

# Editorial

SINCE this issue we welcome to our editorial team a new associate editor, Prof. Ramón Silva Ortigoza, an expert in control of mechatronic systems, control of mobile robots, control in power electronics, and geometric optics, who has authored a number of books and over 30 research papers in these areas.

This issue of Polibits includes ten papers by authors from ten different countries: Colombia, Cuba, India, Italy, Japan, Mexico, Norway, Romania, UK, and USA. The papers included in this issue are devoted to such topics as sensor networks, service robotics, control of mechatronic systems, business process modeling, cross-language information retrieval, unsupervised word sense disambiguation, generation of assessment tests in education, large-scale text classification, knowledge discovery in datasets, and automatic text summarization.

**B. Anjum** and **C. L. Sabharwal** from **USA** in their paper “Filtering Compromised Environment Sensors Using Autoregressive Hidden Markov Model” propose a simple and computationally inexpensive method of identifying compromised sensors in a sensor network. In their experiments on artificial and real datasets, their model shows very high accuracy on a large sensor network, but, even more importantly, it provides high accuracy when learning from a small sensor networks.

**A. Vanzo** et al. from **Italy** in their paper “Robust Spoken Language Understanding for House Service Robots” address one of the most important parts of the interface between the end user and service robot, namely, speech recognition module. They describe a robust method for re-ranking the hypothesis generated by the speech recognition module adapted for use in the context of service robotics and specifically suited for recognition of typical commands. They show that their method outperforms general-purpose automatic speech recognition programs.

**M. G. Villarreal-Cervantes** et al. from **Mexico** in their paper “PC Based Open Control Architecture for Mechatronic Systems” describe a low-cost, flexible, reconfigurable, and versatile open control architecture for mechatronic systems implemented in an ordinary personal computer. This architecture will be useful both for modeling in design of mechatronic systems and for teaching design of mechatronic systems in classroom.

**H. Ordoñez** et al. from **Colombia** in their paper “Business Process Models Clustering Based on Multimodal Search, K-means, and Cumulative and No-Continuous N-Grams” present a method for indexing, searching, and grouping business

processes models in order to facilitate the use of large process repositories. Their method, based on linguistic and behavioral information, outperforms existing approaches in terms of precision and recall.

**R. Prasath** and **S. Sarkar** from **Norway** and **India** in their paper “Cross-Language Information Retrieval with Incorrect Query Translations” improve the performance of information retrieval under cross-language setting, that is, when the user does not know English well enough to formulate the query for a search engine and thus formulates it in his or her own language (Tamil in their experiments), but the documents to be retrieved are in English. Moreover, they consider the situation when the query cannot be reliably automatically translated into English, which is a common phenomenon due to colloquial language style typically used by the Internet users when searching for information.

**S. Torres-Ramos** et al. from **Mexico** in their paper “Unsupervised Word Sense Disambiguation Using Alpha-Beta Associative Memories” extends the classical Lesk algorithm for unsupervised word sense disambiguation with a novel word-similarity measure based on alpha-beta associative memory operators. They show that the new algorithm is especially effective for dealing with inflective and derivational forms of words without the need for stemming procedure. This can be especially useful for languages for which no good stemming procedure has been developed. To the best of my knowledge, this is the first use of alpha-beta associative memories as a similarity measure in natural language processing.

**D. Popescu Anastasiu** et al. from **Romania** in their paper “A Method Based on Genetic Algorithms for Generating Assessment Tests Used for Learning” present a method for automatically generating an optimal sequence of tests, out of a given repository, for use in educational settings for evaluation of the performance of the students on a given topic. The method can be configured by the teacher via the use of specific keywords that describe the topics of interest for the evaluation procedure.

**M. G. Sohrab** et al. from **Japan** in their paper “IN-DEDUCTIVE and DAG-Tree Approaches for Large-Scale Extreme Multi-label Hierarchical Text Classification” propose a novel method for hierarchical text classification applicable to very large document collections such as Wikipedia. The method is based on large-scale hierarchical inductive learning and deductive classification. The authors evaluate their method on standard evaluation datasets provided as part of the PASCAL Challenge on Large-Scale Hierarchical Text Classification.

**Jarvin A. Antón-Vargas** et al. from **Cuba** and **Mexico** in their paper “Instance Selection to Improve Gamma Classifier” present an improvement for the Gamma classifier used in the pre-processing stage of noise filtering in knowledge discovery in datasets, thus alleviating the problem of the presence of misclassified or non-representative instances in the training data. The authors introduce a novel similarity function for the Gamma classifier. Experimental results are presented on fifteen different datasets.

**N. Sanchan** et al. from **UK** in their paper “Understanding Human Preferences for Summary Designs in Online Debates Domain” study user preferences in generating summaries for online debates. While for some other domains such as news or scientific papers summarization it is clear what information should be included in a good summary, in many domains, including online debates, this is not clear and needs a separate research. With the help of sixty independent evaluators, the

authors show that the best summary types for this domain are chart summary and side-by-side summary. This finding will guide future development of automatic summarization systems for this domain.

This issue of the journal will be useful to researchers, students, and practitioners working in the corresponding areas, as well as to general public interested in advances in computer science, artificial intelligence, and computer engineering.

Alexander Gelbukh  
Instituto Politécnico Nacional,  
Mexico City, Mexico  
Editor-in-Chief