

Modèles et Réseaux de confiance

Droits et Devoirs d'Agents Autonomes -

Confiance et modèles de confiances pour la régulation des droits
et devoirs d'agents autonomes

dimitri.melaye@imag.fr, yves.demazeau@imag.fr

Laboratoire Leibniz, Equipe MAGMA

Objectif

Objectifs

- Nécessité de pouvoir agir dans un environnement **ouvert**, **dynamique**, et **imprévisible**, d'y garantir la sécurité, et d'offrir les meilleurs services possibles aux autres services tout comme à **l'utilisateur**.
- D2A2 : régulation entre agents autonomes au sein d'**organisations**, d'institutions.

- 1 Objectifs
- 2 La confiance
- 3 Réseaux de confiance
 - Réseaux de confiance : principes
 - Quel modèle de confiance ?
- 4 Construction, maintenance et exploitation de réseaux de confiance
- 5 Conclusion

La confiance

- Croyance à propos du comportement d'un autre dans un domaine ou un contexte donné.
- Probabilité qu'un agent attende d'un autre une action donnée dont dépend son bien-être [Gam00].
- Mécanisme de coordination en situation d'ignorance.
- Mécanisme d'intégration sociale.

[Rou00]	Social	Individuel
Instantané	Présentation	Engagement dans l'action
Historique	Réputation	Familiarité

La confiance

Représentation et mesure

- Une valeur dans $[-1,1]$ ou $[0,1]$, mais quid de l'absence d'information et de la défiance ?
- Sémantique ? Intervalle discret et calcul d'un seuil pour se ramener une mesure binaire de la confiance.

Dynamique

- Influence des interactions et des observations sur le niveau de confiance.
- Inertie de la confiance : vitesse de basculement de la confiance.
- Erosion de la confiance en absence d'information.

Modèles un à un

Modèles fondés sur la théorie des jeux et le calcul d'intérêt

- Probabilités sur des résultats de jeu.
- Mais...hypothèse de connaissance du système.
- Mais...Répétitions des actions en communs : les agents doivent être inclus dans le système.

Modèles fondés sur une approche cognitive [CF98]

- Processus cognitif.
- Structure de croyance et de buts (croyance de compétence, de bonne volonté, de dépendance).

Modèles n à un centrés réputation

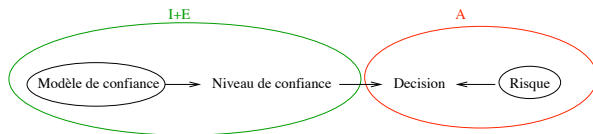
- Système d'évaluation de contenu et des évaluateurs dans un environnement ouvert (web).
- Systèmes fondés sur la transitivité.
- Système ReGreT [SS02]
 - Dimension individuelle (moyenne pondérée des évaluations des interactions)
 - Dimension sociale (réputation de témoignage et choix des témoins, réputation de voisinage, réputation systémique)
 - Dimension ontologique

- 1 Objectifs
- 2 La confiance
- 3 Réseaux de confiance**
 - Réseaux de confiance : principes
 - Quel modèle de confiance ?
- 4 Construction, maintenance et exploitation de réseaux de confiance
- 5 Conclusion

Principes

Modèles 1 à 1 et n à 1

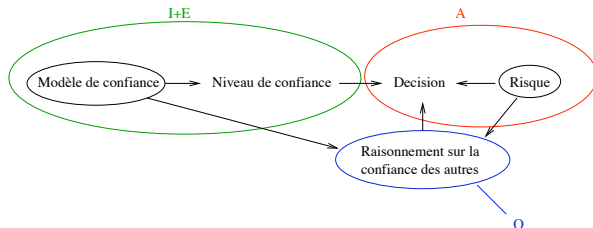
- Modèle de confiance "un à un" (individuel) ou "n à un" (réputation) permettant de déterminer un niveau de confiance.



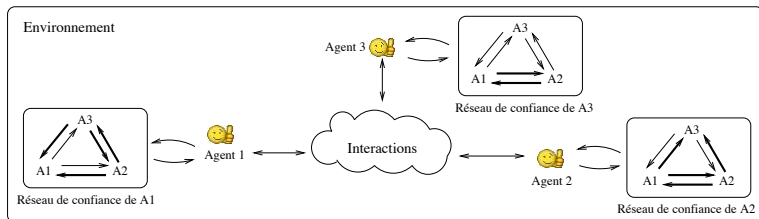
Principes

Objectif : passer aux modèles n à p

- Modèle n à p non centré-agent : proposer un modèle de confiance fondé sur une approche sociale et cognitive en considérant la mise en réseau des relations de confiance.



Quel modèle de confiance pour les réseaux de confiance ?



- Réseau de confiance "naïf"

$$\mathcal{T}_i = \{M_{ijk} \mid 1 \leq j \leq N_i, 1 \leq k \leq N_i, j \neq k\}$$

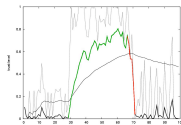
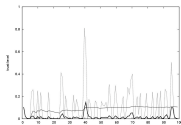
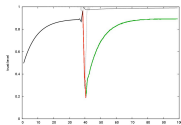
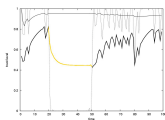
- Graphe orienté valué par les modèles

→ Quel modèle de confiance un-à-un et n-à-un ?

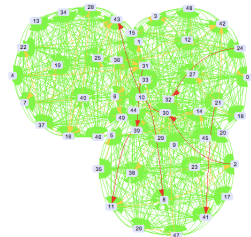
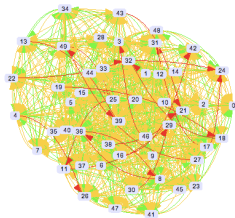
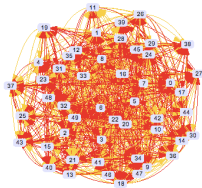
Exemple

Mis en réseau de modèle 1-à-1

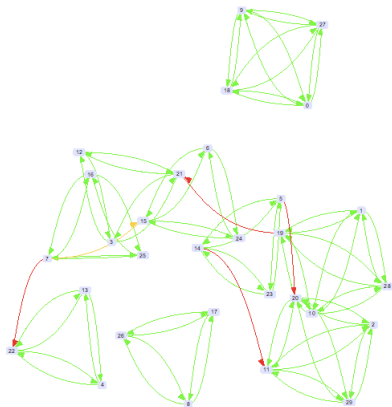
- Modèle de confiance dynamique "un à un" : réseau bayésien avec filtre de Kalman [MD05]
- Observer dynamiquement les relations de confiance entre les agents d'un système, au fur-et-à-mesure de leurs interactions.



Exemple



Exemple



Interaction : observabilité, intelligibilité, manipulabilité

- Observabilité : déduire les relations de confiance existant entre les agents à partir d'observations et d'évaluations

$$observability_{ijk} : \{O_{njk}\}_{n,j,k \in \mathcal{A}_i} \times \{E_{njk}\}_{n,j,k \in \mathcal{A}_i} \rightarrow T$$

- Intelligibilité : décomposition de la confiance

$$observability_{ijk} = g_{ijk} \circ understand_{ijk}$$

$$understand_{ijk} : \{O_{njk}\}_{n,j,k \in \mathcal{A}_i} \times \{E_{njk}\}_{n,j,k \in \mathcal{A}_j} \rightarrow \mathcal{B}_i$$

- Manipulabilité : raisonnement sur les composantes de la confiance

$$observability_{ijk} = manage_{ijk} \circ h_{ijk}$$

$$manage_{ijk} : \mathcal{B}_i \rightarrow T$$

Agent : exploitabilité sociale

Exploitabilité sociale : prise de décision en prenant en compte plusieurs autres agents

$$social_exploitability_i : T_i^n \rightarrow A_i \text{ with } |n| > 1$$

- 1 Objectifs
- 2 La confiance
- 3 Réseaux de confiance
 - Réseaux de confiance : principes
 - Quel modèle de confiance ?
- 4 Construction, maintenance et exploitation de réseaux de confiance
- 5 Conclusion

Vers une approche symbolique

- Le modèle de confiance et sa mise en réseau ne sont pas indépendants
- Intelligibilité et manipulabilité : approche symbolique
- Observabilité et exploitabilité sociale : approche sociale [CF98]

Formalisme ?

→ Logique modale

Logique modale de confiance

Justification

- Meilleure compréhension des composantes de la confiance par une représentation et une manipulation explicite (intelligibilité/manipulabilité).
- Représentation explicite des autres (exploitabilité sociale).
- Traitement automatique possible.

Quelle logique ?

- [Lia03] : logique BIT portant sur les conséquences de la confiance sur les croyances et sur le transfert d'information
 $T_{ij}\phi$: "l'agent i fait confiance en l'agent j pour ϕ "
- [Dem04] : définition de types de confiance "from scratch", pas d'exploitation

La confiance ne consiste pas uniquement en des croyances
→ observations, interactions, temps et engagement/dépendance

Quelle logique ?

Quelle(s) logique(s) ?

- logique de l'action, logique doxastique, logique temporelle.

Croyances communes

- Caractériser un réseau de confiance mutuelle : validité de la formule $B_i T_{jk}$ pour tous les agents i, j, k du réseau testé
- Trouver un réseau de confiance mutuelle : satisfaire la formule $B_i T_{jk}$.

Conclusion

- La prise en compte des organisations nécessite la mise en réseau de la confiance.
- Le modèle de confiance ne peut pas s'abstraire de sa mise en réseau.



C. Castelfranchi and R. Falcone.

Principles of trust for mas : cognitive anatomy, social importance, and quantification.

In *International Conference of Multi-Agent Systems (CMAS'98)*, pages 72–79, Paris, July 1998.



Frederc Cuppens.

Roles and deontic logic.

In A. J. I. Jones and M. Sergot, editors, *Second International Workshop on Deontic Logic in Computer Science*, 1994.



R. Demolombe.

Reasoning about trust : A formal logical framework.

In *iTrust 2004*, pages 291–303, 2004.



Diego Gambetta.

Can we trust trust ?

In Diego Gambetta, editor, *Trust : Making and Breaking Cooperative Relations*, pages 213–237. Department of Sociology, University of Oxford, 2000.



CJ. Liau.

Belief, information acquisition, and trust in multi-agent systems- a modal logic formulation.

Artificial Intelligence, 149(1) :31–60, September 2003.



Dimitri Melaye and Yves Demazeau.

Bayesian dynamic trust model.

In springer, editor, *Proceedings of CEEMAS'05*, volume LNCS 3690, 2005.



I. Porn.

Action theory and social science ; some formal models.

Synthese Library, 120, 1977.



J. Rouchier.

La confiance à travers l'échange.

PhD thesis, Montpellier, 2000.



Jordi Sabater and Carles Sierra.

Reputation and social network analysis in multi-agent systems.

In Proceedings of the first international joint conference on

Autonomous agents and multiagent systems, pages 475–482.

ACM Press, 2002.