

Chapitre 6

Le dialogue explicatif

6.1 Introduction

Les recherches sur le **dialogue explicatif** prolongent directement nos travaux de thèse sur la modélisation du dialogue pédagogique en langue naturelle (Joab, 1990). Nous proposons un module de dialogue fondé sur le modèle genevois d'analyse du discours (Moeschler, 1985, Roulet, et al., 1985) pour un Environnement Interactif d'Apprentissage avec Ordinateur (Joab, 1990). Ce module, indépendant du module pédagogique, permettait de séparer la gestion du dialogue proprement dite de la mise en œuvre des stratégies pédagogiques. Ainsi, pour générer une intervention du système d'EIAO, le module de contrôle était en mesure d'arbitrer entre plusieurs interventions potentielles sur la base de préférences explicites fondées, soit sur des critères pédagogiques, soit sur des critères liés à l'interaction. Dans le cadre du groupe GENE (Génération d'Explications Négociées) du PRC-GDR Intelligence Artificielle, nous avons élargi notre recherche au dialogue explicatif pour permettre à un utilisateur humain d'interagir avec un système à base de connaissances.

Un système expert de première génération était conçu pour fournir à son utilisateur, une solution du problème et, éventuellement, des explications construites à partir de la trace de la résolution (Clancey, 1983). Les utilisateurs ne pouvaient, ni discuter les solutions du système, ni les confronter à leurs solutions propres. Pour parvenir à une meilleure adaptation aux besoins des utilisateurs, il est apparu nécessaire d'accroître leurs capacités en matière de dialogue homme-machine. Ainsi, les utilisateurs peuvent jouer un rôle actif dans toutes les phases de la résolution. Pour ce faire, le dialogue homme-machine doit intégrer des fonctions simulant certaines propriétés des dialogues inter-humains : négocier pour parvenir à un accord, argumenter pour soutenir ou contrer une « thèse » et expliquer pour transmettre une information adaptée à son interlocuteur.

Pour gérer des dialogues homme-machine finalisés, centrés sur la réalisation d'une tâche complexe, c'est-à-dire qui requiert des capacités de résolution de problèmes, le groupe GENE propose d'intégrer des fonctions de négociation, d'argumentation et d'explication au dialogue homme-machine. La spécification de chaque fonction, leur intégration, le choix d'un corpus d'études et la validation expérimentale des systèmes de dialogue homme-machine sont des problèmes de recherche difficiles.

Le premier problème concerne la spécification de modèles prenant en compte chacun des phénomènes de négociation, d'argumentation et d'explication, pris isolément, et leur opérationnalisation dans un système de dialogue homme-machine. Ces modèles, obtenus à

partir de la formalisation d'études en sciences du langage, ne sont que partiellement opérationnalisables. En effet, leur opérationnalisation requiert des capacités de compréhension sémantique fine des énoncés qui constituent, en elles-mêmes, un problème de recherche.

Le deuxième problème consiste à intégrer les phénomènes de négociation, d'argumentation et d'explication dans un même modèle. En effet, négociation, argumentation et explication constituent des phénomènes observables dans les dialogues réels entre humains. Ils surviennent séparément ou de manière conjointe sans que les conditions de leur apparition soient simples à identifier. La spécification d'un modèle unifié requiert de déterminer les interrelations entre ces phénomènes et en particulier les conditions de leurs co-occurrences.

Le troisième problème concerne la méthode de travail elle-même qui consiste à étudier des corpus de dialogues entre humains pour spécifier un système de dialogue homme-machine. Il s'agit alors de reproduire automatiquement certaines régularités observées dans ces dialogues. Deux situations expérimentales ont été exploitées pour le recueil de corpus : des dialogues oraux, recueillis dans des interactions naturelles et des pseudo dialogues homme-machine recueillis par la technique du « Magicien d'Oz ». Nous examinerons dans la section 6.2 les problèmes spécifiques du choix d'un corpus pour l'étude des dialogues explicatifs.

Le dernier problème est celui de la validation expérimentale d'un système de dialogue homme-machine doté de capacités de négociation, d'argumentation et d'explication. Aujourd'hui, les méthodologies d'évaluation ne couvrent pas ces phénomènes du dialogue. Zeiliger, Caelen, et Antoine proposent des tests DRQ (Déclaration, Question, Réponse) pour une évaluation qualitative du dialogue oral (Zeiliger, et al., 1997). Pour évaluer les capacités de dialogue, sont pris en compte, l'interprétation du type d'acte illocutoire (le but du locuteur sous-jacent), la reconnaissance des intentions du locuteur, la pertinence de la réponse du système et la pertinence de la stratégie de réponse du système. Un test DQR¹ consiste, par exemple, à faire valider par l'utilisateur l'interprétation de sa requête produite par le système.

La méthode de travail de GENE consiste à confronter plusieurs types d'analyse sur le même objet d'études, un corpus de dialogues (voir, par exemple, dans (Baker, et al., 1997)).

Dans le groupe GENE, Baker axe ses travaux sur les relations entre explication, argumentation et négociation dans des dialogues finalisés (Baker, 1992, Baker, 1994, Baker,

¹ Un exemple de test DQR

D (usager)	:	Un billet pour Paris SVP
Q (système)	:	Est ce une demande ?
R (usager)	:	Oui

1999). Pour leur part, Safar et Schlienger centrent leur recherche sur les stratégies d'explication coopératives (Safar & Schlienger, 1999). Charnay propose un modèle d'acte de dialogue explicatif Adex (Charnay, 1996, Charnay & Vilnat, 1999). Il décompose dans Adex différents aspects regroupés usuellement dans les actes de langage (Austin, 1970, Searle, 1982) et ajoute les dimensions explicative et argumentative. Adex intègre dans un modèle d'acte de dialogue unique toutes les dimensions pour répondre aux besoins de l'explication, de la négociation et de l'argumentation. Notre contribution personnelle porte sur l'intégration de la dimension explicative dans la gestion du dialogue homme-machine (Joab & Rossari, 1999).

La section 6.3 compare les principaux modèles de dialogues homme-machine. Nous présentons dans la section 6.4 notre travail sur le dialogue explicatif fondé sur le modèle genevois du discours. La section 6.5 présente le bilan des résultats obtenus par GENE et les perspectives de ce travail.

6.2 Le choix d'un corpus d'études

Polity, Francony, Palermi, Falzon, et Kazma soulignent que l'étude de dialogues oraux n'est pas suffisante pour concevoir des dialogues homme-machine car rien ne garantit que les phénomènes observés dans les interactions humaines ne se reproduisent dans des interactions homme-machine (Polity, et al., 1990). Ils proposent d'utiliser la technique du « Magicien d'Oz » pour recueillir des données sur le dialogue homme-machine. Cette technique place des sujets humains face à une pseudo-machine. Les sujets humains pensent interagir avec un système informatique qui, en réalité, est simulé par un expérimentateur humain, « le compère ». Des corpus de dialogues oraux et écrits ont été ainsi recueillis pour expérimenter une situation d'interaction homme-machine encore à venir. La technique du « Magicien d'Oz » permet de simuler une interaction en langue naturelle entre une future application informatique et ses utilisateurs à seule fin d'observer leurs comportements. Ces auteurs soulignent que les consignes données au compère influencent l'expérimentation. En effet, le sujet humain se construit une image des compétences de la machine à partir des réponses de celle-ci et adapte ses interventions en conséquence. Par exemple, Zoltan-Ford observe que, face à un langage restreint de la machine, le sujet humain restreint lui aussi son vocabulaire et simplifie sa syntaxe (Zoltan-Ford, 1984).

Le groupe GENE a étudié les corpus de dialogues de (Polity, et al., 1990). Ces corpus rassemblent des dialogues oraux entre humains et des dialogues « Magicien d'Oz ». Baker, Dessalles, Joab, Raccach, Safar, et Schlienger ont constaté que les explications sont totalement absentes des dialogues « Magicien d'Oz » mais sont très nombreuses dans les dialogues oraux (Baker, et al., 1994a). L'absence d'explications dans les dialogues « Magicien d'Oz » s'explique par le fait que l'utilisateur attribue des compétences explicatives restreintes au système.

Pour obtenir un corpus de dialogues écrits en langue naturelle présentant des explications, une nouvelle expérimentation par la technique du « Magicien d'Oz » a été entreprise par Kassel et son équipe (Kassel, et al., 1994). Le corpus rassemble des dialogues écrits en langue naturelle entre un pseudo système à base de connaissances SATIN, portant sur le diagnostic médical des maladies néo-natales et des internes en médecine. Les dialogues du corpus SATIN sont des dialogues de validation. En effet, la situation d'interaction, où l'interne doit valider le diagnostic émis par la pseudo-machine, favorise l'émergence d'explications. L'utilisateur est amené à confronter ses diagnostics avec ceux de la machine, d'où l'occurrence d'explications mutuelles. Les dialogues du corpus SATIN sont longs : en moyenne composés de 12 tours de parole. Ils sont complexes à cause, d'une part, de la nature de la tâche de diagnostic et, d'autre part, de l'occurrence de négociations sur ce diagnostic.

Le corpus SATIN a été au centre d'un atelier de recherches pluridisciplinaire organisé par le groupe GENE (Baker, et al., 1994b). Cet atelier a permis de comparer des approches et des méthodes, en psychologie, sciences du langage et en intelligence artificielle grâce à un objet d'études commun, le corpus SATIN. Les travaux de l'atelier ont été approfondis dans un numéro spécial de la revue « Psychologie de l'Interaction », édité par Baker, Joab, Safar et Schlienger, à paraître en 1999 (Baker, et al., 1999).

6.3 La modélisation du dialogue homme-machine

Nous présentons deux exemples de travaux sur la modélisation du dialogue. La section 6.3.1 présente un modèle logique de dialogue développé au CLIPS à Grenoble (Caelen, 1997, Villasenor & Caelen, 1998). La section 6.3.2 décrit le modèle genevois d'analyse du discours sur lequel s'appuient de nombreux travaux en informatique linguistique qui se déroulent au LIMSI (Grau, et al., 1994, Vilnat, 1994, Vilnat, 1997), au LIUM, (Lehuen & Luzzati, 1994, Lehuen & Luzzati, 1999, Luzzati, 1989, Luzzati, 1993), à l'IRISA (Bilange, 1992) et au CRIN (Pierrel, 1987). Nous avons choisi de présenter ces deux modèles car ils sont tous deux proches de nos besoins pour le dialogue explicatif (section 6.3.3). Nous ne présenterons pas ici

les travaux sur les chaînes de référence développés à Nancy (Pierrel & Romary, 1997) car la gestion du dialogue n'intervient pas de manière explicite mais au travers de la gestion des thèmes.

6.3.1 Un modèle logique de dialogue (CLIPS-IMAG)

Caelen présente un modèle de dialogue dans un cadre pragmatique qui unifie actions sur le monde et communication. Le modèle de dialogue intègre, dans un même cadre logique, les actions physiques sur l'application (faire) et les actes de langage (faire-faire) (Caelen, 1997, Villasenor & Caelen, 1998). Caelen présente une typologie des actes de dialogue auxquels il ajoute des opérateurs modaux plus classiques : savoir, croire, vouloir, devoir, pouvoir (figure 6.1) (Caelen, 1997). Il précise les principes élémentaires de la communication homme-machine grâce à des ensembles d'axiomes par exemple sur les obligations de la machine.

<i>f : faire</i>	
U f(A,t)	U fait une action A à l'instant t
M f(A,t)	M montre une action A à l'instant t
<i>ff : faire-faire</i>	
U ff(A,t)	U fait faire à M une action A à l'instant t
M ff(A,t)	M fait faire à U une action A à l'instant t
<i>fs : faire-savoir</i>	
U fs(X,t)	U fait savoir X à M (énonce ou répond) à l'instant t
M fs(X,t)	M fait savoir X à U (énonce ou répond) à l'instant t
<i>ffs : faire-faire-savoir</i>	
U ffs(X,t)	U demande X à M à l'instant t
M ffs(X,t)	M demande X à U à l'instant t
<i>fp : faire-pouvoir</i>	
M fp(_,t)	M donne un choix à U
U désigne l'utilisateur	
M désigne la machine	

Figure 6.1 : Typologie des actes de dialogue extraite de (Caelen, 1997)

Caelen formule cinq stratégies, chacune définie par une famille d'axiomes pour gérer le dialogue. Il propose une définition des stratégies directive, réactive, coopérative, négociative, intentionnelle et constructive accompagnées de leurs conditions d'emploi (figure 6.2).

Stratégie	Conditions d'emploi	Résumé
directive	- début de session - guidage d'un utilisateur occasionnel - rattrapage d'un raté de communication	L'initiative appartient à la machine. L'utilisateur répond aux questions ou exécute les ordres.
réactive	- refus d'une autre stratégie - clôture du dialogue	L'initiative appartient à l'utilisateur. Réaction la plus complète possible : exécution immédiate d'une commande ou réponse immédiate à une requête
coopérative	- action incomplètement spécifiée	L'initiative appartient à l'utilisateur. La machine adopte le but de l'utilisateur.
négociative	- positions antagonistes	La machine et l'utilisateur prennent tour à tour l'initiative.
intentionnelle	- en fin d'échange	La machine cherche à interpréter les intentions de l'utilisateur. La machine se ramène à une situation prototypique.
constructive	- thème du dialogue constant - aucune action n'a été faite.	La machine cherche à attirer l'attention de l'utilisateur sur des thèmes voisins pour enrichir le dialogue.

Figure 6.2 : Stratégies d'interaction

La gestion du dialogue suit le cycle suivant à partir d'une intervention de l'utilisateur : recherche du type d'acte, recherche du but, ordonnancement des buts en attente, calcul du plan d'actions, calcul de la stratégie de dialogue, génération de l'acte de la machine, mise à jour.

6.3.2 Le modèle genevois

Le modèle d'analyse du discours développé à Genève propose des outils pour rendre compte des structures sous-jacentes à tout dialogue (Roulet, 1991, Roulet, 1995, Roulet, et al., 1985). Ce modèle postule que les constituants du dialogue entretiennent des relations de dépendance hiérarchique établies sur la base de caractéristiques sémantiques et pragmatiques.

L'analyse hiérarchique découpe une séquence en trois types de constituants : l'échange, l'intervention et l'acte discursif¹. L'intervention est l'unité réursive ; toute intervention peut être composée d'actes, d'interventions ou d'échanges. L'analyse hiérarchique s'appuie sur une grammaire non contextuelle (Moeschler, 1989) (figure 6.3). Les relations hiérarchiques existent uniquement entre les constituants de l'intervention. La dépendance hiérarchique est

¹ Le modèle de Genève s'appuie sur le concept d'acte discursif qui remplace celui de l'acte de langage (Austin, 1970) (Searle, 1982). Roulet considère que l'acte discursif peut être réalisé aussi bien par un syntagme prépositionnel que par une proposition (Roulet, 1991, Roulet, 1995). Nous limiterons ici la segmentation aux actes de langage.

établie sur la base du critère de la suppression du constituant subordonné et sur la nature de la relation entre deux constituants de l'intervention.

R1	E	→	I+SI
R2	SI	→	I I+SI
R3	I	→	I+I A
R4	I	→	CS+CD CD+CS
R5	CD	→	I _p A _p
R6	CS	→	I _s A _s E _s
I	intervention		I _p intervention principale
SI	suite d'interventions		I _s intervention subordonnée
A	acte discursif		A _p acte principal
CD	constituant directeur		A _s acte subordonné
CS	constituant subordonné		

Figure 6.3 : Grammaire du dialogue

Les constituants sont liés par des relations de deux types : les **relations illocutoires**, qui lient les différents constituants d'un échange, et les **relations interactives**, les constituants de l'intervention.

Parmi les **relations interactives** susceptibles d'être explicitées par des connecteurs¹, on distingue : des relations **argumentatives** qui mettent en rapport un argument et une conclusion, des relations **contre-argumentatives** qui mettent en rapport un contre-argument avec une conclusion et des relations **reformulatives** (Rossari, 1994) qui mettent en rapport deux constituants exprimant deux points de vue distincts. Dans les deux premiers types de relations, l'argument (resp. le contre-argument) est subordonné à la conclusion. Dans le dernier type, le constituant reformulé est subordonné au constituant reformulant.

Les constituants reliés par des **relations illocutoires**. Une intervention est initiative si elle se situe en début d'échange, réactive/initiative si elle se situe en cours d'échange, réactive si elle clôt l'échange. Les relations illocutoires ont indépendantes de l'acte illocutoire propre à chaque intervention bien qu'il puisse y avoir des recoupements : par exemple, les demandes d'informations sont souvent des initiatives, et les remerciements, des réactives. Nous utiliserons le terme de fonction illocutoire de type X² pour désigner le type X de l'acte

¹ Un terme est un connecteur s'il permet de lier deux ou plusieurs propositions ou phrases entre elles pour former des propositions ou des phrases plus complexes. Les connecteurs consistent en des conjonctions, locutions, adverbess sans signification référentielle (Moeschler & Reboul, 1994).

² La fonction illocutoire désigne le but dans lequel l'acte a été énoncé. Par exemple, l'énoncé « Peux tu descendre la poubelle ? » est un acte de langage indirect dont la fonction illocutoire correspond à un requête et non à une demande d'information.

illocutoire propre à une intervention initiative, réactive/initiative ou réactive. Ainsi, nous dirons qu'une intervention est pourvue d'une fonction illocutoire de demande d'information.

Deux contraintes permettent de gérer le développement d'un dialogue : la *complétude interactive* qui régit la clôture de l'intervention et la *complétude interactionnelle* qui régit celle de l'échange (Roulet, et al., 1985). La complétude de l'intervention est calculable *a posteriori*. Une intervention sera jugée comme complète, si elle suscite de la part du destinataire une intervention réactive conforme à sa fonction illocutoire.

Pour qu'un échange puisse être clos, il faut que deux interventions successives manifestent *un double accord* dans deux cas. Dans le premier cas, l'acte illocutoire de l'intervention qui a ouvert l'échange est satisfait, deux interventions successives manifestent ce double accord¹ (par exemple, demande d'information, réponse positive, acceptation). Dans le deuxième cas, le locuteur qui a ouvert l'échange renonce à ce que l'acte illocutoire de son intervention initiative soit satisfait : les deux interventions qui manifestent le double accord font alors acte de cette renonciation (par exemple, demande d'information, réponse négative, acceptation).

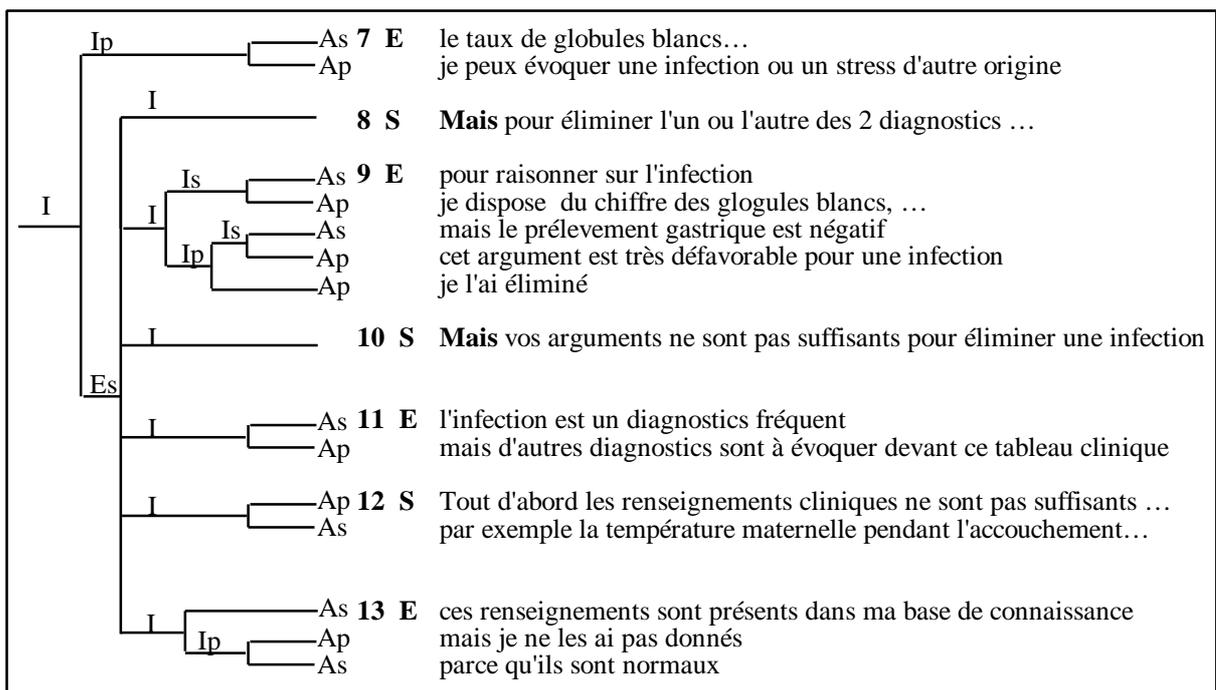


Figure 6.4 : Analyse d'un extrait du dialogue 3 du corpus SATIN

¹ La deuxième de ces interventions peut ne pas être réalisée linguistiquement.

Si on passe d'une intervention initiative à une intervention réactive, la complétude interactive est satisfaite. Ce n'est pas le cas si on passe d'une intervention initiative à l'ouverture d'un échange subordonné.

Si on passe d'une intervention réactive à une intervention initiative d'un nouvel échange, la complétude interactionnelle est satisfaite. Ce n'est pas le cas, si on passe d'une intervention initiative à une intervention réactive de relance.

La figure 6.4 montre l'analyse d'un extrait du corpus SATIN (Kassel, et al., 1994) grâce au modèle genevois. Justifions rapidement cette structure. Le connecteur *mais* est interprété comme introduisant un échange subordonné (en 8S) ou une intervention réactive négative (en 10S). 8S a une fonction illocutoire de requête qui consiste en une demande de clarification consécutive à 7E (qui sera développée jusqu'en 13E), exprimée par le point d'interrogation. Pour cette raison, 8S est analysé comme initiant l'ouverture d'un échange subordonné et non comme une intervention négative de même niveau que 7E. En revanche 10S a une fonction illocutoire d'assertion. Elle est en opposition avec le refus du diagnostic d'infection exprimé en 9E. Il s'agit donc d'une réaction négative. Pour cette raison 10S est au même niveau que l'intervention 9E.

La structure hiérarchique du dialogue, en établissant les liens entre les interventions des locuteurs, permet de montrer la progression du dialogue, d'en isoler des phases successives et de détecter si les attentes d'un locuteur sont remplies au moins temporairement. Maintenir la structure du dialogue permet de mémoriser les éléments sur lesquels les participants pourront éventuellement revenir. La structure hiérarchique du dialogue reflète ainsi la compréhension de l'interaction vue par un observateur distinct des deux locuteurs.

6.3.3 Adéquation des modèles au dialogue explicatif

Les deux modèles de dialogue que nous avons analysés ici font partie des modèles étudiés par le rapport DALI (Dialogue Adaptatif : Langue et Interaction)¹ (Sabah, et al., 1997). Ses auteurs développent l'idée que les différences des modèles s'expliquent par le fait qu'ils répondent à des besoins distincts. Ils constatent que le besoin de prévision de l'énoncé suivant est plus fort pour l'oral que pour l'écrit. Le modèle genevois est difficilement prédictible car

¹ Le projet DALI regroupe les équipes du CRIN et de l'INRIA-Lorraine (Nancy), de l'INPG (Grenoble), de l'IRISA (Lannion), de l'IRIT (toulouse) et du LIMSI (Orsay).

l'analyse hiérarchique est plus aisément applicable *a posteriori*, à la fin du dialogue, qu'au fur et à mesure des interventions. En revanche, le modèle logique a une capacité de prédictibilité plus forte car il s'appuie sur la formation de paires adjacentes d'actes de dialogue.

Le rapport DALI suggère que les modèles de dialogue peuvent être mis en relation avec une catégorie de tâches. Il oppose, d'une part les dialogues de commande et de conception et, d'autre part, les dialogues de recherche d'informations. Un dialogue de commande serait plutôt relié au modèle logique, alors qu'un modèle de recherche d'informations serait relié au modèle genevois.

Une différence importante entre les deux modèles est constituée par l'utilisation d'une structure du dialogue explicite ou implicite. Dans le modèle genevois, la structure du dialogue est une arborescence explicite, alors qu'elle reste implicite dans le modèle logique. Nous considérons que la structure permet d'assurer un suivi global de l'interaction. La structure est d'autant plus utile que l'interaction est longue et qu'elle présente des sous-structures enchâssées témoignant de sa complexité. Nous rejoignons la position exprimée dans (Lehuen & Luzzati, 1994) (p. 111), qui notent que « l'intérêt du modèle pour l'explication est la structuration sémantique qu'il apporte... Le propos du modèle n'est pas d'augmenter directement le niveau de compréhension du système... un système de dialogue doit pouvoir se tromper, à condition qu'il gère suffisamment bien l'interaction pour être à même de récupérer erreurs ou lacunes ».

L'intérêt du modèle de CLIPS est de représenter les stratégies de dialogue dans le même formalisme logique que les actes de dialogue. Ainsi, le modèle ne demande pas de faire communiquer des sources de connaissances hétérogènes comme c'est le cas dans les systèmes de dialogue fondés sur le modèle genevois.

Nous précisons les besoins pour un dialogue explicatif entre un système à base de connaissances et son utilisateur par les contraintes suivantes : une interaction longue, complexe et le plus généralement écrite. Ces besoins nous ont guidé vers l'utilisation du modèle genevois. La structuration *a posteriori* est un inconvénient que nous levons en partie dans la solution algorithmique que nous proposons dans la section 6.4.

6.4 La gestion de dialogues explicatifs

Nous avons travaillé plus particulièrement sur le modèle structurel du dialogue et ses liens avec le développement de phases explicatives. Ce dernier point a fait l'objet d'une collaboration avec Corinne Rossari, une linguiste de l'université de Genève (équipe d'Eddy

Roulet) (Joab & Rossari, 1999). Nous avons dégagé la notion de séquence explicative, qui définit les propriétés d'un segment de dialogue à valeur explicative. Nous avons analysé dans le corpus SATIN (Kassel, et al., 1994), l'occurrence de ces séquences.

Dans cette étude, notre approche est celle du recueil d'expertise. En analysant les dialogues du corpus SATIN, nous avons modélisé, à partir de l'expertise linguistique, des règles de structuration du dialogue et de reconnaissance des séquences explicatives.

L'automatisation de la reconnaissance des séquences explicatives présente des difficultés importantes liées à l'analyse automatique de la langue. Notre contribution porte sur la conception d'un algorithme de structuration dynamique qui exploite la base de connaissances issue du recueil d'expertise linguistique pour reconnaître les séquences explicatives. Nous présentons la notion de séquence explicative (section 6.4.1) et enfin la faisabilité de l'automatisation de la reconnaissance des séquences explicatives (section 6.4.2). Les sections 6.4.1 et 6.4.2 sont reprises de (Joab & Rossari, 1999).

6.4.1 Les séquences explicatives

Nous proposons de distinguer *séquence explicative* et *explication*. En effet, plutôt que de chercher à décrire les processus cognitifs qui caractérisent l'explication, nous limitons notre étude à la description des propriétés de segments de dialogues qui ont une valeur explicative. Nous les définirons comme des séquences explicatives. Les critères pour identifier ces séquences sont de quatre ordres : hiérarchique, fonctionnel, énonciatif et sémantique.

L'analyse hiérarchique permet d'obtenir 4 types de combinaisons entre les unités du discours dans des structures de nœud I (voir figure 6.5).

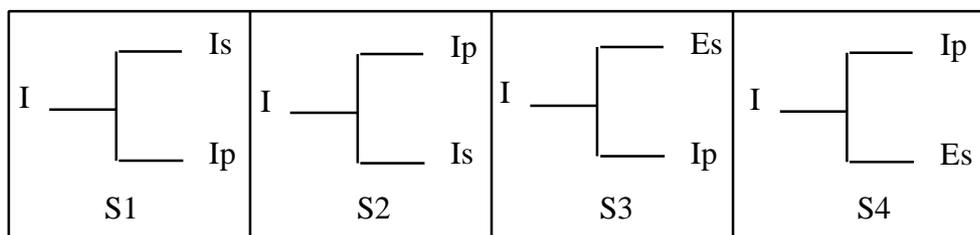


Figure 6.5 : Différentes structures pour une intervention I

Au niveau hiérarchique les séquences explicatives seront définies par des structures du type S4 (figure 6.5), au niveau fonctionnel par le fait que la première intervention de l'échange

subordonné est pourvue d'une fonction illocutoire de demande d'information¹, au niveau énonciatif par le fait que cette première intervention émane d'un autre locuteur que le locuteur d'Ip, et au niveau sémantique par le fait que cette demande d'information est de nature causale.

Une séquence explicative est le produit d'une interaction particulière conduisant un des locuteurs à demander un éclaircissement à propos du raisonnement qui a amené l'interlocuteur à transmettre une certaine information. S1 et S2 sont exclues parce qu'elles n'ont pas de composante dialogique. En particulier, les explications spontanées (Karsenty & Falzon, 1992), fournies par un locuteur qui anticipe une demande d'explication correspondent à ces structures. S3 est exclue parce qu'elle ne comporte pas d'intervention sur laquelle greffer une intervention à fonction de demande d'information du type défini.

Les exemples présentés ci-dessous sont extraits du corpus SATIN (Kassel, et al., 1994). La séquence **SEQ est une séquence explicative** : 2S a une fonction illocutoire de demande d'information formulée en *pourquoi vous avez posé* qui concerne donc bien la motivation du diagnostic. 4S a une fonction illocutoire de demande d'information formulée en *comment vous expliquez*. C'est donc une demande d'explication exprimée littéralement.

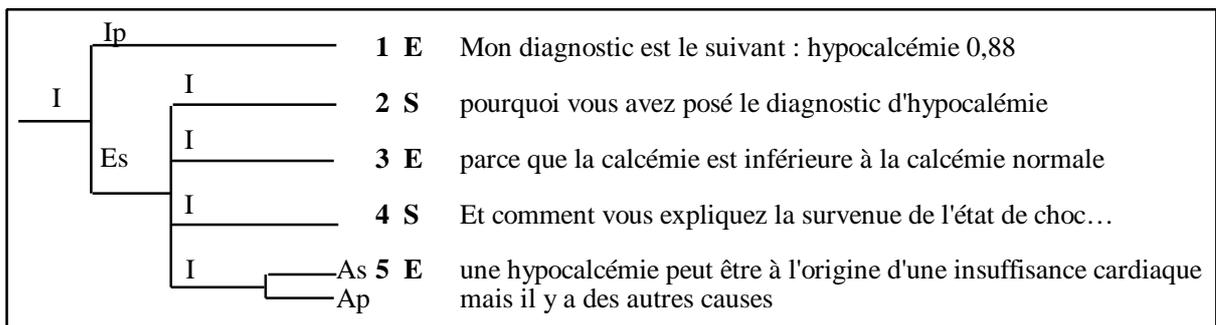


Figure 6.6 : séquence SEQ de type S4 extraite de D3

6.4.2 Automatisation du modèle

Nous précisons comment ce modèle peut être automatisé de façon à faciliter la gestion de dialogues homme-machine finalisés. Nous présentons ici les aspects du modèle pour lesquels une automatisation partielle est possible.

Problèmes liés à la construction de la structure

¹ La fonction illocutoire de demande d'information est associée à la plupart des énoncés interrogatifs.

La construction automatique de la structure pose en elle-même un ensemble de difficultés que l'automatisation n'est pas toujours en mesure de résoudre : l'identification des indices de structuration et la construction dynamique.

Identification des indices de structuration

La construction de la structure d'un dialogue se fait à deux niveaux : l'intervention et l'échange. Le seul critère de la suppression ne permet pas de détecter automatiquement le constituant subordonné. En effet, il faudrait être en mesure d'apprécier l'importance relative de l'apport d'informations de chaque constituant. En revanche, les connecteurs pragmatiques constituent de bons indices pour la structuration. Lorsque ces mêmes connecteurs sont implicites, seules les connaissances du domaine permettent de former des hypothèses sur la nature des relations interactives qui lient les constituants de l'intervention.

Le calcul de la complétude interactive est fondé sur la compatibilité des fonctions illocutoires des interventions successives. L'ouverture d'un échange subordonné est marquée par le fait que la complétude interactive n'est pas atteinte. La détection des relations illocutoires est indispensable pour distinguer l'ouverture d'un échange subordonné d'une simple intervention réactive.

La reconnaissance de la fonction illocutoire des interventions constitue une difficulté essentielle pour appliquer le modèle. De nombreux travaux en analyse automatique de la langue ont porté sur la reconnaissance automatique de la fonction illocutoire des actes de langage (voir par exemple (Allen & Perrault, 1980).

La complétude interactionnelle signale la clôture de l'échange. Le double accord des locuteurs est souvent implicite dans le corpus SATIN. En effet, un échange subordonné n'est jamais clos par un accord explicite des locuteurs. Cette propriété marque une différence entre les dialogues écrits « Magicien d'Oz » et des dialogues oraux entre humains.

Analyse dynamique des séquences explicatives

Une autre difficulté du modèle est liée à son caractère statique. Il est en effet plus facile de proposer une structure du dialogue *a posteriori* plutôt qu'au fur et à mesure de la production des énoncés. Or, pour assurer le suivi du dialogue et gérer une interaction homme-machine, la structuration doit être dynamique. Pour le dialogue homme-machine, le processus de structuration est lancé à chaque nouvelle intervention de l'utilisateur, les structures de type S1 ou S2 sont donc analysées *a posteriori*. Seules les structures de type S3 ou S4 couvrent plusieurs tours de parole et sont difficiles à analyser dynamiquement.

Dans (Grau, et al., 1994), les auteurs proposent un mécanisme de réinterprétation qui recalculé la structure du dialogue dans un gestionnaire de dialogue. L'inconvénient principal de cette solution réside dans le fait que la fiabilité de la gestion du dialogue repose essentiellement sur celle des mécanismes de compréhension automatique de la langue naturelle. En revanche, elle présente l'avantage de ne maintenir que la structure courante.

De façon plus générale, le problème essentiel de la structuration consiste à repérer dynamiquement l'ouverture et la fermeture d'un échange subordonné dans des structures de type S3 ou S4. Nous proposons de considérer ici le problème de la reconnaissance de la fonction illocutoire des interventions comme résolu pour mettre en évidence des règles de structuration dynamique. Une typologie des fonctions illocutoires initiatives et réactives et des couples de fonctions en adéquation figure dans (Moeschler, 1985) (p 95). Nous présentons une forme synthétique de critères de structuration (§ 2) par des règles de production.

On note I_{ENCOURS} l'intervention courante à analyser. Soit I_{ANTE} l'intervention qui précède I_{ENCOURS} . I_{ANTE} fait partie d'un échange E_{ENCOURS} . On désigne par f_A (respectivement f_C) la fonction illocutoire de I_{ANTE} (respectivement I_{ENCOURS}). On désigne par I_{POST} une intervention postérieure à I_{ENCOURS} .

Nous précisons quelques règles de structuration (proches de (Grau, et al., 1994, Joab, 1990).

•**RÈGLE 1A : POURSUITE DE L'ÉCHANGE EN COURS**

Si f_A est initiative et f_C réactive
et f_A et f_C sont en adéquation
et I_{ANTE} et I_{ENCOURS} ont le même thème
Alors I_{ANTE} et I_{ENCOURS} sont au même niveau de l'échange E_{ENCOURS}

•**RÈGLE 2 : OUVERTURE D'UN ÉCHANGE SUBORDONNÉ**

Si f_A est initiative et f_C initiative
et I_{ANTE} et I_{ENCOURS} ont des thèmes différents en relation
Alors I_{ENCOURS} ouvre un échange subordonné E_{SUB} .

Deux types de structure sont alors possibles : S3 ou S4. Dans les deux cas, il ne sera possible de trancher qu'à la clôture de l'échange E_{SUB} . La clôture de l'échange subordonné E_{SUB} ne peut être détectée au même moment pour les deux types de structures (figure 6.7).

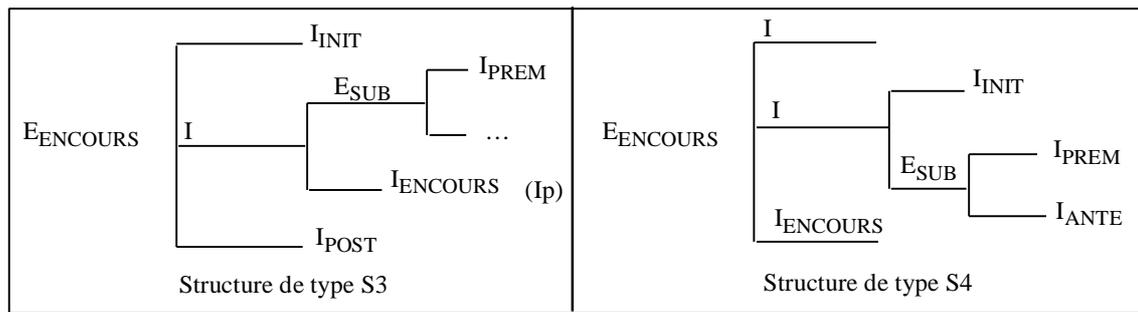


Figure 6.7 : Rattachement d'un échange subordonné lors de sa clôture

A la clôture de l'échange E_{SUB} , pour une structure de type S3, l'intervention $I_{ENCOURS}$ doit présenter les caractéristiques suivantes : sa relation interactive avec E_{SUB} doit pouvoir être marquée, sa fonction illocutoire est transmise à I et doit donc être en adéquation avec celle de I_{INIT} , et enfin elle doit être produite par le même locuteur que l'intervention I_{PREM} qui a ouvert E_{SUB} . Enfin, l'occurrence de l'intervention suivante I_{POST} confirme la clôture de l'échange E_{SUB} si son thème est différent de celui de E_{SUB} .

•RÈGLE 3A : CLÔTURE D'UN ÉCHANGE SUBORDONNÉ

Si $I_{ENCOURS}$ et E_{SUB} entretiennent une relation interactive et les fonctions illocutoires de I_{INIT} et $I_{ENCOURS}$ sont en adéquation et $I_{ENCOURS}$ et I_{PREM} sont issues du même locuteur alors E_{SUB} est clos
 marquer la relation interactive entre E_{SUB} et $I_{ENCOURS}$
 construire une structure de type S3 à confirmer par I_{POST} (E_{SUB} et I_{POST} ont des thèmes différents en relation).

Dans le cas d'une structure de type S4, la clôture de l'échange E_{SUB} n'est constatée qu'après la clôture effective, lors de l'occurrence de l'intervention $I_{ENCOURS}$. I_{ANTE} (qui clôt l'échange E_{SUB}) est produite par l'un ou l'autre des locuteurs. En revanche $I_{ENCOURS}$ et I_{PREM} sont produits par le même locuteur dans le cas des séquences explicatives (cf. critère énonciatif). La clôture de l'échange se manifeste *a posteriori* par le changement de thème opéré entre E_{SUB} et $I_{ENCOURS}$. Lorsque l'échange est clos, il est nécessaire de vérifier la relation interactive entre I_{INIT} et E_{SUB} .

•RÈGLE 3B : CLÔTURE D'UN ÉCHANGE SUBORDONNÉ

Si $I_{ENCOURS}$ et E_{SUB} présentent des thèmes différents et $I_{ENCOURS}$ et I_{PREM} de E_{SUB} sont issues du même locuteur alors E_{SUB} est clos
 marquer la relation interactive entre E_{SUB} et I_{INIT}
 construire une structure de type S4.

Choix des traitements informatiques

Pour mettre en œuvre les règles de structuration précédentes, il est indispensable de choisir les techniques informatiques les plus appropriées.

Plusieurs types de traitement se présentent :

construire une seule structure sujette à réévaluation, ou maintenir différentes hypothèses de structures en parallèle,

construire la structure intervention par intervention, ou différer la structuration jusqu'à disposer des informations nécessaires.

Un ensemble de techniques ont été développées en Intelligence Artificielle pour répondre au problème de l'analyse dynamique des productions de l'utilisateur d'un système à base de connaissances. La construction d'une structure unique du dialogue que l'on modifie à chaque réinterprétation est une solution intéressante si on maîtrise les modifications en cascade induites par les réinterprétations : chaque nouvelle occurrence d'un échange subordonné provoque éventuellement une réinterprétation. La difficulté de restructuration s'accroît du fait de l'emboîtement possible des échanges. Nous n'avons pas de justifications théoriques pour limiter le degré d'emboîtement des échanges. La seule observation empirique des corpus étudiés ne nous paraît pas suffisamment étayée pour proposer un tel seuil. Pour ces raisons, nous opterons plutôt pour l'examen de différentes structures à valeur d'hypothèses.

Les résultats courants de l'analyse peuvent être gérés comme des hypothèses et mémorisés dans une structure de données, « le tableau noir » (Engelmore & Morgan, 1988). Cette solution logicielle nécessite de maintenir en parallèle dans le tableau noir les différentes hypothèses et éventuellement de les réinterpréter. L'avantage de cette solution est d'examiner toutes les hypothèses courantes et ainsi de pouvoir repérer les meilleures interprétations en prenant en considération la totalité de l'historique de l'interaction. L'inconvénient est évident : il est lié au coût de calcul du mécanisme de réinterprétation si le nombre d'hypothèses croît trop rapidement. Cette solution n'est viable que si on parvient à limiter le nombre d'hypothèses courantes. Or, dans l'analyse du discours que nous avons présentée dans cette étude, la confirmation de certaines hypothèses de structuration est liée à la présence explicite de marqueurs. Il est à craindre qu'en l'absence de ces marques, il ne faille maintenir un nombre trop important d'hypothèses et que cette approche ne comporte un risque élevé d'explosion combinatoire.

La deuxième option concerne le moment où la structure sera construite : intervention par intervention ou en différé. Nous avons montré ci-dessus que la clôture d'un échange ne pouvait

être détectée qu'*a posteriori* (règle R3B) ou bien détectée par l'intervention courante et confirmée par l'intervention suivante (règle R3A). Ce phénomène montre qu'on ne peut effectuer une construction incrémentale qu'en la remettant en cause ultérieurement ou en gérant en parallèle les différentes hypothèses.

Nous proposons une solution logicielle qui pallie les insuffisances des approches précédentes et qui en conserve les propriétés intéressantes. Le principe général adopté est de différer certains choix jusqu'au moment où une décision de structuration sera suffisamment étayée pour ne pas risquer une remise en cause ultérieure.

Le problème se pose essentiellement lors de la détection de l'ouverture d'un échange subordonné et de sa clôture. En effet, dès lors qu'il s'agit de poursuivre l'échange en cours, la structuration peut s'opérer. Nous proposons un traitement différé du rattachement de l'échange subordonné jusqu'à la détection de la clôture de ce même échange. Cependant la structure interne de l'échange subordonné peut être construite de façon incrémentale. Une pile d'échanges subordonnés en attente, associée aux interventions I_p potentielles permet de traiter les échanges emboîtés. Un tel algorithme présente l'intérêt de construire à la fois une structure locale à l'échange subordonné et, en différé, une structure globale du dialogue fiable. Ainsi la gestion du dialogue opère au niveau local comme au niveau global.

Séquences explicatives

Les séquences explicatives ont été définies par une structure de type S4 où la première intervention de l'échange subordonné est énoncée par un autre locuteur que celui de I_p et dont la fonction illocutoire est une demande d'information de nature causale.

Les techniques de structuration dynamique du dialogue sont préalables à la reconnaissance de telles structures, mais elles ne donnent aucune indication sur la nature fonctionnelle et sémantique de $I_{ENCOURS}$ (la première intervention de l'échange subordonné). Une fois reconnue une structure de type S4, afin de conférer à la séquence son statut explicatif, il est nécessaire de pouvoir repérer automatiquement une demande d'information de nature causale. Dans le corpus étudié, les marqueurs qui explicitent une demande d'information causale sont le plus souvent présents. Nous considérons que la détection de ces marqueurs contribue à la reconnaissance des séquences explicative.

L'identification des séquences explicatives sur la base des structures à deux conséquences importantes pour la conception d'un système de dialogue homme-machine.

- Lors d'une séquence explicative initiée par l'utilisateur (le sujet, dans la situation expérimentale du corpus), la clôture de l'échange subordonné est repérée et permet de

conclure à l'accord de l'utilisateur et du système et donc d'inférer que la séquence explicative a atteint son but et satisfait l'utilisateur (au moins temporairement) ;

- Lors d'une séquence explicative initiée par le système (l'expert, dans la situation expérimentale du corpus), la clôture de l'échange subordonné constitue une décision stratégique du système.

Nous avons montré les avantages et les difficultés de l'analyse dynamique du dialogue pour la détection des séquences explicatives. La compréhension de la langue reste l'obstacle essentiel à la mise en œuvre extensive du modèle linguistique présenté ici. Cependant la structuration du dialogue permet une meilleure compréhension d'une interaction longue et complexe, dans la mesure où elle donne une image globale du dialogue.

6.5 Conclusion

Les travaux menés au sein du groupe GENE ont permis d'enrichir progressivement notre modélisation d'un dialogue explicatif. Pour tester son modèle, le groupe GENE a choisi de transposer certaines fonctions d'un dialogue explicatif en langue naturelle à la gestion d'une interface graphique interactive (Baker, et al., 1997, Baker, et al., 1996a, Baker, et al., 1996b). Une première maquette d'une interface graphique, destinée à l'utilisateur du SBC SATIN, a été développée. Nous avons choisi une méthode de développement itérative de façon à tester et à valider expérimentalement nos modèles. Cette démarche est indispensable pour évaluer la qualité de la communication entre le système et l'utilisateur via l'interface graphique.