

Un modèle fédérateur pour le dialogue Homme-machine robuste en langage naturel écrit : le modèle de l'énonciation du projet ESPRIT II / PLUS

Violaine Prince et Didier Pernel
(1) 69 85 80 18 (1) 69 85 80 04

LIMSI (CNRS)
B.P 133
91403 Orsay Cédex

Abstract

The work we present here is produced within the Plus/ Esprit II project—it deals with the design of a written Man-machine dialogue system. The discourse model, which is the proper topic of the paper, federates several knowledge sources and makes them collaborate. These sources derive from independant works which emphasize different points of view—dialogue structure, user knowledge and beliefs, move intention and communicative goal, dialogue topic and structuring. They interact within a dialogue history which is no more a pure chronological mark but also a common communicative area to all the processes. The history can be seen here as well as representing the models states as identifying the events.

Key-words : Man-machine written dialogue, natural language, discourse model.

Résumé

Le travail que nous présentons ici est réalisé au sein du projet Esprit II / PLUS—conception d'un système de dialogue écrit Homme-machine. Notre modèle de l'énonciation fédère et fait collaborer plusieurs sources de connaissances issues de travaux indépendants mettant l'accent sur des points de vue différents—structure du dialogue, connaissances et croyances de l'utilisateur, intentionnalité et but communicatif de l'intervention, thème et articulation du dialogue. Elles interagissent au sein d'un historique du dialogue qui n'est plus simplement une trace chronologique, mais aussi une zone de communication commune à tous les processus. Il peut être considéré aussi bien comme représentant l'état des modèles que comme repérant les événements.

Mots-clés : dialogue Homme-machine écrit, langage naturel, modèle de discours.

1. INTRODUCTION

Le travail que nous présentons est réalisé au sein d'un projet européen de conception d'un système d'interrogation des pages jaunes de l'annuaire en langage naturel écrit. La particularité de ce projet, que nous introduisons dans le paragraphe 2, est de s'appuyer fortement sur la pragmatique, dimension de l'analyse linguistique, pour assurer une certaine souplesse dans l'interaction Homme-machine.

Cette souplesse s'exprime en termes d'adaptabilité à l'expression de l'utilisateur, qui doit être en langage naturel non contraint. Les questions rédigées peuvent posséder certaines particularités syntaxiques, sémantiques et référentielles : ellipses, anaphores, référence à des objets du domaine, phrases mal formées, implicites et sous-entendus sont autant de cas qu'il faut examiner pour s'assurer d'une interaction souple.

De la même manière, le système doit émettre ses réponses avec la même latitude linguistique, mais en suivant des consignes en matière de communication en langue : maxims conversationnelles issues des propositions de Grice (1975), bonne tenue de la thématique du dialogue [Grosz et Sidner 1986], gestion appropriée de la conversation avec l'utilisateur [Vilnat 1984] [Luzzatti 1989] sont autant d'éléments qui doivent exister dans le modèle complexe et sous-jacent de l'interaction Homme-machine en langue, en vue de l'obtention d'informations.

Dès lors, nous nous sommes plus particulièrement intéressés à la mise au point d'un modèle comprenant à la fois la représentation des énoncés linguistiques et la gestion du dialogue, pour rendre compte de cette double compétence que doit posséder le système. Nous appelons ce modèle "**modèle de l'énonciation**". Nous lui consacrons le paragraphe 3, dans lequel nous développons aussi les tendances inspiratrices et les solutions que nous avons personnellement tenté de promouvoir auprès de nos partenaires du consortium.

Enfin, ce modèle est la base d'un système réel, pour lequel une architecture a été définie [Black 1991]. Compte tenu de la modularité de sa composition, et de l'existence de la notion de supervision, ou de *contrôle des flux d'information* et des traitements que cette architecture permet de réaliser (inspirée des travaux de Sabah (1990-b et 1991)), nous pensons que cela influence les structures mêmes dans lesquelles le modèle s'exprime. En particulier, nous exprimons la nécessité de l'existence non seulement d'une trace du dialogue poursuivi, mais de toute une structure élaborée et complexe, que nous nommons "**historique du dialogue**" (à la suite d'autres travaux dont [Vilnat et Sabah 1985] et [Vilnat 1988]). Nous essayons de montrer la particularité de notre vision de l'historique, en mettant en avant son aspect à la fois synchronique (lieu de communication de plusieurs modèles coopérants) et diachronique (capacité de retrouver des états antérieurs et de retracer des "trajectoires"). L'ensemble de ces réalisations est examiné dans le paragraphe 4.

2. LE PROJET PLUS

Le projet européen PLUS a pour cadre ESPRIT II (P5254) et a commencé officiellement en novembre 1990. Y participent en plus du LIMSI, Cap Gémini Innovation (Paris), ITK (Tilburg - Pays Bas), Omega Generation (Bologne - Italie), Cap Gémini SCS (Hambourg - Allemagne), UMIST (Manchester - Angleterre), l'Université de Bristol (Angleterre), celle de Göteborg (Suède) et celle de Pise (Italie) pour un total de 150 années-homme.

Il a pour but de créer un système de dialogue Homme-machine, orienté par la tâche, en langage naturel écrit. Il doit permettre un réel dialogue entre l'Homme et la machine, et être robuste en ayant une grande tolérance aux entrées erronées ou incomplètes de l'utilisateur. Cette robustesse est acquise en intégrant tous les types de connaissances aussi bien syntaxiques, sémantiques que pragmatiques. Le contexte intervient à tous les niveaux de l'analyse et de l'interprétation des énoncés de l'utilisateur afin de les situer dans le déroulement de l'interaction. Le système doit non seulement résoudre les problèmes

classiques de linguistique (résolution d'anaphores, d'ellipses, rattrapage d'erreurs, de mots inattendus) mais faire aussi des interprétations contextuelles (dérivation d'*implicatures*, prise en considération des croyances de l'utilisateur...).

Le domaine d'application, retenu pour montrer la faisabilité d'un tel système, est l'interrogation des pages jaunes de l'annuaire par l'intermédiaire d'un Minitel.

Le système est constitué de deux parties □ un moteur de langage naturel et un gestionnaire de dialogue. Le moteur de langage naturel est formé de deux modules dont l'un sert à faire l'analyse syntaxico-sémantique des propos de l'utilisateur en restant au niveau du sens littéral et l'autre à engendrer ceux du système. Le gestionnaire de dialogue, quant à lui, est composé essentiellement de trois modules □

1 □ l'analyseur cognitif qui doit découvrir le but réel de l'utilisateur □

2 □ le formulateur de but qui doit déterminer le meilleur but du système en réaction □

3 □ le planificateur de la réponse qui doit créer la représentation sémantique des énoncés du système. Nous détaillerons les entrées et sorties de ces modules dans le paragraphe 4.

L'analyseur cognitif, par exemple, se sert de la structure du dialogue, des croyances de l'utilisateur, des thèmes courants et passés. De même, que les autres modules, il doit mettre à jour aussi bien les connaissances et les croyances de l'utilisateur que la structure même du dialogue, noter l'apparition ou la fermeture d'un thème, la stratégie qu'il a choisie d'adopter pour satisfaire au mieux la requête de l'utilisateur. Dès lors, nous voyons que *les connaissances manipulées par le système sont multiples* et de natures variées. Nous définissons un modèle de l'énonciation qui va nous permettre de gérer et de faire coopérer toutes ces connaissances dynamiques.

3. LE MODELE DE L'ENONCIATION

Nous avons surtout participé, dans le cadre du projet, à l'élaboration d'un modèle concernant **l'énonciation** ("*discourse*" en anglais). L'énonciation est un processus qui produit un **énoncé en langue** (élément discursif exclusivement) et affecte à cet énoncé une position structurelle dans le dialogue (intervention au sens de Moeschler dans [Moeschler 1985]). Un énoncé a les propriétés suivantes □

1 □ Son contenu propositionnel,

2 □ l'intention du locuteur qui le soutient,

3 □ Son contexte d'apparition.

Dès lors, le modèle de l'énonciation du projet PLUS comprend □

1 □ la modélisation de l'expression linguistique de l'énoncé en langue ;

2 □ la modélisation du dialogue en tant qu'échange finalisé d'interventions, obéissant à des règles implicites que nous avons mentionnées dans le paragraphe 1 sous le nom de "maximes conversationnelles".

Il est évident que ces deux modélisations ne sont pas indépendantes, dans la mesure où la situation de dialogue est susceptible de modifier l'expression linguistique (abondance d'ellipses, d'anaphores, de références à des situations déjà exprimées dans la conversation). Réciproquement, l'expression de l'intervention peut modifier la structure du dialogue □ une intervention à caractère vague (indétermination ou sous-détermination sémantique, voire référentielle) suscitera des échanges d'explicitation (Moeschler, op.cit.).

C'est pourquoi, nous avons voulu un modèle de l'énonciation, plus général qu'un modèle de dialogue.

3.1 La coopération de plusieurs modèles

Comme nous l'avons cité en introduction, dans le projet PLUS, nous avons milité en faveur de plusieurs modèles coopérants afin de réaliser un modèle générique de l'énonciation. La raison principale est que, même en étant intégrateur, le modèle de l'énonciation se décline en sous-structures dont les apports sont différenciés. Il est nécessaire aussi bien de représenter les particularités morpho-syntaxiques et sémantiques

de l'énoncé, que de juger de la portée dialogale de son contenu. Et cette dernière est elle-même un objet composé que nous décrivons ci-après.

3.1.1 La portée dialogale

On peut définir la **portée dialogale** comme étant le résultat complexe de l'analyse d'une intervention, dont les facteurs actifs sont

1- Les **but(s) communicatif(s)** décelables dans les contenus propositionnels de l'intervention (et ceux-ci peuvent être nombreux, que l'intervention comprenne ou non plusieurs actes de langage¹) : ces buts communicatifs peuvent être vus comme des croyances du système sur les intentions de l'utilisateur ; nous voyons un **but communicatif** comme exprimant non seulement une intention sur un objet du monde, mais une intention sur la fonction de communication [Allwood 1976] ce qui rajoute une dimension à l'intention au sens courant. C'est pourquoi, la notion de but communicatif suppose l'existence d'un interlocuteur, alors que la notion **d'intention** (de l'utilisateur) s'adresse plus particulièrement au contenu même de l'expression.

2- Les **croyances de l'utilisateur** sous-jacentes aux buts communicatifs décelés : ce que celui-ci sait du monde, de l'application et du système et qui a été sollicité par lui lorsqu'il a formulé son intervention de la manière linguistique déterminée et avec le(s) but(s) communicatif(s) décelé(s) (on peut proposer la réalisation symétrique lorsqu'il s'agit du système).

3- La **situation dans le dialogue** la chronologie des interventions n'est pas seulement dépendante de leur contenu intentionnel exprimé par les deux premiers points [Allen et Perrault 1980], mais, comme on peut le voir dans [Carberry 1988] elle semble être le résultat du développement d'un plan de la part de chaque interlocuteur. En conséquence, la position structurelle d'une intervention est un indicateur du plan développé, voire un indice de l'articulation du plan. Elle permet alors d'anticiper sur la réplique (de l'autre partenaire) sous forme "**d'attente**" d'actes de langage plausibles.

3.1.2 Un ensemble coopérant et dynamique

Le premier point que nous avons décrit montre que le modèle de l'énonciation comprend des apports relativement distincts en provenance de formes variées dans l'appréhension du discours.

Le deuxième point qui nous paraît important dans la problématique de PLUS est qu'un discours en situation de dialogue est typiquement un processus dynamique. Cela signifie que :

1- Si contenu(s) propositionnel(s) il y a, ce(s) dernier(s) s'insère(nt) dans une trame plus vaste : les liens de référence à un objet réel ou à une structure antérieure du discours (phrase, syntagme nominal) en font un élément dépendant d'une situation donnée par conséquent, il ne peut être retenu tel quel, mais doit être assorti de liens avec ces états antérieurs ou courants ;

¹Voir la définition d'Austin (1962) "acte de langage": unité monologale correspondant à un seul contenu propositionnel. Une intervention peut en comprendre plusieurs. La notion de contenu propositionnel semble plutôt être définie à partir d'un prédicat verbal non modal, moyennant ou non un verbe modal. On ne sait pas bien comment on rend compte de la pluralité propositionnelle en présence de certaines propositions relatives.

Exemple : *Jean mange des fruits verts*

Je pense que Jean mange des fruits verts.

Jean mange des fruits que je crois verts.

pourraient correspondre à un même contenu propositionnel. (voir Moeschler 1985).

Mais comment considérer : *Jean mange des fruits dont le moins que l'on puisse dire c'est qu'ils n'ont pas encore assez mûri ?*

La littérature n'est pas très claire sur le sujet.

2 □ Tout élément déterminé en tant que tel, qu'il soit *intention* (se fondant sur des croyances), *but communicatif* (intention en situation de communication) ou *position structurelle* (se fondant sur la chronologie de l'énonciation), n'est pas suffisant à lui tout seul pour rendre compte du dialogue dans sa globalité aussi bien que dans ses éléments locaux.

3.2 Les modèles retenus dans PLUS

Les modèles qui ont été retenus comme composantes du modèle de l'énonciation dans le projet PLUS sont :

- 1 □ Des modèles de connaissances pragmatiques sur l'environnement du discours □
- 2 □ Un modèle de l'utilisateur □
- 3 □ Un modèle structurel du dialogue □
- 4 □ Un modèle de représentation sémantique des actes locutoires.

Les **modèles de connaissances pragmatiques sur l'environnement du discours** se décomposent en :

- 1.1 un modèle du monde (connaissances générales) ;
- 1.2 un modèle de l'application (pages jaunes de l'annuaire) ;
- 1.3 un modèle de la thématique □ ce dernier est vu comme un élément de détermination des thèmes du dialogue en fonction des éléments de l'application. Fondé sur les travaux de McCoy et Cheng (1990), ce modèle rend compte de l'évolution thématique d'une part, et d'autre part se fonde sur elle pour venir préciser les articulations structurelles et intentionnelles.

Le **modèle de l'utilisateur** retenu est typiquement un modèle de croyances issu de [Bunt, 1985] et [Bunt 1989] : il s'exprime en terme de souhaits ou d'intentions (l'opérateur modal *Want*), de croyances faibles (l'opérateur *WeakBelief*) et de connaissances assimilées à des croyances fortes (l'opérateur modal *Know*). Les contenus propositionnels d'une intervention de l'utilisateur sont exprimés en terme de croyances ou d'intentions, moyennant un examen de la situation en cours des croyances de l'utilisateur.

Le **modèle structurel du dialogue** est inspiré des travaux de Roulet et de Moeschler (voir [Roulet et al.1985] et [Moeschler op.cit.]), mais aussi en fonction de travaux d'adaptation au dialogue Homme-machine plus récents [Vilnat 1989] et relevant d'autres projets ESPRIT [Bilange 1991].

Le **modèle de la représentation sémantique** des actes locutoires s'appuie sur une représentation à base de traits à l'aide de formalismes décrits dans [Bunt 1985] qui en est la source d'inspiration principale.

3.3 Les modèles généraux inspirateurs : ressemblances et différences

Certes, l'ensemble des composantes du modèle de l'énonciation sont directement ou indirectement inspirées de modèles actuellement solidement établis dans le domaine, comme nous avons essayé de le montrer dans le paragraphe précédent. Cependant nous souhaitons argumenter ici les ressemblances et les différences entre notre modèle et les autres travaux ayant les mêmes exigences pluralistes. Les deux modèles auxquels nous nous référerons ici sont d'une part celui de Grosz et Sidner (1986) et d'autre part le modèle de Grau, Luzzatti, Sabah et Vilnat (1992).

1- Le modèle de Grosz et Sidner :

Fondé sur trois composantes principales, ce modèle n'est pas véritablement fédérateur au sens où des modèles complets coopèrent les uns avec les autres, mais néanmoins, il nous paraît être le premier à rendre compte de la nécessité d'inférences en provenance de plusieurs types de connaissances sur un même objet. Ainsi la *composante intentionnelle* est fondée sur les notions de croyances mises en avant par Allen et Perrault (1980). La *composante linguistique* rend compte de la position de l'élément référé (rappelons que ce modèle s'intéressait largement à la représentation de la référence, et que les principales applications ont été la résolution d'ambiguïtés référentielles dues à l'anaphore et à certaines

ellipses). Enfin, la *composante attentionnelle*, prépondérante dans le modèle, notait que les deux autres composantes étaient subordonnées à un empilement thématique. Dès lors, ce modèle, qui propose déjà au moins trois des sources d'information, posait comme condition la hiérarchisation des sources, et mettant la thématique en avant.

2. Le modèle de Grau, Luzzatti, Sabah et Vilnat :

Fondé sur des travaux préalables de cette même équipe, ce modèle se décline déjà en *modules interagissants*, ce qui renforce probablement l'aspect dynamique, et la notion de modèles (au pluriel) est déjà citée. Les auteurs en proposent trois, dont ils font coopérer les résultats. Un *module thématique*, fondé sur les travaux de Grau (1983), s'intéresse à la gestion des changements thématiques, qui sont renforcés par les attentes du deuxième module, le module structurel. Ce dernier, nommé *module intentionnel* par les auteurs, est dominant dans ce modèle. L'intentionnalité n'est pas introduite sous forme de contenu, mais par le biais d'une structuration inspirée des travaux de Roulet (1985). Cette structuration impose la notion d'"attente" associée à une intervention, dont nous avons adopté le principe dans le modèle structurel de PLUS. Enfin, cette approche possède un module interactionnel, qui gère la communication en tant qu'ensemble de comportements (travaux de Luzzatti, op.cit.).

Fortement appuyé sur ces deux modèles, le modèle de l'énonciation de PLUS s'en distingue cependant par deux aspects : premièrement, dans notre modèle il n'y a pas de composant ou de modèle dominant. Tous les modèles, comme d'ailleurs les modules qui les traitent, sont vus comme ayant des compétences différentes et nécessaires. Donc nous préconisons une modélisation distribuée plutôt que hiérarchisée, en relation avec une notion de **contrôle** telle que la définit Sabah (1990-b), dans laquelle un "superviseur" gère les communications et les priorités entre les éléments d'information (processus et bases de connaissances).

Deuxièmement, le modèle de l'énonciation de PLUS ne s'attache pas plutôt à l'énoncé (comme Grosz et Sidner) ou plutôt à la portée dialogale (comme Grau et al.). Les deux éléments cohabitent dans la vision que nous tentons de promouvoir. Ce qui nous permet de "mettre à plat" ainsi l'ensemble des domaines de compétence, c'est l'existence *d'un aspect architectural fondé sur la notion de contrôle, par un module dédié, de la bonne marche de l'interaction*. Comme l'a montré Sabah (1991-b), la robustesse d'un système de dialogue Homme-machine peut être obtenue par le biais de la coopération, à la mesure de chacune, des sources d'information multiples qu'exige une interprétation pragmatique. Nous pensons que notre approche du modèle de l'énonciation est une application de ce principe.

4. MATERIALISATION DU MODELE DE L'ENONCIATION

Le modèle de l'énonciation se matérialise par le biais de bases de connaissances, de modules d'inférences qui alimentent ces bases ou s'appuient sur leur contenu, et que nous détaillons dans les prochains paragraphes. Afin d'appuyer notre description des différents acteurs de notre architecture, nous nous appuyerons sur un exemple de dialogue recueilli dans le cadre d'une acquisition de corpus pour le projet [Pernel 1991]□

U1 J'ai besoin d'une voiture de location à Atalante ?
S1 Où est Atalante ?
U2 Atalante est près de Rennes.
S2 Il y a 15 compagnies répondant à votre requête. Voulez-vous les voir toutes ou préférez-vous affiner votre requête ?
U3 Y en a-t-il à Atalante même ?
S3 Non, mais vous pouvez apporter des précisions supplémentaires. Par exemple les pages jaunes connaissent les critères suivants : gamme importante, aller simple, kilométrage illimité.
U4 Aller simple.
S4 Où voulez-vous la rendre ?
U5 A l'aéroport de Rennes.
S5 Voici la liste...

4.1 Bases de connaissances pragmatiques coopérantes

Elles vont permettre de représenter tous les modèles fédérés au sein du modèle de l'énonciation.

4.1.1 Bases de connaissances générales sur le monde et spécifiques à l'application

Ces deux bases sont statiques dans le sens où elles ne sont pas modifiées au cours d'une interaction Homme-machine.

Les **connaissances générales du monde** contiennent les concepts généraux nécessaires pour comprendre correctement tous les types d'interactions orientées par la tâche. Par exemple, les concepts de voyage, de location, de service, d'action, d'événement font parties des connaissances générales. Un langage objet (CML) les représentent à l'aide d'une hiérarchie "est-un" où chaque noeud est une classe qui contient à la fois des propriétés et des règles. Ces dernières modélisent le "sens commun" elles vont permettre au système de faire des inférences pour mettre à jour d'autres bases pragmatiques.

```

CLASS HireEvent
ISA Event
WITH
  hirer: IN IntelligentAgent;
  object: IN HireObject;
  provider: IN ServiceSupplier;
  collect_location: IN Address IS SINGLE
  return_location: IN Address IS SINGLE
  cost : IN INTEGER
RULES
default_collect_location:
  DESCRIPTION "Par défaut, les objets sont loués à l'adresse du fournisseur"
  PREMISSES "_a in HireEvent and _a:provider:address = _adr and
    consistent (_a:collect_location = adr)"
  CONCLUSION "_a:collect_location = _adr"
default_return_location:
  DESCRIPTION "Par défaut, les objets loués sont rendus à l'adresse du fournisseur"
  PREMISSES "_a in HireEvent and _a:provider:address = _adr and
    consistent (_a:return_location = adr)"
  CONCLUSION "_a:return_location = _adr"
cost:
  DESCRIPTION "Détermine le coût de location"
  PREMISSES "_a in HireEvent and _a:object:rate = _r and _a:temporal_duration = _t and
    _c == _r * _t"
  CONCLUSION "_a:cost = _c"
ENDCLASS

```

Les **connaissances spécifiques à l'application** sont de trois types□

1□□□Un modèle de la base de données définit la structure et le contenu de la base de l'application (ici, les pages jaunes) en terme de concepts pour en obtenir une description sémantique□

2□□□Un modèle d'accès à la base de données de l'application. Il définit le moyen de retrouver l'information stockée dans la base de données au moyen d'une taxinomie des entêtes des rubriques des pages jaunes. Il existe en plus, un type de lien transversal "*relatif à*" qui associe deux rubriques sémantiquement proches□

3□□□Un modèle d'unification des concepts à la taxinomie des pages jaunes. Il définit la structure des informations disponibles dans la base de données.

```

INSTANCE 02050600 IN Heading
WITH
  superheading = 02050000;
  related_to = 02050300, 08202400;
  name = apiculture;
ENDINSTANCE

```

4.1.2 Bases des thèmes du dialogue

La notion de **thème** peut être définie à deux niveaux□

1□□□Niveau de l'intervention□ il correspond à l'idée que le locuteur y développe, aux articulations intentionnelles et argumentatives□

2□□□Niveau de l'échange□ il s'agit ici davantage du foyer, autrement dit, de quel(s) objet(s) l'échange a traité. C'est surtout à ce niveau que l'évolution des thèmes est importante à saisir pour découvrir les véritables intentions de l'utilisateur, c'est-à-dire les relations sémantiques qui existent entre les thèmes et les différents échanges.

Intervention U1-U2 : S1-S2 est une clarification de U1.
Intervention S2 : S2a est une justification de S2b.
Intervention U1-U4 : S2a-U4 est une spécification de U1-U2.

Echange S1-U2 : localisation d'Atalante.
Echange S2-U4 : compagnies de location de voitures.
Echange U3-S3 : compagnies de location de voitures à Atalante.

4.1.3 Modèle de l'utilisateur

Il définit la base de connaissances qui contient les hypothèses explicites faites par le système sur tous les aspects de l'utilisateur. Tout est décrit comme un *jugement que porte le système sur l'utilisateur*.

Il contient deux sous-bases²

- 1² Une sous-base qui rassemble les hypothèses faites par défaut,
- 2² Une sous-base qui rassemble les déductions qu'opère le système.

Chaque sous-base se divise en unités plus petites qui sont constituées d'informations sur²

- 1² Les intentions et les buts de l'utilisateur,
- 2² Les croyances et les connaissances de l'utilisateur,
- 3² Les croyances et les connaissances de l'utilisateur sur le système.

Les intentions de l'utilisateur permettent de définir une typologie des buts de l'utilisateur. De manière symétrique, le système a besoin de connaître tous les buts possibles qu'il peut accomplir pour satisfaire ceux de l'utilisateur. A chacun de ses tours de parole, le système doit en choisir un ou plusieurs en fonction de tous les événements passés.

Croyances du système après U1 :

```
BEL (sys, I (user, BEL (sys, P2))),  
BEL (sys, BEL (user, P)),  
BEL (sys, P),  
BEL (sys, I (user, BELV (user, select (COMPS, lambda (_8664, object,  
exist (OBJECTS, lambda (_8671, object, and (Car (_8671), Supplies (_8664,  
_8671))))))))))
```

*But de l'utilisateur*²

```
INSTANCE ug1 IN Request  
WITH  
  objecttype = ServiceSupplierType1  
  owner = user  
  infodirection = required  
ENDINSTANCE
```

4.1.4 Modèle structurel du dialogue

Ce modèle va définir dynamiquement une structure du dialogue en cours, à l'aide d'une "grammaire de dialogue". Il va attacher aux *actes illocutoires*, un **acte de dialogue** (terminologie issue de [Bunt 1989]) qui est le seul élément terminal de la structure. Ils sont organisés au sein d'interventions et d'échanges.

De manière traditionnelle dans la littérature, les constituants d'un échange possèdent une fonction illocutoire qui indique clairement leur rôle dans l'échange. Elle peut prendre trois valeurs essentiellement² initive, réactive, évaluative.

Les constituants d'une intervention entretiennent des liens hiérarchiques, autrement dit, l'un d'entre eux est dit *directeur* et les autres *subordonnés*. Les constituants subordonnés

² P est la représentation contextuelle de la phrase.

4.2 Modules d'inférence

Les modules du système de traitement et de production du dialogue se décomposent comme suit [Black 1991] :

1. Un analyseur morphosyntaxique et sémantique correspondant à une technique d'analyse par *chart* ;

2. Un module de gestion du dialogue, que nous avons introduit dans §2, qui reçoit en entrée la représentation sémantique de l'intervention de l'utilisateur (proposition par proposition), et produit en sortie la représentation sémantique de la réponse du système,

3. Un générateur de surface qui transforme la représentation sémantique fournie par le gestionnaire de dialogue en phrases de la langue.

Comme nous l'avons déjà mentionné dans le paragraphe §2, les modules composant le gestionnaire du dialogue du système sont :

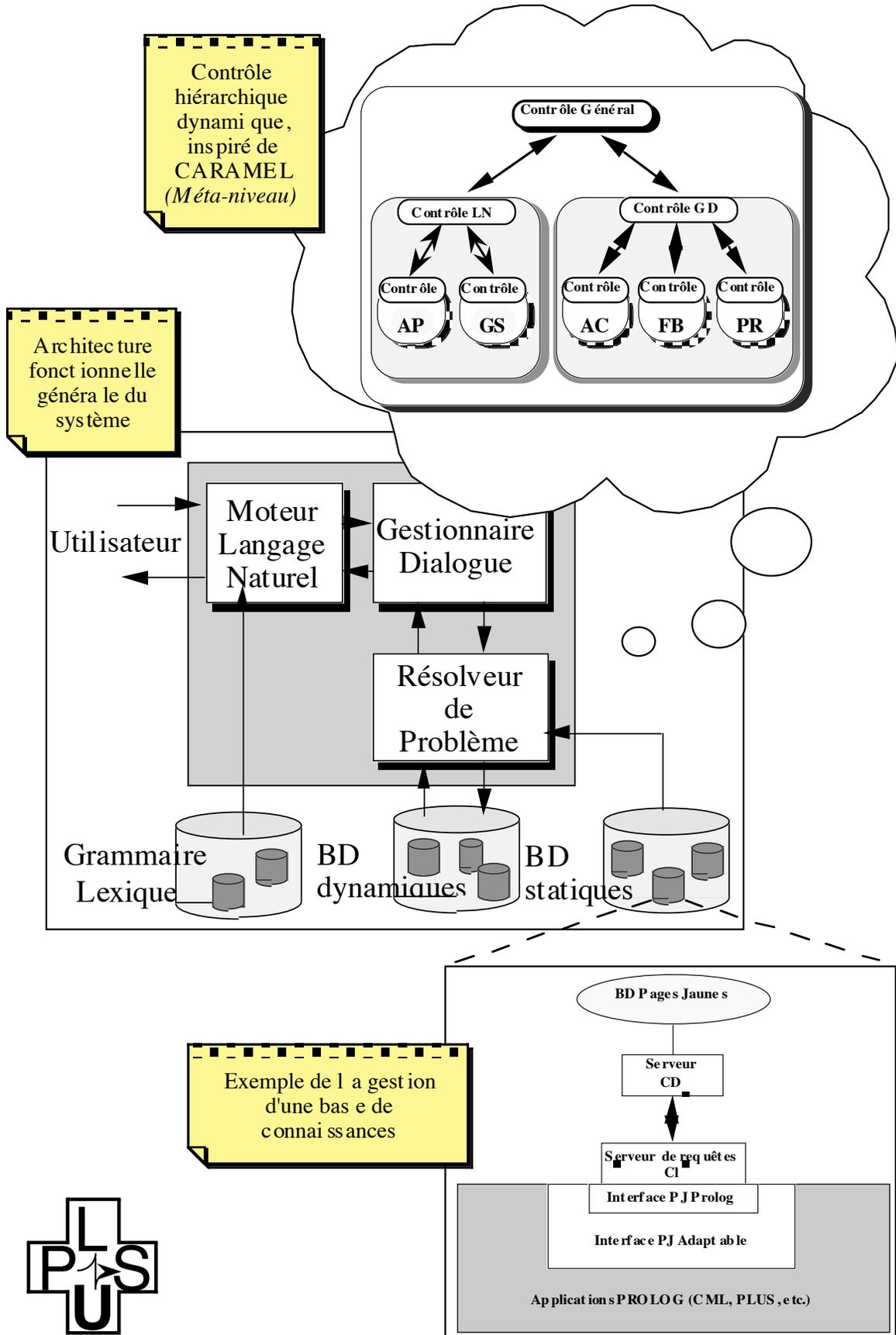
1. Un analyseur cognitif, qui reçoit en entrée une représentation sémantique de l'acte de langage émis en situation de dialogue Homme-machine (auquel il sera attribué le nom d'**acte de dialogue**), lorsque cet acte est réalisé par l'utilisateur. Il le transforme en :

- l'intention ou croyance dans le modèle de l'utilisateur, en fonction de l'état des croyances en cours ;
- l'objectif communicatif, articulation du plan de l'utilisateur, en fonction duquel le système pourra planifier sa propre intervention, moyennant l'état de la structure du dialogue ;
- le thème ou foyer local, permettant d'indiquer éventuellement des modifications par rapport à l'attente du système s'il y a eu changement vis-à-vis du thème en cours.

2. Le formulateur de buts, dont le rôle planificateur est relativement central. Recevant en entrée les sorties de l'analyseur cognitif, il doit engendrer, en sortie, le but communicatif du système pour répondre au mieux à la requête de l'utilisateur, selon son but ultime. Sa tâche première est déterminée a priori par la tâche globale du système ; en l'occurrence, fournir les noms et les coordonnées d'entreprises. Pour faire cela, il recherche le véritable plan de l'utilisateur et décide alors de la meilleure stratégie à adopter, pour que ce plan se réalise.

3. Le planificateur de la réponse, dont le rôle est symétrique de celui de l'analyseur cognitif. A partir des buts communicatifs du système engendrés par le formulateur de buts, et en fonction des croyances actuelles de l'utilisateur, de l'état du dialogue en cours (thématique, positionnel), il engendre une représentation sémantique à partir de laquelle sera produite la forme de surface.

4. Enfin, et un module transversal de contrôle, fondé sur des travaux en matière d'architecture pour le traitement du langage naturel [Sabah 1990-b]. Il doit permettre un fonctionnement robuste des divers modules du gestionnaire du dialogue. Sa fonction est de réaliser une planification dynamique de la séquence des modules à solliciter [Sabah 1991-a] [Sabah et al. 1991] dans la mesure où des interventions de l'utilisateur peuvent nécessiter une révision de la chronologie {analyseur, analyseur cognitif, formulateur de but, planificateur de réponse, générateur de surface}. Cela est vrai en particulier pour des interventions mal formées, hors sujet, intempestives, etc.



Contrôle hiérarchique dynamique, inspiré de CAMEL (Méta-niveau)

Architecture fonctionnelle générale du système

Exemple de la gestion d'une base de connaissances



4.3 Le lieu de rencontre des processus et des bases de connaissances □ l'historique du dialogue

4.3.1 La notion d'historique du dialogue

La notion d'historique du dialogue telle que nous la concevons, c'est-à-dire en tant que véritable chronique de la conversation, a été introduite de manière explicite par [Vilnat et Sabah 1985] à la suite de [Vilnat 1984]. Dans cette approche, l'historique était conçu comme un carnet de notes relatant l'ensemble des découvertes importantes du gestionnaire du dialogue et lui permettant de s'appuyer sur elles pour calculer le plan de ses interventions. Le modèle dominant étant un modèle structurel, les informations enregistrées servaient à compléter les positions par des éléments de contenu. Rappelons que cette approche représentait un premier pas dans l'appréhension d'un historique qui soit différent d'une simple "récapitulation" d'états antérieurs, faite à posteriori. L'historique y était envisagé a priori, et utilisé comme support pour la gestion du dialogue.

Des travaux ultérieurs ont permis, ainsi que nous l'avons mentionné, aux auteurs de se pencher sur le problème d'un modèle de dialogue pluraliste (voir Grau et al., 1992 op.cit.). Dès lors, dans ces travaux, l'historique s'est développé pour devenir davantage un enregistrement d'états ponctuels de diverses sources d'information nécessaires à la réalisation de la tâche de dialogue. La conjugaison du modèle de l'interaction de Luzzatti avec le modèle de structure tel que Vilnat l'avait adapté a conduit les auteurs à imaginer un historique reflétant l'état concomittant des deux modèles.

Pour le projet PLUS, nous reprenons et étendons cette idée à un historique devenant le lieu de rencontre des divers modèles coopérants. En effet, les différents modèles alimentés et sollicités doivent être mis en correspondance afin que les modules de traitement puissent y puiser les diverses notions dont ils ont besoin. Par ailleurs, un historique devenu prépondérant permet de nuancer les rôles des différents modèles : ainsi, il n'y a pas dominance d'une approche (structurale ou intentionnelle) sur l'autre. Enfin, nous pensons que seul un historique aussi central peut efficacement servir l'extension d'un modèle de l'énonciation depuis des modèles élaborés du discours et des partenaires vers des modèles dus à des modalités communicatives différentes (discours écrit, oral, partiellement complété par des graphiques, etc.).

Pour Grau et al., aussi bien que pour nous, l'historique conserve les propriétés que lui confère l'informatique traditionnelle □ c'est un ensemble d'informations de type cumulatif (l'ensemble ne subit pas d'opérations de rétraction). Autrement dit lorsque le système veut remettre en cause la valeur d'un champ, il doit commencer par dupliquer la totalité de l'enregistrement et seulement après il peut modifier la valeur. Des procédures de restauration d'états antérieurs peuvent rendre de nouveau active une situation où le système a fait un choix qu'il remet en question.

Chez les auteurs cités, le contenu de chaque structure de cet ensemble correspond à des pointeurs sur des états des modèles en cours d'élaboration, et plus particulièrement il comprend le thème en cours, la feuille de l'arbre du dialogue engendré par le modèle structurel et le niveau dans le modèle interactionnel.

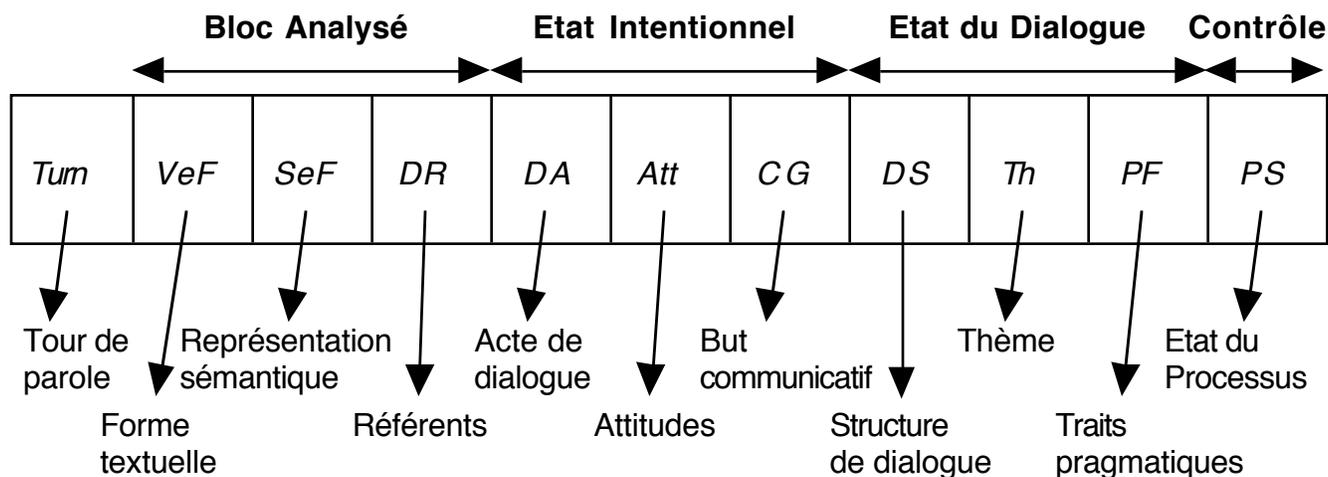
Dans [Prince et al. 1991), nous définissons **l'historique** comme étant un ensemble structuré correspondant à une collection d'événements, chaque événement correspondant à un acte de dialogue. L'événement est représenté par une structure, elle-même composée d'index et de variables, dont le but est de représenter □

- 1 □ une structure de l'événement dialogal lui-même ;
- 2 □ ensemble des états des modèles coopérants produits par l'occurrence de l'événement □
- 3 □ éventuellement, une "appréciation" de l'événement (dévolue au module de contrôle) en terme de succès de traitement (par tel ou tel module faisant des inférences) ou en terme d'indications stratégiques à conserver en cas d'échec partiel [Sabah et al. 1991]

4.3.2 Matérialisation de l'historique proposé pour PLUS

Voici la structure d'un enregistrement de l'historique implanté dans PLUS □

- 1 □ un lien vers l'acte locutoire qui est une forme d'une grammaire HPSG □
- 2 □ un lien vers une structure se rapportant aux objets définis dans le discours □
- 3 □ un lien vers le noeud courant de l'arbre des thèmes □
- 4 □ un lien vers les hypothèses qui viennent d'être déduites sur l'utilisateur □
- 5 □ un lien vers le but communicatif courant □
- 6 □ un lien vers le type d'acte de langage réalisé □
- 7 □ un lien vers les traits pragmatiques □ satisfaction ou non, désir... □
- 8 □ un lien vers le noeud courant de la structure du dialogue □
- 9 □ un lien vers l'appréciation faite par le contrôle.



5. CONCLUSION

Les points pour lesquels nous pensons avoir présenté une argumentation et sur lesquels nous voudrions conclure s'énoncent par les propositions suivantes :

(i) *pour une application réellement robuste, il y a nécessité de plusieurs modèles coopérants.*

Nous pensons en effet que l'originalité de notre modèle de l'énonciation ne tient pas en ce qu'il approche l'interaction Homme-machine d'une manière différente des autres modèles, mais qu'au contraire, il essaie d'intégrer l'ensemble des démarches actuellement en cours en matière de dialogue Homme-machine. Nous nous fondons sur le fait que l'apport de chaque tendance dans les modèles de dialogue est original et nous paraît nécessaire si l'on veut rendre compte de tous les aspects de l'interaction. Or, si un système se veut robuste, c'est-à-dire tolérant et non contraint, alors il nous semble crucial de pouvoir bénéficier de tout apport d'information et d'inférences pour la gestion de l'interaction. Dès lors, on s'aperçoit qu'il y a nécessité d'une modélisation multiple, car chaque tendance correspond elle-même à un modèle élaboré, avec ses paradigmes, ses systèmes de contraintes et ses résultats.

(ii) *Ces modèles multiples peuvent aussi s'étendre à la représentation des modalités multiples.*

Une des réflexions que nous nous sommes faites tout au long de notre travail théorique sur le modèle de l'énonciation est la suivante. Etant donné l'autonomie des modèles (en dépit de leur coopération constante) qui composent le modèle de l'énonciation, il est tout à fait possible qu'ils obéissent aux contraintes de modes multiples (mode oral, mode écrit, mode graphique) lorsque ceux-ci doivent coopérer au sein d'une interaction donnée. Etant donné que la coopération est elle-même la base de la mise en œuvre de ces modèles, alors le modèle de l'énonciation nous paraît aussi un bon fondement pour la spécification des caractéristiques d'un modèle multimodal.

(iii) *Le principe de coopération des modèles multiples doit passer par la matérialisation d'une zone de communication à la fois synchronique et diachronique : l'historique du dialogue.*

Notre thèse est que l'historique joue le rôle de pilier d'une configuration autorisant la coopération de modèles et de modules multiples et spécialisés. Au-delà de son aspect chronologique, c'est surtout en terme de repérage événementiel, de capacité de recul dans les résultats des différents éléments que l'historique peut être exploité. C'est donc une zone privilégiée de communication entre les processus, en même temps qu'une mémoire de l'état des modèles. Cette double appartenance en fait un carrefour important du modèle de l'énonciation.

BIBLIOGRAPHIE

[Allen & Perrault 1980] Allen J. et C.R. Perrault *Analysing Intention in Utterances*". In *Artificial Intelligence*, vol 15. Pp. 148-178.1980

[Allwood 1976] Allwood J. *Linguistic communication as action and cooperation*. Gothenburg University Monographs in Linguistics, vol 1.1976

[Austin 1962] Austin J. *How to do things with words*. Clarendon Press. Oxford.1962

[Bilange 1991] Bilange E. *Modélisation du dialogue oral finalisé personne-machine par une approche structurelle*, Thèse de l'Université de Rennes I.1991

[Black 1991] Black B (ed.) *Functional Design*, version2.1, PLUS deliverable. Dec. 1991

[Bunt 1985] Bunt H.C. *Mass terms and model-theoretic semantics*. Cambridge University Press.

[Bunt 1989] Bunt H.C. *Information Dialogues as communicative action in relation to partner modelling and information processing*. In *The Structure of multimodal dialogue*, MM.Taylor, F. Néel & D.G. Bowhuis ed. North Holland.

[Carberry 1989] Carberry S. *Modelling the user's plans and goals*. *Computational Linguistics*, 14:23-27, 1988.

[Grau 1983] Grau B. *Analyse et Représentation d'un texte d'après le thème du discours*. Thèse de doctorat de 3ème cycle. Université Paris VI.

[Grau et al. 1992] Grau B., Luzzatti D., Sabah G. et A. Vilnat *A pragmatic driven Man-machine Dialogue System*. Document de travail LIMSI. mars 1992

[Grosz & Sidner 1986] Grosz B. et C. Sidner *Attention, intentions and the structures of discourse*, *American journal of computational linguistics*, 12, 3, pp 175-204.1986

[Luzzatti 1989] Luzzatti D. *Recherches sur le dialogue homme-machine : modèles linguistiques et traitements automatiques*. Thèse de doctorat d'état ès lettres, Université de Paris III.1989

[Moeschler 1985] Moeschler J. *Argumentation et conversation, éléments pour une analyse pragmatique du discours*, Hatier, Paris. 1985

[Pernel 1991] Pernel D. Le Corpus "PLUS". Notes et Documents LIMSI : 91-19. 1991

[Pierrel 1987] Pierrel J.M. *Dialogue oral homme-machine*, Hermès Paris.1987

[Prince et al. 1991] Prince V, Pernel D. et C. Godin *Discourse model & dialogue history*, Internal Paper PLUS.sept. 1991

[Roulet et al. 1985] Roulet E., Auchlin A., Moeschler J., Rubattel C., et M. Schelling *L'articulation du discours en français contemporain*. Ed. Lang. Berne.1985

[Sabah 1990-a] Sabah G. *L'intelligence artificielle et le langage (tome 2: Processus de compréhension)*. Hermès. Paris. 1990

[Sabah 1990-b] Sabah, G. *CARAMEL : a flexible model for interaction between cognitive processes underlying natural language understanding*, in COLING-90. Helsinki.

[Sabah 1991-a]Sabah G. *Control in the PLUS architecture* document interne du projet ESPRIT II /PLUS. Juillet 1991.

[Sabah 1991-b] Sabah G. "*Dialogue Homme-Machine : Pragmatique et Robustesse*" Tutorial congrès RFIA-91. AFCET ed. Lyon.Nov. 1991

[Sabah et al. 1991] Sabah G., Godin C. et A. Derain "*Proposition for the Control Architecture of PLUS*" Document interne du projet ESPRIT II/PLUS in "*Deliverable 1.2□ FUNCTIONAL DESIGN*"(Bill Black ed.). Décembre 1991

[Vilnat 1984] Vilnat A. *L'élaboration d'interventions pertinentes dans une conversation homme-machine*, thèse de 3e cycle, Université Pierre et Marie Curie.1984

[Vilnat 1988] Vilnat A.*Le dialogue homme-machine en recherche d'information*. Séminaire ATALA. Paris.1988

[Vilnat 1989] Vilnat A. *Relevant responses in man-machine conversation*. In *The Structure of multimodal dialogue*, MM.Taylor, F. Néel & D.G. Bowhuis ed. North Holland. Pp 399-406.1989

[Vilnat & Sabah 1985] Vilnat A. et G. Sabah *Be brief, be to the point...be seated or Relevant responses in man-machine conversation*. IJCAI-85. Los Angeles.