Raisonner avec une politique d'échange d'informations incomplète

Laurence Cholvy[1] Stéphanie Roussel[1,2]

[1]ONERA Centre de Toulouse

[2]SUPAERO, Toulouse

JIAF, 2-3 Grenoble 2007

- Système multi-agents dans lesquels les agents doivent se coordonner pour réaliser une tâche globale
- Les agent doivent donc échanger de l'informations pour partager une vision commune de la situation
- Echanges régulés par une politique (réglementation)

Les propriétés (majeures) attendues sont :

- **Cohérence** : la politique ne doit jamais placer un agent face à une contradiction
  (1) *Dès qu'un agent reçoit une information de type T_1, il doit la diffuser à tous les autres agents*.
  (2) *L'agent A a l'interdiction de diffuser les informations de type T_2 qu'il reçoit*.
- **Complétude** : la politique prescrit l'attitude de tout agent dans toute situation relative à l'échange d'informations.
  (1) *Toute information de type T_1 doit obligatoirement être diffusée à tous les agents*.
  Que faire si une information de type T_2 arrive ?

Langage du premier ordre avec égalité qui permet de raisonner avec des notions déontiques.
Soit $\mathcal{A}$ les axiomes :

- $(Ax1) \quad \forall x \quad P(x) \leftrightarrow \neg O(not(x)).$
- $(Ax2) \quad \forall x \quad F(x) \leftrightarrow O(not(x)).$
- $(Ax3) \quad \forall x \quad T(x) \leftrightarrow P(x) \wedge P(not(x))$
- $(D) \quad \forall x \quad O(not(x)) \rightarrow \neg O(x).$

Example

(R_0) : *Tout agent qui reçoit une information de type T_1 a l'interdiction de la diffuser*
$(R_0) \quad \forall(x, y, i) \quad Recoit(x, i) \wedge Theme(i, T_1) \wedge Ag(y) \wedge \neg(x = y) \rightarrow F(dire(x, y, i))$

Definition (Cohérence dans un monde - version simplifiée)

Soit $\mathcal{P}$ une politique et W un état du monde (interprétation) dans lequel elle est appliquée. On dit que $\mathcal{P}$ est cohérente dans W ssi $W \models \mathcal{P} \wedge \mathcal{A}$.

Example

$W_0 = \{Ag(a), Ag(b), Theme(i_1, T_1), Theme(i_2, T_2), Recoit(a, i_2)\}$.
$\mathcal{P}_0 = \{ \quad \forall(x, y, i) \quad Recoit(x, i) \wedge Theme(i, T_1) \wedge Ag(y) \wedge \neg(x = y) \rightarrow F(dire(x, y, i))\}$.
$W_0 \models \mathcal{P}_0 \wedge \mathcal{A}$.
Donc, $\mathcal{P}_0$ est cohérente dans le monde W_0.

Definition (Cohérence)

$\mathcal{P}$ est cohérente si et seulement si elle est cohérente dans tout état du monde.

Definition (Complétude dans un monde)

Soit $\mathcal{P}$ une politique et W un état du monde dans lequel elle est appliquée. $\mathcal{P}$ est complète pour $\models$ dans W ssi pour tout x, y, i

$$W \models Recoit(x, i) \wedge Ag(y) \wedge \neg(x = y) \Rightarrow$$

$$W \models \mathcal{P} \wedge \mathcal{A} \rightarrow O(dire(x, y, i)) \text{ ou}$$

$$W \models \mathcal{P} \wedge \mathcal{A} \rightarrow F(dire(x, y, i)) \text{ ou}$$

$$W \models \mathcal{P} \wedge \mathcal{A} \rightarrow T(dire(x, y, i))$$

Example

$W_0 \models Recoit(a, i_2) \wedge Ag(b) \wedge \neg(a = b)$ mais
$W_0 \not\models \mathcal{P}_0 \wedge \mathcal{A} \rightarrow O(dire(a, b, i_2))$
$W_0 \not\models \mathcal{P}_0 \wedge \mathcal{A} \rightarrow F(dire(a, b, i_2))$
$W_0 \not\models \mathcal{P}_0 \wedge \mathcal{A} \rightarrow T(dire(a, b, i_2))$
$\mathcal{P}_0$ est incomplète pour $\models$ dans W_0.

- **Idée** : Adapter la CWA de Reiter définie pour les bases de données incomplètes.
- **Rappel (CWA)** Si un littéral l ne peut pas être déduit d'une base de données, alors on admet que sa négation $\neg l$ peut être déduite.
- **Justification de la CWA** : Dans le monde réel, on a $l \otimes \neg l$ (tiers exclu).
- **En ce qui concerne les notions déontiques**, on a $\mathcal{A} \models O(l) \otimes F(l) \otimes T(l)$.
- On définit des règles de complétion paramétrées par des conditions E_i.

NOTATION : On note $''\mathcal{P}, W$ *incomplet pour* $(x, y, i)''$ à la place de : $W \models Recoit(x, i) \wedge Ag(y) \wedge \neg(x = y)$ et $W \not\models \mathcal{P} \wedge \mathcal{A} \rightarrow O(dire(x, y, i))$ et $W \not\models \mathcal{P} \wedge \mathcal{A} \rightarrow T(dire(x, y, i))$ et $W \not\models \mathcal{P} \wedge \mathcal{A} \rightarrow F(dire(x, y, i))$

Definition (Règles de complétion)

$$(R_{E_F}) \qquad \frac{\mathcal{P}, W \textit{ incomplet pour } X, \quad W \models E_F(X)}{F(dire(X))}$$

$$(R_{E_T}) \qquad \frac{\mathcal{P}, W \textit{ incomplet pour } X, \quad W \models E_T(X)}{T(dire(X))}$$

$$(R_{E_O}) \qquad \frac{\mathcal{P}, W \textit{ incomplet pour } X, \quad W \models E_O(X)}{O(dire(X))}$$

NOTATION :
On note $\models_*$ l'inférence définie par $\models + R_{E_F} + R_{E_T} + R_{E_O}$

Résultat

Soit $\mathcal{P}$ une politique et W un état du monde. $\mathcal{P}$ est cohérente pour $\models_*$ dans W et complète pour $\models_*$ dans W ssi

$$\forall X = (x, y, i), \mathcal{P}, W \text{ incomplet pour } X \Rightarrow \\ W \models E_F(X) \otimes E_T(X) \otimes E_O(X)$$

Example

- $E_F(x, y, i) = Theme(i, T_1)$,
  $E_T(x, y, i) = false$,
  $E_O(x, y, i) = Theme(i, T_2)$.
- $\mathcal{P}_0, W_0$ incomplet pour le triplet (a, b, i_2).
- $W_0 \models E_F(a, b, i_2) \otimes E_T(a, b, i_2) \otimes E_O(a, b, i_2)$. La politique $\mathcal{P}_0$ est donc complète et cohérente pour $\models_*$ dans W_0.

Résultat (plus faible)
Soit $\mathcal{P}$ une politique et W un état du monde.
Si $\forall X = (x, y, i)$

$$W \models Recoit(x, i) \wedge Ag(y) \wedge \neg(x = y) \rightarrow E_F(X) \otimes E_T(X) \otimes E_O(X)$$

alors
$\mathcal{P}$ est cohérente et complète pour $\models_*$ dans W

Exemples de E_i simples

- $E_F = \mathit{True}$, $E_T = \mathit{False}$ et $E_O = \mathit{False} \Rightarrow$ (sauf mention explicite, l'envoi d'informations est interdit)
  Applicable à des réglementations pour un système hautement sécurisé où chaque transmission d'information doit être explicitement autorisée avant d'être effectuée.
- $E_F = \mathit{False}$, $E_T = \mathit{True}$ et $E_O = \mathit{False} \Rightarrow$ (sauf mention explicite, l'envoi d'informations est autorisé)
  Applicable à des réglementations pour un système faiblement sécurisé où chaque transmission d'information, sauf précision, est implicitement autorisée.
- $E_F = \mathit{False}$, $E_T = \mathit{False}$ et $E_O = \mathit{True} \Rightarrow$ (sauf mention explicite, l'envoi d'informations est obligatoire)
  Applicable à des serveurs mail qui doivent laisser passer tous les mails sauf les spams.

Relation avec les défauts

Il existe une reformulation des règles d'inférence précédentes sous forme de défauts.

$$(d_F) \quad \frac{Recoit(x,i) \wedge Ag(y) \wedge \neg(x=y) \wedge E_F(x,y,i) \ : \ F(dire(x,y,i))}{F(dire(x,y,i))}$$

$$(d_T) \quad \frac{Recoit(x,i) \wedge Ag(y) \wedge \neg(x=y) \wedge E_T(x,y,i) \ : \ T(dire(x,y,i))}{T(dire(x,y,i))}$$

$$(d_O) \quad \frac{Recoit(x,i) \wedge Ag(y) \wedge \neg(x=y) \wedge E_O(x,y,i) \ : \ O(dire(x,y,i))}{O(dire(x,y,i))}$$

Conclusion

- Formalisation des politiques d'échange d'informations
- Propriétés de cohérence et de complétude
- Proposition de règles de complétion
- Extensions à la prise en compte du temps
- Généralisation à d'autres réglementations