

Modélisation incrémentale du dialogue

Stagiaire : Adèle Mortier

Encadrant : Jonathan Ginzburg

Lieu : Laboratoire de linguistique formelle de Paris Diderot

mardi 7 août 2018

Le contexte général

La modélisation du dialogue vise à intégrer les différents aspects de la hiérarchie linguistique :

- la phonologie (reconnaissance de la parole) ¹
- la syntaxe (*parsing*, construction d'un arbre syntaxique)
- la sémantique (détermination du sens des énoncés)

Le but de cette approche est de pouvoir rendre compte, de façon formelle, de phénomènes langagiers liés à l'interaction et à l'oralité, comme l'approbation ("hmm", "d'accord"), la demande de clarification, les ellipses (en particulier, le *sluicing*), les disfluences (en particulier, l'auto-correction). Ces phénomènes peuvent apparaître à la fin d'un énoncé ou de façon anticipée, d'où la nécessité d'un traitement incrémental des énoncés.

Deux approches pour ce traitement sont actuellement développées : une approche intégrée où syntaxe et sémantique sont déterminées et même temps, et une approche plus traditionnelle (DS-TTR IP) où l'interprétation sémantique est construite sur la base d'un arbre syntaxique (iRMRS).

Le problème étudié

Les méthodes actuelles proposent un cadre général pour le traitement des énoncés à la volée (mot par mot). Par ailleurs, les phénomènes propres au dialogue ont été largement caractérisés sur la base de corpus. Le problème est

¹cet aspect n'est pas au cœur du stage, compte tenu du fait qu'il pose moins de problèmes que les deux autres points, en ce qui concerne l'incrémentalité notamment

désormais de définir des règles conversationnelles sous la forme préconditions-postconditions, qui puissent exploiter les informations obtenues par *parsing* incrémental (cf.) de sorte que l'état global du dialogue soit mis à jour d'une façon cohérente. Par "état global du dialogue", on entend une structure de données (typiquement un *record*) susceptible de stocker de façon organisée les informations corrélatives au dialogue : qui parle ? Qu'est-ce qui a été dit ? Quelles sont les questions discutées ? Quelles informations supplémentaires sont apportées par l'environnement ? Les règles conversationnelles traitent des phénomènes langagiers cités plus haut.

La contribution proposée

Qu'avez vous proposé comme solution à cette question ? Attention, pas de technique, seulement les grandes idées ! Soignez particulièrement la description de la démarche *scientifique*.

Les arguments en faveur de sa validité

Qu'est-ce qui montre que cette solution est une bonne solution ? Des expériences, des corollaires ? Commentez la *robustesse* de votre proposition : comment la validité de la solution dépend-elle des hypothèses de travail ?

Le bilan et les perspectives

Et après ? En quoi votre approche est-elle générale ? Qu'est-ce que votre contribution a apporté au domaine ? Que faudrait-il faire maintenant ? Quelle est la bonne *prochaine* question ?