Using Ontologies for Improving the Communication Process in a Dialogue System

Marta Gatius and Meritxell Gonzalez
Talp Research Center of Technical University of Catalonia, Barcelona

November 17, 2004

1 Introduction

Developing and maintaining friendly dialogue systems, adapted to the application functionality is expensive. In this paper we propose the use of ontologies for improving the communication process and engineering features in dialogue systems. Using ontologies facilitates the separation between application knowledge, generic dialogue interaction rules and the linguistic resources. Domain ontologies have been used for years in natural language(NL) systems for interpretation and generation. There are also recent dialogue systems using ontologies for controlling the interaction ([BMR94]). We present a dialogue system using an ontology for improving the interaction, the processing of user's interventions and the generation of system's messages. The system has been designed to provide NL access in several languages to applications in the web. The NL components of the system were adapted from those of a NL Interface System we had previously developed ([GR99]). Next sections describe the use of the ontology in the different stages of the dialogue.

2 The ontology and the structure dialogue

The dialogue system follows the Information State Update approach. The information states are obtained from the ontology, where all information the application needs from the user is represented by a set of concepts. Concepts are related to each other by the basic relations: `is_a` and `instance_of`. The application concepts involved in communication are described in the ontology by a set of attributes. These attributes represent the information the application needs from the user (or gives the user). For instance, if the application needs the user's `name`, `address` and `telephone`, a concept described by these three attributes will be represented in the ontology. Ontology concepts may include peconditions. These preconditions govern the information that has to be asked to the user at each state of the dialogue.

The ontology can include taxonomies providing more information about the domain. This information could be used to avoid asking the user difficult questions which answer can be inferred. It can also be used to detect an hiperonym or hipononym in his interventions. Let us consider the application for collecting large objects from houses. This application needs information about the specific type of object the user wants to throw away. A taxonomy classifying the different types of objects could be used. For example, if the user says `I want to throw away an appliance` and the concept `appliance` is classified in the objects taxonomy as an hiperonym of the information the application needs, the dialogue system would ask the user to be more specific. The taxonomy of objects can also be used to determine if an object is pollutant or it is not. The application needs this information and the user may not know it.

3 The ontology and the semantic processing

The components of the system involved in the processing of user interventions are the grammar, the lexicon and the parser. The semantic interpretation associated with the linguistic resources is based on lambda calculus. Lambda calculus allows a simple and efficient interpretation process. The semantic interpretation associated with each lexical entry consists of a lambda function or a lambda value. Lexical entries including application information are incorporated into the general lexicon. These entries can be generated automatically from the concepts, attributes and values in the ontology. They basically represent the parameters of the operations accessing the application. Dynamic entries are incorporated into the lexicon to improve the efficiency and friendliness of the communication. There are two different types of dynamic entries: Those entries representing a value the user will introduce (a proper noun or a number) and those entries associated with a menu that will be displayed on the screen.

Grammar rules representing the different forms of expressing the information related to a particular application have also to be incorporated into the general definite clause grammar. Semantic information is associated with each grammar rule to indicate the order of interpretation of its constituents. This information consists of a list of the numbers representing the constituents. Preconditions can be incorporated into the grammar rules to dynamically adapt the linguistic resources to the application requirements. Syntactic and semantic features can also be incorporated into the categories to improve the space and time for processing user interventions. The specific information representing the identifiers of the concepts and attributes in the ontology can be incorporated into the semantic features associated with the categories.

The left-corner parser performs the syntactic and semantic analyses in parallel. Once the categories in a rule have been recognized, they are interpreted according to the semantic list. The semantic analysis consists of applying the lambda functions over the lambda values following the order indicated in the rule. The semantic interpretation process returns a list of words representing operations and their parameters (concepts, attributes and values).

4 The ontology and the system's interventions

The system messages are also obtained from the application knowledge represented in the ontology. Because all information the application needs or gives is represented as attributes describing the concepts, the system interventions consist of sentences asking or giving the attribute values. In order to facilitate the automatic generation of these sentences a syntactic-semantic classification of attributes has been used. All attributes in the ontology have to be classified in the syntactic-semantic taxonomy according to their linguistic behavior. Each attribute class is associated with a set of general patterns. The attribute classification is based on Spanish linguistic distinctions. However, because most distinctions considered in defining the classes are close to surface regularities but are beyond specific realizations (following [MB04]), this taxonomy could be reused with few changes in other languages. As an example of this, the same classification has been used to generate messages in Catalan.

5 Conclusions and future work

In this paper we have described the use of ontologies for improving the communication process and engineering features in a dialogue system. Representing the application knowledge, the generic dialogue interaction rules and the linguistic knowledge in separated bases facilitates the adaptation of the system to different applications and languages. The current system implementation has been applied to support Spanish dialogues for an application for collecting large objects. Future work will include adapting the system to other applications and languages.

References

[BMR94] J. Bateman, B. Magnini and F. Rinaldi. The Generalized *Italian,German,English* Upper Model. *ECAI Workshop*, 1994.

[GR99] M. Gatius and H. Rodrguez. Adapting general linguistic knowledge to applications in order to obtain friendly and efficient NL interfaces. *VEXTAL Conference*, 1999.

[MB04] D. Milward and M. Beveridge. Ontologies and the structure of dialogue. *Eigth Workshop on the Semantics and Pragmatics of Dialogue Catalog*, 2004.