod(E)$ but V_{σ_2} does not correspond to any extension of Δ.

These two simple examples show that the defaults can not be applied in any order for a general DT. There are however a special class of DT in which the sπ-restrictions correspond exactly to extensions.

Theorem 3 *Let Δ be a DT as specified in Theorem 1, then each sπ-restriction of $Tr(\Delta)$ is equal to $Mod(E)$ for some extension E of Δ.*

This is a specialization of Theorem 1 of [3] where any normal DT is allowed. Because the defaults can be applied infinitely many times there, it is unnecessary to require that $W \models \alpha_i$ for all $1 \leq i \leq m$.

4.3 The parallel application scheme

This is another scheme exemplified in [7]. The basic principle of the scheme is to consider the applicability of all defaults simultaneously according to the knowledge base F and if the application of some default will block the applicability of other defaults, then none of the defaults are applied. Let us present the scheme in the following way. The ΠDT Δ_π is called *simultaneously satisfiable* iff $\bigcap_{i=1}^m (F \cap \delta_i)$ is normalized[1] and *order-irrelevant* iff for any two permutations σ_1 and σ_2, the sπ-restrictions $V_{\sigma_1} = V_{\sigma_2}$.

Definition 4 *The parallel possibilistic restriction (pπ-restriction) of Δ_π is defined as $V_p \subseteq U$ so that*

1. $V_p = \bigcap_{i=1}^m (F \cap \delta_i)$, *if Δ_π is both simultaneously satisfiable and order-irrelevant.*
2. $V_p = F$, *otherwise.*

A ΠDT is called crisp if the sets F, A_i, B_{ij}, and C_i involving in Eq. (3) are all crisp. There is an interesting relation between sπ-restrictions and the pπ-restriction of a crisp ΠDT.

Theorem 4 *If Δ_π is a simultaneously satisfiable and order-irrelevant crisp* ΠDT, *then its pπ-restriction is equal to its unique sπ-restriction.*

4.4 An auxiliary scheme

Comparing the different application scheme mentioned above, we can find that the most appealing one is the sequential application scheme. This is because it can serve as not only a credulous reasoner but also an agnostic one (See [1] for a comparison of these two types of reasoners in default logic). Since we can find the extensions of a DT from its sπ-restrictions (Theorem 2), we can intersect all of them to form an agnostic reasoner. The only remaining problem is that there may be too many undesirable sπ-restrictions. However, the problem can easily be circumvented by supplying a filter with it. The following fixed point characterization achieves such filtering.

Definition 5 *Let Δ_π be a* ΠDT, *associate a mapping $\Omega : \widetilde{\wp(U)} \to \widetilde{\wp(U)}$ with it so that for each $V \in \widetilde{\wp(U)}$, $\Omega(V)$ is the largest (least specific) fuzzy subset of F such that for each d_i in Eq. (3) and $u \in U$,*

$$\Omega(V)(u) \leq \max(Poss[\overline{A_i}|\Omega(V)], 1 - Poss[B_{i1}|V], \ldots, 1 - Poss[B_{in_i}|V], C_i(u)).$$

A fuzzy set V is called a possibilistic restriction (π-restriction) of Δ_π iff $\Omega(V) = V$.

[1] Recall that a fuzzy set A is called normalized iff $\sup_{u \in U} A(u) = 1$.

From the definition, we can have the following theorem.

Theorem 5 *Let Δ be a DT, then V is a crisp π-restriction of $Tr(\Delta)$ iff $V = Mod(E)$ for some extension E of Δ.*

A careful reader may have found the word "crisp" is used to describe the π-restriction in this theorem. Indeed, even though our ΠDT involve only crisp sets, we may have many fuzzy π-restrictions in addition to the crisp ones. To eliminate these fuzzy π-restrictions, we can use the sequential application scheme. By combining Theorem 2 and the preceding one, we have the following corollary.

Corollary 1 *Let Δ be a DT, then V is both an $s\pi$-restriction and π-restriction of $Tr(\Delta)$ iff $V = Mod(E)$ for some extension E of Δ.*

Generalizing the result to the fuzzy case, the following definition can be given for any ΠDT.

Definition 6 *Let Δ_π be a* ΠDT*, then a fuzzy subset $V \subseteq U$ is called its generated π-restriction iff V is both of its $s\pi$-restriction and π-restriction.*

The definition is less ad hoc than direct use of crisp π-restriction because it is unjustified to require the π-restrictions of a general ΠDT to be crisp but we can indeed require that the π-restrictions should be generated from a sequential application scheme. Though, this does not mean that the π-restrictions not generated from a sequential scheme is of no use. All π-restrictions should be equally valuable and the requirement of generatedness is just for the pragmatic purpose.

5 Conclusion

While Yager's proposal was mainly outlined and exemplified in some literatures, we analyze it in an rigorous way in this paper. Different application schemes are discussed and their implications are investigated when multiple defaults are represented. We present the correspondence theorems between ΠDT and DT when different application schemes are used. The scope within which these theorems are valid is delimited and some examples are used to show that there is much mismatch between ΠDT and DT if it is outside the scope. To avoid the mismatch, the fixed point characterization of default logic is suggested as an auxiliary scheme to filter the unintended restrictions produced by the sequential application scheme.

Acknowledgements

The research is partially supported by National Science Council of ROC under the grant number NSC 82-0113-E-001-015-T. I am grateful to T. F. Fan for her patient and extensive discussion with me on this topic.

References

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

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

[3] W. Łukaszewicz. "Two results on default logic". In *Proccedings of the 9th International Joint Conference on Artificial Intelligence*, pages 459–461, 1985.

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

[5] R. R. Yager. "Possibilistic qualification and default rules". In B. Bouchon and R. R. Yager, editors, *Uncertainty in Knowledge-Based Systems*, LNCS 286, pages 41–57. Springer-Verlag, 1987.

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

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

[8] R. R. Yager. "Nonmonotonic Inheritance Systems". *IEEE Transactions on Systems, Man, and Cybenetics*, 18(6):1028–1034, 1988.

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