Editorial

IT is my pleasure to present the readers a new issue of the Polibits journal. This issue of Polibits includes 10 papers by authors from 10 different countries: Australia, Chile, Ecuador, Germany, Hungary, Mexico, Portugal, Spain, Sweden, and USA. The majority of the papers included in this issue are devoted to various topics within Artificial Intelligence, probably the widest and most actively growing area of computer science nowadays.

The first five papers of this issue are devoted to one of the most fundamental problem in Artificial Intelligence and in general in science: logical reasoning and knowledge representation.

David Sundgren and Alexander Karlsson from Sweden address a problem of reasoning under uncertainty, which is of great importance in Artificial Intelligence. They analyze the phenomenon of second-order probability: uncertainty, in the form of probability, about the probability of an event. First-order probability reflects our knowledge about an event. For example, suppose we have a fair die; then in the next throwing all numbers are equally probable. Suppose now we have an unfair die that always shows the same number, but we don't know which. Then again, the probability of seeing all numbers at the next throwing are all equal. Now, second-order probabilities reflect our knowledge about the distributions of events. In our example, in the case of a fair die there is only one option: an equal distribution of numbers. However, in the case of unfair die, there are six options: it can always show the number 1, or always the number 2, etc.; since we have no information about this particular die other than that it is unfair, all six options are equally probable. The authors study the uncertainty levels that appear in reasoning with second-order probabilities.

Chaman L. Sabharwal et al. from the USA study intersection of triangles in frame of qualitative spatial reasoning. Triangles are basic shapes used both in mathematics (triangulation in the homology theory) and computer science to represent more complex spatial objects. Detecting intersections between spatial objects is useful for their computational modeling, for example, in CAD / CAM systems. The author describe rather complex logic resulting from intersections of these basic shapes in spatial combinatorics.

Irosh Fernando and Frans A. Henskens from Australia introduce an algorithm for the use of the Select and Test reasoning model in medical expert systems. They give a detailed pseudo-code for their algorithm and even implementation of its most important parts in Java. Their algorithm involves a bottom-up and recursive process with logical inferences, abduction, deduction, and induction. An example small knowledgebase is also given.

Guida Gomes et al. from Portugal analyze various factors and events that can lead to deterioration of buildings and are thus important to know in order to determine a correct strategy for repairing. They use logic programming approach for knowledge representation and reasoning about these events and factors. Specifically, they extend the Eindhoven Classification Model and adapt it to the area of conservation and maintenance of buildings and its causal tree.

Juan Carlos Nieves and Helena Lindgren from Sweden show how to consolidate heterogeneous knowledge sources for decision making and reasoning, for example, about medical diagnosis. They present an algorithm capable to merge deductive and abductive knowledge bases. They explore an argumentation context approach, which follows the way medical professionals typically reason in order to merge two basic kinds of reasoning approaches: deductive and abductive inferences. For this, they introduce two kinds of argumentation frameworks: deductive argumentation frameworks and abductive argumentation frameworks, and merge the corresponding knowledge sources using an approach based on argumentation context systems.

The next two papers show how computer modeling can be usefully applied in economy domain, such as forecasting of economic activity and modeling of economic processes.

Nibaldo Rodriguez and Jose Miguel Rubio L. from Chile and Lida Barba from Ecuador present a forecasting strategy based on stationary wavelet transform combined with radial basis function (RBF) neural network. As a case study, they apply this strategy to improve the accuracy of 3-month-ahead hake catches forecasting of the fisheries industry in the central southern Chile. Their forecasting model decomposes the raw data set into an annual cycle component and an inter-annual component by using 3-levels stationary wavelet decomposition. The components are independently predicted using an autoregressive RBF neural network model. The utility of the proposed model is demonstrated on hake catches data set for monthly periods from 1963 to 2008.

Borja Ponte et al. from Spain address an important economic problem, which is a major concern for companies nowadays: supply chain management. The Bullwhip Effect, related to the amplification of the demand supported by the different levels, is a major cause of inefficiency in the supply chain. The authors present an application of simulation techniques to the study of the Bullwhip Effect in comparison to modern alternatives such as the representation of the supply chain as a network of intelligent agents. They show that the

supply chain simulation is a particularly interesting tool for performing sensitivity analyses in order to measure the impact of changes in a quantitative parameter on the generated Bullwhip Effect. A sensitivity analysis of safety stock illustrates the relationship between Bullwhip Effect and safety stock.

The last three papers are devoted to yet another major area of artificial intelligence: natural language processing. Two of these papers address an emerging and very active area or web opinion mining, and the remaining one introduces a new kind of feature useful for classification tasks.

Melanie Neunerdt et al. from Germany present a technique to train a part-of-speech tagger on real comments left by the users on social media websites. The importance of the analysis of user-contributed contents in social media stems from the huge quantity of such texts, from which user's opinions about products, companies, political parties, or events can be successfully mined. Knowing this information presents better quality of life for consumers, better income for businesses, and real-time democracy for the governments. However, the grammar and style of such texts greatly differs from the grammar and style of traditional sources such as books or newspapers. The majority of existing natural language processing tools is tailored for traditional language and not for web social media. The authors show how a part-of-speech tagger can be trained on real web social media texts. The work described in this paper has received third place best paper award at the 12th Mexican International Conference on Artificial Intelligence, out of 284 submissions from 45 countries.

Grigori Sidorov from Mexico discusses an extension of the syntactic n-gram feature space suggested earlier by him and his co-authors. The extension consists in allowing bifurcations in the traversal of the syntactic tree when forming syntactic n-grams. Syntactic n-grams has been previously shown to be a useful tool in classification tasks such as author identification or plagiarism detection. Numerous examples of forming the syntactic n-grams with bifurcations are given, and a way of their representation in plain text is described.

Finally, István Endrédy and Attila Novák from Hungary present a novel algorithm for removing unimportant information from webpages. Huge body of useful information can be mined from webpages, such as user's opinions about products or political parties. However, analysis of webpages is hindered by a large amount of text and images, such as advertising, formatting, styling, pointers to other articles and webpages, etc., unrelated to the main contents of the webpage and usually automatically added by the webserver. Removing such overhead and leaving only the important contents of webpages for subsequent analysis is a very important practical task. The authors improve over existing algorithms for this task. They also present a new gold standard corpus for evaluation of text cleaning algorithms.

Ildar Batyrshin

Research Professor,
Instituto Mexicano del Petróleo, Mexico
Treasurer,
Mexican Society of Artificial Intelligence