

Paul Sabatier

Juin 2011

Magazine scientifique - numéro 22



Dossiers

Raisonnement
et décision

La chimie



www.ups-tlse.fr

Avec la participation de

Délégation
Midi-Pyrénées du CNRS



Délégation régionale
Midi-Pyrénées, Limousin de l'Inserm

Instituts
thématiques



Inserm

Institut national
de la santé et de la recherche médicale

Théorie des jeux, émotion et confiance : des sciences sociales à l'intelligence artificielle



>>> Dominique LONGIN et Emiliano LORINI, chargés de recherche CNRS à l'IRIT (unité mixte UPS/CNRS/UT1/UT2/INP).

Pour modéliser les interactions sociales, la théorie des jeux commence à prendre en compte des motivations et des émotions de plus en plus complexes, comme la confiance, la culpabilité ou la honte. Des concepts utiles pour concevoir des robots proches des humains.

Pendant le siècle dernier la théorie des jeux s'est imposée comme paradigme dominant en sciences sociales pour modéliser et expliquer l'interaction sociale entre agents humains et entre agents économiques (états, entreprises, banques, etc.). But de cette théorie : expliquer et prédire les choix des acteurs sociaux dans un contexte d'interaction stratégique, c'est-à-dire, quand l'utilité du choix d'un agent donné dépend de ce que les autres agents décideront de faire. La théorie des jeux classique est fondée sur un cadre conceptuel très simple incluant les notions de préférence et d'action. Mais, à partir des travaux des économistes John Charles Harsanyi et Robert Aumann, la théorie des jeux a commencé à intégrer les notions de connaissance et de croyance pour pouvoir modéliser des situations d'interaction stratégique avec information imparfaite. Plus récemment, les concepts d'émotion et de confiance sont devenus des éléments centraux de cette théorie. En fait, les évidences empiriques et certaines théories psychologiques montrent que des émotions comme la culpabilité, la honte et le regret affectent les décisions stratégiques des acteurs humains. En outre, il a été prouvé que la confiance joue un rôle crucial pas simplement dans la décision individuelle, mais aussi dans l'interaction sociale en favorisant la coopération.

Raisonnement à partir des émotions

Dans les dernières années, la théorie des jeux est devenue la base théorique la plus utilisée dans le domaine des systèmes multi-agent (SMA). Les SMA constituent un domaine de l'Intelligence Artificielle (IA) dont le but est de développer des modèles d'interaction entre agents autonomes artificiels (ex. modèles de coopération et de coordination, modèles de négociation). De même, l'émotion et la confiance sont devenues des thématiques centrales dans ce domaine de l'IA. Il existe des modèles computationnels d'agents cognitifs autonomes capables de raisonner à partir des émotions d'un utilisateur humain, et dont les décisions sont affectées par leurs émotions. Plusieurs modèles de confiance ont été proposés dans le domaine des SMA : modèles de confiance de type statistique, modèles sociocognitifs de confiance qui se basent sur des théories psychologiques, modèles de réputation. Ces modèles fournissent les spécifications



>>> Tableau « Les Tricheurs » de Caravaggio. Kimbell Art Museum.

formelles et abstraites pour le développement de différentes applications comme les services web, les systèmes de réputation (par ex. EBay), le web sémantique, les agents conversationnels animés et la robotique cognitive.

Expressions du visage

Les travaux réalisés par l'équipe LILAC de l'IRIT ont pour but de développer des modèles logiques de l'interaction sociale, basés sur la théorie des jeux et sur les théories psychologiques de l'émotion et de la confiance. Ces modèles peuvent être exploités comme base pour l'implantation d'agents artificiels capables de raisonner à partir des concepts d'émotion et de confiance dans un contexte d'interaction avec un utilisateur humain ou avec d'autres agents, et dont les décisions stratégiques sont affectées par leurs émotions. Un exemple de travail réalisé est le développement d'un modèle logique des émotions basé sur le raisonnement contrefactuel comme le regret et la culpabilité. Ce modèle a été récemment utilisé dans le contexte du projet ANR CECIL comme base pour l'implantation d'un agent conversationnel animé capable d'exprimer telles émotions d'une façon multimodale (par ex. par les expressions du visage et les gestes).

Pour en savoir plus : - Emiliano LORINI, François SCHWARZENTRUBER. A logic for reasoning about counterfactual emotions. Artificial Intelligence, Vol. 175 N. 3-4, p. 814-847, 2011.

Contacts : Emiliano.Lorini@irit.fr,
Dominique.Longin@irit.fr