

# Un système basé sur les types logiques pour la compréhension de la parole

Jeanne Villaneau-Berthélemy

VALORIA (EA2593), Université de Bretagne Sud (Lorient-Vannes)  
e-mail Jeanne.Berthelemy@univ-ubs.fr

---

## Résumé

Cet article présente un système de compréhension d'une requête orale dans le cadre d'un dialogue homme-machine. Si le dialogue est finalisé, son domaine est néanmoins relativement vaste : la construction d'une représentation sémantique de l'énoncé exige une analyse linguistique plus fine que celle utilisée dans les applications classiques de CHM<sup>1</sup>. Le système présenté ici, entièrement lexicalisé, utilise à différents niveaux les types logiques : au niveau syntaxique, il reconstitue les groupes de mots par composition de  $\lambda$ -termes ; à un niveau plus sémantique, il compose les différents constituants pour construire une formule logique qui correspond à la représentation sémantique de la requête. Les tests réalisés semblent valider la méthode.

---

## 1. Introduction et présentation générale du système

Les progrès accomplis ces dernières années dans la reconnaissance de la parole ouvrent des perspectives dans le domaine de la compréhension automatique de la parole (CAP) appliquée à la CHM. Pour appréhender de manière robuste les manifestations inhérentes à la parole spontanée (hésitations, répétitions, corrections, etc.), les systèmes de CAP actuels reposent sur une extraction de mots ou de séquences-clefs qui réduit la compréhension à la détection d'un sous-ensemble de mots pour représenter le sens "utile" de l'énoncé correspondant. Ces systèmes donnent des résultats intéressants dans des domaines très finalisés (Minker, 1999). On peut cependant penser qu'élargir le cadre d'application de telles méthodes risque de poser problème (Antoine J-Y., 1999).

Le système de CAP présenté ici a pour objectif de proposer une analyse plus fine mais néanmoins robuste d'énoncés oraux spontanés, dans un domaine finalisé mais néanmoins plus étendu que celui des applications classiques : celui des renseignements touristiques. Il s'inscrit dans une chaîne de traitement classique avec, en amont, une étape de reconnaissance de parole qui extrait du signal de parole une (ou plusieurs : analyse N-best) séquence de mots sensée correspondre à l'énoncé prononcé. La CAP travaille donc sur cet énoncé. En aval, un module de gestion de dialogue assure la liaison avec la base de données (élaboration de la requête et de la stratégie de dialogue avec l'utilisateur).

---

1. CHM : Communication Homme-Machine.

Pour la représentation sémantico-pragmatique de l'énoncé, le formalisme choisi est une formule logique obtenue par composition de ses différents constituants. L'analyse elle-même utilise le  $\lambda$ -calcul typé (Warren, 1983), équivalent des grammaires catégorielles (Moorgat, 1997). Les différents constituants de la phrase sont obtenus par étapes de compositions successives, un peu comme dans l'analyse en cascades proposée par Abney (Abney, 1996). La souplesse du système vient de sa tolérance aux incorrections grammaticales et aux variations dans l'ordre des mots ainsi qu'à la possibilité de ne pas prendre en compte certains d'entre eux.

Le langage de programmation utilisé est  $\lambda$ Prolog. Comme son nom le suggère,  $\lambda$ Prolog<sup>2</sup> est un langage de programmation logique dont les termes sont les  $\lambda$ -termes simplement typés munis des  $\alpha$ ,  $\beta$  et  $\eta$ -équivalences (Ridoux, 1998) ; ses formules permettent également l'utilisation du connecteur de l'implication intuitionniste. Ces deux extensions de Prolog élargissent le champ d'applications classique de Prolog pour l'analyse grammaticale (Miller D.A., 1986).

## 2. Formule sémantique et types logiques

Les phrases du corpus Pariscorp composé de demandes d'informations touristiques sur la ville de Paris ont servi de base de travail (Rosset S., 1997). Pour cette raison, l'analyse s'est jusqu'à présent essentiellement focalisée sur l'analyse des requêtes. Les formules et les objets ont respectivement les types atomiques *requete* et *objet*. Le type d'une question portant par exemple sur un objet est donc (*objet*  $\rightarrow$  *requete*). Si l'on se réfère aux formules de Montague (Montague, 1974), les objets jouent le rôle des groupes nominaux alors que les différentes questions correspondent aux verbes. La construction des formules considérées comme correctes nécessite donc préalablement l'élaboration d'un système des objets du discours et des différentes questions.

**Les objets :** ils sont obtenus par application d'une *étiquette* à des *propriétés*. Dans l'objet (*hotel (et (etoiles 2) (proche\_de (musee (nom "Louvre"))))*), *hotel* et *musee* sont des étiquettes, (*etoiles 2*) , (*nom "Louvre"*) et (*proche\_de (musee (nom "Louvre"))*) sont des propriétés. Le *et* permet la coordination des propriétés. Le type *proprietes* est atomique, le type d'une étiquette est donc (*proprietes*  $\rightarrow$  *objet*). L'absence de propriété particulière est désignée par la propriété *qq*. Ainsi (*hotel qq*) désigne l'ensemble de tous les hôtels.

La relation binaire *est\_sous\_objet\_de* est définie sur les étiquettes des objets. Elle établit entre ceux-ci une relation sémantique de dépendance : ainsi un objet d'étiquette *chambre* est lié à un objet d'étiquette *hotel* alors qu'un objet d'étiquette *reservation* peut être lié à un objet dont l'étiquette appartient à une liste qui comporte entre autres l'étiquette *chambre*. Ces dépendances entre objets peuvent être représentées par un graphe acyclique et reposent bien évidemment sur des considérations sémantiques. Quant aux propriétés, chacune d'entre elles est sémantiquement liée à la liste des étiquettes des objets auxquels elle peut être attachée.

**Les questions :** elles peuvent porter sur des objets ou sur des propriétés liées à ces objets. Par exemple, la question *tarifs\_q* porte sur un seul objet. Ainsi la phrase exemple :

"Une chambre dans un hotel deux étoiles pas loin du Louvre est-ce que vous pouvez me dire combien ça coûte?" est traduite par :

(*tarifs\_q (chambre (de (hotel (et (etoiles 2) (proche\_de (musee (nom "Louvre"))))*))))

---

2. La version de  $\lambda$ Prolog utilisée est Prolog/MALI, implémentation réalisée à l'IRISA de Rennes.

### 3. Les différentes étapes de l'analyse

L'analyse de la phrase se fait en deux étapes : la première consiste à reconnaître les différents constituants non récursifs suivant des principes comparables à ceux du *chunking* (Abney, 1991). La deuxième est la construction de la formule logique à partir de ces constituants. Le principe, entre autres, que l'ordre des mots dans un constituant est beaucoup plus stable qu'entre les constituants semble particulièrement adapté à l'expression orale (Blanche-Benveniste C., 1990). Par ailleurs, des considérations sémantiques et pragmatiques sont progressivement prises en compte au fil de l'analyse.

**Première étape :** pour préserver la finesse et la robustesse du système, le problème est de prendre en compte le plus grand nombre possible des mots de la phrase sans pour autant bloquer l'analyse. La première phase de cette première étape a pour objet la reconstitution de groupes de mots qui, pris isolément, risqueraient d'être considérés comme non essentiels à la compréhension. À ce stade, les regroupements sont essentiellement syntaxiques et l'ordre des mots est très rigoureusement pris en compte. La deuxième phase est la reconstitution des objets sur lesquels portent la requête : elle met déjà en œuvre un certain nombre de considérations sémantiques.

**Groupelements syntaxiques.** Le lexique donne, pour chacun des lexèmes<sup>3</sup> reconnus, une liste d'équivalents lexicaux<sup>4</sup>. Certains de ces équivalents lexicaux ne sont pris en considération que s'ils sont directement adjacents à d'autres équivalents lexicaux avec lesquels ils peuvent se composer (composition de  $\lambda$ -termes). C'est le cas en particulier pour les articles, démonstratifs et prépositions qui ne seront conservés que s'ils précèdent certains autres équivalents correspondant par exemple à des étiquettes des objets du discours. Pour éviter des "fausses adjacences", les mots non reconnus laissent en général une trace (liste vide) alors que les marques d'hésitation par exemple ("euh", "bon", "disons" etc.) sont effacés sans laisser de trace. Ainsi, l'un des équivalents lexicaux du mot "pas" est (*adjd*  $\lambda x.(pas\ x)$ ), celui de "loin" est (*adj* (*pas* *proche*)). Suivant un principe de calcul très proche du calcul de Lambek (Moorgat, 1997), le prédicat de composition comporte une clause selon laquelle la composition de (*adjd* *F*) avec (*adj* *A*) a pour résultat (*adj* (*F* *A*)). Ainsi, "pas loin" a pour équivalent lexical (*adj* (*pas* (*pas* *proche*))).

À ce stade de l'analyse, les compositions sont prioritaires ; lorsque deux équivalents lexicaux adjacents peuvent se composer, l'équivalent obtenu par composition élimine les équivalents parallèles qui ne correspondent pas à des compositions. Le principe heuristique ainsi mis en œuvre est le suivant : lorsqu'un groupe de deux mots "élémentaires" qui se suivent a une signification particulière, les autres sens de chacun des deux mots pris séparément ne sont pas pris en compte. Cependant, si plusieurs compositions sont possibles, les composés obtenus restent en parallèle pour la poursuite de l'analyse.

Lorsque toutes les compositions envisagées ont été opérées, les équivalents lexicaux considérés comme incomplets ( $\lambda$ -termes, le plus souvent) sont éliminés. Ainsi dans la phrase exemple, à l'issue de cette première phase, le regroupement des mots correspond à :

[Une chambre] [dans un hotel] [pas loin] [du Louvre] [est-ce que] [pouvez] [dire] [combien] [coûte]

---

3. Certaines expressions telles que *tout de suite*, *à partir de*, *y a-t-il...* sont reconnues comme des lexèmes uniques.

4. Le terme d'*équivalent lexical* désigne un type générique de résultat partiel de l'analyse

**Reconnaissance des objets.** La deuxième phase de l'étape consiste à reconnaître des groupes d'équivalents lexicaux correspondant aux différents objets du discours. L'analyse des propriétés sémantiques et la prise en compte du contexte de lieu et de temps font partie intégrante de l'analyse ; ils permettent d'établir des liens entre les différents éléments (Sleator D., 1991). Il est évidemment fait appel aux relations de dépendances définies entre les objets ainsi qu'à celles entre les propriétés et les objets. Ainsi une propriété ne peut être interprétée que si elle peut être rattachée à un objet de la phrase ou à défaut du contexte<sup>5</sup>. C'est aussi dans cette deuxième phase qu'un certain nombre de marques d'hésitation, de reprise ou de correction sont analysées (étude et comparaison d'objets ou de propriétés juxtaposés, avec ou sans mots de coordination ou de correction).

À la fin de cette étape, la phrase est sous la forme d'une liste de listes d'équivalents lexicaux. Chacun d'eux correspond soit à la définition d'un objet, soit à la définition d'une propriété de l'un de ces objets, soit à la définition de la nature même de la requête. Pour la phrase exemple, deux étapes de composition ont concerné les deux propriétés (*etoiles 2*) et (*proche\_de ...*): elles ont d'abord été interprétées comme telles puis ont été incorporées à *hotel* car situées immédiatement derrière. La liste résultat est (seul le premier élément de chaque liste est donné) :  
 [[(objet chambre qq pasdeprep), ...],  
 [(objet hotel (et (etoiles 2) (proche\_de (objet musee (nom "Louvre")))) dans), ...],  
 [(mot\_quest "oui\_non"), [(mot\_req "recherche"), [(mot\_req "recherche")]],  
 [(mot\_quest "combien"), [(mot\_req "prix")]].

**Deuxième étape :** elle consiste en la construction de la formule logique. Elle commence par une étude des équivalents lexicaux qui permettent de définir la nature de la requête. La reconstruction choisie est alors guidée par la nature de la requête ou à défaut, elle est guidée à la fois par la nature de la requête et par les objets.

#### **L'analyse de la nature de la requête.**

Les mots qui permettent d'analyser la nature de la requête sont de deux sortes :

- les mots-questions tels que *est-ce que, faut-il, peut-on, quand, où*, etc. Leur présence assure que la phrase analysée est une requête, la réciproque étant fausse.
- les mots-requêtes tels que *prix, heure, aller, adresse, trajet*, etc.

Un relation d'ordre partiel est définie sur chacune de ces deux catégories. L'analyse de la nature de la requête consiste à faire opérer les relations d'ordre puis à composer la sélection du (des) mot(s)-question(s) avec le(s) mot(s)-requête(s). Dans la phrase exemple, le mot-question "comment" l'emporte sur le mot-question "oui-non" et le mot-requête le plus fort est "prix". La composition de ces deux résultats correspond à la question (normalisée) *tarifs\_q*. Deux mots requêtes incomparables correspondent soit à une requête multiple (lorsqu'ils n'admettent pas de majorants), soit à une question plus élaborée (durée d'un trajet par exemple) qui est la borne supérieure des deux questions.

#### **La construction de la formule sémantique.**

Deux cas se présentent alors :

1. L'étude précédente détermine une (ou à la rigueur 2) question(s) précise(s) : la construc-

5. L'analyse d'une phrase conduit à la conservation de la question et des objets concernés pour l'analyse de la phrase suivante afin d'éviter les recours inutiles au module de dialogue ; l'implication intuitionniste de  $\lambda$ Prolog est utilisée dans ce but.

tion de la formule est alors guidée par la question. On recherche prioritairement dans le reste de la phrase ou (et) dans le contexte les objets et les attributs correspondant, avec vérification que les objets sont compatibles avec la question posée. Dans la phrase exemple, *chambre* et *hotel* sont tous les deux compatibles avec la question *tarifs\_q*. Mais le système cherche la chaîne d'objets la plus longue et *chambre* se compose avec *hotel* (c'est un sous-objet), d'où la formule résultat donnée précédemment.

2. Très souvent, les résultats de l'étude précédente sont trop vagues pour correspondre à une analyse du type précédent. Les phrases suivantes :

*"Je voudrais aller dans un cinéma qui passe des films en VO".*

*"Je voudrais aller en métro au Louvre"*

ne se distinguent pas par la forme sous laquelle est exprimée la nature de la requête. Ce sont les objets qui permettent de penser que la première phrase correspond à une recherche et la deuxième à une demande de trajet. Pour construire la formule sémantique, la stratégie consiste à composer l'analyse de la nature de la requête avec celle de la nature des objets et des propriétés de la phrase. L'analyse de la nature de la requête ne fait qu'autoriser certains choix inspirés par l'analyse des objets et des propriétés.

Il se peut également que l'absence ou la quasi-absence de mot-requête soit due à une référence à la requête précédente. Pour distinguer si l'analyse correspond à ce cas ou au précédent, la nature des objets et de leurs relations joue une fois de plus un rôle prépondérant.

Les requêtes multiples sont identifiées

- soit parce que la phrase est la coordination de deux requêtes :

*Quels sont les horaires d'ouverture du Louvre et comment on peut faire pour y aller d'ici ;*

- soit parce que l'analyse de la nature de la requête rend deux ou plusieurs résultats non comparables :

*Pouvez-vous me donner l'adresse et les tarifs de l'hotel Caumartin?.*

Dans les deux cas, la formule rendue est une coordination de deux formules-requêtes. Dans le premier exemple, l'objet (*musee (nom "Louvre")*) est ajouté au contexte (implication intuitionniste) pour l'analyse de la deuxième requête. Les analyses par extraction de mots-clefs ont beaucoup de mal à reconnaître ce genre de requêtes multiples.

#### 4. Conclusion - Perspectives

Le système couvre actuellement la quasi-totalité du corpus. Il a fait l'objet de tests préliminaires sur un oral simulé non réel, dans le cadre de l'Action de Recherche Concertée de l'AUF. L'évaluation DCR porte sur la qualité sémantique des représentations (Antoine J-Y., 2000). Les résultats sont très encourageants : appelé LAMBDA COMP à cette occasion, le système a généré un pourcentage d'erreurs de 10,5% sur les 251 phrases testées<sup>6</sup>. Il convient aussi de l'étendre à d'autres types de phrases et de le confronter à des phrases spontanées réelles. Les perspectives d'amélioration de la méthode portent sur les deux points suivants :

- Le raffinement de l'analyse. La prise en compte progressive de considérations sémantiques semble une stratégie pertinente pour concilier la finesse de l'analyse (prise en considération d'un maximum de mots) sans pour autant nuire à sa robustesse grâce à une série de filtres sémantiques. Néanmoins, il conviendrait de mieux maîtriser les différentes

---

6. L'analyse de ces erreurs est faite dans (Antoine J-Y., 2000).

étapes de composition et leur coordination avec la détection des répétitions, hésitations et auto-corrrections.

- L'élaboration des formules normalisables destinées à la représentation sémantique de l'énoncé. Dans les formules présentées ci-dessus, les définitions des questions et du système d'objets dépendent du contexte de l'application. L'objectif actuel est de ne plus faire dépendre les questions de ce contexte, les formules étant générées par le graphe des objets du discours et des questions génériques (telles que *liste*, *recherche*, *obligation\_de* etc.); ainsi des concepts tels que *trajet* et *tarif* considérés comme des questions dans le système présenté deviendraient des objets dans le système en cours d'étude. L'analyse serait en grande partie fondée sur l'étude de "chaînes d'objets". Elle pourrait devenir ainsi beaucoup moins dépendante de l'application et gérer plus facilement les énoncés incomplets (intrinsèquement ou par suite d'erreurs de la reconnaissance de la parole).

À plus long terme, un lien pourrait être établi avec des études en cours pour la généralisation logique de l'analyse des concepts formels (Ferré S., 1999). On peut imaginer que les formules logiques produites par l'analyse linguistique soient directement utilisées pour la navigation dans la base de données.

## Références

- ABNEY S. (1991). Parsing with Chunks. In C. T. R. BERWICK, S. ABNEY, Ed., *Principle-Based Parsing*. Kluwer Academic Publishers.
- ABNEY S. (1996). Partial Parsing via Finite-State Cascades. In J. CARROLL, Ed., *Workshop on Robust Parsing ESSLI'96*, p. 8–15.
- ANTOINE J-Y. G. D. (1999). Méthodes hybrides issues du TALN et du TAL Parlé : état des lieux et perspectives. In *TALN'1999, Cargèse*, p. 1–17. D. Genthial.
- ANTOINE J-Y. S. J. . C. J. . V. J. . G. J. . A. M. (2000). Obtaining Predictive Results with an Objective Evaluation of Spoken Dialogue Systems : Experiments with the DCR Assessment Paradigm. In *LREC'2000, Athènes*, p. à paraître.
- BLANCHE-BENVENISTE C. B. M. . R. C. . V. D. E. K. (1990). *Le Français parlé ; études grammaticales*. CNRS éditions.
- FERRÉ S. R. O. (1999). *Une généralisation logique de l'analyse des concepts formels*. INRIA, Rapport de recherche, 3820.
- MILLER D.A. N. G. (1986). Some Uses of Higher-Order Logic in Computational Linguistics. *24<sup>th</sup> Annual Meeting of the Association for Computational Linguistics*, p. 247–255.
- MINKER W. (1999). *Compréhension automatique de la parole spontanée*. Paris: L'Harmattan.
- MONTAGUE R. (1974). *Formal Philosophy*. New Haven, USA: Yale University Press.
- MOORGAT M. (1997). Categorical Type Logics. In E. S. B.V., Ed., *Handbook of Logic and Language*, p. 93–177. J. van Benthem and A. ter Meulen.
- RIDOUX O. (1998).  *$\lambda$ Prolog de A à Z*. INRIA, Rapport d'habilitation.
- ROSSET S. L. L. . B. S. . D. L. . G. J.-L. (1997). Corpus oral de renseignements touristiques. AUPELF-UREF ACR-ILOR-B2, technical report.
- SLEATOR D. T. D. (1991). *Parsing English with a Link Grammar*. USA: Rapport interne, CMU-CS-91-196, CMU.
- WARREN D. (1983). Using  $\lambda$ -calculus to Represent Meanings in Logic Grammars. In *21st Annual Meeting of the Association for Computational Linguistics*, p. 51–56. Cambridge, Ma, USA.