

SEMANTIC RANGE OF SPOKEN DIALOGUE SYSTEMS

Roman Mouček and Miloslav Konopík

*University of West Bohemia in Pilsen, Czech Republic
moucek | konopik@kiv.zcu.cz*

Abstract: The aim of this paper is to revise the world of possibilities within development of computerized spoken dialogue systems and to discuss if the “better” dialogues systems than the existing ones can be created at all. More generally, its aim is to provide a different view on possibilities and limits of semantics processing and human-computer interaction. The used dialogue strategies are discussed and several theories concerning especially objective reality and semantic context are briefly presented to clear up this view.

1 Introduction

The history of spoken dialogue systems development is quite long. The crucial task during building these kinds of systems is to model the process of understanding to user queries. There have been presented a lot of theories introducing general dialogue strategies, general rules working with large ontologies, rules for extraction of relevant semantics, etc. Nowadays the very popular techniques are mainly based on the stochastic approaches to semantic elaboration. They are also described many possibilities how to construct computerized spoken dialogue systems or how to process spoken queries.

However, if we look at really working computerized spoken dialogue systems using semantic elaboration, we find out that these systems work in very narrow domains. Thus we can use the highly specialized dialogue systems providing e.g. information about train arrivals/departures or enabling the plane tickets booking. Moreover, a lot of dialogue systems are based only on one-way natural communication between human and computer (computer uses up the spoken language, a human reacts e.g. by pressing the keys on keyboard).

The aim of this paper is not to discuss how to solve the gap between possibilities (theories), which describe the general processing of semantics, and reality, which introduces systems working in simple, contextually very limited domains. This paper also does not want to think out or evolve another theory contributing to the possible development of next generation of dialogue systems. The aim of this article is to revise the world of possibilities and to discuss if the “better” computerized spoken dialogues systems than the existing ones can be created at all. More generally, its aim is to provide a different view on possibilities and limits of semantics processing and human-computer interaction. The used dialogue strategies are discussed and several theories concerning especially objective reality and semantic context are briefly presented to clear up this view. We have to emphasize that we will focus only on process of understanding, the problems connected e.g. with speech recognition are not mentioned in this paper.

2 Spoken Dialogue Strategies

Several spoken dialogue strategies have been introduced in order to increase a level of understanding between human and computer. These dialogue strategies have been proposed to respect and partly solve the problems arising during communication between human and computer.

The strategies usually differ in number and type of questions, which react on user queries. The verification questions and complementary question are used frequently.

2.1 Verification Questions

The aim of verification questions is to confirm or correct the recognized user utterances. It means that e.g. the slips of the tongue, recognition errors or some semantic misunderstandings can be successfully corrected. The verification of user queries is always judged according to narrow, highly specialized context. The problems arise if user speaks in a way, which is not incorporated in dialogue system as an allowed context. The system is not able to fill in the semantic container (structure like semantic frame, semantic network, etc.) in this case. Then the user can be persuaded that he provided the relevant information to the dialogue system, whereas a computer is not able to give a relevant answer (computer does not understand).

This situation can be marginally solved by “extension of context”. However, this approach does not mean that computer is able to process larger context, but it includes more simple semantic extraction. The user utterances are processed by searching keywords and mapping these keywords into semantic container. The related syntactic-semantic analysis is restrained. On one hand, this approach extends a set of possible user utterances, which a computer is able to process; on the other hand, the larger amount of total misunderstandings is produced. The number of misunderstandings increases because of effort to process the user utterances, which are out of allowed context.

2.2 Complementary Questions

The aim of complementary questions is to fill in the semantic container. The sufficient content of semantic container is finally the most important condition for generation of answer, which should satisfy a dialogue system user. Complementary questions are usually asked separately. Then one complementary question finds out only one piece of information from user in order to simplify the filling of semantic container.

2.3 Overview of Dialogue Strategies

The dialogue strategies are customized to computer to enable dialogue conduction at all. Then the only effort, how to make a dialogue more pleasant to user, is not to disturb him/her too much. However, this approach is often just in opposite to computer needs. The following dialogue strategies are usually used:

- confirm initiative strategy
 - system repeats all the information (how this information has been understood) given by user,
 - system asks complementary questions to fill in at least all the compulsory slots in the semantic container,
- confirm alone strategy
 - system verifies correctness of information provided by user in one question,
 - system asks complementary questions to fill in at least all the compulsory slots in the semantic container,
- separate confirmation strategy
 - system verifies correctness of information provided by user in separate questions,
 - system asks complementary questions to fill in at least all the compulsory slots in the semantic container,
- no confirmation strategy
 - system does not verify correctness of provided information,
 - system asks complementary questions if they are necessary; the default values are implicitly used in some cases.

2.4 Dialogue Strategies and System Users

The introduced dialogue strategies have different efficiency, which cannot be simply computed or estimated. The confirmation strategies often lead to higher success of whole dialogue because the semantic container is filled more correctly. However, these strategies also disturb more people because the dialogue can be prolonged uselessly. There is an effort to solve this problem by adjusting a dialogue strategy to the user experience. However, there are still groups of people, who refuse to use the computerized dialogue system at all and require services of human operator. The reason of this behavior we can find in the dialogue strategy limits and characteristics of these people expressed by preferences of stimuli ([1]).

The (probably unsolvable) problem is that people more oriented on life dynamics and efficiency of relationships between people are so restricted by used dialogue strategies that they refuse this kind of communication in principle. On the other hand, people oriented on stability and usability of stimuli do not have problem with this kind of communication (the most of dialogue system developers probably belong to this group). The relationships between these groups of people can suffer from mutual misunderstanding. It can be also reason why the computerized dialogue systems are not used as it has been expected.

3 Sharing of Knowledge and Domain Narrowness

To understand a user query we have to rely not only on shared knowledge (conceptual or cognitive meaning described in [2]), but also on coherent expression of this knowledge by natural language means (connotation meaning in [2]). Then the narrowness of domain increases the probability of sufficient amount of shared knowledge. The extension of domain, widely discussed in theoretical studies, leads to the situation that we have to rely on unshared knowledge and unique semantic context of individual, which is influenced by whole life experience. If we concede the idea that unique life experience and corresponding semantic context play the most important role in the process of mutual language understanding in the larger domains (a larger domain is considered as a domain consisting of more than one simple semantic container), then the computerized elaboration of very narrow domains is the only practical solution.

4 Boundaries of Objective Reality

The essential problem of spoken dialogue systems is closely connected with the perception of world. The philosophical theories argue about dependence or independence of objective reality on our perception and understanding. The prevailing answer is that the world exists objectively and it is independent on our individual perception. That is why we can subsequently understand it.

However, the observations in science in the last fifty years support also an opposite opinion. We are only human beings imprisoned in our brains; everything what we can say about the outer world is only the result of our brain processes. These conceptions of the world are based on the notion of complexity – everything acts altogether (they are not isolated objects, there is no objective reality).

How to work with these opposite theories and how does it relate to the problem of understanding in spoken dialogue systems? There is a possibility to connect these theories within a shared view; the complexity of the world can be reduced by individual into image of objective reality. This image of objective reality is then expressed by the means of natural language.

If the image of objective reality is very simple and language expressions of various individuals describing this simple image are similar, then we are able to construct a computerized spoken dialogue system, which can be usable by a larger group of users. Then a successful effort to build up such a dialogue system can lead us to some kind of proof that an objective reality really exists.

Moreover, the limits in development of this system can determine a fuzzy but practical border between objective reality and subjective image of world. From the considerations given above we can also suppose that there is no sense to try to construct dialogue systems covering large domains because perception of reality and language expression of this reality cannot be transferable to other person without substantial misunderstandings.

4.1 Development of the System within Boundaries of Objective Reality

The hardly eliminable problem during development of computerized spoken dialogue system is (and probably has to be) a danger of self dupey of system developer. If a developer thinks about the inner system functioning, hierarchy of semantic concepts, dialogue strategies etc., there is no other possibility then to incorporate his/her own semantic view, own vision of world, etc. in the system. The method how to reduce this subjectivity is to collect domain information from a large number of various system users.

Then another phenomenon frequently arises. There is no possibility to make a semantic consensus (there is not a sufficient amount of shared knowledge or the language expression of the knowledge is different). This result then implies a basic decision about domain. If there is not possibility to construct a general semantic container, which would be accepted by most of the users, there is no sense to build a dialogue system covering this domain. If we determine more than one shared semantic container, but we are not able to decide, what kind of container will be used for specific utterance, there is also no sense to construct a dialogue system, because the number of misunderstandings would be not acceptable.

5 Theory of Communication

If we take a look on the theory of communication and focus on two of its essential ideas, we can support our ideas concerning the necessity of elaboration of simple domain covered by simple semantic container:

- only semantic context is a bearer of meaning,
- the full understanding considers knowledge of all possible contexts.

It means that we can understand only to words, descriptions etc. of our dialogue partner but we cannot understand e.g. to his/her life events.

These two theses also support a necessity to keep the domain area semantically simple (the semantic context is shared again). An uncontrollable domain expansion implies also expansion of possible context and consequent inapplicability of computerized dialogue system. Even in a very limited domain we can detect several possible semantic contexts. Then the shared knowledge (context) has to be satisfactory to construct a successful dialogue system.

In general, there is no possibility to know all the contexts of the user. The experiments carried out in various workplaces, which have tried to note all the life experience of tested person in a short time, have been notably unsuccessful. To find out some layers of dialogue partner context means to verify his/her utterances. That is why the confirmation dialogue strategies are used within human – computer interaction. The usage of these strategies decreases not only the number of mistakes caused e.g. by incorrect speech recognition but it also helps to understand what the user really wanted to say.

6 Conclusion

This paper offers a more general view on the problem of development of computerized spoken dialogue systems. It does not describe any particular technique how to construct a successful dialogue system. Its effort is to introduce a more complex view on possibility of practical development of computerized spoken dialogue systems.

The paper formation has been also invoked by the authors' perception of semantic gap between "technically oriented people", philosophers and linguists. The next important motivation of authors originates from their experience with construction of semantic module of spoken dialogue system; hence from their practical experience with limits and constraints of these systems.

The complexity of the topic and mixture of general and often opposite ideas would lead to a very extensive list of references. Since a correct detection of original sources of general ideas is often impossible, only a short list of necessary references is added.

Acknowledgements

This work was supported by the grant of the Czech Ministry of Education No. 2C06009 COT-SEWing.

References

- [1] Plamínek, J.: Synergický management, Argo, Czech Republic, 2000.
- [2] Machová, S., Švehlová, J.: Sémantika & Pragmatická lingvistika (Semantics & Pragmatic linguistics), Charles University, Prague, Czech Republic, 2001.