

Proposition de thèse

Gestion opportuniste du dialogue dans les conversations sociales avec des agents virtuels et des robots

Nicolas Sabouret – LIMSI-CNRS

Charles Rich - Worcester Polytechnic Institute

Contexte

L'équipe Cognition, Perception, Usages (CPU) du LIMSI-CNRS travaille depuis plusieurs années sur la réalisation d'agents virtuels intelligents capables d'interagir avec des humains et d'adapter leur comportement non-verbal au contexte socio-affectif de l'interaction. Une collaboration débutée en 2013 avec le Pr. Charles Rich du Worcester Polytechnic Institute (WPI) aux États-Unis a permis de développer de nouvelles recherche autour du comportement dialogique des agents, en particulier à travers le système Disco [RS12] basé sur les HTN (Hierarchical Task Network).

Cette thèse se situe dans la poursuite de cette collaboration internationale.

Présentation générale

Les recherches dans le domaine du dialogue homme-machine ont produit de grandes avancées dans le contexte des dialogues à initiative mixte guidés par la tâche [ALL95], par exemple lorsqu'un agent virtuel ou un robot collabore avec un humain pour la manipulation d'un équipement électronique ou mécanique complexe [RS07], ou pour rechercher des informations (voir par exemple les travaux de [LDC12] dans le domaine médical). Dans ces dialogues, la structure conversationnelle est fortement guidée par la structure de la tâche, avec des blocs et des sous-blocs de dialogue correspondant aux buts et aux sous-butts de la tâche (retirer une pièce donnée, recherche un complément d'information...).

Cependant, les conversations sociale, comme on peut en avoir au café avec des amis ou des collègues, ne revêtent pas cette structure organisée: les sujets vont et viennent au gré des opportunités et des associations d'idées. Par exemple, lorsque deux amis discutent d'un film, si l'un des deux comprend que l'autre déteste l'acteur principal, il va essayer de changer de sujet rapidement, et parler des dernières élections. Modéliser ce type de flexibilité et d'adaptation au contexte dans les conversations est un problème complexe et non résolu dans le contexte de la gestion de dialogues. Pourtant, avec l'émergence des compagnons virtuels [RAP14], il apparaît comme un objectif majeur pour les agents virtuels et les robots.

L'étude de ce problème requiert une approche pluridisciplinaire qui combine des travaux en psychologie des émotions et de la personnalité avec des travaux en intelligence artificielle. Par exemple, le changement de sujet dans la discussion à propos d'un film va dépendre du ressenti vis-à-vis de l'interlocuteur, du sentiment de malaise dans la conversation et de la personnalité des intervenants. La modélisation informatique de ces éléments pour la construction d'un raisonnement se fait classiquement en utilisant des modèles de théorie de la décision [MGP10] ou des logiques modales [AHL09, DL12].

L'objectif de cette thèse est de définir un modèle de dialogue pour les conversations sociales, en combinant différents modèles d'intelligence artificielle.

Sujet détaillé : HTNs et modification de plans

L'approche que nous proposons dans cette thèse repose sur deux postulats. Premièrement, les conversations sociales, bien que plus souples que les conversations guidées par la tâche, possèdent aussi une structure. Cette structure conversationnelle est orientée vers des buts qui sont ici d'ordre social [CAS98]. Ainsi, de même qu'un but dans une tâche est généralement défini par un ensemble d'états désirés dans le monde réel (par exemple, mettre un bloc sur un autre), les buts sociaux dans le contexte dialogique peuvent être vus comme des états mentaux souhaités chez l'interlocuteur (comme le proposent les chercheurs en linguistique à travers la théorie des actes de langage [SEA69, SID94]). Par exemple, on peut avoir comme but de faire en sorte que l'interlocuteur nous apprécie, ou pense que nous sommes intelligents, ou s'amuse... Pour formaliser ce type de but, nous proposons d'utiliser des modèles logiques pour représenter et raisonner sur l'état mental de l'interlocuteur. On parle alors, en intelligence artificielle, de théorie de l'esprit (Theory of Mind, ou ToM, en anglais) [PWM13]. Plusieurs modèles logiques ont déjà été proposés dans ce sens [BMT11, AHL09, HL02, BS14].

Notre second postulat est qu'il est difficile (et probablement impossible pour les 10 prochaines années) de définir des modèles *complets* de théorie de l'esprit qui prennent en compte les émotions, la personnalité, les croyances, les désirs et les intentions, avec une axiomatisation complète du domaine. Une telle axiomatisation permettrait effectivement de planifier les actions dialogiques en fonction des états mentaux de l'interlocuteur, donc de contrôler exactement la gestion du dialogue. Cependant, les travaux récents dans le domaine [DL12] ont montré que ce problème difficile constituera l'un des principaux challenges de l'intelligence artificielle pour les prochaines années dans le domaine des agents intelligents. C'est pourquoi nous partons du principe que l'adaptation du dialogue en fonction de l'état mental souhaité de l'interlocuteur ne pourra pas se faire en utilisant des méthode de planification qui nécessitent une telle axiomatisation.

Notre approche se situe donc sur trois niveaux. Premièrement, nous proposons d'utiliser un réseau hiérarchiques de tâches (Hierarchical Task Networks, ou HTN en anglais [RIC08]) défini manuellement pour décrire la structure du dialogue social, comme cela se fait dans le modèle Disco [RS12]. Ce modèle a montré de bons résultats pour les dialogues orientés tâche. Dans notre approche, il joue le rôle de descripteur des stratégies de dialogue génériques sur les sujets sociaux (le cinéma, la politique). Deuxièmement, nous voulons combiner cette structure avec un modèle de représentation des connaissance spécifique au domaine pour l'utiliser dans un cadre concret (par exemple, pour discuter d'un film, une ontologie OWL [MCG04] pourrait fournir les éléments de discussion que sont les acteurs, la réalisation, etc). Troisièmement, notre objectif est de combiner ce HTN avec des méthodes algorithmiques de modification de plans qui s'appuient sur une théorie de l'esprit incomplète ou superficielle, décrite à l'aide de logique propositionnelle ou de logique modale.

Ainsi, nous voulons *combiner* un dialogue structuré pour gérer les discussion sur les films à partir de buts et de sous-buts (parler des acteurs, de l'intrigue, de la mise en scène) avec un ensemble de buts sociaux (être poli avec mon interlocuteur, lui faire plaisir) décrits dans un modèle logique, et munir l'ensemble d'un planificateur simple (type STRIPS) pour modifier ou réparer dynamiquement le HTN lors que l'agent comprend, à l'aide des observables du dialogue, qu'un sujet (*i.e.* un sous-but de discussion) va entrer en conflit avec les objectifs sociaux (par exemple, l'interlocuteur déteste l'acteur principal).

Nous proposons donc dans cette thèse de définir une approche hybride comme solution pratique pour gérer l'adaptation automatique dans les dialogues sociaux. Cette approche combine des HTN comme modèle d'exécution, des modèles sémantiques pour décrire le domaine de discussion et des logiques d'action pour les buts sociaux. Le premier objectif de la thèse, qui se situe dans la continuité du travail de stage de Mme Lydia Ould Ouali effectué au LIMSI en 2014, consiste à

étudier la combinaison des HTN avec les modèles de planification pour la modification de plans en utilisant des théories incomplètes. Le deuxième objectif de la thèse est d'étendre ces modèles à des logiques d'actions pour représenter une théorie de l'esprit et les buts sociaux. Ce travail contribuera aussi au développement d'une formalisation logique de la théorie de l'esprit, qui reste un objectif à long terme de notre équipe. Ces travaux s'appuieront sur les recherches menées par N. Sabouret au LIMSI.

En parallèle, il faudra définir le méta-modèle de dialogue à l'aide d'un HTN et le coupler avec des ontologies de domaines. Ce travail s'appuiera sur les recherches menées au WPI par C. Rich et C. Sidner. Il faudra ensuite proposer les algorithmes d'adaptation compatibles avec cette nouvelle théorie. L'objectif final de la thèse sera la mise en œuvre de ce modèle et des ces algorithmes dans un agent virtuel dialoguant capable de s'adapter à différents contexte d'interaction sociale, dans le cadre de la définition de compagnons artificiels.

Références

- [AHL09] C. Adam, A. Herzig, and D. Longin. A logical formalization of the OCC theory of emotions. *Synthese*, 168(2):201-248, Feb. 2009.
- [ALL95] J. Allen et al. The TRAINS project: A case study in building a conversational planning agent. *Journal of Experimental & Theoretical Artificial Intelligence* 7(1):7-48, 1995.
- [BMT11] T. Bosse, Z. A. Memon, and J. Treur. A recursive BDI agent model for Theory of Mind and its applications. *Applied Artificial Intelligence*, 25(1):1-44, 2011.
- [BS14] M. Belkaid and N. Sabouret. A logical model of Theory of Mind for virtual agents in the context of job interview simulation, in *Proc. Second International Workshop on Intelligent Digital Games for Empowerment and Inclusion at IUI 2014 (IDGEI 2014)*.
- [CAS98] C. Castelfranchi. Modelling social action for AI agents. *Artificial Intelligence*, 103(1):157-182, 1998.
- [DL12] M. Dastani and E. Lorini. A logic of emotions: From appraisal to coping. In *Proceedings of the 11th International conference on Autonomous Agents and Multiagent Systems*, pages 1133-1140, 2012.
- [HL02] A. Herzig and D. Longin. A logic of intention with cooperation principles and with assertive speech acts as communication primitives. In *Proc. First Int. Joint Conf on Autonomous Agents and Multiagent Systems*, pages 920-927, ACM, 2002.
- [LDC12] A. Loisel, G. Dubuisson Duplessis, N. Chaignaud, J.-P. Kotowicz, A Conversational Agent for Information Retrieval Based on a Study of Human Dialogues, *Proceedings of the 4th International Conference on Agents and Artificial Intelligence (ICAART)*, pp.312-317, 2012.
- [MCG04] D. McGuinness and F. Van Harmelen. OWL web ontology language overview. *W3C recommendation*, 10:2004-03, 2004.
- [MGP10] S. Marsella, J. Gratch and P. Petta, *Computational models of emotion. A Blueprint for an Affectively Competent Agent: Cross-Fertilization Between Emotion Psychology, Affective Neuroscience, and Affective Computing*. Oxford University Press, Oxford, 2010
- [PWM13] D. V. Pynadath, N. Wang, and S. C. Marsella. Are you thinking what I'm thinking? An Evaluation of a Simplified Theory of Mind. In *Intelligent Virtual Agents*, pages 44-57. Springer,

2013.

[RAP14] J. Rivière, C. Adam and S. Pesty, Un ACA sincère, affectif et expressif comme compagnon artificiel. *Revue d'Intelligence Artificielle* 28(1):67-99, 2014.

[RIC08] C. Rich, Building task-based user interfaces with ANSI/CEA-2018, *IEEE Computer*, Vol. 42, No. 8, August 2009, pp. 20-27.

[RS07] C. Rich and C. Sidner, DiamondHelp: A generic collaborative task guidance system, *AI Magazine*, Vol. 28, Summer 2007.

[RS12] C. Rich and C. Sidner, Using collaborative discourse theory to partially automate dialogue tree authoring, *Proc. Int. Conf. on Intelligent Virtual Agents*, Santa Cruz, CA, 2012.

[SEA69] J. R. Searle, *Speech acts: An essay in the philosophy of language*, Vol. 626. Cambridge University Press, 1969.

[SID94] C. Sidner, An artificial discourse language for collaborative negotiation, *Proc. 12th National Conf. on Artificial Intelligence*, Seattle, WA, pp. 814–819, Aug. 1994.