Editorial

Setsuya Kurahashi and Kazuhiko Tsuda

Graduate School of Business Sciences,
University of Tsukuba,
Tokyo, Japan
Email: kurahashi@gssm.otsuka.tsukuba.ac.jp
Email: tsuda@gssm.otsuka.tsukuba.ac.jp

Masao Fuketa and Jun-ichi Aoe

Information Solution,
Institute of Technology and Science,
University of Tokushima, 2-1 Minami-Josanjima-Cho,
Tokushima-Shi, Tokushima 770-8506, Japan
Email: fuketa@is.tokushima-u.ac.jp Email: aoe@is.tokushima-u.ac.jp

Biographical notes: Setsuya Kurahashi is an Associate Professor at Graduate School of Business Sciences at University of Tsukuba, Japan. Since 1996, he has been doing research on multi-agent systems, data mining and knowledge discovery. He completed his PhD degree from University of Tsukuba, Japan, in 2002.

Kazuhiko Tsuda has been working as a Professor at the Graduate School of Systems Management, University of Tsukuba,Tokyo, Japan since 1998. He has also worked with Mitsubishi Electric Corporation from 1986 to 1990 and with Sumitomo Metal Industries Ltd. from 1991 to 1998. He is also a Member of The Information Processing Society of Japan and The Institute of Electronics, Information and Communication Engineers.He received his BA and PhD, both in engineering, from Tokushima University, Japan, in 1986 and 1994, respectively. His research interests include natural language processing, database, and human-computer interaction.

Masao Fuketa is an Associate Professor at the Department of Information Science and Intelligent Systems at the University of Tokushima where he was a Research Assistant from 1998 to 2000. He is also a member of the Information Processing Society of Japan and the Association for Natural Language Processing of Japan. He received his BSc, MSc and PhD in Information Science and Intelligent Systems from the University of Tokushima, Japan in 1993, 1995 and 1998, respectively. His research interests are sentence retrieval from huge text databases and morphological analysis.

Jun-ichi Aoe is currently a Professor at the Department of Information Science and Intelligent Systems, Tokushima University, Japan. He is the editor of Computer Algorithm Series of the IEEE Computer Society Press. He received his BSc and MSc degrees in Electronic Engineering from the University of Tokushima, Japan, in 1974 and 1976, respectively and the PhD degree in communication engineering from the University of Osaka, Japan. Since 1976, he has been working with the University of Tokushima. His research interests include design of an automatic selection method of key search algorithms based on expert knowledge bases and natural language processing.

1 Introduction

Knowledge-based systems provide intelligent assistance in solving any problems. They can be used not only in engineering, but also in management, marketing, internet, communication, networking, psychological, educational, etc., systems. Moreover, this issue covers a broad spectrum of disciplines working towards enabling intelligent systems to interact with humans using natural language, and towards enhancing human–machine communication through services such as automatic translation, information retrieval, text summarisation and information extraction. The research and development of these systems, which exploit knowledge in the target domain, is at the forefront of modern research.

2 The papers in this issue

The first paper in this issue, 'A fuzzy-rough case-based learning approach for intelligent die design', by Chi Zhou et al., attempts to present a fuzzy-rough approach of mining rules from existing successful die designs, which improve

learning capability of intelligent stamping die design systems. The core of the learning mechanism includes: (1) a feature-based case representation; (2) fuzzification of feature attributes; (3) a fuzzy classification method to partition the cases into clusters based on similarities; (4) a rough set theory based approach to compute attribute reduct and mine rules.

The second paper, 'Integrated evaluator extracted from infringement lawsuits using extension neural network accommodated to patent assessment', by Yi-Hsuan Lai and Hui-Chung Che, provides the basis of patent law and proposes an integrated evaluator constructed by a revolutionary evaluation model for patent assessment. Regardless of stock performance or revenue generated by the enterprise, the damage award of a patent infringement lawsuit is deemed to be a legal value of a patent in view of the patent law.

The third paper, 'Improved genetic algorithm for optimal design of fuzzy classifier', by Ganesh Kumar and Devaraj, presents a Genetic Algorithm (GA) approach to obtain the optimal rule set and the membership function. To develop the fuzzy system, the membership functions and rule set are encoded into the chromosome and evolved simultaneously using a GA. Advanced genetic operators are applied to improve the performance of the GA in designing the fuzzy classifier and take the advantage of the polar coordinates system, which takes the orientation and distance as elementary consequently can well keep the relative hereditary factors in the design varies.

The fourth paper, 'A condition-based maintenance policy for intelligent monitored system', by Wenzhu Liao, Ershun Pan and Lifeng Xi, proposes a sequential Condition-Based Maintenance (CBM) policy for intelligent monitored system based on cost and reliability prioritisation. It is assumed that system's reliability could be continuously monitored, whenever it reaches the threshold R, scheduled maintenance activity is performed to restore the system. This maintenance policy differs from other policies in taking into consideration of influences from the frequency of maintenance activities and operating time on system's failure rate function subject to a deterioration process.

The fifth one, 'Digital watermarking for relational databases using traceability parameter', by Rao and Prasad, presents a mechanism for proof of ownership based on the secure embedding of a robust imperceptible watermark in the relational database. It is achieved by formulating a watermarking method, which will watermark only the numeric attributes and introduced the traceability parameter in the watermark detection technique. This will monitor the modifications made to the relational databases by the attacker.

The sixth paper, 'A virtual layout system integrated with polar coordinates-based genetic algorithm', by Ya-Bo Luo, proposes a new method to solve the problem of constringency and takes the virtual reality technology to solve the problem of virtual detailed location. First, the construction of optimisation model for MLP is proposed, and then based on which the polar coordinates-based GA methodology is presented to improve the efficiency of constringency of GA. Second, the virtual layout tactics are proposed to realise the accuracy location considering the human factor, which take the results from GA as initial values. Contrastive experiments demonstrate the combination of the polar coordinates-based GA and the virtual layout has improved the efficiency of constringency and the practical function of detailed locations.

The seventh paper, 'Analysing the behavioural intention of using telehomecare from a management viewpoint – an application of artificial neural network', by Jui-Chen Huang, discusses interviews with 369 samples in Taiwan. The results reveal that the most effective way to enhance BI to use telehomecare is to improve the Perceived Benefits (PB) for potential users. In addition, based on the results, the conclusion is made that the application of ANN in analysing data on BI to use telehomecare is a feasible approach. Further, this study made contributions on modelling and identifying key factors in terms of both business operation and management. These findings may offer significant reference for subsequent studies.

The eighth paper, 'A technique for retrospective computer validation of drug manufacturing software', by Takahashi, proposes a technique for validating that in-service Drug Manufacturing Software (DMSW) is adequate. This is called Retrospective Computer Validation (RCV). When conducting RCV for the first time, one validates the adequacy of DMSW's functions and performance by collecting existing documents and operational records. When we use RCV conducted DMSW, some modifications occur. In this case, modification of documents and enforcement of additional tests are required.

The ninth paper, 'Feature string-based intelligent information retrieval from Tamil document images', by Abirami and Manjula, proposes a simple and effective method to extract the text and perform intelligent information retrieval from Tamil document images without OCR. This methodology generates a feature string for every word image by extracting its features. This relies on their basic characteristics or shapes of letters instead of recognising the letters like OCR. The strength of this technique lies in extracting the text based on their basic features such as lines and black and white disposition rates in characters which is almost same for the characters across various font sizes and font faces.

The tenth paper, 'Ontology-based relevance analysis for automatic reference tracking', by Mahalakshmi, Sendhilkumar, Irulappan, Mirinda and Gnanasekaran, establishes that the objective of reference tracking is to aid the research of the scholars. Therefore, analysing the relevance of the tracked publication with respect to the seed paper is of utmost importance. This paper tries to address the issue of ontology-based relevance analysis of research publications. In comparison, we have attempted to experiment the relevance calculations based on 'WordNet'.

The eleventh paper, 'Exploring optimised route selection strategy in tree- and mesh-based multicast routing in MANETs', by Baburaj and Vasudevan, explores the intelligent GA-based On Demand Multicast Routing Protocol (GA-ODMRP), and GA-based Multicast Ad-hoc On demand Distance Vector Protocol (GA-MAODV) which improve the control overheads and packet delivery ratio in the routing messages.

The twelfth paper, 'Comparative analysis of regression and machine learning methods for predicting fault proneness models', by Singh, Kaur and Malhotra, examines and compares Logistic Regression (LR), SVM and DT models with ANN models predicted in an analogous study using same dataset. The goal of this study is to compare the results of machine learning methods with regression methods. These methods are explored empirically to find the effect of object-oriented metrics given by Chidamber and Kemerer on the fault proneness of object-oriented system classes.

The thirteenth paper, 'A novel framework for optimised privacy preserving data mining using the innovative desultory technique', by Indumathi and Uma, proposes a novel three-tier architecture for combining three best methods for privacy preservation, namely access control limitation technique, randomisation and Privacy Preserving Clustering (PPC). The access control protects the data content from illegal access preventing grave tribulations. A randomisation technique christened as the desultory technique is an economical and efficient approach for Privacy Preserving Data Mining (PPDM). PPC gives a further impetus for striking an effectual and efficient balance between privacy and data utility. Thus, the framework amalgamates and gives an efficient control system comprising authentication, authorisation and access for each database application; efficiency and economical benefits of randomisation and advantages of PPC.

The fourteenth paper, 'Analysis of the effect of Headline News in financial market through text categorisation', by S. Takahashi, H. Takahashi and Tsuda, analyses the relation between stock price returns and Headline News. Headline News is a very important source of information in asset management, and is sent in large quantities every day. The paper examines the effect of more than 13,000 Headline News sent from JIJI Press. It classifies Headline News using text categorisation and analyses the reaction of a stock price return for every type of news.

The fifteenth paper, 'The discovery of history using inverse simulation', by Kurahashi, analyses a particular family line, which has so many successful candidates, who have passed the very tough examinations of Chinese government officials for over 500 years. This paper also investigates what would happen in a Chinese historical family line. First, this approach studies the genealogical records 'Zokufu' in China. Second, based on the study, it implements an agent-based model with the family line network as an adjacency matrix and the personal profile data as an attribution matrix. Third, using an 'inverse simulation' technique, it optimises the agent-based model in order to fit the simulation profiles to the real profile data.

The sixteenth paper, 'Architecture for effective personalised web search', by Sehdhilkumar and Geetha, suggests context- and content-based search to effectively eliminate the most appropriate irrelevant pages according to the current context of search, and recommends pages that were left unvisited. Personalisation using such conceptual graphs can produce better results than keyword-based searching by providing conceptual links between visited and unvisited pages, and thus pages that are unvisited but relevant can also be recommended to the users.

The seventeenth paper, 'A method to retrieve telops based on the distance of character image features between queries and telops', by Shishibori, Nishikawa and Kita, presents a new video scene retrieval method based on telop characters. In order to specify suitable scenes, this method recognises the only telop characters that correspond to the query keyword, not all characters. This method calculates the distance between each image feature of telop characters and template image features of query keyword.

The eighteenth paper, 'A unified approach for determining the underlying causes of non-stationary distrubances', by Pankajakshan, presents a method directed at developing the right framework and for attaining it eventually. A fundamental but important part is the segmentation of the disturbances from the captured signal using either a Kalman filter or a Multiresolution Signal Decomposition (MSD) technique.

The nineteenth paper, 'A method to implement effective My-page service system using three-dimensional vectors', by Kessoku et al., proposes a method to construct My Page, which is an effective one-to-one marketing method for Internet Service Providers (ISPs). One of the challenges in constructing My Page is the heavy load on the hardware when extracting customer preference information owing to the enormous number of customers and the extremely broad areas that the customers might be interested in.

The twentieth paper, 'Ontology-based query processing for understanding intentions of indirect speech acts in natural-language question answering', by Mima, Ota and Nagatsuna, proposes a scheme for ontology-based query processing to infer intentions from indirect speech-act that do not express users' real intentions explicitly. In the method, the real intentions of the indirect speech-act are classified into: (1) refusal, (2) reversal, (3) restriction, (4) benefit and (5) disability.

The twenty-first paper, 'An automatic extraction method of word tendency judgement for specific subjects', by Morita, Atlam, Fuketa, Iwabu and Aoe, focuses on word tendencies in documents and suggests an automatic extraction method for specific subjects. Field judgement is conducted by using field association words, and similarity among word tendencies and other word tendencies computed with field information.

The twenty-second paper, 'Relevant estimation among fields using field association words', by Tanaka, Atlam, Morita, Tsukuda, Fuketa and Aoe, presents a method of relevant estimation between fields by using field association words. Two methods propose a method of extraction co-occurrence among fields and a method of judgement similarity between fields as the methods of relevant estimation between fields.

The last paper, 'An efficient search method of e-government information using knowledge of search history', by Mitani, Takahashi and Tsuda, looks at reference methods for special terms and business world terms when searching e-government material. It also looks at expressions of vague numerical information often found in government reports and proposes how these can be taken into consideration and efficiently searched.