# A Solution for the Problem of Interactive Disambiguation

Hervé BLANCHON

GETA, Institut IMAG (UJF & CNRS)
BP 53X, 38041 Grenoble Cedex, France
e-mail : blanchon@imag.fr

## Extended abstract

After the experiences of dialogue based MT systems with ITS [9], N-Tran [6] and KBMT-89 [5], the LIDIA project aims at the construction of a mock-up of a personal MT system for a monolingual user. One major aspect of the LIDIA project is thus, the study of a dialogue of standardization and disambiguation between the system and the user so as to produce a high quality translation. This dialogue satisfies two properties: its questions are explicit, so no linguistic knowledge is required; its questions are monolingual, so no foreign language knowledge is needed. Here, we focus on one part of the disambiguation process: the disambiguation of the structure produced by the analyser.

The structure produced by our analyser is called MMC (Multisolution, Multilevel and Concrete). Multisolution means that the analyser produces every analysis fitting with the syntagmatic, syntactic and logico-semantic model of the grammar (an example is shown fig. 1). Multilevel means that the same structure consists of three levels of linguistic interpretation, namely the level of syntactic and syntagmatic classes, the level of syntactic functions and the level of logic and semantic relations. Finally, the structure is said to be concrete because the original utterance can be found back by a simple left-to-right reading of the structure.

We have taken into account three kinds of differences between the solutions produced for one sentence, and each kind of difference is associated with the name of an ambiguity. We have defined ambiguities of syntactic classes (cf fig. 2), ambiguities of geometry (cf fig. 3) and ambiguities of syntactic, logic and semantic decoration (cf fig. 4). We have also defined three principles (§ III.1) to order the questions if there is more than one to be asked. The first principle is: first of all, find out the right segmentation into simple sentences. The second principle is: for each common predicate in the MMC structure, find out the right subject, objects and adjuncts. The last principle is: for each simple sentence, find the right structure.

With those principles we are able to define a strategy (cf fig. 5). We have also isolated some patterns in the three classes of ambiguity. The class of ambiguities of syntactic classes needs no refinement (§ III.3.1). On the other hand we create four patterns of ambiguity of geometry (§ III.3.2) called: verbal coordination, argument structure of the verb, non verbal coordination, subordination; and three patterns of ambiguity of syntactic, logic and semantic decoration (§ III.3.3) called: logico-semantic labelling, argument order of direct transitive verbs, syntactic labelling.

Here is an example with the interpretations for each pattern we have chosen:

*Problem of class.* Le pilote ferme la porte: The firm pilot carries her. The pilot shuts the door.

*Problem of verbal coordination.* Il regarde la photo et la classe: He looks at the photograph and the class. He looks at the photograph and files it.

*Problem of the argument structure of the verb.* Il parle de l'école de cuisine: He talks about the cooking school. He talks from the cooking school. He talks from the school about cooking.

*Problem of non-verbal coordination.* Il prend des crayons et des cahiers noirs: He takes pencils and black notebooks. He takes black pencils and black notebooks.

*Problem of subordination.* L'école de cuisine lyonnaise est fermée: The lyonnaise cooking school is closed. The school of lyonnaise cooking is closed.

*Problem of logico-semantic labelling.* Pierre fait porter des chocolats à Lucie: Pierre lets Lucie carry chocolates. Pierre gets chocolates to be delivered to Lucie.

*Problem of argument order of direct transitive verbs.* Quel auteur cite ce conférencier: Which author this lecturer is quoting? Which lecturer this author is quoting?

*Problem of syntactic labelling.* Il parle de la tour Eiffel: He is talking about the Eiffel Tower. He is talking from the Eiffel Tower.

For each pattern we have defined a method to produce the appropriate dialogue (§ III.3). These methods use two kinds of processing: projection and paraphrase. To build paraphrases we use basically three operators: an operator of semantic replacement of occurrence, an operator of permutation of groups of occurrences and an operator of distribution of occurrences. The examples (§ IV) give an idea.

In conclusion we can say that our method is quite simple but fixed once and for all. We are going to study two points in the near future. The first one is to reduce the number of analysis and thus, by getting information from the user, reduce the time to spend on the disambiguation. The second is to try to build tools which will allow the linguist, designer of the linguistic part of the LIDIA system, to define its own methods of disambiguation.

## Keywords

Computer Aided Translation, Personal MT, Interactive Disambiguation, Dialogue Production