

Preface

LANGUAGE technologies is an area of computer science, more specifically, of artificial intelligence, devoted to the computational analysis of human language, in its written or spoken form, and the corresponding practical applications in which computers reasonably deal with language data.

Applications of natural language technologies include search, translation, or summarization of texts; automatically obtaining answers to questions basing on the information present in huge collections of documents or in Internet, detecting trends and prevailing opinions or sentiments of people expressed in social networks with respect to a company's products or a government's actions, interfaces between the users and computers using normal language, detection of plagiarism, especially important in educational or academic settings, as well as numerous other applications.

Such applications save people time on reading or looking for information, improve quality of life of the users by, say, recommending them appropriate products or making it easy for people to use computers or other devices, increase the income for businesses by helping them better understand the consumers' needs, and improve democracy by timely informing decision makers, political parties and government of the citizens' opinions.

Many of these applications are based on a varying degree of understanding of the contents of the corresponding texts. Computational semantics is a branch of natural language processing that studies the techniques for extraction of meaning from natural language texts or building computational structures that reflect elements of the meaning of the text important for its computational treatment in practical applications.

This issue of *Polibits* features a special section on computational semantics and language technologies, which consists of the following six papers.

Arda Çelebi and Arzucan Özgür from Turkey introduce a novel concept on *n*-gram parsing and show that it improves the performance of a classical constituency parser. In addition, *n*-gram parsing is capable of producing partial parsing even if the whole sentence is not grammatically correct. This will be especially useful in applications related to user-contributed contents such as comments or blogs or to social networks such as Facebook and Twitter, given that in many cases texts from such sources cannot be parsed with classical parsers.

Marzieh Fadaee, Hamidreza Ghader, Hesham Faily, and Azadeh Shakery from Iran present a statistical-based method for automatic construction of WordNet-like dictionaries. WordNet is a dictionary that provides information vital for most of the current language processing technologies.

However, the original WordNet was developed for English. While WordNet-like dictionaries of varying quality and completeness exist for a small number of major languages, the great majority of languages, spoken by a considerable part of humankind, lack such a very important resource. Fadaee et al. show that for such a language as Persian, their method produces highly accurate WordNet-like dictionary.

Yanhui Gu, Zhenglu Yang, Miyuki Nakano, Masaru Kitsuregawa from Japan aim at boosting both the effectiveness and efficiency of measuring sentence similarity by combining various existing sentence similarity measuring techniques. They also introduce a number of optimization strategies and give insights into balancing of effectiveness and efficiency of measuring sentence similarity, a task of crucial importance for many natural language processing applications such as information retrieval, question answering, information extraction, text mining, text summarization, and machine translation, to name only a few.

Carlos Cobos and Martha Mendoza, and Elizabeth León from Colombia, Milos Manic from the USA, and Enrique Herrera-Viedma from Spain address the task of web page clustering. Web page clustering improves the user experience in information retrieval by presenting the user with a whole cluster of roughly equivalent or similar documents instead of individual documents, which helps the user to grasp all the alternatives at a glance and not dig into a long list of very similar pages. Cobos et al. report that 90% of the users agreed on that their method improves the search process.

Prashant Mathur, Nick Ruiz, Marcello Federico from Italy show how automatic translation of text segments can be used to help human translators in manual translation. Generally, automatic translation output needs extensive editing effort from human translators to reach the quality of manual translation. In many cases it is faster for the translator to type the translated text than to correct a suggestion from an automatic system; in other cases it is faster to correct the suggestion than to type the translation from scratch. Reading each suggestion and deciding whether it is easier to correct it or to discard it takes even more time and effort. Mathur et al. show how to automatically decide in which cases suggestions are likely to be useful and thus are to be shown to the translator for editing, and in which cases it is better to allow the user to type the translation without asking him or her explicitly.

Amitava Kundu, Dipankar Das, and Sivaji Bandyopadhyay from India address the problem of automatic understanding of the contents of movies. This task is important on the one hand for contents-based recommender systems in order to

recommend specific movies to users, and, on the other hand, it can be a source of information on real-world situations for computers to automatically learn common-sense knowledge. Kundu et al. present a method for detection the change of scene by analyzing the movie script.

The next four papers are regular papers on diverse topics of artificial intelligence, computer science, and automatic control.

Amir Elalouf and Eugene Levner from Israel and T.C.E. Cheng from Hong Kong discuss the problem of routing mobile agents in computer networks. Congestion in the network environment can cause unpredictable delays in information transfer and, in particular, in routing of mobile agents, which can be undesirable in time-sensitive applications. Elalouf et al. consider a problem of finding a path in the network that with some probability guarantees that the information will be passed within the specified time. They show that this is an NP-hard problem, and present algorithms for its exact or approximate solution.

María Bárbara Calva-Yáñez, Paola Andrea Niño-Suárez, Miguel Gabriel Villarreal-Cervantes, Gabriel Sepúlveda-Cervantes, and Edgar Alfredo Portilla-Flores from Mexico consider a problem of optimal control or a four-bar mechanism with spring and damping forces. They show how this dynamic optimization problem can be solved to with evolutionary techniques, specifically, a differential evolution algorithm. The efficiency of the proposed method is verified with a simulation experiment.

Mailyn Moreno Espino, Alternán Carrasco Bustamante, Alejandro Rosete Suárez and Marta Dunia Delgado Dapena from Cuba propose software development patterns for agent-based programs. Agents are active software components, which, in contrast to objects, not only react to external requests but perform actions when they deem appropriate judging on the state of the system and the internal state of the agent. While object-oriented programming is now a mature and stable area with decades of experience, agent-oriented

programming is still novel and lacks the variety of well-established methodologies that object-oriented programming has. Moreno Espino et al. show how agent-oriented elements can be incorporated into well-established object-oriented methodology.

L.E. Gómez, J.H. Sossa, R. Barrón, F.J. Cuevas, J.F. Jiménez from Mexico present their improvements to music recommendation, which is an application with very high commercial demand and which affect the quality of life of a considerable part of humankind. Music recommendation systems help people choose, or just automatically offer, the music pieces or songs to listen in their everyday environment, basing on the knowledge of their preferences extracted from various sources, such as the history of songs the user has chosen in the past. Most often the selection is done basing on the metadata associated with the song, such as genre, author, singer, year, etc., which are numbers or text strings. Gómez et al. show that with the user of artificial neural networks, namely dynamic neural networks, it is possible to compare and recommend music basing on the acoustic properties only and not on metadata. This can lead to better recommendation of songs lacking metadata at all, and to improved recommendation of songs that have them.

This issue of Polibits will no doubt be useful for researchers, students, and general public interested in artificial intelligence, especially computational semantics and natural language processing, as well as other areas of computer science and automatic control.

Efstathios Stamatatos

Assistant Professor,
Department of Information and
Communication Systems Engineering,
University of the Aegean, Greece