Proceedings of the Twenty-Second International FLAIRS Conference (2009)

Prime Implicants and Belief Update

Laurent Perrussel
IRIT - Université Toulouse I
Toulouse - France
laurent.perrussel@univ-tlse1.fr

Jerusa Marchi
DAS - UFSC
Florianópolis - SC - Brazil
jerusa@das.ufsc.br

Guilherme Bittencourt
DAS - UFSC
Florianópolis - SC - Brazil
gb@das.ufsc.br

Adaptation is one of the most important capabilities that an intelligent agent should present. In order to provide this capability to logical-based agents, the *Belief Change* area proposes methods to incorporate new information in a belief set. Properties like consistency and minimal change guide the change process. The notion of minimal change is usually associated with a closeness criterion based on a distance between the models of the belief set and the models of new information (Katsuno and Mendelzon 1991). In (Dalal 1988), Dalal has proposed a notion of distance that considers as the minimal belief unit a propositional symbol. This notion has been widely used by *belief revision* (Dalal 1988) and *belief update* (Forbus 1989; Winslett 1988) methods. Both areas focus on a semantic definition of belief sets.

In this paper we investigate how we can obtain syntactic belief update operators by representing beliefs as *prime implicants*. The belief update is performed over *sets of models* instead of models.

Prime implicants are defined as a special case of DNF form. This form consists of the smallest set of terms closed for inference, without any subsumed terms and without any pair of contradictory literals. We consider that beliefs are expressed with respect to propositional language $\mathcal{L}(P)$, where $P = \{p_1, \ldots, p_n\}$ is a finite set of propositional symbols and $\{L_1, \ldots, L_{2n}\}$ is the set of associated literals. A term is a *conjunction* of literals $D = L_1 \wedge \cdots \wedge L_{k_D}$. In the sequel, terms are seen as sets and $\mathcal{D}$ the set of all possible terms. We define IP_ψ as a disjunction of all prime implicants of ψ such that $\psi \equiv IP_\psi$. Since we consider each prime implicant as representing a belief state, instead of a world, we have to assume that we are able to express preferences over a set of terms. Let $\leqslant$ be a pre-order representing these preferences. Let Γ be a set of terms and $\min(\Gamma, \leqslant)$ be the subset of minimal terms w.r.t. preference relation $\leqslant$ such that $\min(\Gamma, \leqslant) = \{D \in \Gamma \mid \nexists D' \in \Gamma \text{ s.t. } D' \leqslant D \text{ and } D' \nleqslant D\}$. Among all the possible terms we characterize the set of "relevant" terms which will help us to calculate the updated belief set. Those terms are constructed with the help of prime implicants IP_ψ and IP_μ as follows: for every D_ψ and D_μ, a new term is obtained by adding to D_ψ all the literals of D_μ which are not conflicting with the literals to D_ψ. Let Γ be a function that computes this set of relevant terms:

Definition 1 (Γ) *Let IP_ψ and IP_μ be two sets of prime implicants. Let $\Gamma : \mathcal{L}(P) \times \mathcal{L}(P) \mapsto 2^{\mathcal{D}}$ be a function that combines terms as follows:*

$$\Gamma(\psi, \mu) = \{D_\mu \cup (D_\psi - \overline{D_\mu}) | D_\psi \in IP_\psi \text{ and } D_\mu \in IP_\mu\}$$

Terms characterize sets of models and the following condition (**PI-T**) relates preferences over sets of models. That is, we constrain preferences in such a way that, removing some models from a set of models (represented by a term D_u), can not lead to a new set (represented by D_v) that is more preferred. In other words, this constraint will help us to enforce minimal change by avoiding preferences that favor too specific terms. Let $D_u \in \mathcal{D}$ and $\leqslant$ be a pre-order; condition (**PI-T**) holds iff

$$\forall D_v \in \mathcal{D}\ ((D_u \subseteq D_v) \Rightarrow (D_u \leqslant D_v \text{ and } D_v \nleqslant D_u)) \quad \textbf{(PI-T)}$$

Our aim is to state a theorem that describe belief update operation in terms of preferences over terms. First we introduce constraint (**P1-T**) which constrains preferences.

(P1-T) For all $D_u, D_v \in \Gamma(\psi, \mu)$, if $D_u \neq D_v$ then $D_u <_{D_u} D_v$

As shown by constraint **(P1-T)**, preferences are now indexed by terms instead of worlds. In order to accommodate our own notion of change, we slightly reformulate the *KM* postulates (Katsuno and Mendelzon 1991). That is, postulates have to reflect that update should be viewed as applying change to each *prime implicant* of the initial belief set. It follows that our update notion differs from Katsuno-Mendelzon definition. In our context, the minimal object of change is the minimal set of literals that entails the belief set while Katsuno-Mendelzon considers as minimal object of change a model of the belief set. It means that we consider a prime implicant as a minimal unit of interpretation of a belief set rather than a world; in fact we consider as minimal unit the answer to the question "what is the minimal set of literals that is required to entail a belief set?"; i.e., what is the set of relevant literals that entail the belief set. Let ψ and μ be two propositional formulas; $\psi \diamond_{IP} \mu$ denotes the updated belief base. The following postulates characterize operator $\diamond_{IP}$:

(U1-T) $\psi \diamond_{IP} \mu$ implies μ.

(U2-T) If ψ implies μ then $\psi \diamond_{IP} \mu$ is equivalent to ψ.

(U3-T) If both ψ and μ are satisfiable then $\psi \diamond_{IP} \mu$ is also satisfiable.

(U4-T) If $\psi_1 \equiv \psi_2$ and $\mu_1 \equiv \mu_2$ then $\psi_1 \diamond_{IP} \mu_1 \equiv \psi_2 \diamond_{IP} \mu_2$.

(U5-T) $(\psi \diamond_{IP} \mu) \wedge \varphi$ implies $\psi \diamond_{IP} (\mu \wedge \varphi)$.

(U6-T) If $\psi \diamond_{IP} \mu_1$ implies μ_2 and $\psi \diamond_{IP} \mu_2$ implies μ_1 then $\psi \diamond_{IP} \mu_1 \equiv \psi \diamond_{IP} \mu_2$.

(U7-T) If $IP_\psi = \{D_\psi\}$ then $(\psi \diamond_{IP} \mu_1) \wedge (\psi \diamond_{IP} \mu_2)$ implies $\psi \diamond_{IP} (\mu_1 \vee \mu_2)$.

(U8-T) $(\psi_1 \vee \psi_2) \diamond_{IP} \mu \equiv (\psi_1 \diamond_{IP} \mu) \vee (\psi_2 \diamond_{IP} \mu)$.

All postulates are identical to *KM* postulates **(U1)–(U8)** except postulate **(U7-T)**. This postulate rephrases condition "ψ has to be complete" as "ψ has to be represented by only one prime implicant". In other words, this postulate enforces to focus changes on the set of relevant symbols (defined by the implicants). Preferences are only defined over a specific set of terms given by Γ. As mentioned, our unit of interpretation is a prime implicant and thus the definition of the output given by $\diamond_{IP}$ will be obtained by selecting minimal elements w.r.t. to preference relations indexed by terms rather than by worlds. The remaining question is "what is the relevant set of terms that has to be considered for extracting minimal elements?". Since the selection is relative to a prime implicant D_ψ, the relevant set of terms will only be defined w.r.t. D_ψ, that is set $\Gamma(D_\psi, \mu)$. Hence, a key difference between the operator $\diamond$, proposed by Katsuno and Mendelzon, and $\diamond_{IP}$ is that the result given by $\diamond_{IP}$ is a subset of $\Gamma(D_\psi, \mu)$ while the result given by $\diamond$ is a subset of $\mathcal{W}$, i.e. the set of all possible interpretations. Following (Katsuno and Mendelzon 1991), we now show that whenever constraints **(P1-T)** and (**PI-T**) hold, the eight update postulates are satisfied:

Theorem 1 (Update operator) *Let IP_ψ and IP_μ be two sets of prime implicants. Update operator $\diamond_{IP}$ satisfies* **(U1-T)–(U8-T)** *if and only if for all $D_\psi \in IP_\psi$ (i) relation $\leqslant_{D_\psi}$ defined over the subset of terms $\Gamma(D_\psi, \mu)$ is a pre-order, (ii) constraint* **(P1-T)** *holds, (iii) condition* (**PI-T**) *holds w.r.t. $\Gamma(D_\psi, \mu)$ and $\leqslant_{D_\psi}$ and*

$$DNF_{\psi \diamond_{IP} \mu} = \bigcup_{D_\psi \in IP_\psi} \min(\Gamma(D_\psi, \mu), \leqslant_{D_\psi})$$

We omitted the proof due to the lack of space. Even if the two update operators $\diamond$ and $\diamond_{IP}$ are different, they obey almost identical postulates. Thus, if we connect preferences over terms and preferences over models, we are able to establish a link between update operators $\diamond$ and $\diamond_{IP}$. We relate preferences by formulating a constraint denoted (**KPW**). This constraint states that, for each $D_\psi \in IP_\psi$, for all $D_\mu \in IP_\mu$, and for all D_u and $D_w \in \Gamma(\psi, \mu)$, at least one model in which D_u holds is preferred to all models in which D_w holds:

$$(\forall D_u, D_w \in \Gamma(\psi, \mu), \forall D_\psi)\ D_u \leqslant_{D_\psi} D_w \iff \exists u_0 \in [\![D_u]\!], \forall w \in [\![D_w]\!]\ u_0 \leqslant_{D_\psi} w \quad \textbf{(KPW)}$$

Suppose that constraint (**KPW**) holds. It entails that our update operator produces less models than the initial ones since preferences are indexed by prime implicants; i.e., each prime implicant represent a set of worlds. It follows that in terms of strength, we get that operator $\diamond_{IP}$ is *stronger* than the initial operator $\diamond$: an operator $\diamond_1$ is said to be *stronger* than a second operator $\diamond_2$ if the set of models of $\psi \diamond_1 \mu$ is included in the set of models of $\psi \diamond_2 \mu$ (Herzig and Rifi 1999):

Theorem 2 *Let ψ and μ be two formulas. Assume preferences $\leqslant_{D_\psi}$ such that* **(P1-T)** *holds for every $\leqslant_{D_\psi}$ and for all $D \in \Gamma(\psi, \mu)$, condition* (**PI-T**) *holds w.r.t. $\Gamma(\psi, \mu)$ and every $\leqslant_{D_\psi}$; assume preferences $\leqslant_w$ such that for all w, $w' \neq w$, constraint $w <_w w'$ holds for every $\leqslant_w$ and constraint* (**KPW**) *holds. The set of models of $[\![\psi \diamond_{IP} \mu]\!]$ is included in $[\![\psi \diamond \mu]\!]$.*

The remaining question is to set specific conditions, so that both update operators produce equivalent results. First if ψ is inconsistent or μ is inconsistent then the resulting belief sets are inconsistent and thus equivalent. Second, *KM* postulate **(U2)** and postulate **(U2-T)** entail that if $\psi \wedge \mu$ is consistent, then the resulting belief set is consistent and equivalent. Finally, resulting belief sets are also equivalent if ψ is complete.

Proposition 1 *Let ψ and μ be two propositional formulas. Assume update operator $\diamond_{IP}$ s.t.* **(U1-T)–(U8-T)** *hold and update operator $\diamond$ s.t.* **(U1)–(U8)** *hold.*

$$[\![DNF_{\psi \diamond_{IP} \mu}]\!] = [\![\psi \diamond \mu]\!]$$

holds if (i) ψ is inconsistent or μ is inconsistent; or (ii) $\psi \wedge \mu$ is consistent; or (iii) ψ is complete and constraint (**KPW**) *holds.*

The update operator proposed focuses on the set of relevant literals, those that are involved in the change operation. If the belief base is viewed as a set of options, prime implicants represents which are the relevant literals in an option and, in that context, update operator $\diamond_{IP}$ can be viewed as a process of option *revision*. The output is then the union of the revised options.

References

Dalal, M. 1988. Investigations into a theory of knowledge base revision: Preliminary report. In *Proc. of 7th AAAI*, volume 2, 475–479. Menlo Park, California: AAAI Press.

Forbus, K. 1989. Introducing actions into qualitative simulation. In *Proc. of IJCAI'89*, 1273–1278.

Herzig, A., and Rifi, O. 1999. Propositional belief base update and minimal change. *Artificial Intelligence* 115(1):107–138.

Katsuno, H., and Mendelzon, A. 1991. On the difference between updating a knowledge base and revising it. In *Proc. of KR'91*, Morgan Kaufmann. 387–394.

Winslett, M. 1988. Reasoning about action using a possible models approach. In *Proc. of 7th AAAI*, 89–93. Menlo Park, California: AAAI Press.