Aspects linguistiques du traitement automatique du dialogue

Claire Gardent* — **Jean-Marie Pierrel****

** LORIA – UMR 7503 CNRS – INRIA – Universités de Nancy*
Laboratoire Lorrain de Recherche en Informatique et ses Applications
Campus scientifique, BP 239, F-54506 Vandœuvre-lès-Nancy cedex

Claire.Gardent@loria.fr

*** ATILF – UMR 7118 CNRS – Université Nancy 2*
Analyse et Traitement Informatique de la Langue Française
44, avenue de la Libération, BP 30687, F-54063 Nancy cedex

Jean-Marie.Pierrel@inalf.fr

RÉSUMÉ. *Dans cet éditorial, nous commençons par retracer brièvement l'historique du traitement automatique du dialogue. Ce faisant, nous identifions les aspects linguistiques qui sont spécifiques à l'interaction homme-machine et doivent être abordés lors de la modélisation du dialogue et de l'implantation de systèmes. Nous concluons en situant les articles sélectionnés par rapport à cette problématique.*

ABSTRACT. *In this editorial, we start by sketching the history of dialog processing. In so doing, we identify the linguistic issues which are specific to man-machine interaction and must be addressed when modeling dialog and implementing systems. We conclude by situating the selected articles with respect to these issues.*

MOTS-CLÉS : *historique, dialogue, modélisation linguistique.*

KEYWORDS: *History, Dialog, Linguistic modeling.*

Un dialogue entre êtres humains peut faire intervenir des processus plus ou moins complexes suivant le nombre de participants qu'il implique et la relation de pouvoir qui les relie (plus les participants se trouvent sur un pied d'égalité, plus il pourra y avoir de négociation avant d'atteindre un objectif donné et plus les processus de raisonnement engagés seront complexes) ; suivant le nombre et le type des modalités qui y sont employées (langage seul ou en combinaison avec d'autres modalités telles que le geste ou le graphisme) ; et suivant sa finalité (échange d'information, accomplissement d'une tâche, négociation, etc.).

De même, les systèmes de communication homme-machine sont d'une complexité variable.

Ainsi les premiers systèmes des années 1970 tels ELIZA [WEI 67] et BARRY [COL 71], n'admettent que deux participants et une modalité unique : la langue naturelle écrite. Dans ces systèmes, l'initiative de l'utilisateur est, en outre, restreinte par une stratégie de contrôle qui privilégie les questions du système : plus le dialogue est guidé par le système et moins le risque est grand que l'utilisateur ne pose de questions, ou n'obtienne de réponses, qui démasquent ses limites.

Si ces premiers systèmes peuvent prétendre simuler un être humain en tant que partenaires conversationnels (de fait, le *turing test* montra que ces systèmes arrivaient à induire en erreur certains utilisateurs qui pensaient dialoguer avec un être humain), ils restent limités dans leur capacité et leurs modalités d'interaction. Au cours des deux dernières décennies, on a ainsi assisté au développement de systèmes et d'axes de recherche visant à augmenter ces capacités.

D'une part, les systèmes interactifs basés sur la parole se sont largement développés avec en particulier le système GALAXY du MIT [GOD 94], le système d'information ferroviaire de Philips [AUS 95] et plus récemment le système téléphonique de traduction VERBMOBIL [WAH 00]. L'emphase porte dans ce contexte sur la construction de systèmes à partir de modules existants (modules de reconnaissance/production de la parole, analyseurs, générateurs et gestionnaires de dialogue) l'idée étant de produire des applications qui utilisent et mettent en avant les apports scientifiques et technologiques fait dans le domaine du traitement de la parole. Les aspects linguistiques propres au dialogue ne sont que peu pris en compte.

Par contraste, diverses approches se sont développées qui visent à simuler certaines des caractéristiques du dialogue humain.

Ainsi, une direction importante de recherche vise à modéliser les processus complexes de raisonnement qui permettent de reconstruire suivant un contexte continuellement changeant, l'interprétation de segments linguistiques qui sont fréquemment sous-spécifiés, sémantiquement et/ou pragmatiquement. Dans ce domaine, deux grands types d'approches peuvent être identifiés : les approches qui utilisent les techniques de plannification pour reconnaitre l'intention de l'utilisateur et interpréter son énoncé en conséquence [ALL 82, TRA 96] et celles [COO 98] plus restreintes au plan pragmatique mais moins coûteuses en calcul, qui se basent sur les résultats de l'analyse conversationelle et une représentation dynamique du contexte.

Plus récemment, la recherche s'est également orientée vers le traitement de dialogues impliquant plus de deux agents (systèmes multiagents, [ELI 99, RIC 01, XU 00]), plusieurs tâches (systèmes multitâches, [LEV 98, ROS 95]) ou plusieurs modalités (*i.e.*, la langue naturelle mais aussi la vision et le geste, [BOL 80, TEN 83, WAH 89]).

Dans ce numéro, notre objectif était de faire un bilan des études et des recherches qui concernent la modélisation et la simulation d'aspects linguistiques spécifiques au traitement automatique du dialogue homme-machine. Les articles sélectionnés illustrent différents aspects de cette problématique.

Trois articles portent sur l'interaction entre gestion du dialogue et interprétation.

Le premier article de Raquel Fernández et Jonathan Ginzburg s'appuie sur une étude de corpus et montre l'importance des énoncés non verbaux dans le dialogue. De tels énoncés, sans verbe, pourraient à première vue être considérés comme mal formés ou a-syntaxiques et être classés comme des inattendus dans la langue. Mais leur importance dans les dialogues, que ce soit dans des réponses courtes à des questions fermées ou ouvertes, à des demandes de clarification ou de précision, ou dans des demandes de clarification et dans des énoncés de validation, nécessite de les étudier plus à fond, de les classifier et de proposer des méthodes linguistiques pour en assurer tout à la fois l'analyse et l'interprétation. C'est précisément l'objet de leur contribution dans laquelle ils montrent en particulier que, sur la base d'une classification fine et à l'aide d'analyses issues des travaux récents en HPSG, il est possible de proposer des traitements efficaces et valides d'un point de vue linguistique pour ce type d'énoncés incontournable dans le traitement automatique de dialogue.

Les deux contributions suivantes abordent les aspects de modélisation du dialogue. La première, de Laurent Prévot, Philippe Muller, Pascal Denis et Laure Vieu montre comment la sémantique formelle, dans sa version dynamique (SDRT), permet de combiner les apports de la sémantique formelle et de la rhétorique du discours pour traiter d'aspects propres au dialogue telle la gestion des références communes entre les interlocuteurs. Puis elle aborde la représentation de la structure du dialogue en enrichissant la SDRT pour prendre en charge, d'une part, les actes de langage d'Austin et Searle et, d'autre part, à la suite des travaux de Nicolas Asher, un niveau cognitif. Dans une dernière partie enfin, les auteurs montrent les apports de leurs propositions pour traiter des « fragments phrastiques », donnant ainsi un éclairage complémentaire au premier papier traitant des énoncés non verbaux.

La deuxième contribution, de Nicolas Maudet, nous propose une analyse critique des modélisations de dialogue qui se réduisent à celle de la « structure intentionnelle » des interlocuteurs. Il montre pourquoi de son point de vue une approche combinant des modèles intentionnels et conventionnels (ces derniers étant essentiellement issus des jeux de dialogue) est sans doute plus prometteuse : les jeux de dialogue permettant d'intégrer des notions sociales au sein des approches intentionnelles et ainsi de mieux prendre en compte les conventions d'usage des interactions langagières.

Le quatrième article porte sur la multimodalité et son effet sur l'interprétation dialogique. Cette contribution, proposée par Frédéric Landragin, Susanne Salmon-Alt et Laurent Romary, aborde pour sa part les problèmes d'interprétation de la référence dans le contexte de dialogues multimodaux avec support visuel. Ce type de situation est particulièrement courant pour bon nombre d'applications. En dehors des serveurs vocaux téléphoniques, la plupart des applications potentielles de dialogue homme-machine s'appuie en effet sur une représentation visuelle du monde. Les auteurs montrent dans leur papier qu'il est illusoire dans ce cas de fonder les processus d'interprétation des expressions référentielles uniquement sur des contraintes linguistiques mais qu'il faut, de plus, s'appuyer sur une représentation des contextes intégrant les principes de fonctionnement de la perception visuelle, de la tâche et de la mémoire de l'utilisateur.

La cinquième contribution, proposée par Oliver Lemon, Alexander Gruenstein et Stanley Peters, aborde quant à elle les aspects d'architecture d'un gestionnaire de dialogue faisant intervenir plusieurs tâches. Après avoir rappeler les caractéristiques d'un système de dialogue homme-machine multimodal d'interaction avec un robot, les auteurs décrivent l'architecture générique de systèmes de dialogue multimodal qu'ils proposent. Leur proposition s'appuie, pour représenter les informations dynamiques du dialogue, sur une représentation du contexte conversationnel fondé sur deux structures de données arborescentes complémentaires : l'une pour modéliser l'état courant du dialogue, le Dialogue Move Tree, l'autre pour l'état de la tâche, l' Activity tree.

Enfin la dernière contribution, proposée par Laurence Devillers, Hélène Maynard et Patrick Paroubek, évoque les aspects de méthodologie d'évaluation de systèmes de dialogue. Après avoir rappeler l'état de l'art en ce domaine en dégageant sept axes méthodologiques – objectifs, taille de contexte, objets mesurés, nature des mesures, mode d'évaluation, nature de l'évaluation, généricité – et en présentant les grands paradigmes d'évaluation utilisés à ce jour et leurs résultats, ils proposent un nouveau paradigme PEACE (Paradigme d'Evaluation Automatique de ComprEhension) dont l'objectif est de permettre une évaluation automatique et comparative pour la compréhension littérale et la compréhension en contexte dialogique. Un tel paradigme devrait permettre de mieux évaluer les aspects linguistiques du traitement automatique de dialogue.

1. Bibliographie

[ALL 82] ALLEN J., « Recognising Intentions from Natural Language Utterances », BRADY M., BERWICK R., Eds., *Computational Models of Discourse*, MIT Press, 1982.

[AUS 95] AUST H., OERDER M., SEIDE F., STEINBISS V., « The Philips automatic train timetable information system », *Speech Communication*, vol. 17, 1995, p. 249-262.

[BOL 80] BOLT R., « Put that there : Voice and Gesture at the User Interface », *Computer Graphics*, vol. 14, n° 3, 1980, p. 262-270.

[COL 71] COLBY K., « Artificial paranoia », *Artificial Intelligence*, vol. 2, 1971.

[COO 98] COOPER R., LARSSON S., « Dialog Moves and Information States », Gothenburg papers in computational linguistics, 1998, Goteborg University.

[ELI 99] ELIO R., HADDADI A., « On abstract task models and conversation policies », *Proceedings of the Workshop on specifying and implementing conversation policies, Autonomous agents'99*, 1999.

[GOD 94] GODDEAU D., BRILL E., GLASS J., PAO C., PHILLIPS M., POLIFRONI J., SENEFF S., ZUE V., « GALAXY : A Human-Language Interface to On-line Travel Information », *Proceedings of the International Conference on Spoken Language Processing*, 1994.

[LEV 98] LEVIN L., THYME-GOBBEL A., LAVIE A., RIES K., ZECHNER K., « A discourse coding scheme for Conversational Spanish », *Proceedings of the International Conference on Speech and Language Processing (ICSLP)*, 1998.

[RIC 01] RICH C., SIDNER C., LESH N., « Collagen : Applying collaborative discourse theory to Human-Computer Interaction », *AI Magazine*, vol. 22, n° 4, 2001, p. 15-25.

[ROS 95] ROSÉ C., EUGENIO B., LEVIN L., ESS-DYKEMA C., « Discourse Processing of Dialogues with Multiple Threads », *Proceedings of the Association for Computational Linguistics*, 1995.

[TEN 83] TENNANT H. R., ROSS K. M., SAENZ R. M., THOMPSON C. W., MILLER J. R., « Menu-based natural language understanding », *Proceedings of the 21st Annual Meeting of the ACL*, 1983, p. 151-158.

[TRA 96] TRAUM D., ALLE J., HEEMAN P., HWANG C., MARTIN N., M. P., SCHUBERT L., « Integrating Natural Language Understanding and Plan Reasoning in the TRAINS-93 Conversation System », Computer Science Department, University of Rochester, 1996.

[WAH 89] WAHLSTER W., « User and Discourse Models for Multimodal Communication », SULLIVAN J., TYLER S., Eds., *Intelligent User Interfaces*, p. 45-67, CUP, 1989.

[WAH 00] WAHLSTER W., « Mobile Speech-to-Speech Translation of Spontaneous Dialogs : An Overview of the Final Verbmobil System », WAHLSTER W., Ed., *Verbmobil : Foundations of Speech-to-Speech translation*, p. 3-21, Springer, 2000.

[WEI 67] WEIZENBAUM J., « ELIZA - A Computer Program for the Study of Natural Language Communication between Man and Machine », *CACM*, vol. 10, 1967.

[XU 00] XU W., RUDNICKY A., « Task-Based Dialog Management using an Agenda », *Proceedings of ANLP/NAACL 2000 Workshop on Conversational Systems*, 2000, p. 42-47.

Nous remercions les auteurs pour la qualité de leurs soumissions, le comité de rédaction de la revue TAL et le comité de lecture spécifique à ce numéro :

– Jean-Yves Antoine, Université de Bretagne (France)
– Niels Ole Bernsen, Université de Southern-Denmark (Danemark)
– Johan Bos, Université d'Edimburg (Royaume Uni)
– Jean Caelen, CNRS, Grenoble (France)
– Jean Carletta, Université d'Edimburg (Royaume Uni)
– Claire Gardent, CNRS, Nancy (France)
– Daniel Luzati, Université du Mans (France)
– David Milward, SRI, Cambridge (Royaume Uni)
– Jacques Moeschler, Université de Genève (Suisse)
– Jean-Marie Pierrel, ATILF & LORIA, Nancy (France)
– Massimo Poesio, Université d'Essex (Royaume Uni)
– Owen Rambow, ATT Labs (USA)
– Norbert Reithinger, DFKI (Allemagne)
– Hannes Rieser, Université de Bielefeld (Allemagne)
– Laurent Romary, INRIA, Nancy (France)
– Gérard Sabah, LIMSI (France)
– Paul Sabatier, LIM (France)
– Donia Scott, ITRI, Brighton (Royaume Uni)
– Jacques Siroux, IRISA (France)
– Kees Van Deemter, ITRI, Brighton (Royaume Uni)
– Henk Zeevat, Université d'Amsterdam (Pays-Bas)