



The Dialogue Systems Group

Department of Chinese, Translation and Linguistics
City University of Hong Kong

Leveraging technology to access expert knowledge

Every day people need answers to important questions, about their health for example. Yet they can't always ask an expert, and information on the Internet or elsewhere may be unreliable, outdated, or difficult to search. Yet usually the answer exists somewhere. What if we could provide reliable answers to questions "on demand" in a matter of seconds? Soon we may be able to, through research being carried out by the Dialogue Systems Group (DSG) at the City University of Hong Kong.

The DSG is a multidisciplinary team of researchers who are developing innovative solutions to improve our ability to leverage knowledge through the use of computers. Their work falls under a field known as Natural Language Processing (NLP), which combines computer science and linguistics to bridge the gap between humans and computers by developing machine-understandable models of human communication. The DSG conducts research in speech generation, voice recognition, database models, knowledge bases, corpus linguistics, lexicography and terminology.

The DSG is currently developing a prototype system (<http://dsg.ctl.cityu.edu.hk/index.php/resources/53>) that allows a user to obtain medical information from a massive encyclopedic database through a telephone. Voice recognition technology enables the system to understand the user's question. The question is further interpreted based on internally coded specialized terminologies and medical classifications. The system then queries a database of reliable medical information to find the answer, and, using speech synthesis, delivers an audible response over the telephone. The dialog between the user and the computer is interactive. Instead of a telephone, the user can also query the system from a personal computer.

Such a system could be deployed in a variety of real-world situations. Medical personnel, even hospital interpreters, could access reliable medical information quickly. And people at risk with various medical issues could get answers to their questions, in areas such as substance abuse, mental health, chronic ailments such as diabetes and hypertension, early detection of communicable diseases, nutrition, and so forth. By simply picking up the phone, people can get answers to some questions without having to consult a health care professional.

Researchers with a common purpose

The DSG brings together researchers from different parts of the world with expertise in multiple disciplines. Their common purpose is to use NLP and computing technologies to develop a multi-faceted modular system that can extract knowledge from large repositories of texts (called corpora) in response to a real-world need for information. The core members of the DSG are:

- **Dr. Alex Fang**, Director of the DSG, is a specialist in corpus linguistics, computational linguistics and natural language processing.
- **Prof Jonathan Webster**, a core member of the DSG, is a systemic functional linguist specialising in text analysis.
- **Dr Wei Zhang**, core member of the DSG, is a specialist in conversation analysis, discourse, and sociolinguistics.
- **Dr. Sebastian Fleissner**, a research fellow, is a software engineer with expertise in system architectures and programming paradigms. Originally from Germany, Sebastian is also a computational linguist specializing in dialogue systems.
- **Xiaoyue Liu**, a senior researcher and PhD candidate, has a background in computer science. Her current research focusses on speech recognition and the design of dialogue systems.
- **Cecilia Cao**, a PhD candidate in corpus linguistics, specializes in using linguistic properties to derive information about terms in corpora.
- **John Li**, a PhD candidate in corpus linguistics, has a background in lexicography and e-learning. His current research investigates properties of words in corpora.
- **Maggie Zhang**, a PhD candidate in computational linguistics, is a specialist in automatic term recognition and extraction from corpora.
- **Masayuki Hirata**, a PhD candidate, is an experienced lexicographer from Japan. He is currently researching the use of bilingual corpora to aid in the development of bilingual dictionaries.
- **Kara Warburton**, a PhD candidate, is an experienced terminologist from Canada. She has developed international standards for managing terminology in databases.



The Dialogue Systems Group

The DSG also comprises affiliated experts from outside of Hong Kong who collaborate in specific areas:

- **Professor Harry Bunt**, Professor of Linguistics and Computer Science, Universiteit van Tilburg (Netherlands). Specialist in dialogue acts.
- **Dr. David Traum**, USC Institute of Creative Technologies, California. Specialist in discourse and dialogue.
- **Dr. Jonathan Ginsburg**, Kings College London and Université Paris-Diderot. Specialist in dialogue systems.
- **Dr. Mark Huckvale**, University College London. Specialist in speech recognition and synthesis.
- **Professor Nancy Ide**, Vassar College, New York. Specialist in linguistic annotation.
- **Professor Kiyong Lee**, Korea University. Linguist active in the development of standards for language resources.
- **Professor Alan Melby**, Brigham Young University. Computational linguist with a specialization in translation technology and terminology.

Future directions

The current prototype utilizes a large corpus of medical data as a testing ground to fine-tune the system for medical scenarios. The dialogue system can be customized to address specific human interactions and to deliver appropriate responses. For example, health care professionals need expert-level information, whereas teens and seniors need easy-to-understand answers to questions about topics such as diabetes and substance abuse respectively. The system could also be applied to other needs by incorporating data from other domains. For example, a version could be developed for pre-teens as a self-help tool for early intervention in situations of stress and social pressures. At any point in time during the conversation with the system, the teen could press a key on the handset to be connected to a "real" counsellor.

Expressions of interest

Expressions of interest in the research and development conducted by the DSG are invited.

Contact - Dr. Alex Fang

Department of Chinese, Translation and Linguistics
City University of Hong Kong
Tat Chee Avenue, Kowloon
Hong Kong SAR
acfang@cityu.edu.hk
3442 8793