

Fusion par R-ensembles : une méthode complète de fusion

Julien Hué¹, Odile Papini², Eric Würbel¹

¹ LSIS UMR CNRS 6168 Université du Sud-Toulon Var, Av. de l'université
BP20132, 83957 La Garde Cedex {hue,wurbel}@univ-tln.fr

² LSIS UMR CNRS 6168 Université de la méditerranée, 163 avenue de Luminy,
13288 Marseille Cedex 9. odile.papini@esil.univmed.fr

Résumé : La fusion d'informations provenant de sources multiples est un problème important en intelligence artificielle. La plupart des propositions ont été faites dans le cadre sémantique qui présente une très grande complexité. De plus, peu d'implantations sont disponibles. Ce papier propose un cadre pour réaliser la fusion de bases de croyances propositionnelles. Puis, il décrit une implantation basée sur la programmation logique avec sémantique des modèles stables qui peut prendre en compte n'importe quelle formule bien formée et s'exécuter indépendamment du solveur. Finalement, il présente une étude expérimentale. ¹

1 Introduction

La disponibilité des sources d'informations distribuées ne cesse d'augmenter. Mais celles-ci manquent souvent de structure, elles peuvent être conflictuelles ou redondantes, ce qui rend leur exploitation difficile. La fusion a pour objectif d'obtenir un point de vue global, exploitant la complémentarité des sources, résolvant les conflits existants et supprimant les répétitions. Parmi les nombreuses approches de fusion, les approches logiques ont obtenu un intérêt croissant, spécifiquement dans le cadre propositionnel (Cholvy, 1998), (Revesz, 1997), (Baral *et al.*, 1991). Plusieurs postulats permettant de caractériser le comportement d'un opérateur de fusion rationnel ont été proposés ainsi que de nombreux opérateurs (Konieczny & PinoPérez, 1999). Plus récemment, des approches sémantiques se basant sur la distance de Hamming (Konieczny, 2000) ou sur des propriétés morphologiques comme la dilatation ou l'érosion ont été définies (Bloch & Lang, 2002). Une implantation basée sur les BDD a récemment été proposée et une expérimentation sur des opérateurs sémantiques a été conduite (Gorogiannis & Hunter, 2008).

¹Ce travail a été réalisé avec le soutien de la Communauté européenne au sein du projet VENUS (IST-034924) du 6ème programme cadre pour la recherche et le développement (FP6) (Société, Information et technologie)(IST). Les auteurs sont seuls responsables du contenu de ce papier qui ne reflète pas l'opinion de la Communauté européenne. De plus, la Communauté européenne n'est pas responsable de l'utilisation des données apparaissant dans l'article.

Dans (Meyer *et al.*, 2001), des approches syntaxiques ont été proposées pour la logique propositionnelle. Nous avons récemment proposé une approche syntaxique appelée fusion par R-ensembles (ou RSF) pour la fusion de bases de croyances propositionnelles (Hue *et al.*, 2007). Suite à une expérimentation précédente (Bennaïm *et al.*, 2004), nous avons choisi de mettre en œuvre cette approche dans le cadre de la programmation logique avec sémantique des modèles stables (ou ASP) et l’algorithme Smodels a été adapté afin de fournir une implantation.

Nous proposons ici de généraliser RSF afin que la méthode puisse gérer n’importe quelle formule bien formée. Nous montrons que les opérateurs classiques de la fusion peuvent être capturés. Nous proposons également une implantation indépendante du solveur. Ainsi, notre principale contribution est la suivante :

- La notion de R-ensembles, pour simplifier les sous-ensembles de clauses causant l’incohérence, est étendue aux sous-ensembles de formules. Les opérateurs de fusion classiques sont capturés en les traduisant en préférences entre R-ensembles.
- Une implantation de RSF en programmation logique avec sémantique des modèles stables (ASP). Le problème de fusion est traduit en ASP et l’équivalence entre les modèles stables préférés du programme et les R-ensembles est démontrée.
- Une étude expérimentale qui illustre le comportement de RSF pour les stratégies Σ et *Max* est menée et compare RSF à la méthode de Gorogiannis et Hunter.

Le papier s’articule comme suit. La section 2 fixe les notations. La section 3 présente la généralisation de RSF à toute formule bien formée. L’implantation de la méthode en ASP est décrite dans la section 4 ainsi que les propriétés essentielles. La section 5 montre comment RSF peut être mise en œuvre indépendamment du solveur. Enfin, les résultats de l’étude expérimentale sont présentés dans la section 6.

2 Préliminaires

2.1 Notations

Nous considérons un langage \mathcal{L} sur un alphabet fini \mathcal{P} constitué d’atomes propositionnels. Une interprétation est une fonction de \mathcal{P} vers $\{0, 1\}$ et l’ensemble de toutes les interprétations est noté \mathcal{W} . Pour toute interprétation I et toute formule A , on dit que I implique A ($I \models A$) ssi A est vraie pour I . Si A et B sont deux formules, alors $A \models B$ ssi $I \models A$ implique $I \models B$. $mod(A)$ représente l’ensemble des modèles de A .

Une base de croyances K est un ensemble fini de formules propositionnelles. Elle peut être vue comme une formule unique qui est la conjonction de ses formules. Soient K_1, \dots, K_n n bases de croyances qui ne sont pas forcément différentes deux à deux. Nous appelons *profil de croyances* le multi-ensemble E constitué de tous les K_i ($1 \leq i \leq n$). L’ensemble des atomes apparaissant dans E est noté $Atome(E)$. L’union sur les multi-ensemble est représentée par \sqcup et \sqsubseteq dénote l’inclusion multi-ensembliste. On note E^n le multi-ensemble constitué de n répétitions de E . De plus, la conjonction (resp. disjonction) est notée \bigwedge (resp. \bigvee). Pour simplifier, nous notons K le profil de croyances $E = \{K\}$. Pour le reste de cette section, nous supposons que $\forall i$, la base de croyances K_i est cohérente. Nous considérons qu’un profil de croyances $E = \{K_1, \dots, K_n\}$ est cohérent ssi il existe au moins une interprétation I telle que $\forall i, I \models K_i$. Soient

$E_1 = \{K_{1,1}, \dots, K_{1,n_1}\}$ et $E_2 = \{K_{2,1}, \dots, K_{2,n_2}\}$ deux profils de croyances, E_1 et E_2 sont dits équivalents, noté $E_1 \leftrightarrow E_2$, ssi il existe une bijection f de E_1 vers E_2 telle que $\models f(K_{1,i}) \leftrightarrow K_{2,j}$.

Un pré-ordre sur \mathcal{W} est une relation transitive et réflexive sur les éléments de \mathcal{W} . Un pré-ordre est total ssi $\forall I, \forall J$ on a soit $I \leq J$, soit $J \leq I$. On définit $<$ comme suit : $I < J$ ssi $I \leq J$ et $J \not\leq I$. On définit également $=$ par : $I = J$ ssi $I \leq J$ et $J \leq I$. Une interprétation I appartient aux modèles minimaux de K par rapport à \leq (noté $I \in \min(\text{mod}(K), \leq)$) ssi $I \models K$ et $\forall J \in \text{mod}(K), I \leq J$.

Une opération de fusion syntaxique Δ est définie comme une fonction qui associe à tout profil de croyances E une base de croyances cohérente notée $\Delta(E)$.

2.2 La programmation logique avec sémantique des modèles stables

Un *programme logique normal* est un ensemble de règles de la forme $c \leftarrow a_1, \dots, a_n, \text{not } b_1, \dots, \text{not } b_m$ où $c, a_i (1 \leq i \leq n), b_j (1 \leq j \leq m)$ sont des atomes de la logique propositionnelle et le symbole *not* représente la négation par échec. Soit r une règle, on introduit $\text{head}(r) = c$ et $\text{body}(r) = \{a_1, \dots, a_n, b_1, \dots, b_m\}$. De plus, $\text{body}^+(r) = \{a_1, \dots, a_n\}$ représente l'ensemble des éléments positifs du corps de la règle et $\text{body}^-(r) = \{b_1, \dots, b_m\}$ l'ensemble de ses éléments négatifs. r^+ représente la règle $\text{head}(r) \leftarrow \text{body}^+(r)$ obtenue à partir de r . Un *programme basique* est un programme logique ne comportant pas de négation par échec.

Un ensemble d'atomes X est *clos sous* un programme basique Π ssi pour n'importe quelle règle $r \in \Pi$, $\text{head}(r) \in X$ lorsque $\text{body}(r) \subseteq X$. Le plus petit ensemble d'atomes qui est clos sous un programme basique Π est noté $CN(\Pi)$. La *réduction* (Gelfond & Lifschitz, 1988), Π^X d'un programme Π relativement à un ensemble X d'atomes est définie par $\Pi^X = \{r^+ \mid r \in \Pi \text{ et } \text{body}^-(r) \cap X = \emptyset\}$. Un ensemble d'atomes X est un *modèle stable* de Π ssi $CN(\Pi^X) = X$.

Ces dix dernières années, la programmation logique avec sémantique des modèles stables (ASP) a été considérée comme étant un outil pratique pour le raisonnement non-monotone et certains solveurs ASP se sont montrés efficaces : Smodels (Niemelä & Simons, 1997), NoMore (Anger *et al.*, 2002), CLASP (Baral *et al.*, 2007).

Afin d'améliorer son expressivité, l'ASP a été étendue avec de nouvelles règles et instructions : les *définitions de domaines* permettent de coder de manière compacte les valeurs possibles d'un domaine. Par exemple, les règles $\#domain\ possible(X). possible(1 \dots n)$ remplacent chaque apparition de la variable X par n règles où X est remplacé par l'ensemble des valeurs de 1 à n . Les *restrictions de domaines* : par exemple, la règle $short(X) \leftarrow size(Y), X < Y$ spécifie que la règle ne peut s'appliquer (et donc n'est instancié) que pour les valeurs de X et Y tel que $X < Y$. Les *contraintes de cardinalité* permettent d'exprimer qu'au moins (resp. au plus) un certain nombre d'atomes d'un ensemble doivent apparaître. Par exemple, $h \leftarrow k\{a_1, \dots, a_n\}l$ exprime qu'au moins k atomes et au plus l atomes parmi les $a_i (1 \leq i \leq n)$ doivent être vrais pour que h soit vrai. Les *instructions d'optimisation* permettent de sélectionner parmi les modèles stables seulement ceux qui contiennent le moins possible ou le plus possible d'atomes parmi un ensemble donné. Par exemple, l'instruction $minimize\{a_1, \dots, a_n\}$

permet de ne sélectionner parmi les modèles stables que ceux qui contiennent un nombre minimaux d'atomes parmi les $a_{i(1 \leq i \leq n)}$. L'inverse s'effectue au travers de l'instruction $maximize\{\}$.

3 La fusion par R-ensembles

L'idée principale consiste à identifier, puis retirer, un sous-ensemble adéquat de formules afin de restaurer la cohérence dans l'union des bases de croyances. Dans un premier temps, l'ensemble des sous-ensembles de formules est considéré avant d'y effectuer une sélection grâce à l'une des stratégies que nous décrirons *infra*.

Définition 1

Soit $E = \{K_1, \dots, K_n\}$ un profil de croyances tel que $K_1 \sqcup \dots \sqcup K_n$ est incohérent. Soit X un sous-ensemble de formules de $K_1 \sqcup \dots \sqcup K_n$. X est un R-ensemble potentiel de E ssi $(K_1 \sqcup \dots \sqcup K_n) \setminus X$ est cohérent.

Le nombre de R-ensembles potentiels est exponentiel par rapport au nombre de formules. Il faut donc faire un choix parmi ceux-ci et ne garder que les plus pertinents. Cependant la notion de minimalité dans le cadre de la fusion n'est pas unique et plusieurs stratégies existent qui permettent d'obtenir de bonnes solutions selon le critère de minimalité choisi. Soient X et Y deux R-ensembles potentiels. Pour chaque stratégie P , un pré-ordre total \leq_P sur les R-ensembles potentiels est défini. $X \leq_P Y$ signifie que X est préféré à Y selon la stratégie P . On définit $<_P$ comme étant le pré-ordre strict associé à \leq_P (i.e. $X <_P Y$ ssi $X \leq_P Y$ et $Y \not\leq_P X$).

Définition 2

Soit $E = \{K_1, \dots, K_n\}$ un profil de croyances tel que $K_1 \sqcup \dots \sqcup K_n$ est incohérent. Soit P une stratégie de fusion. $X \subseteq K_1 \sqcup \dots \sqcup K_n$ est un R-ensemble de E selon P ssi (1) X est un R-ensemble potentiel de E ; (2) Il n'existe pas de Y qui soit un R-ensemble potentiel de E tel que $Y <_P X$.

La collection des R-ensembles de E selon la stratégie P est notée par $\mathcal{F}_P \mathcal{R}(E)$. La fusion par R-ensembles est définie par :

Définition 3

Soit $E = \{K_1, \dots, K_n\}$ un profil de croyances. L'opération de fusion $\Delta^P(E)$ est définie par : $\Delta^P(E) = \bigvee_{X \in \mathcal{F}_P \mathcal{R}(E)} \{(K_1 \sqcup \dots \sqcup K_n) \setminus X\}$

L'opportunité de l'union des bases pourrait être discutée. Ce choix peut permettre de définir, dans un second temps, une sélection (partial meet, full meet, ...) sur l'union. Les opérateurs de fusion classiques (Σ , Max , $Gmax$) sont capturés dans notre cadre.

3.1 Différents opérateurs capturés par RSF

Dans un premier temps, nous allons donner les définitions formelles des pré-ordres associés à ces opérateurs avant, dans un second temps, de donner un exemple permettant d'illustrer leur comportement.

Définition 4

Soient X et Y deux R -ensembles potentiels de E . Les pré-ordres selon les différentes stratégies sont :

$$\Sigma : X \leq_{\Sigma} Y \text{ ssi } \sum_{1 \leq i \leq n} |X \cap K_i| \leq \sum_{1 \leq i \leq n} |Y \cap K_i|.$$

$$Card : X \leq_{Card} Y \text{ ssi } |X \cap s(E)| \leq |Y \cap s(E)|.^2$$

$$Max : X \leq_{Max} Y \text{ ssi } \max_{1 \leq i \leq n} |X \cap K_i| \leq \max_{1 \leq i \leq n} |Y \cap K_i|.$$

Gmax : Pour tout R -ensemble potentiel X et chaque base de croyances K_i , on définit $p_X^{K_i} = |X \cap K_i|$. Soit L_X^E la séquence $(p_X^{K_1}, \dots, p_X^{K_n})$ rangée dans l'ordre décroissant. $X \leq_{Gmax} Y$ ssi $L_X^E \leq_{lex} L_Y^E$.

La stratégie Σ minimise le nombre de formules à retirer de E . Elle correspond à l'opérateur *intersection* défini dans (Konieczny & PinoPérez, 1999). La stratégie *Card* tente, tout comme Σ , de minimiser le nombre de formules retirées. Cependant, elle ne tient pas compte des formules exprimées plusieurs fois. La stratégie *Max* essaye de répartir au mieux les formules à retirer parmi les bases de croyances. Elle tente cela en retirant le moins possible de formules dans la base de croyances où elle en retire le plus. La stratégie *Gmax* est un raffinement lexicographique de la stratégie *Max*.

Exemple 1

On considère le profil de croyances suivant $E = \{K_1, K_2, K_3\}$ avec $K_1 = \{-d, s, s \vee o\}$, $K_2 = \{\neg s, d \vee o, \neg(d \wedge o)\}$ et $K_3 = \{s, d, o\}$. Les R -ensembles préférés de E stratégie par stratégie, sont :

\mathcal{W}	K_1	K_2	K_3	Σ	<i>Card</i>	<i>Max</i>	<i>Gmax</i>
(000)	$s \vee o, s$	$d \vee o$	s, d, o	6	5	3	321
(001)	s		s, d	3	2	2	210
(010)	$s \vee o, \neg d, s$		s, o	5	4	3	320
(011)	$\neg d, s$	$\neg(d \wedge o)$	s	4	3	2	211
(100)		$d \vee o, \neg s$	d, o	4	4	2	220
(101)		$\neg s$	d	2	2	1	110
(110)	$\neg d$	$\neg s$	o	3	3	1	111
(111)	$\neg d$	$\neg(d \wedge o), \neg s$		3	3	2	210

Ainsi, les R -ensembles et le résultat de la fusion des bases de E , pour chaque stratégie, seront :

$$\Sigma : \mathcal{F}_{\Sigma} \mathcal{R}(E) = \{\{\neg s, d\}\}; \Delta^{\Sigma}(E) = \{\{\neg d, s \vee o, s, d \vee o, \neg(d \wedge o), o\}\}.$$

$$Card : \mathcal{F}_{Card} \mathcal{R}(E) = \{\{\neg s, d\}, \{s, s, d\}\}; \Delta^{Card}(E) = \{\{\neg d, s \vee o, s, d \vee o, \neg(d \wedge o), o\}, \{\neg d, s \vee o, \neg s, d \vee o, \neg(d \wedge o), o\}\}.$$

$$Max : \mathcal{F}_{Max} \mathcal{R}(E) = \{\{\neg d, \neg s, o\}, \{\neg s, d\}\}; \Delta^{Max}(E) = \{\{s \vee o, s, d \vee o, \neg(d \wedge o), d\}, \{\neg d, s \vee o, s, d \vee o, \neg(d \wedge o), o\}\}.$$

$$Gmax : \mathcal{F}_{Gmax} \mathcal{R}(E) = \{\{\neg s, d\}\}; \Delta^{Gmax}(E) = \{\{\neg d, s \vee o, s, d \vee o, \neg(d \wedge o), o\}\}.$$

Nous devons maintenant comment effectuer un calcul efficace des R -ensembles.

4 Mise en œuvre de la méthode RSF

Cette section présente une mise en œuvre de l'opération de fusion syntaxique. Soit $E = \{K_1, \dots, K_n\}$ un profil de croyances et P une stratégie de fusion, ce problème est traduit en un programme logique Π_E^P dont les solutions permettent de déterminer

²Soit E un profil de croyances, on note $s(E)$ la base de croyances créée à partir de E où toutes les formules apparaissant plusieurs fois sont réduites à un singleton.

les R-ensembles. Dans un premier temps, nous construisons un programme Π_E dont les solutions correspondent aux R-ensembles potentiels. Puis, nous introduisons des atomes représentant chaque règle permettant de comparer les R-ensembles potentiels. Enfin, nous ajoutons des instructions permettant de sélectionner les R-ensembles.

4.1 Générer les R-ensembles potentiels

Soient f, f^1, \dots, f^n des formules de K_i . L'ensemble des littéraux positifs (resp. négatifs) de Π_E est noté V^+ (resp. V^-). Soit $f \in K_i$ une formule, on note r_f^i l'atome de règle qui correspond à f . L'ensemble de tous les atomes représentant les formules, appelés atomes de règle, est défini par $R^+ = \{r_f^i \mid f \in K_i\}$. La traduction du problème nécessite la création d'atomes intermédiaires représentant les sous-formules de f . On note ρ_f^j l'atome intermédiaire représentant f^j qui est une sous-formule de $f \in K_i$. De plus, on note $F_O(r_f^i)$ la formule de K_i correspondant à r_f^i dans Π_E . En d'autres termes $\forall r_f^i \in R^+, F_O(r_f^i) = f$. La création de Π_E se déroule en deux étapes :

(i) Pour tout atome $a \in V^+$, la première étape introduit les règles $a \leftarrow \text{not } a'$ et $a' \leftarrow \text{not } a$ où a' représente la négation de a . On construit ainsi une correspondance entre les interprétations de V^+ et les modèles stables de Π_E . (ii) La seconde étape introduit les atomes de règle qui permettent de déterminer les R-ensembles potentiels associés aux interprétations. La présence de l'atome de règle r_f^i dans le modèle stable S signifie que f est dans le R-ensemble potentiel correspondant. Quelle que soit la formule f , les règles suivantes sont introduites selon la syntaxe de f : (1) Si $f \equiv a$, la règle correspondante est $r_f^i \leftarrow \text{not } a$; (2) Si $f \equiv \neg f^1$, la règle correspondante est $r_f^i \leftarrow \text{not } \rho_{f^1}$; (3) Si $f \equiv f^1 \vee \dots \vee f^m$, la règle correspondante est $r_f^i \leftarrow \rho_{f^1}, \dots, \rho_{f^m}$; (4) Si $f \equiv f^1 \wedge \dots \wedge f^m$, il est alors nécessaire de rajouter plusieurs règles au programme. Les règles correspondantes sont $\forall 1 \leq j \leq m, r_f^i \leftarrow \rho_{f^j}$.

Exemple 2

Soit $E = \{K_1, K_2, K_3\}$ le profil de croyances constitué des bases $K_1 = \{a \vee b, b\}$, $K_2 = \{a \leftrightarrow b, b\}$ et $K_3 = \{\neg a \wedge \neg b, \neg a \vee \neg b\}$. Le programme logique correspondant à E, Π_E , est le suivant :

$\Pi_E = \{a \leftarrow \text{not } a', a' \leftarrow \text{not } a, b \leftarrow \text{not } b', b' \leftarrow \text{not } b, r_{a \vee b}^1 \leftarrow \text{not } a, \text{not } b, r_b^1 \leftarrow \text{not } b, r_{a \leftrightarrow b}^2 \leftarrow \rho_a^2, r_{a \leftrightarrow b}^2 \leftarrow \rho_b^2, \rho_a^2 \leftarrow a, \text{not } b, \rho_b^2 \leftarrow \text{not } a, b, r_b^2 \leftarrow \text{not } b, r_{\neg a \wedge \neg b}^3 \leftarrow \rho_{\neg a}^3, r_{\neg a \wedge \neg b}^3 \leftarrow \rho_{\neg b}^3, \rho_{\neg a}^3 \leftarrow a, \rho_{\neg b}^3 \leftarrow b, r_{\neg a \vee \neg b}^3 \leftarrow a, b.\}$

4.2 Sélectionner les R-ensembles

4.2.1 Sélection off the shelf

Afin de choisir parmi ces modèles stables ceux qui correspondent aux R-ensembles selon la stratégie P , nous introduisons la notion de modèles stables préférés selon P . On note I_X comme étant : $I_X = \{a \mid a \in X\} \cup \{\neg a \mid a' \in X\}$.

Définition 5

Soient Π_E un programme logique et X et Y deux ensembles d'atomes de Π_E . X est un modèle stable préféré de Π_E selon la stratégie P ssi (1) X est un modèle stable de Π_E ; (2) $\forall Y$ modèle stable de Π_E , Y n'est pas préféré à X selon P .

Proposition 1

Soit ρ un atome de règle ou un atome intermédiaire. $\rho \in CN(\Pi_E^X)$ ssi $I_X \not\models F_O(\rho)$.

Proposition 2

Soit $E = \{K_1, \dots, K_n\}$ un profil de croyances. Soit X un ensemble d'atomes. X est un modèle stable de Π_E ssi X correspond à une interprétation I_X de V^+ qui satisfait $(K_1 \sqcup \dots \sqcup K_n) \setminus F_O(X \cap R^+)$.

Les résultats suivants établissent une correspondance entre les modèles stables préférés de Π_E et les R-ensembles de E selon la stratégie employée.

Proposition 3

Soit $X \subseteq (K_1 \sqcup \dots \sqcup K_n)$ tel que $(K_1 \sqcup \dots \sqcup K_n) \setminus X$ est cohérent. Alors, il existe un ensemble d'atomes S tel que S est un modèle stable de Π_E et $F_O(S \cap R^+) \subseteq X$

Proposition 4

Soit $X \subseteq (K_1 \sqcup \dots \sqcup K_n)$ un ensemble d'atomes. X est un R-ensemble de E selon Σ , $Card$ et Max . ssi il existe un modèle stable préféré de Π_E tel que $F_O(S \cap R^+) = X$.

Nous détaillons maintenant l'implantation qui nous permet de sélectionner les R-ensembles.

4.2.2 L'implantation

Une méthode pour réaliser la fusion par R-ensembles (RSF) à l'aide de n'importe quel solveur ASP consiste à rajouter au programme Π_E , des règles et des instructions dont le but est de compter les atomes de règle présents dans chaque modèle stable.

Pour la stratégie Σ , le principe est relativement simple puisqu'il suffit de compter le nombre total d'atomes de règle et de sélectionner les R-ensembles en contenant le moins. Cette sélection est réalisée par l'intermédiaire de l'instruction *minimize*. Ainsi, $\Pi_E^\Sigma = minimize \{r_f^i \mid r_f^i \in R^+\} \cup \Pi_E$. La méthode employée pour la stratégie $Card$ est similaire au détail près que les doublons sont préalablement enlevés du profil. Ainsi $\Pi_E^{Card} = minimize \{r_f^i \mid r_f^i \in R^+\} \cup \Pi_{s(E)}$ La proposition suivante est vérifiée :

Proposition 5

L'ensemble de modèles stables de Π_E^Σ (resp. Π_E^{Card}) est l'ensemble des modèles stables préférés de Π_E selon la stratégie Σ (resp. $Card$).

Exemple 3

Afin d'illustrer les définitions, nous reprenons l'exemple 2 et, stratégie par stratégie, montrons la traduction du problème. Pour la stratégie Σ (resp. $Card$) $\Pi_E^\Sigma = \Pi_E \cup$ (resp. $\Pi_{s(E)} \cup$) $minimize\{r_b^1, r_{a \leftrightarrow b}^2, r_b^2, r_{\neg a \wedge \neg b}^3, r_{\neg a \vee \neg b}^3\}$. La formule b qui était présente deux fois dans E n'apparaît qu'une seule fois dans $s(E)$.

La stratégie Max requiert une autre utilisation de l'instruction *minimize*. Soit X un modèle stable. Une première étape calcule $max_{1 \leq i \leq n} (|\{r_f^i \mid r_f^i \in X\}|)$ pour chaque X . La seconde étape utilise ce calcul et l'instruction *minimize* pour conserver les modèles stables minimaux au sens de Max . Pour savoir combien de formules sont retirées dans la base où il y a le plus de retraits, il est nécessaire de commencer par

calculer combien de formules sont retirées dans chaque base de croyances. Ce calcul est effectué par :

$$\Pi_E^{max,size} = \{\delta_0 : \#domain\ possible(U). \delta_1 : \#domain\ base(V). \delta_2 : possible(1..m). \delta_3 : base(1..n). \alpha : size(U) \leftarrow U\{r_f^V | F_0(f) \in \varphi_V\}U.\}$$

où m est la taille de la plus grande base de croyances. Les règles δ_i sont les domaines de définition pour la règle α . S'il existe une base φ_V où le nombre d'atomes r_f^V vrais dans X est égal à U , alors $size(U)$ sera vrai dans X . La plus grande valeur U pour laquelle $size(U)$ est vrai est calculée grâce à $\Pi_E^{max,bound}$:

$$\Pi_E^{max,bound} = \{\delta_4 : \#domain\ possible(W). \beta_1 : negmax(W) \leftarrow size(U), U > W. \beta_2 : max(U) \leftarrow size(U), not\ negmax(U). \}$$

β_1 détermine tous les entiers W pour lesquelles il existe $U > W$ tel que $size(U)$ est vrai. Alors, $max(U)$ est vrai pour la plus grande valeur de U telle que $size(U)$ est vrai. Pour la stratégie Max , soit $\Pi_E^{Max} = \Pi_E^{max,size} \cup \Pi_E^{max,bound} \cup \Pi_E \cup minimize[max(1) = 1, \dots, max(m) = m]$. La proposition suivante est vérifiée :

Proposition 6

L'ensemble des R-ensembles de Π_Ψ^{Max} est l'ensemble des modèles stables préférés de Π_E selon Max .

Exemple 4

$$\Pi_E^{Max} = \Pi_E \cup \{\#domain\ possible(U). negmax(W) \leftarrow size(U), U > W. \#domain\ base(V). size(X) \leftarrow U\{r_b^1\}U. \#domain\ possible(W). size(U) \leftarrow U\{r_{a \leftrightarrow b}^2, r_b^2\}U. base(1..3). size(U) \leftarrow U\{r_{-a \wedge -b}^3, r_{-a \vee -b}^3\}U. possible(1..2). max(U) \leftarrow size(U), not\ negmax(U). minimize[max(1) = 1, max(2) = 2]\}$$

La stratégie $Gmax$ compare les R-ensembles potentiels d'après la séquence des $|X \cap K_i|$ rangée dans l'ordre décroissant. Comme pour Max , il faut connaître le nombre de formules retirées dans chacune des bases. Ce calcul est représenté par la règle suivante où m représente la taille de la base φ_V .

Si $size(V, U)$ est vrai, dans la base φ_V , le modèle stable en cours contient U atomes de règle. Une fois calculées pour toutes les bases K_V , ces valeurs sont ordonnées par $\Pi_E^{gmax,bound}$ où X_1 est la plus grande valeur de $size()$ et X_n la plus petite. Finalement, le polynôme $X_n + X_{n-1} \times m + \dots + X_1 \times m^{n-1}$ est optimisé.

$$\Pi_E^{gmax,size} = \{\gamma_1 : size(V, U) \leftarrow U\{rem(V, 1), \dots, rem(V, m)\}U.\} \quad \Pi_E^{gmax,bound} = \{\alpha_0 : max(X_1, \dots, X_n) \leftarrow size(Y_1, X_1), \dots, size(Y_n, X_n), X_1 \geq X_2, \dots, X_{n-1} \geq X_n, neq(Y_1, \dots, Y_n). \alpha_1 : max_1(X_1) \leftarrow max(X_1, \dots, X_n), X_1 \geq X_2, \dots, X_{n-1} \geq X_n. \alpha_i : \dots \alpha_n : max_3(X_3) \leftarrow max(X_1, \dots, X_n), X_1 \geq X_2, \dots, X_{n-1} \geq X_n.\}$$

Pour la stratégie $Gmax$, soit $\Pi_E^{Gmax} = \Pi_E^{gmax,size} \cup \Pi_E^{gmax,bound} \cup \Pi_E \cup minimize[max_n(1) = 1, max_n(2) = 2, \dots, max_i(1) = m^{n-i}, max_i(2) = 2 \times m^{n-i}, \dots, max_1(n) = n \times m^{n-1}]$. La proposition suivante est vérifiée :

Proposition 7

L'ensemble des R-ensembles de Π_E^{Gmax} est l'ensemble des modèles stables préférés de Π_E pour la stratégie $Gmax$.

Exemple 5

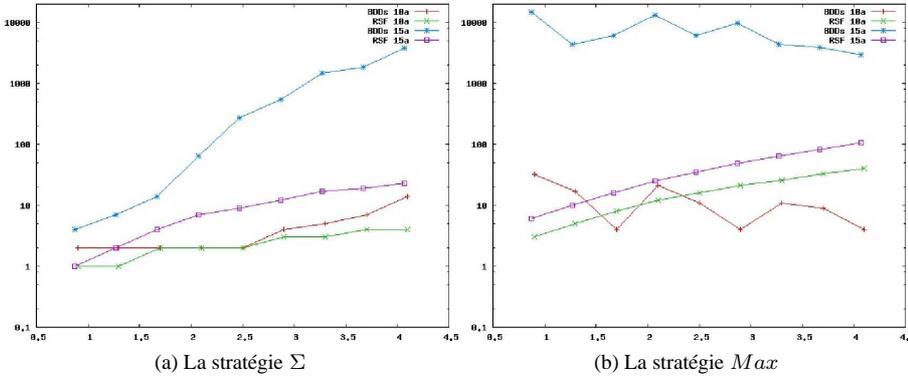
$$\Pi_E^{Gmax} = \Pi_E \cup \{\#domain\ possible(U). possible(1..2). \#domain\ base(V). base(1..3). \#domain\ possible(X_1). size(1, U) \leftarrow U\{r_b^1\}U. \#domain\ possible(X_2). size(2, U) \leftarrow$$

$U \{r_{a \leftrightarrow b}^2, r_b^2\} U$. $\#domain\ possible(X_3)$. $size(3, U) \leftarrow U \{r_{\neg a \wedge \neg b}^3, r_{\neg a \vee \neg b}^3\} U$.
 $\#domain\ base(Y_1)$. $max_1(X_1) \leftarrow max(X_1, X_2, X_3)$. $\#domain\ base(Y_2)$. $max_2(X_2) \leftarrow$
 $max(X_1, X_2, X_3)$. $\#domain\ base(Y_3)$. $max_3(X_3) \leftarrow max(X_1, X_2, X_3)$. $max(X_1, X_2, X_3) \leftarrow$
 $size(Y_1, X_1), size(Y_2, X_2), size(Y_3, X_3), Y_1! = Y_2, Y_1! = Y_3, Y_2! = Y_3, X_1 >= X_2, X_2 >= X_3.$

5 Etude expérimentale et comparaison

Nous présentons maintenant les résultats de l'étude expérimentale pour les stratégies *Max* et Σ et nous comparons les approches implantées en ASP avec les approches sémantiques avec BDD présentées dans (Gorogiannis & Hunter, 2008). Les tests ont été menés sur un Centrino Duo cadencé à 1.73GHz et équipé de 2Go de RAM.

FIG. 1: Tests pour les stratégies *Max* et Σ



Nous utilisons un protocole proche de celui décrit dans (Gorogiannis & Hunter, 2008). Chaque profil de croyances possède 7 bases de croyances. Les bases de croyances sont générées aléatoirement, selon le nombre d'atomes (na) et le nombre de formules (nf) souhaitées. Les bases contiennent des formules de tout type dont la profondeur ne dépasse pas 3. Les tests ont été effectués pour les stratégies *Max* et Σ .

La figure 1 montre le temps d'exécution en fonction du ratio $(nf)/(na)$. Les temps d'exécution sont donnés en centièmes de seconde. Les courbes ont été tracées pour 10 et 15 atomes pour les deux stratégies. Sur les deux stratégies, les temps d'exécution pour les BDD sont très instables, ce qui explique la forme du graphe.

Ce qui ressort de cette étude est que RSF obtient de meilleurs résultats lorsque le nombre d'atomes augmente. D'un autre côté, les BDD obtiennent de meilleurs résultats lorsque le nombre de formules augmente. Néanmoins, la différence fondamentale entre ces deux approches réduit la valeur de cette comparaison. Malgré cette différence, cette étude a le mérite de pointer les difficultés propres aux deux approches.

6 Conclusion

Ce papier généralise l'approche RSF aux formules quelconques et montre que les opérateurs classiques de fusion peuvent être exprimés dans ce cadre. Ce papier propose une implantation en ASP indépendante du solveur. Enfin, une expérimentation a été conduite et les résultats ont été comparés avec l'approche (Gorogiannis & Hunter, 2008).

Une expérimentation plus large doit être menée sur une application réelle afin de donner une meilleure idée des performances de RSF. Cette expérimentation sera faite dans le cadre du projet européen VENUS dans le contexte de l'information archéologique. Un travail futur détaillera la caractérisation sémantique de RSF.

Références

- ANGER C., KONCZAK K. & LINKE T. (2002). Nomore non-monotonic reasoning with logic programs. In *JELIA*, p. 521–524.
- C. BARAL, G. BREWKA & J. SCHLIPF, Eds. (2007). *Proc. of the Ninth International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR07)*, volume 4483 of *Lecture Notes in Artificial Intelligence*.
- BARAL C., KRAUS S., MINKER J. & SUBRAHMANIAN V. S. (1991). Combining knowledge bases consisting of first order theories. In *ISMIS*, p. 92–101.
- BENNAIM J., BENFERHAT S., PAPINI O. & WÜRBEL E. (2004). An answer set programming encoding of prioritized removed sets revision : application to gis. In *JELIA*, p. 604–616.
- BLOCH I. & LANG J. (2002). Towards mathematical morpho-logics. p. 367–380.
- CHOLVY L. (1998). Reasoning about merging information. *Handbook of DRUMS*, **3**, 233–263.
- GELFOND M. & LIFSCHITZ V. (1988). The stable model semantics for logic programming. In *Proc. of the 5th Int. Conf. on Log. Prog.*, p. 1070–1080.
- GOROGIANNIS N. & HUNTER A. (2008). Implementing semantic merging operators using binary decision diagrams. *International Journal of Approximate Reasoning*, p. 234–251.
- HUE J., PAPINI O. & WÜRBEL E. (2007). Syntactic propositional belief bases fusion with removed sets. *Proceedings of ECSQARU*, p. 66–77.
- KONIECZNY S. (2000). On the difference between merging knowledge bases and combining them. In *KR*, p. 135–144.
- KONIECZNY S. & PINOPEZ R. (1999). Merging with integrity constraints. *LNCS*, **1638**, 233–244.
- MEYER T., GHOSE A. & CHOPRA S. (2001). Syntactic representations of semantic merging operations. In *PRICAI '02 : Proceedings of the 7th Pacific Rim International Conference on Artificial Intelligence*, p. 620.
- NIEMELÄ I. & SIMONS P. (1997). An implementation of stable model and well-founded semantics for normal logic programs. In S. VERLAG, Ed., *Proc. of LP-NMR'97*, volume 1265 of *LNAI*, p. 420–429.
- REVESZ P. Z. (1997). On the semantics of arbitration. *Journal of Algebra and Computation*, **7**(2), 133–160.