

L. R. I.

Les grammaires sémantiques, outil puissant pour
interroger les bases de données en langage naturel

Alain BONNET

Septembre 1979

Rapport de Recherche n° 52



UNIVERSITE DE PARIS-SUD
Centre d'Orsay
LABORATOIRE DE RECHERCHE EN INFORMATIQUE
Bât. 490
91405 ORSAY (FRANCE)

LES GRAMMAIRES SEMANTIQUES, OUTIL PUISSANT POUR
INTERROGER LES BASES DE DONNEES EN LANGAGE NATUREL

Alain BONNET

RESUME

Les bases de données fournissent généralement un cadre dans lequel les relations entre objets sont suffisamment précises pour qu'il n'y ait pas besoin d'une analyse complète (i.e syntaxique) pour interpréter les requêtes du demandeur. Les grammaires sémantiques prennent en considération la structure des questions sans s'intéresser explicitement aux catégories syntaxiques classiques telles que noms, verbes... Les catégories intervenant dans les règles sont plutôt fortement orientées vers le domaine considéré. L'intervention de sémantique dans la grammaire rend l'analyse des questions fortement déterministe et partant, très efficace. De plus, de telles grammaires peuvent être implantées en très peu de temps du fait que l'écriture d'une clause formelle résultant d'une requête (correspondance "pattern" syntaxique / expression réponse) est quasi-immédiate, même pour des phrases de structure relativement complexe.

ABSTRACT

Data bases usually provide a framework in which the relations between objects are sufficiently precise that there is no need for a complete analysis (i.e syntactic) in order to interpret the requests of the user. Semantic grammars take into account the structures of the questions without explicitly using conventional syntactic categories such as nouns, verbs... Instead, the categories intervening in the rules emphasize the domain of application. The use of semantics in the grammar makes the parsing process deterministic and therefore very efficient. Furthermore, such grammars can be rapidly implemented. This is because writing the formal clause resulting from the request (syntactic pattern / response expression correspondance) is fairly straightforward, even for those queries which have a relatively complex structure.

This research has been supported by the Advanced Research Project Agency of the Department of Defense under contract N00039-78-C-0060 with the Naval Electronic Systems Command.

INTRODUCTION

Ce papier introduit la notion de grammaire sémantique, justifie l'absence d'une phase syntaxique compte tenu des buts poursuivis, donne un aperçu du "package" LIFER [Hendrix, 1976] [Hendrix, 1977] puis décrit la grammaire implantée pour la version française d'INLAND ¹ d'interrogation de bases de données concernant les navires. Enfin, il discute les limites de couverture du langage, et compare le programme présenté ici avec d'autres travaux proches. Le programme a été implanté par l'auteur en Interlisp [Teitelman, 1975] sur le DEC KL-10 du SRI International, Menlo Park, Californie. ² L'aspect analyse du langage naturel est le seul considéré ici, c'est à dire que, ni la recherche d'une réponse dans la base, ni la génération des réponses n'est abordée.

1 PUISSANCE ET LIMITES DES GRAMMAIRES SEMANTIQUES

1.1 Definition

Une grammaire sémantique peut se définir par le fait que les catégories apparaissant dans la partie droite des règles de production (si le formalisme utilisé est la forme BNF) ou sur les arcs (cas des réseaux de transition) désignent des classes sémantiques telles que <pays>, <oiseau>, <verbe-ingestion>, ou des concepts utiles à la représentation des connaissances dans une base de données tels que <objet>, <attribut>, <valeur>. Ceci diffère notablement des catégories syntaxiques telles que <groupe verbal>, <nom>, <adjectif> définies par leur rôle et/ou leur position relative dans la phrase.

Le constructeur de la grammaire a toute liberté de définir des catégories qui lui

¹ Inland est décrit dans [Sacerdoti, 1977] ainsi que le projet global LADDER. L'accès à la base de données est décrit dans [Sagalowicz, 1977].

² Je remercie Gary Hendrix, Earl Sacerdoti, Daniel Sagalowicz et Jonathan Slocum pour leurs conseils au cours de ce travail.