HAL
open science

Workshop Proposal for Young Researchers' Roundtable on Spoken Dialogue Systems 2011

Sai Qian

To cite this version:

Sai Qian. Workshop Proposal for Young Researchers' Roundtable on Spoken Dialogue Systems 2011. Young Researchers' Roundtable on Spoken Dialogue Systems 2011, Jun 2011, Portland, United States. hal-00646845

HAL Id: hal-00646845

https://inria.hal.science/hal-00646845v1

Submitted on 30 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sai Qian

LORIA & INRIA Nancy Grand-Est
BP 239 - 54506 Vandœuvre-lès-Nancy Cedex
Nancy 1 University
24-30 Rue Lionnois - BP 60120 | 54003 Nancy Cedex

sai.qian@loria.fr

1 Research Interests

Frege's Principle, also known as the principle of compositionality, is one important concept in areas such as philosophy, linguistics, logic and computer science. In the 1970s, Richard Montague combined principle of compositionality, First Order Logic, λ-calculus, and type theory into the first formal natural language semantic system, people call it Montague Grammar (MG). However, MG (Montague, 1974) was designed to handle single sentence semantics. Later on, some linguistic phenomena, such as anaphora, donkey sentences, and presupposition projection motivated another branch in the field: dynamic semantics. That is where my research interests mainly lie. More specifically, topics such as **event semantics**, **rhetorical relation**, and **variable accessibility** in discourse and dialogue are what I am working on.

1.1 Dynamic Semantics & Discourse

According to MG, the meaning of a sentence is represented as its truth conditions, that is, the circumstances in which the sentence is true. Different from that, dynamic semantics evaluate the sentence meaning as its **context change potential**. In other words, meaning is viewed as a function that always builds new information states out of the old ones by updating the current sentence. Some of the representative works include File Change Semantics, Discourse Representation Theory (DRT), and Dynamic Predicate Logic (DPL).

With the notion of context dynamics, researchers noticed that rhetorical relations were crucial for analyzing discourse semantics (Mann and Thompson, 1974; Asher, 1993). Those theories declare an internal structure existing in the discourse. The Segmented Discourse Representation Theory (SDRT), a variant of DRT (Kamp, 1981) with rhetorical structure implemented, is able to explain various dynamic phenomena in discourse and dialogue, such as anaphora, temporal relation among events, and presupposition.

Recently in (de Groote, 2006), the author proposed a new framework, which integrates the concept of context within traditional MG. The framework is only based on Church's simply-typed λ calculus and some standard computational techniques, which remedies the computational drawbacks of other dynamic approaches, e.g., the variable renaming in DRT. However, same as DRT, the framework in (de Groote, 2006) does not take into account the discourse structure. Thus one of my research topics is to combine SDRT with this new framework, then to further investigate the **accessibility** of discourse variables for applications such as anaphora resolution. This topic is also highly related with the dialogue system, where appropriate responses should base on a set of mature resolution mechanisms.

1.2 Humor in Dialogue System

Humor has always been an important element in the human society. There are many different ways to express humor, obviously, language is the most popular one. However, the implementation of humor in language technologies is never a simple task, because humor is considered as a highly advanced linguistic behavior. A large amount of works on humor can be found in the literature of linguistics (Attardo, 1994).

For spoken dialogue systems, not only will humor brings laughters to the user, but also it will increase the **naturalness** and add bonus to the **interaction** between the user and the system. Recently I am also planning to work on **analyzing and implementing humor** in real-world dialogue and discourse corpus. The general steps are as following. First of all, humor will be defined semantically. Then a hierarchy for humor will be constructed. The hierarchy will be able to evaluate different level of humor in the context (based on one culture probably). Finally, modules for automatic humor detection and generation can be built based on previous steps.

2 Future of Spoken Dialog Research

Since the emergence of computer, people have always been trying to "communicate" with machines. A very first chatterbot example is ELIZA, which used few background information, but provided a certain level of human-like reactions. Nowadays, more and more di-

alogue systems can be found everywhere throughout the world. For example, the client service department of many companies (Apple, Microsoft, etc.) already changed their phone operators from human to talking robots. Also, there are lots of multi-function service dialogue systems for different kind of events (e.g., the "Hai Bao Robot" in Expo 2010, Shanghai China).

Currently, the state of the art of automatic speech recognition (ASR) already reached a considerable industrial level, I think the future development of spoken dialogue system will focus on deep linguistic layers, namely the semantic and pragmatic level. The followings are some of my expectations on this area.

- With the precondition of a nearly perfect ASR, some fancy semantic features will be added into the dialogue system. For instance, the system may acquire emotions, which can be expressed by either the content of spoken text or the change of pitch for the output. Another example is the implementation of humor. Not only the system will be able to tell jokes, which are predefined, but also it shall demonstrate its sense of humor by creating instant humorous response according to the user's input.

- The symbolic strategy will play a more and more important role, compared with its status nowadays. Many semantic theories (e.g., presupposition, dynamic semantics) will become mature enough for industrial implementation. Thus the dialogue system will "understand" natural language on a much decent level. Also some machine learning techniques will be involved, which render the system the ability to evolve throughout the whole dialogue.

- In the next 5 to 10 years, I think the applications of spoken dialogue system will flourish mainly in the game industry. Most computer games will embed such intelligent chatting system, which might change the direction of development in the game industry. In addition, as far as I am concerned, the dialogue system might be applied to handle more administrative stuffs in governments or universities. Finally I believe dialogue system will definitely take part in people's daily life to a larger extent (e.g., the future PC operation system could even be dialogue-oriented).

3 Suggestions for Discussion

- Dialogue System Complexity: Thanks to a variety of applications for spoken dialogue system, such as in-car system, robot interface, and etc., we are in need of a deeper understanding of each application from its origin. It will be more efficient and precise for system development with a **standard classification** of all applications based on the complexity of the involved dialogue systems.

- Naturalness & Fluentness: As described in Section 1.2, humor is a way to improve naturalness and fluentness of a dialogue system. One topic for discussion can be what are the other techniques to make a system response more like a human being during the conversation?

References

Hans Kamp. 1981. *A theory of truth and semantic representation*. Wiley Online Library.

Nicholas Asher. 1993. *Reference to abstract objects in discourse*. Springer.

Philippe de Groote. 2006. *Towards a montagovian account of dynamics*. Proceedings of Semantics and Linguistic Theory XVI.

Richard Montague. 1974. *English as A Formal Language*. Formal philosophy, 188–221. Yale: Yale University Press.

Salvatore Attardo. 1994. *Linguistic Theory of Humor*. Mouton de Gruyter, Berlin.

William C. Mann and Sandra A. Thompson. 1988. *Rhetorical structure theory: Toward a functional theory of text organization*. Text-Interdisciplinary Journal for the Study of Discourse 8(3): 243–281.

Biographical Sketch

Sai Qian was born in Jiang Xi Province, southeast of PR China. He was studying in Beijing University of Posts and Telecommunications (BUPT) from 2003 to 2007 for bachelor degree. During bachelor, he majored in communication engineering, especially, signal processing and speech recognition. From October 2007, he joined the European Master program LCT (Language and Communication Technologies) to study natural language processing. He spent one year (2008) in Saarland University, Germany, and one year (2009) in Nancy 2 University, France. Currently he is doing his PhD (from October 2009) in computational linguistics at Nancy 1 University and LORIA, INRIA Nancy Grand-Est, France. His main research interests include syntax-semantic interface, coreference resolution, compositionality, accessibility constraints, and discourse, dialogue semantics.