

DE LA MODÉLISATION DE L'ACTIVITÉ CONVERSATIONNELLE DES SYSTÈMES DE DIALOGUE PERSONNE-MACHINE

Thierry LEMEUNIER

Laboratoire d'Informatique de l'Université du Maine
Avenue Olivier Messiaen — F72085 Le Mans Cedex 09
Mél : Thierry.Lemeunier@univ-lemans.fr

Résumé

Cet article traite des systèmes de dialogue personne-machine en langue naturelle écrite. Il porte sur la modélisation du dialogue en lui-même et des rapports entre le dialogue et l'application. En m'appuyant sur des travaux de la psychologie de l'interaction, je montre que l'activité conversationnelle est soumise à une logique interlocutoire tandis que l'activité applicative est soumise à une logique formelle. Cette différence importante me pousse ensuite à proposer un modèle informatique de l'activité dialogique fondé sur une représentation du dialogue vue par le système qui permet le partage d'éléments de signification entre ses différentes activités. Cette représentation est constituée de structures d'unités de différents états créées et modifiées au cours du dialogue. Les intentions de communication de la machine sont générées à partir de la forme et de l'état de ces structures.

Abstract

This paper is about human-computer dialog systems with written interactions. It deals with dialog modeling and with the relations between dialog and task-oriented applications. Following several research works in interactional psychology, I show that the conversational activity is managed by an interlocutory logic whereas applicative activity is managed by a formal logic. This important distinction leads me to propose a computational model of the dialogical activity based on a dialog representation viewed by the system. This representation allows to share meaning elements between the various activities of the system. The dialog representation is composed of tree diagrams of units in different states created and modified during conversation. Communicative intentions of the system are generated according to the tree diagrams.

Riassunto

Quest'articolo tratta dei sistemi di dialogo uomo-macchina in lingua naturale scritta. Il testo riguarda la modellizzazione del dialogo in sé e i rapporti fra il dialogo e l'applicazione. Prendendo ispirazione da alcuni lavori di psicologia dell'interazione, l'articolo dimostra che l'attività conversazionale si sottopone ad una logica interlocutoria mentre l'attività applicativa si sottopone ad una logica formale. Di conseguenza, quest'importante differenza condurrà alla proposta di un modello informatico dell'attività dialogica basato su una rappresentazione del dialogo visto dal sistema, che permette di dividere le unità di significato fra le diverse attività del sistema. Questa rappresentazione è costituita da strutture d'unità

create e modificate durante il dialogo. Le intenzioni comunicative della macchina sono generate dalla forma e dagli stati di queste strutture.

Resumen

Este artículo trata de sistemas de diálogo persona-máquina en lengua natural escrita. Conciérne a la modelización del diálogo en sí y de las relaciones entre el diálogo y la aplicación. Apoyándome en los trabajos de la psicología de la interacción, muestro que la actividad conversacional está sometida a una lógica interlocutoria mientras que la actividad aplicativa está sometida a una lógica formal. Esta diferencia importante me incita a proponer un modelo informático de la actividad dialógica basada en una representación del diálogo vista por el sistema, que permite compartir los elementos de significación entre sus diferentes actividades. Estructuras de unidades de diferentes estados creados y modificados a lo largo del diálogo forman dicha representación. Se generan las intenciones de comunicación de la máquina a partir de la forma y del estado de dichas estructuras.

Resumo

Este artigo trata dos sistemas de diálogo pessoa-máquina em língua natural escrita. Ele diz respeito ao modelamento do diálogo ele mesmo e às relações entre o diálogo e a aplicação. Apoiando-me sobre os trabalhos da psicologia da interação, mostro que a atividade conversacional é submissa a uma lógica interlocutória enquanto que a atividade aplicativa é submissa a uma lógica formal. Essa importante diferença me leva em seguida a propor um modelo computacional da atividade dialógica fundamentado sobre uma representação do diálogo vista pelo sistema que permite a partilha de elementos de significação entre suas diferentes atividades. Essa representação é constituída por estruturas de unidades de diferentes estados, criadas e modificadas durante o diálogo. As intenções de comunicação da máquina são geradas a partir da forma e do estado dessas estruturas.

1. Introduction

Les chercheurs de la communauté savent que le développement d'un système de dialogue personne-machine n'est pas une chose aisée. À la complexité technologique inhérente aux grands systèmes logiciels viennent s'ajouter les difficultés théoriques et conceptuelles des thèmes abordés. Faute d'avoir jusqu'à présent un modèle général et viable d'une machine dialoguant de façon vraisemblable, nous sommes amenés à juxtaposer avec plus ou moins de succès les différents modèles nécessaires à une telle entreprise. Pour développer un système de dialogue il faut en effet au minimum que ce système connaisse la langue naturelle de l'utilisateur (ou une langue commune imposée) pour en analyser les énoncés et générer ses propres énoncés, dialogue avec l'utilisateur pour être en mesure d'établir l'intercompréhension nécessaire au déroulement de l'application, maintienne à jour un modèle du locuteur/utilisateur pour s'adapter harmonieusement à celui-ci et, enfin, suive un modèle de l'application pour fournir les services attendus.

Cette complexité peut encore être augmentée si l'on s'intéresse au dialogue oral. Le dialogue oral fait en effet intervenir un autre facteur qui est celui du taux de reconnaissance de l'entrée vocale. Les erreurs du module de la reconnaissance de la parole doivent être prises en compte dans la gestion du dialogue. Ce qui est présenté dans cet article ne concerne que le dialogue écrit, ce qui, j'en suis

conscient, m'éloigne d'un certain nombre de difficultés (mais celles qui restent sont déjà assez nombreuses).

Dans cet article, je vais m'intéresser à la modélisation du dialogue (écrit) et des rapports entre le dialogue et l'application. Dans la deuxième section, je vais tenter de montrer les différences entre ces deux aspects d'un système de dialogue, chacun d'eux étant géré par une logique propre différente l'une de l'autre. Je serai ainsi amené à présenter dans la troisième section une logique interlocutoire pour modéliser l'activité conversationnelle en m'appuyant sur des résultats de la psychologie de l'interaction. Dans la quatrième section, je proposerai ensuite un modèle de fonctionnement de l'activité conversationnelle de la machine qui s'inspire de la logique interlocutoire. Ce modèle s'appuie sur la définition d'une représentation du dialogue dans laquelle des structures d'éléments de différents états sont créées et modifiées au cours du dialogue. Je montrerai ensuite comment les intentions de communication de la machine sont générées à partir de ces structures. Enfin, dans une cinquième section, je présenterai un exemple complet issu d'une implémentation du modèle présenté dans cet article.

2. Les logiques d'un système de dialogue personne-machine

Un système de dialogue personne-machine finalisé est un système matériel et logiciel dans lequel la machine fournit un service ou un ensemble de services à une personne humaine par l'entremise d'une communication dialogale en langue naturelle orale ou écrite. Un tel système présente donc au moins deux aspects devant être simultanément pris en compte. Le premier aspect est celui du dialogue, c'est-à-dire des interactions langagières entre l'utilisateur et le système, le second est celui des services, c'est-à-dire de l'application informatique. Outre les difficultés de chacun des deux aspects pris isolément, d'autres difficultés viennent s'ajouter du fait de leurs fortes inter-relations : d'une part le dialogue porte essentiellement sur l'application (même s'il ne faut pas oublier de prendre en considération les métadiscours), et d'autre part l'application s'appuie sur le dialogue pour rendre efficacement les services proposés.

2.1. Aspect dialogique

Considérons le premier aspect, celui des interactions langagières. Trognon et les membres de son équipe, psychologues de l'interaction au GRC de Nancy¹, ont montré dans leurs travaux un certain nombre des caractéristiques des conversations humaines qui me semblent intéressantes à prendre en considération pour modéliser le dialogue homme-machine. L'idée centrale est que « la rationalité telle qu'elle s'exerce dans les conversations est une rationalité pragmatique ».

Dans (Trognon, 1997), Trognon relate les résultats d'un certain nombre d'expériences menées pour étudier les interactions. Il montre d'abord comment les sujets expérimentaux utilisent par défaut un raisonnement pragmatique. Un individu, compétent dans les règles de la communication, se trouve de fait plongé dans une activité conversationnelle dès lors qu'il est en relation avec un autre individu. En conséquence de quoi, ce ne serait pas sur des *propositions logiques* que

¹ <http://www.univ-nancy2.fr/RECHERCHE/Grc.html>

raisonneraient naturellement les gens mais sur des *illocutions* et selon la rationalité qui gouverne leur usage. En particulier, cela signifie que lorsque les sujets raisonnent et lorsqu'il n'apparaît pas explicitement qu'ils doivent utiliser un raisonnement formel, leurs raisonnements suivent le principe de pertinence énoncé par Sperber et Wilson (1989).

Trognon montre ensuite comment les conversations sont gouvernées à la fois communicationnellement et logiquement. Dans l'expérience de la conservation du volume, un expérimentateur présente à un enfant deux verres identiques contenant la même quantité de liquide. Il lui demande s'ils contiennent la même quantité. Si la réponse est positive, l'expérimentateur prend alors l'un des verres et le vide dans un troisième verre de forme différente, puis il repose la question. Dans de nombreux cas, l'enfant donne alors une réponse non conservante : il dit que le volume a changé alors qu'il connaît parfaitement le principe de la conservation des quantités². D'après l'auteur, ce type de comportement s'explique bien si on adopte une nouvelle fois le point de vue pragmatique : « la répétition de la question peut parfaitement impliquer que la première réponse de l'enfant était incorrecte. Il interprétera donc la nouvelle question comme une demande de rectification et proposera cette fois une réponse non conservante ». Pour suivre son raisonnement pragmatique, le sujet a passé outre sa connaissance logique des choses du monde. Ce comportement diminue pour les sujets plus âgés lorsqu'ils ont appris à se méfier des questions pièges.

Enfin, dernier élément de la rationalité pragmatique, Trognon montre comment les problèmes survenant dans le dialogue sont gérés au fur à mesure qu'ils apparaissent. Ceci vaut aussi bien pour les problèmes communicationnels (liés à la transmission d'informations), conversationnels (liés au maintien de la cohérence sociale de l'interaction) et logiques (liés à la cohérence de l'argumentation). Le principe qui apparaît ici est celui de la séquentialité, c'est-à-dire le fait que dans toute interaction les problèmes ne peuvent être et ne sont traités que les uns après les autres et seulement si leurs traitements sont nécessaires à la poursuite de la conversation. Le traitement d'un problème n'est d'ailleurs pas exhaustif mais juste suffisant à débloquer la situation.

Les résultats des travaux en psychologie montrent ainsi que la conversation est une activité essentiellement non déterministe et opportuniste, car les effets des actions langagières du locuteur sur l'interlocuteur ne sont jamais entièrement prévisibles, même si le locuteur, à la suite de son énonciation, peut s'attendre à un certain nombre d'effets sur l'interlocuteur. La rationalité des conversations doit donc être traitée par une logique propre aux interactions langagières, une *logique interlocutoire* que nous présentons en troisième section.

2.2. Aspect dialogique et aspect logique

Considérons maintenant le second aspect d'un système de dialogue homme-machine évoqué précédemment, celui des services applicatifs rendus. La machine doit fournir un service exempt de dysfonctionnement. L'idée est ici de profiter de la

² La preuve de cette connaissance chez les sujets est d'ailleurs apportée expérimentalement. Lorsque l'expérimentateur qui a posé la question et fait le transvasement est « appelé » dans une autre pièce et qu'il est remplacé par un second expérimentateur qui repose la question sans être censé savoir que le premier l'a déjà posée, le nombre de réponses conservantes augmente beaucoup.

puissance de calcul des machines et de leur grande vitesse d'exécution. Ce second aspect fait ainsi apparaître un autre thème de recherche bien différent du précédent qui porte lui sur la modélisation des connaissances sur les services à effectuer par la machine. Des techniques issues de l'intelligence artificielle telles que la planification (Cohen *et al.*, 1990) ou plus récemment le formalisme tâche-méthode peuvent alors être utilisées à bon escient (Trichet et Tchounikine, 1999). La logique utilisée dans ce cadre doit être une logique formelle pour s'assurer d'un fonctionnement efficace.

Nous sommes donc en présence de deux contraintes qui sont de par leur nature assez contradictoires. Je pense que la contrainte d'une rationalité pragmatique imposée par les interactions en langue naturelle doit être respectée pour que le système soit socialement accepté et effectivement utilisable et utilisé. Cette première contrainte rentre en conflit avec la seconde contrainte d'une logique formelle qui elle est imposée par le but applicatif du système. Deux logiques doivent donc co-exister au sein du système de dialogue homme-machine.

Pour autant, il me semble que depuis quelques années l'aspect applicatif a pris une part plus prépondérante que l'aspect interactionnel dans les recherches sur le dialogue homme-machine. Les chercheurs de la communauté ont ainsi développé de nombreux modèles de dialogue fondés sur la notion de plan. Par exemple, en France, des systèmes ont été développés sur ce principe au LIMSI (Vilnat, 1997) et à l'IRISA (Guyomard *et al.*, 1998). Dans cette approche, qui s'appuie par ailleurs sur la théorie des actes de langage dont Austin fut le principal initiateur (Austin, 1962), les actions langagières sont modélisées dans le même formalisme que les actions non langagières. Cette uniformisation de la notion d'acte a permis de grandes améliorations dans les techniques de raisonnement logique mais conduit aussi à un défaut majeur qui est celui d'une trop grande rigidité des échanges. Cette rigidité est la conséquence du *manque d'autonomie* de l'activité langagière par rapport à l'activité applicative du système (Lemeunier, 2000). Tout se passe comme si, en dehors du strict déroulement des questions-réponses imposées par la recherche d'informations, le système devenait muet.

Cette dernière critique a été faite par certains auteurs qui rejettent une approche purement logique et préconisent plutôt des approches mixtes. Maudet affirme que :

« L'influence des circonstances locales n'est pas prise en compte de manière satisfaisante par le formalisme des plans, et il n'est guère envisageable de replanifier au cours du dialogue à chaque intervention de son interlocuteur » (Maudet, 2001).

L'auteur s'intéresse aux approches mixtes conciliant approche intentionnelle de type BDI³ et approche conventionnelle impliquant d'autres attitudes mentales, comme les obligations, liées à la dimension sociale des interactions langagières. Il propose ainsi un modèle de jeux de dialogue à partir de la notion d'engagement, un jeu de dialogue étant une macrostructure d'interaction.

Pour Rosset, une approche mixte est nécessaire pour parvenir à une plus grande souplesse d'utilisation et d'expression :

³ L'approche BDI (Belief Desire Intention) est une approche intentionnelle du dialogue qui utilise une logique multimodale alors que l'approche par plan s'appuie sur des relations de cause à effet entre actions. Ces deux points de vue ne sont pas contradictoires mais complémentaires et souvent utilisés conjointement.

« Ces approches tendent à considérer le dialogue comme une suite d'actions simples et prévisibles. Or le dialogue, dans la mesure où il procède d'une co-construction et d'une collaboration entre une machine (pour un système de DOHM) et un humain, ne peut être que partiellement planifié, les utilisateurs étant susceptibles de ne pas suivre le plan fixé a priori » (Rosset, 2001).

L'auteur a donc développé une approche mixte : le gestionnaire de dialogue s'inspire du modèle de dialogue de Luzzati (1995) et le modèle de la tâche s'appuie sur une approche « classique » orientée plan. Ceci a ainsi permis au système ARISE de dialogue oral de passer d'un taux de succès de dialogue de 53 % à 85 %.

Comme ces auteurs, et compte tenu du constat précédent sur la nécessité de la co-existence des deux logiques, je pense également qu'une approche mixte est essentielle pour améliorer les systèmes actuels. Établir un plan puis effectuer les actions qui le composent est nécessaire pour atteindre un but, et c'est la meilleure méthode pour y parvenir si le monde est parfaitement connu. Cependant, selon moi, mettre les actions langagières (de l'activité conversationnelle) et les actions non langagières (de l'activité applicative) sur un même niveau est une erreur de modélisation, car je pense que les relations de cause à effet identifiables pour les actions sur le monde n'existent pas pour les actions langagières qui, les travaux en psychologie l'ont montré, ne sont pas soumises à de telles lois.

Le point de vue que j'ai pris est différent des deux auteurs précédents. Le modèle de la tâche s'inspire des approches par plan (planification dynamique) mais le modèle de dialogue s'inspire de la logique interlocutoire que je vais maintenant présenter.

3. La logique interlocutoire

La logique interlocutoire tente de rendre compte des enchaînements dialogiques observés dans les conversations humaines. Issue des travaux en psychologie de l'équipe de Trognon, elle a maintenant été reprise en informatique par Nicolle (Nicolle et Saint-Dizier De Almeida, 1999). Ces derniers auteurs en proposent une version identique dans son principe mais plus opératoire et qui s'appuie sur une redéfinition d'un certain nombre de concepts et en ajoute un certain nombre. En particulier, afin d'explicitier les mécanismes de raisonnements pragmatiques, les auteurs introduisent la notion bien connue en psychologie cognitive de terrain commun (Caron, 1997). De plus, nous verrons que le modèle initial de Trognon s'appuie sur la théorie des actes de langage et de la logique illocutoire de Vanderveken (Vanderveken, 1990) et que Nicolle en propose une version modifiée.

Aussi, après une présentation du principe de la logique interlocutoire, je présenterai les éléments les plus importants de la version de Nicolle sur lesquels je m'appuie dans la section suivante de l'article pour proposer un modèle de fonctionnement de l'activité dialogique.

3.1. Principe de la logique interlocutoire

La logique interlocutoire tente de rendre compte de la *construction dynamique* et *conjointe* du sens en interaction. Montrons au travers d'un exemple tiré de (Brassac et Pesty, 1999) ce que construction conjointe signifie. Il s'agit de quatre séquences (reproduites tableau 1) se déroulant dans un même contexte.

Tableau 1 — Illustration de la logique interlocutoire tirée de Brassac et Pesty (1999).

Séquence 1		Séquence 2	
E1 Tu as le téléphone ici		E1 Tu as le téléphone ici	
L1 Oui, c'est moderne		L1 Oui, c'est le numéro 04-83-35-36-09	
E2 Ah je n'aurais pas cru		E2 Ah ben je pourrais t'appeler comme ça	
Séquence 3		Séquence 4	
E1 Tu as le téléphone ici		E1 Tu as le téléphone ici	
L1 Oui, c'est moderne		L1 Oui, c'est le numéro 04-83-35-36-09	
E2 (rires) Ah ben j'aurais bien aimé que tu me donnes le numéro		E2 Euh, mais je ne demandais pas le numéro	

Ces quatre séquences peuvent toutes s'être réellement passées car elles sont toutes interprétables⁴. Par exemple, dans les séquences 1 et 3, l'énoncé E1 peut être une requête d'information paraphrasée par « je te demande si tu as le téléphone ». En revanche dans les séquences 2 et 4 l'énoncé E1 doit plutôt être considéré comme une requête d'action. Il est paraphrasé par « je te demande ton numéro de téléphone ». Comment expliquer que ces quatre séquences soient différentes ? Ce qui les différencie est uniquement la manière dont l'interlocuteur L interprète l'énoncé initial E1. S'il l'interprète comme une requête d'information (séquences 1 et 3) son interprétation en actes (énoncé L1) est soit validée (énoncé E2 de la séquence 1) soit invalidée (énoncé E2 de la séquence 3). Il en est de même pour les séquences 2 et 4 : son interprétation de E1 comme requête d'action est validée dans la séquence 2 et invalidée dans la séquence 4. Ce qui revient à dire que dans les quatre cas, l'interprétation de l'énoncé E1 par L n'est qu'hypothétique, et qu'elle nécessite alors d'être évaluée à l'énoncé E2.

En règle générale, le sens d'un énoncé peut nécessiter d'autres négociations ou même ne jamais être totalement partagé. En effet, dans certaines situations le manque d'information peut conduire à ce qu'un observateur extérieur ne puisse pas donner une interprétation acceptable pour lui. La logique interlocutoire permet justement aux énoncés de ne pas prendre un statut fixe. Non seulement un même énoncé peut prendre des sens différents dans des situations différentes, mais, également, le sens co-construit peut différer selon les locuteurs dans des situations similaires.

Finalement, le principe de la logique interlocutoire est résumé par Bange, ethnométhodologue, de la manière suivante (A et B sont les deux participants et B interprète l'énoncé de A) :

« Cette interprétation de B, qui reste le plus souvent implicite dans la réponse qu'il donne à A, est en réalité une offre d'interprétation adressée au locuteur A qui l'examine pour voir si elle est une interprétation acceptable pour lui et son énonciation initiale. Si A n'a pas de raison massive de refuser, il accepte l'offre d'interprétation de B et cette acceptation est le plus souvent implicite dans la réaction de A à la réponse de B. Cette réaction de A à la réponse de

⁴ Interprétable signifie que l'analyste peut expliquer l'enchaînement conversationnel. Seule la première séquence s'est réellement déroulée.

B peut donc être comprise comme comportant une évaluation par A de l'interprétation faite par B sur son énonciation initiale à lui A. [...] C'est dans ce jeu de réciprocité que le sens se constitue » (Bange, 1992).

Autrement dit, dans le meilleur des cas, la valeur conversationnelle de l'acte initial E1 de A n'est effective qu'après la réponse de l'interlocuteur B car elle est fonction de cette réponse. Dans les autres cas, elle est sujette à négociation si l'interprétation en actes de E1 produit par B est trop éloignée de celle intentionnée par A.

3.2. Le terrain commun

Pour expliciter le travail de négociation que mènent les interactants, il est nécessaire d'introduire un concept important, celui de *terrain commun*, qui peut être vu dans une première approximation comme une sorte de « tableau de travail » qui constitue le moyen de l'intercompréhension. Lorsque les éléments de ce tableau sont « marqués » comme étant plus ou moins fortement partagés, le locuteur peut considérer que l'intercompréhension est atteinte.

La notion de terrain commun a été étudiée en particulier par Clark (1992). Il s'agit « d'un ensemble de connaissances, hypothèses et croyances partagées par les partenaires, et présumées telles par chacun » (Caron, 1997). Clark distingue différentes connaissances susceptibles d'être partagées dans le terrain commun :

- les connaissances liées à la co-présence physique de la situation de communication ;
- les connaissances liées à la co-présence linguistique (connaissances propres à la mémoire de ce qui a été dit) ;
- les connaissances communes aux membres d'une même communauté.

Le terrain commun est construit au cours du dialogue. L'image que s'en fait chaque interactant intervient dans l'interprétation des énoncés et s'en trouve modifiée. Il est construit au fur et à mesure que les participants ajustent leur compréhension réciproque en voyant les conséquences de leurs actions. Il s'agit d'un mécanisme itératif : certains éléments préexistent avant même que le dialogue commence. Ces éléments par défaut sont ensuite complétés, modifiés, enlevés, etc. De plus, l'image du terrain commun est subjective puisque chaque participant en a une représentation propre. Enfin, il permet aussi d'éviter ou de lever les ambiguïtés rencontrées.

Le terrain commun est donc une condition nécessaire au fonctionnement du dialogue et il dépend de ce dialogue pour être construit (Vivier, 1996). Ainsi, l'intercompréhension et la construction du terrain commun se conditionnent réciproquement.

3.3. Négociation du sens et valeur interlocutoire

La version initiale du modèle de la logique interlocutoire s'appuie sur la théorie des actes de langage initié par Austin (Austin, 1962) et sur une « dialogisation » de la sémantique formelle générale de Vanderveken (Vanderveken, 1990).

Sans retracer toute l'histoire de cette branche de la philosophie du langage qui est à la base de nombreux modèles de dialogues actuels, je me contenterai de rappeler que, selon cette théorie, parler est la réalisation d'un acte qu'on appelle acte de langage. Selon Austin, un acte de langage peut se décomposer lui-même en trois

actes : l'acte locutoire qui est le fait d'énoncer, l'acte illocutoire qui est l'acte réalisé en produisant l'énoncé pour exprimer ses intentions et l'acte perlocutoire qui sont les conséquences sur l'auditeur de l'acte illocutoire.

Searle et Vanderveken (1985) ont proposé ensuite une logique illocutoire s'appuyant, entre autres, sur une définition des conditions à l'accomplissement réussi d'un acte de langage. Un acte de langage est alors vu comme une certaine force illocutoire qui peut être de cinq types (assertive, déclarative, directive, expressive et commissive) s'appliquant à un contenu propositionnel. Vanderveken a ensuite unifié ces conditions de succès aux conditions de satisfaction d'un acte qui définissent les conditions dans lesquelles un acte peut être satisfait. En effet, un acte réussi n'est pas forcément satisfait. Par exemple, lorsque le locuteur réalise une promesse, il le fait avec succès s'il remplit les conditions de réussite (être dans la capacité de faire en sorte que le contenu propositionnel devienne vrai, être sincère, etc.) mais ces conditions ne prédisent rien sur le fait que la promesse sera tenue ou pas. Ce sont les conditions de satisfaction qui permettent de le dire. Ces notions importantes de réussite et de satisfaction d'un acte de langage sont reprises ci-après et ultérieurement dans le modèle que je propose.

La version de Vanderveken de la théorie des actes de langage, se fonde sur une nature vériconditionnelle du langage naturel. Nicolle pense au contraire qu'un énoncé ne peut pas être réduit et évalué selon une logique binaire : un énoncé peut être faux, vrai, de valeur inconnue, vrai pour une personne et faux pour une autre, vrai un jour et faux un autre jour. De plus, selon l'auteur, ce que les personnes partagent ce sont des états mentaux et des croyances plus ou moins fortes. De ce point de vue, la réalité n'est pas la même pour tous les gens et la notion de vérité doit être au mieux relativisée. C'est cette réinterprétation de la théorie des actes de langage que je présente maintenant dans laquelle intervient la notion de valeur interlocutoire.

En reprenant les notions d'actes locutoires, d'actes illocutoires et d'actes perlocutoires, les auteurs expliquent que l'acte illocutoire est inféré à partir de l'acte locutoire (ensemble d'indices mis dans le terrain commun) et des effets perlocutoires que l'allocutaire imagine être intentionnés par le locuteur. L'allocutaire produit alors une interprétation en actes, résultat de son raisonnement, qui sera évaluée par le locuteur.

Explicitons plus précisément les éléments intervenant dans le travail interprétatif. Tout d'abord l'allocutaire doit avoir connaissance des termes employés par le locuteur. Si ce n'est pas le cas, il peut soit poser une question d'éclaircissement, soit poser une hypothèse de signification. Ensuite, il recherche des éléments d'intentionnalité et de sincérité du locuteur. Ici, les auteurs mentionnent comme indices : les unités lexicales, le mode et le temps, la structure syntaxique, la prédication, etc. Interviennent également des éléments de la relation sociale (type, objectifs et finalité de l'interaction) et des éléments du terrain commun (ce qui s'est passé avant et ce qui risque de se passer après, notamment les fonctions illocutoires et interactives potentielles de l'acte).

Les auteurs reformulent une taxinomie des forces illocutoires : la force assertive lorsque les mots s'ajustent au monde, la force déclarative lorsque, du fait de l'énonciation, les mots et le monde s'ajustent, la force directive lorsque le monde doit s'ajuster aux mots par l'action de l'allocutaire et enfin la force commissive lorsque le monde doit s'ajuster aux mots par l'action du locuteur.

La valeur illocutoire attribuée est transmise par l'allocutaire du fait de ses réactions : si elle est validée par le locuteur, c'est que les effets perlocutoires inférés sont acceptés. Dans le cas contraire, l'interprétation en actes n'est pas ratifiée et une négociation est ouverte explicitement. Dans les deux cas, une intercompréhension doit finalement émerger : c'est *la valeur interlocutoire*.

En reprenant la notion de réussite d'un acte de langage, les auteurs affirment qu'un acte illocutoire est réussi lorsqu'il lui est attribué une valeur interlocutoire. Ils définissent également un ensemble de contraintes sur la satisfaction d'un acte. Elles portent sur l'acceptabilité et la satisfiabilité nécessaire à la satisfaction d'un acte :

[C1] L'acceptabilité. Un acte réussi est acceptable si l'allocutaire accepte la position sociale qu'occupe le locuteur en énonçant l'acte.

[C2] La capacité à satisfaire. Un acte réussi est satisfiable :

[C2a] si les capacités du locuteur (pour un commissif, un assertif) ou de l'allocutaire (pour un directif) permettent la satisfaction de l'acte ;

[C2b] si l'état du monde permet la satisfaction de l'acte.

[C3] La volonté de satisfaire. Un acte réussi est satisfiable si les attitudes mentales de l'allocutaire autorisent la satisfaction de l'acte :

[C3a] si l'allocutaire croit en la sincérité du locuteur ;

[C3b] si l'allocutaire croit en la véracité du contenu communiqué ;

[C3c] si l'allocutaire a le désir et l'intention de satisfaire l'acte.

4. Activité dialogique d'un système de dialogue

Nous venons de voir que la logique interlocutoire pour les interactions dialogiques et la logique formelle pour les services rendus sont deux aspects incontournables pour un système de dialogue homme-machine. La logique interlocutoire intervient dans l'activité conversationnelle et la logique formelle intervient dans l'activité applicative. Je propose que la liaison entre ces deux activités soit assurée par un élément d'interface : *la mémoire du dialogue* vu par le système de dialogue homme-machine.

Avec cette interface, les deux modes de fonctionnement restent indépendants mais sont en liaison *via* les éléments partagés dans la mémoire du dialogue. Il ne s'agit donc pas d'une liaison directe et statique puisque les deux modes de fonctionnement sont différents (l'un suit une logique formelle et l'autre suit une logique interlocutoire), mais d'une liaison indirecte, dynamique et asynchrone car elle se fait *via* les éléments d'interface, seulement lorsque cela est nécessaire et enfin de façon séquentielle (un élément n'est manipulé que par une seule activité à la fois).

L'activité conversationnelle et l'activité applicative ont accès pareillement à la mémoire mais l'usage qu'elles en font est différent. Elles consultent en effet les éléments de la mémoire et prennent en compte l'état et la structuration des éléments présents mais elles effectuent des raisonnements différents : l'une sur l'avancement de l'application et l'autre sur l'avancement du dialogue. On verra que le but de l'activité applicative est de tendre vers une satisfaction des éléments présents dans la mémoire du dialogue alors que le but de l'activité interactionnelle est de tendre vers un partage des éléments présents dans la mémoire du dialogue.

J'ajouterai dans cette introduction que le terme « mémoire » a été choisi non pas pour revendiquer une quelconque relation avec la notion de mémoire de travail en psychologie (Tiberghien, 1997), mais simplement pour désigner le fait qu'il s'agit d'un lieu de stockage et de représentation informatique actif (c'est-à-dire amené à être modifié) d'éléments intervenus au cours du dialogue. Par contre, on verra que la mémoire du dialogue recouvre en partie la notion de terrain commun vue précédemment. Dans le même ordre d'idée, on verra également qu'il ne s'agit pas non plus de représenter une structure éventuelle du dialogue mais bien de fournir un moyen aux processus de l'activité dialogique de s'effectuer.

Je vais donc dans cette partie de l'article définir la mémoire du dialogue en introduisant les éléments de cette mémoire et leurs structurations. Je décris ensuite les processus liés à cette mémoire (créations et modifications des éléments et des structures) en décrivant l'activité dialogique du système. Je montre ainsi les processus d'interprétation puis les processus de génération des intentions de communication de la machine.

4.1. Définition de la mémoire du dialogue

La mémoire du dialogue est un espace de travail permettant la jonction des activités du système en supportant le partage des représentations qui y sont maintenues. Les activités dont il est question ici sont : l'activité langagière (analyse des énoncés de l'utilisateur et la génération des énoncés de la machine), l'activité applicative qui justifie l'existence du système (aide ou renseignement dans un domaine particulier, résolution de problèmes, etc.) et l'activité dialogique (cohérence des échanges, maintien du canal de communication, gestion des erreurs, etc.).

Plus précisément, nous définissons la mémoire du dialogue de la façon suivante : il s'agit d'un ensemble de structures arborescentes de composants élémentaires appelés Unités Minimales de Mémoire (UMM). Ces composants élémentaires sont activés en mémoire, selon le déroulement du dialogue, par les différentes activités du système (ici je ne développe en détail que l'activité dialogique).

4.1.1. Les unités minimales de mémoires (UMM)

L'UMM est une entité symbolisant les éléments de signification activés dans la mémoire du dialogue. On y trouve les éléments du terrain commun mais pas uniquement. Cette unité réunit, sous une même désignation, d'une part les objets que la machine escompte partager, et d'autre part les objets qu'elle suppose être partagés.

Formellement, une UMM est un objet (instance de classe) possédant des champs comme, en particulier, un état pouvant prendre un certain nombre de valeurs, une liste des UMM qui lui sont liées ainsi que d'autres informations nécessaires à la génération des intentions de communications de la machine.

4.1.2. Les structures d'actions

Les UMM sont organisées en *structure d'action* du type action / actant(s) / circonstant(s). Tesnière a défini ce type de structure pour la syntaxe des phrases (Tesnière, 1959). J'ai repris ce modèle de grammaire de dépendance en l'appliquant au niveau pragmatique : les unités d'une structure sont porteuses de significations sémantiques ; les référents et les référés doivent être attachés les uns aux autres ; les anaphores et les ellipses résolues.

Dans ce cadre, les actants sont les sujets ou les objets intervenant dans l'action. Les circonstants sont les circonstances du déroulement de l'action. Enfin, l'action met en relation les actants et les circonstants. Elle fait référence à un schéma d'action qui lui est pris en charge par l'activité applicative (non présentée dans cet article).

Les structures sont des arborescences permettant d'organiser les UMM en fonction des contraintes qui les lient : les éléments obligatoires d'une action sont groupés ensemble et les éléments optionnels également. Ceci se fait en typant les liens entre nœuds pères et nœuds fils des deux manières suivantes :

- Un nœud père maintient une relation de disjonction⁵ avec ses nœuds fils optionnels : l'arborescence est dite *en OU*. Ce type de lien est représenté graphiquement figure 1 et noté linéairement sous forme fonctionnelle infixée : $\langle X \rangle (\langle Y \rangle \vee \langle Z \rangle)$.



Figure 1 — Une arborescence en OU.

- Un nœud père maintient une relation de conjonction avec ses nœuds fils obligatoires : l'arborescence est dite *en ET*. Ce type de lien est représenté graphiquement figure 2 et noté sous la forme : $\langle X \rangle (\langle Y \rangle \wedge \langle Z \rangle)$.



Figure 2 — Une arborescence en ET.

Une structure d'action peut comprendre plusieurs niveaux mais la complexité peut être limitée selon les besoins de l'application. Par exemple, lorsque le système ne comprend pas un mot ou un groupe de mots et qu'il est capable de poser une hypothèse lexicale, il doit confronter ce mot ou ce groupe de mots avec son hypothèse afin d'identifier la catégorie de ce mot. Dans l'application AMI qui simule un standardiste téléphonique (Lemeunier, 2000), cette action est représentée par la structure d'action de la figure 3.

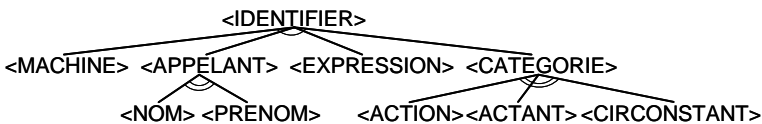


Figure 3 — Exemple de la structure d'action <IDENTIFIER>.

⁵ Le modèle permet de distinguer les liaisons disjonctives exclusives des liaisons disjonctives inclusives. Le modèle permet également de prendre en considération les arbres ET-OU. Ces points ne sont pas développés dans cet article.

L'action <IDENTIFIER> possède quatre actants obligatoires {<MACHINE> ; <APPELANT> ; <EXPRESSION> ; <CATEGORIE>}. L'actant <CATEGORIE> peut prendre l'une des valeurs de l'ensemble {<ACTION> ; <ACTANT> ; <CIRCONSTANT>}. L'actant <APPELANT> peut prendre une ou les valeurs de l'ensemble {<NOM> ; <PRENOM>}.

4.2. Fonctionnement de la mémoire du dialogue

Le système de dialogue fonctionne selon un cycle de deux phases qui se répète tant que la machine a des intentions de communication. Les différentes activités de la machine — dont l'activité dialogique et l'activité applicative — interviennent au cours de ces deux phases. Elles modifient à leur gré la ou les structures d'action activées en mémoire (on peut alors parler de forêt puisqu'il peut y avoir plusieurs arborescences actives au même moment). Les deux phases sont les suivantes :

- Des intentions de communication de la machine sont générées selon l'état et la structuration des UMM, puis des heuristiques de tris et de filtrages sur les intentions générées sont appliquées. Cette partie est assurée par l'activité dialogique. Un énoncé est généré à partir des intentions candidates. La génération de la forme de surface de l'énoncé est assurée par l'activité langagière.
- L'énoncé de l'utilisateur est analysé puis interprété. L'analyse sémantique est assurée au début de cette seconde phase par l'activité langagière. Si l'analyse n'aboutit à aucun résultat et si l'activité langagière est capable de faire des hypothèses d'analyse, elle va créer une structure d'action d'identification des mots non catégorisés. Ce principe hypothético-déductif a été étudié dans la thèse de Lehuen (1997) ; il n'est pas développé dans la suite de cet article. L'interprétation est effectuée par l'activité dialogique tandis qu'en parallèle les activités développent, dès qu'elles peuvent le faire, les structures d'action dont elles sont responsables.

4.2.1. L'interprétation et les états des unités de mémoire

Pour un agent écoutant, l'interprétation consiste à donner une signification plausible à un énoncé compte tenu des éléments connus et inférés du contexte courant. Pour ce qui est d'un système de dialogue, je propose de dire que l'interprétation consiste à confronter le résultat de l'analyse sémantique avec les UMM présentes en mémoire du dialogue. En tenant compte du fait que l'énoncé peut être porteur d'un seul sens ou de plusieurs sens et que ces significations sont ou ne sont pas rattachables aux UMM déjà activées, il existe quatre cas à considérer :

- Le cas le plus favorable est lorsque l'analyseur attribue un seul sens à l'énoncé et que certains éléments de signification de ce sens correspondent aux UMM d'une arborescence. Dans ce cas, il y a *appariement* c'est-à-dire passage d'un état à un autre état des UMM appariées. Les différents états que peut prendre une UMM sont expliqués plus loin.
- Un second cas survient lorsque l'analyseur attribue plusieurs sens à l'énoncé et qu'un seul d'entre eux correspond aux UMM d'une structure d'action. Comme dans la situation précédente, les UMM sont appariées aux éléments de signification coïncidents, mais ce choix ne correspond pas forcément au sens intentionné par l'utilisateur. La réponse de l'utilisateur validera ou invalidera l'interprétation en actes de la machine qui résultera de ce choix.

- Le troisième cas apparaît lorsque l'analyseur attribue un seul sens à l'énoncé mais qu'aucun élément de signification ne peut s'apparier. Je qualifie ce cas d'analyse hors contexte. Cet énoncé doit être quand même pris en compte : des UMM sont créées et elles seront éventuellement attachées à une structure d'action ultérieurement.
- Enfin, le dernier cas survient lorsqu'aucun des sens attribués à l'énoncé de l'utilisateur ne peut être apparié. Il existe dans ce cas deux possibilités : soit le système fait une hypothèse sur l'intention de l'utilisateur, et on revient alors au second cas, soit le système pose directement une question pour choisir parmi les résultats de l'analyseur (cette seconde option est plus simple à gérer et donne une meilleure robustesse mais est moins naturelle).

Le travail interprétatif de la machine aura donc pour conséquence la création de nouvelles UMM ou un changement de l'état des UMM déjà présentes en mémoire. Je vais maintenant parler des états des unités de la mémoire du dialogue. Ceux-ci ont été définis à partir de la notion de valeur interlocutoire et des contraintes de satisfiabilité des actions langagières telles qu'elles sont définies dans la version de Nicolle de la logique interlocutoire présentée à la section suivante.

Une UMM peut prendre successivement trois états différents : l'état *en attente*, l'état *réussi* et l'état *satisfait*. L'état en attente est l'état par défaut à la création d'une unité. Lorsqu'une UMM est créée par une des activités du système cela signifie que le système s'attend à la présence de cette unité dans le terrain commun. La création d'une UMM peut être soit une initiative propre du système soit le résultat d'une analyse hors contexte. *Une UMM <X> en attente est notée X sans signe distinctif tandis qu'une UMM dont l'état est inconnu est notée <X>*.

Un acte de langage est réussi lorsque les interactants lui attribuent une valeur interlocutoire, c'est-à-dire lorsqu'ils se sont mutuellement compris. Cette intercompréhension implique le partage des éléments de signification. Dans notre cas cela signifie concrètement que des UMM en attente passent à l'état réussi. En effet, lorsque la machine s'attend à ce qu'un élément soit dans le terrain commun (cet élément est alors en attente) et que cet élément apparaît dans l'analyse sémantique de l'énoncé de l'utilisateur, cela signifie que la machine et l'utilisateur ont en partage cet élément de signification. *Une UMM réussie <X> est notée [X]*.

Enfin, un acte de langage réussi est satisfiable si les trois contraintes C1 à C3 du modèle de Nicolle présenté ci-avant sont satisfaites. Cet aspect doit être précisé car il ne peut être que partiellement pris en compte dans l'état actuel du modèle de l'activité dialogique.

Tout d'abord, j'ai fait certaines simplifications : premièrement, je considère que la machine accepte toujours la position sociale qu'occupe l'utilisateur, ce qui satisfait C1 ; deuxièmement, je fais la simplification de dire que l'utilisateur est toujours sincère, ce qui satisfait C3a ; troisièmement, le système a, par principe (voir après pour une explication), toujours le but de satisfaire les demandes de l'utilisateur et donc le désir et l'intention de satisfaire l'acte de l'utilisateur, ce qui satisfait C3c.

Ensuite, la contrainte C3b portant sur la croyance en la véridicité du contenu propositionnel de l'énoncé n'est pas du tout prise en compte actuellement. Cela revient à dire que l'utilisateur n'énonce rien qui puisse entrer en conflit avec les connaissances du système. Cette simplification est une hypothèse forte sur l'usager mais ne modifie en rien le principe du modèle tel qu'il est présenté ici.

Enfin, en ce qui concerne la contrainte C2 sur la capacité à satisfaire, elle est satisfaite si l'état de la tâche applicative le permet. Plus précisément, les UMM dans l'état réussi représentant des actants et des circonstants passent dans l'état satisfait si elles sont référencées par le système (ou simplement catégorisées⁶), et les UMM dans l'état réussi représentant des actions passent à l'état satisfait si elles sont faisables par le système compte tenu de ses connaissances sur le monde. Une UMM $\langle X \rangle$ satisfaite est notée \boxed{X} .

Il faut ajouter à tout cela que le passage de l'état réussi à l'état satisfait des UMM d'une arborescence se fait par propagation du bas vers le haut de la structure d'action : une UMM ne peut être dans l'état satisfait que si toutes les autres UMM qui lui sont liées conjonctivement sont dans l'état satisfait ou, au moins, si l'une des UMM qui lui est liée disjonctivement est dans l'état satisfait. Ces dépendances des états entre nœud père et nœud fils entraînent que certaines structurations sont normalement impossibles à obtenir, comme par exemple $\boxed{X}(Y \wedge [Z])$ dans laquelle l'UMM racine $\langle X \rangle$ notée \boxed{X} est satisfaite et possède deux fils Y et [Z] non satisfaits.

4.2.2. Règles de dialogue et intentions de communication

Le système de dialogue poursuit deux buts : un but conversationnel qui consiste à la construction d'un terrain commun avec le locuteur humain pour que ce dernier et la machine parlent des mêmes choses et un but applicatif qui consiste à satisfaire les demandes de l'utilisateur humain. La notion de satisfaction doit être appliquée aussi bien pour les attentes que pour les actes langagiers du locuteur. Lorsque ces derniers sont satisfaits, dans le sens de la satisfaction des actes de langage défini par Vanderveken et repris par Nicolle, c'est que les services attendus sont rendus.

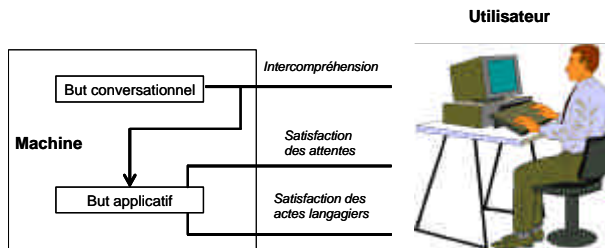


Figure 4 — But conversationnel et but applicatif.

Ce que le système doit dire doit être orienté vers le premier but, celui qui vise à atteindre l'intercompréhension. Le but applicatif n'intervient pas directement dans la génération des intentions de communication même si la satisfaction des demandes de l'utilisateur passe nécessairement par l'accomplissement du but conversationnel. La satisfaction des actes de langage de l'utilisateur n'est aucunement un but en soi. Elle est une conséquence de l'accomplissement du but applicatif.

⁶ Une UMM est catégorisée lorsqu'une catégorie sémantique lui est attribuée. Par exemple « Jean » est catégorisé comme un prénom, « Dupont » comme un nom. En revanche, l'UMM $\langle \text{APPELANT} \rangle$ est catégorisée comme la personne qui appelle et référencée comme l'unique personne s'appelant Jean Dupont.

Afin d'expliciter cet aspect j'ai commencé par définir des règles comportementales et interactionnelles qui justifieront ensuite les choix pris pour le processus de génération des intentions de communication effectué par l'activité dialogique, cette dernière se faisant en toute indépendance des autres activités du système.

La première règle est une règle comportementale générale. Elle régit les croyances et les désirs du système. Elle permet l'accomplissement des actions représentées en mémoire et par effet de bord la satisfaction des actes de langage de l'utilisateur.

Première règle : un système de dialogue homme-machine finalisé collaboratif doit faire tout ce qui est de son ressort pour satisfaire l'utilisateur.

La deuxième et la troisième règle gouvernent le comportement interactionnel du système. Elles sont le « moteur » de l'interaction du système avec son environnement car elles le poussent à dialoguer afin d'atteindre le but conversationnel poursuivi.

Deuxième règle : un système de dialogue homme-machine finalisé ne doit informer l'utilisateur que de ce que le système croit que l'utilisateur ne sait pas.

Troisième règle : un système de dialogue homme-machine finalisé doit s'informer de ce qu'il ne sait pas auprès de l'utilisateur.

Ces deux règles régissent la génération des intentions de communication. Elles forment en quelque sorte une version extrêmement minimaliste des maximes de la conversation de Grice (Moeschler et Reboul, 1994).

La deuxième règle gère la direction de la communication de la machine vers l'utilisateur dans la mesure où c'est de la machine que vient l'information à partager. Généralement, cette règle donne lieu à l'énonciation d'actes de force illocutoire assertive. La troisième règle gère la direction de l'utilisateur vers la machine car l'information est détenue initialement par l'utilisateur. Cette règle donne lieu à l'énonciation d'actes de force illocutoire directive.

4.2.3. La génération des intentions de communication

Les intentions de communication proviennent de l'application des deux règles interactionnelles aux structures d'actions présentes en mémoire du dialogue à un instant donné de l'interaction. La principale difficulté consiste à garantir la pertinence des énoncés porteurs des intentions issues du processus de génération. Les deux règles interactionnelles ne suffisent pas car elles sont générales et non spécifiques à un contexte particulier. L'état et la structuration des UMM doivent être pris en compte en détail.

L'étude des structurations normales pose un problème du fait du nombre de cas à considérer. Il est en effet impossible de déterminer ce nombre *a priori* puisque cela dépend de l'application et du dialogue, et, même si en théorie ce nombre n'est pas infini, il reste tout de même très grand en pratique. Par exemple, en tenant compte des liens optionnels et obligatoires entre père et fils, il existe 25 920 arborescences de hauteur trois dont chaque nœud a trois fils. L'étude de l'ensemble des arborescences étant impossible à réaliser, j'ai limité mon investigation aux arborescences de hauteur zéro et de hauteur un tout en sachant qu'un arbre de hauteur supérieur à un est la superposition d'arbres de hauteur un.

À partir des trois états qu'une UMM peut prendre, le développement de la combinatoire de toutes les arborescences de hauteur zéro (un seul nœud) et de hauteur un (au moins un nœud racine et un nœud fils) donne 228 cas théoriques. Ce chiffre n'est cependant pas atteint du fait des dépendances des états entre nœud père et nœud fils. En éliminant les cas anormaux, nous obtenons un sous-ensemble dont l'étude détaillée a montré qu'il est possible, moyennant des *généralisations* et des *opérations de factorisation*, de résumer à quatre *classes d'équivalence* regroupant onze configurations remarquables.

L'utilisation directe de ces configurations pose cependant parfois un problème de surnombre d'intentions générées. Quelques heuristiques de tri et de filtrage permettent de résoudre ce problème.

Je présente tout d'abord les opérations de factorisation puis les classes d'équivalence pour enfin revenir aux heuristiques de filtrage.

Les opérations de factorisation — Les opérations de factorisation sont des opérations récursives de restructuration effectuées sur les arborescences. D'une part ces opérations sont utilisées dans la définition des classes d'équivalence (voir ci-après) et d'autre part, elles interviennent dans le processus de génération des intentions. Dans ce dernier cas, elles ne sont que temporairement effectuées.

Les opérations de factorisation consistent à remplacer plusieurs UMM de même état par une UMM composée afin de masquer les détails des arborescences avant que ne soient créées les intentions de communication. Ils existent deux opérations de factorisation (voir figure 5) : la *factorisation horizontale* possible entre feuilles de même état et de même père et la *factorisation verticale* possible entre un nœud racine et un nœud fils de même état.

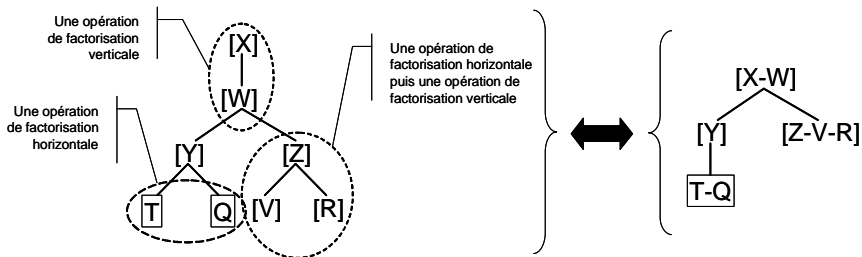


Figure 5 — Exemples d'opérations de factorisation possibles.

Les classes d'équivalence — L'ensemble normalement possible des arborescences de hauteur zéro et de hauteur un (quel que soit le nombre de fils du nœud père) est représenté par quatre classes d'équivalence de onze configurations remarquables génératrices d'intentions de communication. Ce résultat a été obtenu en appliquant les deux règles interactionnelles et les deux opérations de factorisation, en éliminant les cas impossibles et les cas redondants et en généralisant les autres cas.

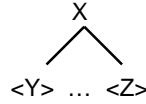
Un cas redondant est une structuration déjà prise en compte moyennant une interversion des fils. Par exemple l'arborescence $X([Y] \vee [Z])$ est équivalente à l'arborescence $U([V] \vee [W])$.

Une généralisation consiste soit à ajouter des fils de même état au nœud père soit à regrouper plusieurs cas conduisant à la même intention de communication. Par exemple, les arborescences $[X](\overline{Y}\vee Z)$ et $[U](\overline{V}\vee W\vee T)$ sont généralisables en $[X](\overline{Y}\vee <Z>+)$ dans laquelle un nœud père réussi est lié disjonctivement à plusieurs fils dont au moins un est satisfait les autres étant non satisfaits (soit dans l'état en attente soit dans l'état réussi).

Classe C1

Cette classe regroupe les configurations remarquables C1a et C1b (voir tableau 2). La configuration C1a est une UMM unique en attente. La configuration C1b est une arborescence dont le nœud racine est en attente et dont l'état et le nombre de fils sont indifférents quelle que soit la liaison avec le nœud racine (conjonctive ou disjonctive). La racine peut être une UMM composée.

Tableau 2 — Classe C1.

Classe		Notation graphique	Notation fonctionnelle	Intention de communication
C1	C1a	X	X	Questionner(X) ou Proposer(X)
	C1b		$X(<W>+)$	

Lorsqu'un nœud racine est en attente, quels que soient l'état et le nombre de ses fils, le système doit agir en sorte de le rendre réussi c'est-à-dire être dans le terrain commun. Ainsi, dans les deux cas, l'intention associée porte sur le nœud racine. Si la racine est une UMM composée, elle s'écrit $X=W(T)$. On reconnaît à nouveau la configuration remarquable C1b. L'intention porte alors sur W. Ce schéma se reproduit autant de fois qu'il y a eu une opération de factorisation verticale effectuée.

L'intention attachée à une UMM en attente est fonction de la direction de la communication. Si la machine a besoin d'une information pour accomplir ses différentes activités (application de la troisième règle interactionnelle), le système pose une question pour qu'ils puissent partager cette information (en partant du principe que l'utilisateur doit avoir une réponse) — Questionner(X). À l'opposé, si le système possède une information non partagée, il fera une proposition (application de la deuxième règle interactionnelle) — Proposer(X). La différence entre ces deux se fait simplement en consultant un champ de l'UMM considérée.

Par exemple, au tableau 3, l'arborescence de gauche représente une action que doit effectuer un apprenant. Cette action est en attente (ainsi que d'autres éléments de l'action) ce qui signifie qu'elle n'est pas considérée comme faisant partie du terrain commun. Par contre les UMM <Apprenant> et <Plat> sont satisfaits car ils sont déjà référencés du fait des interactions précédentes. Lorsque les opérations de factorisation sont appliquées, on obtient l'arborescence équivalente sur laquelle on reconnaît la configuration remarquable C1b. L'intention résultante est dans ce cas une proposition (qui prend la forme d'un ordre dans le cadre applicatif de l'exemple) parce que la machine doit partager le fait (but conversationnel) qu'elle « veut » que

l'apprenant effectue cette action pour la suite du déroulement de l'application (but applicatif).

Tableau 3 — Exemple d'application de la classe C1.

Mémoire	Équivalences et intentions	Exemple d'énoncé
		<p>Mettez le plat au four pendant 2 heures à thermostat 200°.</p>

Classe C2

Cette classe regroupe les configurations C2a, C2b et C2c (voir tableau 4). Il s'agit de trois configurations de hauteur un dont la racine est réussie. Dans C2a, il y a un seul fils en attente ; dans la disjonction C2b, il y a un fils en attente et un fils réussi ; dans la conjonction C2c, il y a un fils en attente et un fils satisfait. Dans les trois cas, les fils peuvent être des UMM composées.

Tableau 4 — Classe C2.

Classe	Notation graphique	Notation fonctionnelle	Intention de communication
C2	C2a 	[X](Y)	Questionner(Y) ou Proposer(Y)
	C2b 	[X](Y∨[Z])	
	C2c 	[X](Y∧[Z])	

Lorsqu'un nœud racine est réussi, on ne peut rien conclure de son état s'il possède un ou plusieurs fils en attente liés conjonctivement ou s'il possède au moins un fils en attente lié disjonctivement. L'intention de communication doit porter sur les fils en attente qu'il faut amener dans le terrain commun. Comme dans la classe précédente, l'intention est alors, selon la direction de la communication, soit une question (Questionner(Y)) soit une proposition (Proposer(Y)). Cependant, cette intention est différente si les fils sont issus de factorisations horizontales :

Si $Y = V \vee W \vee \dots \vee R$ alors $\text{Questionner}(Y) \Rightarrow \text{Questionner}(V \vee W \vee T \vee \dots \vee R)$

Si $Y = V \wedge W \wedge T \wedge \dots \wedge R$ alors $\text{Questionner}(Y) \Rightarrow \text{Questionner}(V \wedge W \wedge T \wedge \dots \wedge R)$

Classe C3

Dans cette classe, nous avons quatre configurations remarquables C3a, C3b, C3c et C3d (voir tableau 5). Il s'agit d'arborescences dont la racine est réussie alors qu'elle devrait être satisfaite au regard des fils présents. Dans C3a, il y a une seule UMM

réussie ; dans C3b, il y a un seul fils satisfait ; dans la disjonction C3c, il y a un fils satisfait et un ensemble de fils non satisfaits (en attente ou réussis) ; enfin, dans la conjonction C3d, il y a un fils réussi et un ensemble de fils non réussis (en attente ou satisfaits). Dans tous les cas la racine peut être une UMM composée.

Tableau 5 — Classe C3.

Classe	Notation graphique	Notation fonctionnelle	Intention de communication	
C3	C3a	[X]	Indiquer(NS(X))	
	C3b			$[X](\overline{Y})$
	C3c			$[X](\overline{Y} \vee \langle Z \rangle +)$
	C3d		$[X]([Y] \vee \langle Z \rangle +)$	Indiquer(CAR(NS(X),NS(Y)))




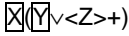
Il existe deux possibilités qu'un nœud racine soit non satisfait : soit cet état est inhérent à l'UMM, soit cet état provient de l'état de ses fils. Les configurations C3a, C3b et C3c relèvent du premier cas, tandis que la configuration C3d relève du second cas. L'intention de communication du système doit être d'informer l'utilisateur de la non satisfaction de la racine (Indiquer(NS(X)) qui se lit indiquer que X n'est pas satisfait), avec, dans le second cas, la possibilité de justifier de cet état (Indiquer(CAR(NS(X),NS(Y))) qui se lit indiquer que X n'est pas satisfait car Y n'est pas satisfait).








Si la racine [X] est une UMM composée, elle provient d'une factorisation verticale et s'écrit $[X]=[W]([T])$. L'expression NS(X) s'écrit alors CAR(NS(W),NS(T)). De même, si [Y] est une UMM composée et s'écrit $[Y]=[U]([V])$ alors NS(Y) s'écrit CAR(NS(U),NS(V)). L'intention se complexifie si les UMM composées sont issues de factorisations horizontales. Par exemple, si $[V]=[A] \vee [B] \vee \dots \vee [C]$ alors NS(V) s'écrit OU(NS(A),NS(B), ..., NS(C)).

Classe C4

Dans cette dernière classe, nous avons les deux configurations C4a et C4b représentant les cas normaux où la racine est satisfaite (voir tableau 6). Cela ne peut se produire que dans trois cas : soit il n'y a pas de fils (cas C4a), soit tous les fils sont satisfaits (cas C4a où \overline{X} est une UMM composée), soit au moins un fils est satisfait dans le cas d'une arborescence disjonctive (cas C4b).

Tableau 6 — Classe C4.

Classe		Notation graphique	Notation fonctionnelle	Intention de communication
C4	C4a			Indiquer(X)
	C4b			

Dans tous ces cas l'intention porte sur le nœud racine et doit être d'informer l'utilisateur de son état — Indiquer(X). Si  est une UMM composée, l'intention continue de porter sur l'UMM la plus haute dans l'arborescence. Par exemple, si =(()) alors l'intention est Indiquer(P).

La génération et les heuristiques de filtrage — Le processus de génération des intentions de communication consiste à *reconnaître* une ou plusieurs configurations remarquables établies précédemment puis à appliquer des heuristiques de tri et de filtrage. La reconnaissance s'effectue après l'exécution des opérations de factorisation qui sont annulées après la création des intentions.

Les intentions créées par reconnaissance sont d'abord triées selon deux critères :

- La troisième règle interactionnelle est prioritaire sur la deuxième règle : le système doit en priorité acquérir les informations nécessaires pour progresser dans ses différentes activités. Les intentions issues de l'application de la troisième règle sont donc plus importantes que les intentions issues de la deuxième règle :
Intérêt(Questionner(X)) > Intérêt(Proposer(Y)) > Intérêt(Indiquer(Z)).
- Une intention déjà créée est moins importante qu'une nouvelle intention. Ceci permet d'éviter les répétitions et laisse le système insister lorsqu'il ne peut pas agir autrement.

La liste d'intentions triées est ensuite soumise aux trois heuristiques suivantes :

- Le nombre d'intentions issues d'une arborescence est limité à un.
- Le nombre d'intentions de même type (questionner, proposer ou indiquer) issue de la forêt est limité à un.
- Le nombre de questions ou de propositions est limité à deux. Par exemple Questionner(U∨V∨W∨Z∨Q) est transformé en Questionner(U∨V).

5. Implémentation et exemple

Le modèle de l'activité dialogique présenté dans la section précédente a été implémenté en CommonLisp dans un système prototypique de dialogue en langue naturelle écrite. Il s'agit du système AMI (Agent de Messagerie Interactif) simulant la fonction de standardiste téléphonique. Le modèle est également utilisé actuellement

dans un système d'information géographique en cours d'élaboration au laboratoire COGIT⁷ de l'IGN (Hubert, 2001).

Face à la complexité des phénomènes langagiers et dialogiques sur lesquels je travaille, j'ai adopté une démarche expérimentale et pragmatique (dans le premier sens du terme). En procédant de manière itérative, comme par exemple Rouillard (2000) pour le système Halpin, nous développons successivement plusieurs versions d'un système qui est amélioré au fur et à mesure en prenant en compte les résultats de l'expérimentation de la version n pour développer la version $n + 1$.

Une première version AMI1 a été d'abord développée en partant de l'analyse d'un corpus réel obtenu dans une expérience de type magicien d'Oz (Castaing, 1993). Cette première version a ensuite été expérimentée et les résultats analysés. Ces derniers et le protocole d'expérimentation sont présentés dans (Lemeunier, 2000). La seconde version AMI2 que nous présentons ici comme exemple d'implémentation du modèle de l'activité dialogique est donc issue des résultats de l'expérimentation de AMI1. Cette dernière version n'a pas été expérimentée (pour des raisons institutionnelles) mais elle a été testée en reprenant les dialogues du corpus de AMI1.

L'architecture générale est détaillée dans (Lemeunier, 2000). Il s'agit d'un tableau noir à huit niveaux et à contrôle procédural autour duquel trois expertises ont été codées dans des sources de connaissances. On trouve une expertise sur l'activité dialogique, une expertise sur l'activité applicative et une expertise sur l'activité linguistique.

M1 : Quelle est votre requête
H1 : je voudrais parler à Jérôme
M2 : Quelle est votre identité ?
H2 : c'est de la part de Pierre
M3 : Je sais que Pierre est un prénom
Lorsque vous dites "c'est de la part de", est-ce que
vous vous présentez ?
H3 : oui je suis Pierre Dupond
(...)

Figure 6 — Exemple de dialogue avec le système AMI2.

L'exemple de la figure 6 est issu d'une session avec le système AMI dans le cas d'une compréhension partielle d'un énoncé de l'utilisateur au cours du dialogue. Ce dernier se passe en deux temps : le premier temps porte sur la détermination de la requête d'action de l'utilisateur et le second temps porte sur l'identification de l'appelant pour satisfaire la requête d'action identifiée.

Cet exemple est maintenant analysé en détail en suivant le cycle des deux phases de fonctionnement de la mémoire du dialogue décrit précédemment. Je rappelle que la première phase est celle de la génération des intentions de communication de la machine et la seconde phase est celle de l'interprétation de l'énoncé de l'utilisateur et des conséquences de cette interprétation sur l'activité applicative.

⁷ <http://recherche.ign.fr/activ/cogit/index.html>

Phase 1

Initialement la mémoire du dialogue comporte une structure d'action <Requête> car le système AMI s'attend au début de chaque dialogue à une requête d'action de l'utilisateur sans préjuger de l'action en question (voir tableau 5). Cette structure est créée par l'activité applicative. On observe que l'UMM <Machine> est satisfaite car on pose l'hypothèse, comme Sperber et Wilson dans la Théorie de la Pertinence (Sperber et Wilson, 1989), que les faits mutuellement manifestes font automatiquement partie du terrain commun et que la machine est capable de « s'auto-référencer ».

La configuration C1b est reconnue sur l'arborescence équivalente obtenue après l'exécution des opérations de factorisation. Une intention portant sur l'attente de plus haut niveau est donc créée : le système a l'intention de demander à l'utilisateur (troisième règle) s'il a une requête particulière dans le but de la connaître. Cette requête sera ensuite satisfaite ou non (première règle). Cette intention produit l'énoncé M1.

Tableau 7 — Premier énoncé de la machine.

Mémoire	Équivalences et intentions	Énoncé M1
		Quelle est votre requête ?

Phase 2

La réponse H1 de l'utilisateur est interprétée comme une requête d'action : l'utilisateur veut entrer en contact avec une tierce personne. Cette interprétation en actes de M1 satisfait la demande initiale du système. Elle se fait de la manière suivante :

- L'activité dialogique change les états des UMM <Requête> <Action> et <Parler> qui deviennent réussies car elles sont appariées avec des éléments de signification de l'analyse sémantique résultat de l'activité langagière.
- L'activité applicative supprime alors les frères de l'UMM [Parler] qui ne servent plus, puis complète cette unité pour former une structure d'action en ajoutant, d'une part, l'UMM <Appelé>, le premier actant obligatoire de l'action dont on connaît le prénom, et, d'autre part, l'UMM <Appelant>, le second actant obligatoire de l'action.
- À ce stade les deux dernières UMM précédentes sont en attente. L'UMM <Appelé> devient réussie car elle est appariée à un élément de signification apparaissant dans l'analyse de H1.
- Puisque l'UMM <Appelé> est réussie, l'activité applicative continue le développement de l'arborescence dans cette branche : les UMM <Prénom> et <Nom> lui sont ajoutées.
- L'UMM <Prénom> devient à son tour réussie (elle correspond à un élément de l'analyse) et aussitôt satisfaite car elle a été catégorisée par l'activité langagière.

- L'UMM [Appelé] devient à son tour satisfaite car elle est référencée par l'activité applicative du fait de la satisfaction de son fils Prénom, la propagation de l'état satisfait se faisant des fils vers le père.
- L'UMM <Appelant> reste en attente car aucun élément de l'analyse de H1 ne lui correspond. En conséquence de quoi les unités [Parler] et [Requête] restent non satisfaites.

La situation finale est celle du début de la phase suivante (tableau 8). L'énoncé H1 est réussi mais ne peut pas être satisfait car l'état du monde ne le permet pas ; en effet, le système ne connaît pas l'identité de l'appelant.

Phase 1

Le système doit maintenant satisfaire la requête d'action de l'utilisateur. De la structure et de l'état de l'arborescence maintenant en mémoire sont reconnues les configurations remarquables C3d et C2c (voir tableau 8). Le système a l'intention de demander à l'utilisateur (troisième règle et configuration C2c) son identité, information dont il a besoin pour satisfaire sa requête. Le système a également l'intention d'informer l'utilisateur (deuxième règle et configuration C3d) que sa requête ne peut pas être satisfaite du fait de la non satisfaction de l'action [Parler].

Du fait de la prédominance de la troisième règle sur la deuxième règle et du fait qu'une arborescence ne peut engendrer qu'une seule intention à la fois (à cause d'une heuristique de filtrage), l'intention C3d est éliminée. Cela aboutit à l'énoncé M2.

Tableau 8 — Deuxième énoncé de la machine.

Mémoire	Équivalences et intentions	Énoncé M2
		<p>Quelle est votre identité ?</p>

Phase 2

L'analyse de la réponse H2 de l'utilisateur échoue en partie car l'expression « c'est de la part de » ne fait pas partie des connaissances linguistiques du système AMI. L'activité langagière intervient donc pour créer une structure d'action d'identification de l'expression inconnue. Comme dans ce cas précis elle n'est pas capable de poser une hypothèse d'analyse, elle va confronter cette expression à l'attente sur laquelle a porté l'énoncé M2 porteur de l'intention précédente. L'arborescence créée est donc : Identifier(Machine^Appelant^[Expression]^Catégorie(Action(Se-Presenter)))

En outre, le système reconnaît « pierre » comme un prénom. Cet élément d'analyse ne correspond à aucune UMM non catégorisée. Une UMM est donc créée pour mémoriser cet élément de signification. La situation finale est donc celle du début de la phase suivante (voir tableau 9).

Phase 1

À partir des arborescences en mémoire sont générées quatre intentions de communication. Il y a les deux intentions sur la structure d'action [Requête] déjà générées précédemment, une question portant sur l'unité <Identifier> puis une indication portant sur l'unité <Prénom>. Sur les trois premières intentions, seule la troisième est gardée car elle est plus importante que la seconde du fait qu'elle n'a pas encore été générée. Cela aboutit à l'énoncé M3.

Tableau 9 — Le troisième énoncé de la machine.

Mémoire		
Équivalences et intentions		Énoncé M3
		<p>Je sais que pierre est un prénom. Lorsque vous dites « c'est de la part de », est-ce que vous vous présentez ?</p>

Phase 2

L'énoncé H3 est analysé sans problème. Il comporte deux parties : une réponse positive à la question M3 du système et une présentation de l'appelant. La confirmation ne correspond à aucune UMM active. Elle est directement prise en compte par l'activité langagière qui modifie ses connaissances en conséquence. De plus, elle supprime la structure de l'action d'identification de l'expression inconnue car elle ne sert plus. L'énoncé H2 est de nouveau analysé et peut être interprété :

- L'unité <Appelant> devient réussie car elle correspond à un élément de signification de l'analyse de l'énoncé H2.
- L'activité applicative continue de développer l'arborescence : les unités <Nom> et <Prénom> sont ajoutées à l'unité [Appelant].
- L'unité <Prénom> précédente devient à son tour réussie (elle correspond à un élément de l'analyse) puis aussitôt satisfaite car elle est catégorisée. L'UMM Prénom qui était déjà présente est supprimée de la mémoire car ces deux UMM sont redondantes.
- En conséquence de la satisfaction de l'unité <Prénom>, l'unité <Appelant> devient satisfaite car elle est complète puis référencée par l'activité applicative.

Le contenu de la mémoire à la fin de cette phase est celui de la figure 7.

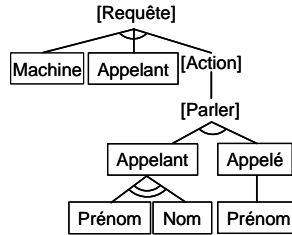


Figure 7 — Arborescence après l'interprétation de H3.

6. Conclusion : apports et limites du modèle

J'ai commencé cet article en exposant les différences entre une activité de dialogue en langue naturelle soumise à une logique interlocutoire et une activité de services informatisés planifiables et soumise à des règles logiquement formalisables, la première activité étant nécessaire pour une interaction souple et naturelle avec l'utilisateur humain et la seconde justifiant l'existence du système informatique.

Les travaux de l'équipe de Trognon du GRC de Nancy en psychologie de l'interaction ont montré que les interactions langagières sont soumises à une logique interlocutoire dans laquelle le sens est co-construit par les interactants qui négocient la valeur interlocutoire des énoncés. La caractéristique principale d'un tel fonctionnement est l'opportunisme des échanges : rien n'est prévu d'avance et tout est négocié. Cette négociation est d'ailleurs la plupart du temps implicite car les interactants arrivent généralement à donner une interprétation valable du premier coup pour eux.

Pour ma part, je considère que l'intérêt principal de la recherche dans le domaine du dialogue homme-machine en langue naturelle réside non pas dans sa dimension applicative mais dans sa dimension dialogique. Il me semble en effet que le but du dialogue homme-machine est de pouvoir un jour parvenir à utiliser un ordinateur en le commandant en langue naturelle (et oralement si possible) et de parvenir à ce qu'il puisse aider l'utilisateur en lui parlant en langue naturelle (oralement si possible). Il ne s'agit pas d'une utopie (comme par exemple tenir des conversations philosophiques avec la machine) mais d'un programme de recherche à très long terme.

Le problème est, toujours selon mon point de vue, que jusqu'à récemment le défaut des recherches dans le domaine a été de ne pas bien distinguer les deux aspects d'un système de dialogue homme-machine négligeant l'un au profit de l'autre ou en les considérant identiquement traitables. Le modèle que j'ai présenté dans cet article, malgré les simplifications qui n'ont pas manqué d'apparaître au lecteur assidu, est un début dans la voie que je propose de suivre. Le modèle doit être affiné et étoffé puis confronté de nouveau à la réalité.

L'approche mixte que je préconise (logique formelle pour la tâche et logique interlocutoire pour le dialogue) et le recours aux théories de la psychologie de l'interaction ne doivent pas faire penser pour autant que les limites du dialogue homme-machine vont être dépassées. Comme je l'ai expliqué dans (Lemeunier, 2000), le mot dialogue appliqué à la communication homme-machine ne peut pas recouvrir la même signification que celle du dialogue naturel tout simplement parce

que le dialogue homme-machine met en présence, d'un côté, un être d'origine naturelle, et, de l'autre, un objet d'origine artificielle. Ce qu'il faut faire, c'est étudier comment ce type de dialogue inédit peut malgré tout « fonctionner » puisque, comme le dit Vivier, les expériences en magicien d'Oz sont là pour attester que cela peut « marcher » (Vivier, 1996).

En résumant le travail présenté ici, je dirais que ma réponse à cette question a consisté à retenir les deux éléments suivants :

- la notion de terrain commun comme élément central et nécessaire (mais non suffisant) de la communication homme-machine pour s'assurer au moins de l'intercompréhension par le partage des informations communiquées ; et,
- la prise en compte dans la gestion du dialogue des notions de réussite et de satisfaction d'un acte de langage pour générer le comportement dialogique du système.

Dans ce cadre, le recours aux résultats de l'analyse des dialogues homme-homme n'est pas exempt de limites, car ce qui vaut pour les conversations humaines ne peut être transposé que de façon hypothétique (et quand cela est possible) aux communications homme-machine en langue naturelle. C'est la raison pour laquelle seul le transfert de la notion de terrain commun au domaine du dialogue homme-machine a été retenu dans ma réponse. Cependant, puisque l'un des partenaires du dialogue est humain, il faut tenir compte du « fonctionnement » de l'homme dialoguant.

Dans le même ordre d'idée, les propos du type « intention de communication de la machine » et « co-construction du dialogue » ne doivent pas être compris comme étant une stricte transposition des concepts issus de l'analyse des interactions humaines. La machine n'a pas réellement d'intention au sens humain du concept. Une nouvelle terminologie doit être inventée.

Parmi les difficultés posées, je voudrais évoquer un problème concernant la définition de la mémoire du dialogue : la question de l'historique du dialogue. Dans mon approche, la notion d'historique disparaît au profit de la notion de mémoire du dialogue. Des problèmes apparaissent alors qui portent sur l'oubli, la remémoration des éléments en mémoire, et plus généralement sur le cycle d'activation des unités de la mémoire. Ces dernières suivent le cycle suivant : après leurs instanciations, elles subissent certains traitements (modification de leur état, structuration), puis doivent disparaître d'elles-mêmes ou laisser la place, être « écrasées » par d'autres éléments ou tout simplement être supprimées. Le problème concernant l'oubli est délicat à traiter. La grande capacité de mémorisation à court terme de la machine par rapport aux capacités humaines peut poser des problèmes : un utilisateur ne se rappelle pas forcément de tout ce qu'il a dit depuis le début de l'interaction, surtout si le dialogue dure depuis plusieurs minutes. La machine peut faire une référence à un objet évoqué au début du dialogue mais l'utilisateur ne sera plus en mesure de la résoudre car le terrain commun ne sera plus le même. Ici deux attitudes sont concevables. Si l'on décide que la machine doit être mise au niveau de l'utilisateur humain, alors ce n'est pas un oubli « sec » et définitif qu'il faut réaliser, mais un oubli progressif, où les unités se transforment pour ne former qu'un « noyau » ne retenant que les éléments les plus saillants, à la manière des humains. Au contraire, si l'on pense que le dialogue homme-machine est un type de dialogue particulier, alors les singularités, qui ne manqueront pas d'apparaître avec une machine incapable d'oublier, doivent être acceptées par les utilisateurs humains.

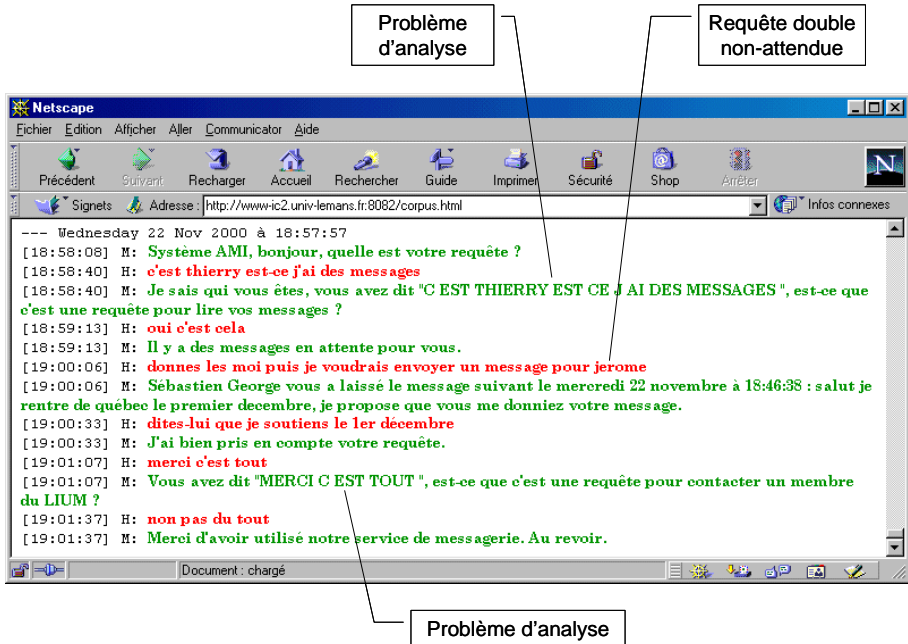


Figure 8 — Exemple de dialogue avec AMI2.

Le lecteur aura noté également que l'application choisie pour implémenter le modèle proposé est relativement simple puisqu'elle porte sur une tâche très finalisée. Cette dernière pourrait être modélisée avec des approches beaucoup moins complexes par exemple une approche par scripts ou par grammaire de dialogue. D'une part ce type d'application est une tâche classique du domaine (Grau *et al.*, 1994 ; Balkanski et Hurault-Plantet, 1997 ; Guyomard *et al.*, 1998) et le premier but était celui de montrer qu'il était possible de parvenir au même résultat, et même mieux si possible, que les approches concurrentes et ainsi d'être plus à même de comparer. D'autre part, je pense qu'il est toujours difficile de montrer l'intérêt d'une nouvelle approche dans un système complexe comme l'est un système de dialogue. Comme je l'ai rappelé en introduction, un système de dialogue se compose d'un ensemble de modules interconnectés. Le travail présenté a porté sur l'activité conversationnelle et non pas, par exemple, sur l'activité langagière (analyse et génération des énoncés). Les limites du module d'analyse qui est également très important, ne doivent pas cacher les avancées du module de dialogue.

Pour ma part, ce que je retiendrai, c'est que le module de dialogue est capable de générer des intentions de communication indépendamment des autres modules et en cela le « contrat » a été rempli. Cela permet effectivement de mettre en place l'autonomie de l'activité conversationnelle. Dans l'exemple de la figure 8, on peut voir que le système est capable de gérer les problèmes d'analyse, les requêtes doubles, et les non-attendus.

Enfin, bien que le système AMI2 n'ait pas pu être expérimenté⁸ comme l'a été AMI1, le savoir-faire et les connaissances en matière de dialogue homme-machine ainsi que le modèle seront réutilisés dans le cadre des interactions apprenants-système dans le domaine des environnements informatiques d'apprentissage humain (EIAH).

Références bibliographiques

- Austin J. L. (1962). *How to do things with words*. Oxford : Oxford University Press.
- Balkanski C., Hurault-Plantet M. (1997). Communicative actions in a dialogue model for cooperative discourse: An initial report. Actes AAAI'97 Fall Symposium on communication actions in humans and machines. Cambridge, MA.
- Bange P. (1992). *Analyse conversationnelle et théorie de l'action*. Paris : Hatier.
- Brassac C., Pesty S. (1999). Simuler la conversation : un défi pour les systèmes multi-agents. In Moulin B., Delisle S., Chaib-draa B. (éd.), *Analyse et Simulation de Conversations — De la théorie des actes de discours aux systèmes multiagents*. Limonest : L'Interdisciplinaire. 317-345.
- Caron J. (1997). Psychologie cognitive et interactions conversationnelles. In Bernicot J., Caron-Pargue J., Trognon A. (éd.), *Conversation, interaction et fonctionnement cognitif*. Nancy : Presses Universitaires de Nancy. 221-237.
- Castaing M.-F. (1993). Corpus de dialogues enregistrés dans un standard téléphonique. *Rapport technique n°93-29*. LIMSI, Orsay.
- Cohen P., Morgan J., Pollack M. (1990). *Intentions in communication*. Cambridge : The MIT Press.
- Clark H.H. (1992). *Arenas of language use*. Chicago : The University of Chicago Press.
- Grau B., Sabah G., Vilnat A. (1994). Pragmatique et dialogue homme-machine. *Technique et science informatique*. 13(1), 269-357.
- Guyomard M., Nerzic P., Siroux J. (1998). Plans, métaplans et dialogue. *Rapport de recherche n°1169 de l'Irisa*. Rennes.
- Hubert F. (2001). Assistance mechanisms use for needs specifications in geographical information on the Web. *Proceedings of the 20th International Cartographic Conference*. Vol 4, 2320-2329.
- Lehuen J. (1997). Un modèle de co-adaptation langagière pour le dialogue homme/machine. In Zreik K. (éd), *Apprentissage par l'interaction*. Europa Productions : Paris. 119-142.
- Lemeunier T. (2000). L'intentionnalité communicative dans le dialogue homme-machine en langue naturelle. Thèse de doctorat. Université du Maine.
- Luzzati D. (1995). *Le dialogue verbal homme-machine*. Paris : Masson.
- Maudet N. (2001). Modéliser les conventions des interactions langagières : la contribution des jeux de dialogue. Thèse de doctorat. Université Paul Sabatier.
- Moeschler J., Reboul A. (1994). *Dictionnaire encyclopédique de pragmatique*. Paris : Éditions du Seuil.

⁸ Un changement de thématique de recherche intervenu entre temps n'a pas permis d'aller jusqu'au bout de la démarche expérimentale.

- Nicolle A., Saint-Dizier De Almeida A. (1999). Vers un modèle des interactions langagières. In Moulin B., Delisle S., Chaib-draa B. (éd.), *Analyse et Simulation de Conversations, De la théorie des actes de discours aux systèmes multiagents*. Limonest : L'Interdisciplinaire. 133-169.
- Rouillard J. (2000). Hyperdialogue sur Internet — Le système HALPIN. Thèse de doctorat. Université Joseph Fourier Grenoble 1.
- Rosset S. (2001). Gestionnaire de dialogue pour un système d'informations à reconnaissance vocale. Actes *TALN'2001*. Tours, France.
- Searle J., Vanderveken D (1985). *Foundations of illocutionary logic*. Cambridge : Cambridge University Press.
- Sperber D., Wilson D. (1989). *La pertinence, communication et cognition*. Paris : Les Éditions de Minuit.
- Tesnière L. (1959). *Éléments de syntaxe structurale*. Paris : Klincksieck.
- Tiberghien G. (1997). *La mémoire oubliée*. Belgique : Mardaga.
- Trichet F., Tchounikine P. (1999). DSTM: A framework to operationalize and refine a problem-solving method modeled in terms of tasks and methods. In *International Journal of Expert Systems With Applications (ESWA)*. Vol. 16, 105-120.
- Trognon A. (1997). Conversation et raisonnements. In Bernicot F., Caron-Pargue J., Trognon A. (éd.), *Conversation, interaction et fonctionnement cognitif*. Nancy : Presses Universitaires de Nancy. 253-282.
- Vanderveken D. (1990). *Meaning and speech acts*. Cambridge : Cambridge University Press.
- Vilnat A. (1997). Quels processus pour les dialogues homme-machine ? In Sabah G. (éd.), *Machine, langage et dialogue*. Paris : L'Harmattan. 129-182.
- Vivier J. (1996). *Psychologie du dialogue en langage naturel*. Paris : Europa Productions.

L'auteur



Thierry Lemeunier est maître de conférences en informatique à l'Université du Maine. Ses recherches au sein du LIUM s'appuient sur les travaux de la psychologie sociale et interactionnelle et portent sur le dialogue homme-machine en langue naturelle. En 2000, il a soutenu au Mans une thèse sur un modèle du fonctionnement d'un agent artificiel dialoguant. Ses travaux se sont maintenant orientés vers l'étude des interactions dialogiques apprenant-système dans le cadre des environnements informatiques d'apprentissage humain (EIAH). Il travaille plus particulièrement sur le projet Cordon Bleu pour l'apprentissage du français pour les étrangers.