Call for Papers

Memetic Computing Journal special issue on **Metaheuristics for Large Scale Data Mining**

Guest editors:

Jaume Bacardit
puter Science and School of Biosciences
University of Nottingham
jaume.bacardit@nottingham.ac.uk

Xavier Llorà
ter for Supercomputing Applications
of Illinois at Urbana-Champaign
xllora@illinois.edu

Submission deadline: **March 31st , 2009**

Aim and Scope

Data mining and knowledge discovery are crucial techniques across many scientific disciplines. Recent developments such as the Genome Project (and its successors) or the construction of the Large Hadron Collider have provided the scientific community with vast amounts of data. Metaheuristics and other evolutionary algorithms have been successfully applied to a large variety of d a ta mining tasks. Competitive metaheuristic approaches are able to deal with rule, tree and

prototype induction, neural networks synthesis, fuzzy logic learning, and kernel machines -to mention but a few. Moreover, the inherent parallel nature of some metaheuristics (e.g. evolutionary approaches, particle swarms, ant colonies, etc) makes them perfect candidates for approaching very large-scale data mining problems.

Although a number of recent techniques have applied these methods to complex data mining domains, we are still far from having a deep and principled understanding of how to scale them to datasets of terascale, petascale or even larger scale. In order to achieve and maintain a relevant role in large scale data mining, metaheuristics need, among other features, to have the capacity of processing vast amounts of data in a reasonable time frame, to use efficiently the unprecedented computer power available nowadays due to

advances in high performance computing and to produce –when possible- human understandable outputs.

Several research topics impinge on the applicability of metaheuristics for data mining techniques: (1) proper scalable learning paradigms and knowledge representations, (2) better understanding of the relationship between the learning paradigms/representations and the nature of the problems to be solved, (3) efficiency enhancement techniques, and (4) visualization tools that expose as much insight as possible to the domain experts based on the learned knowledge.

We would like to invite researchers to submit contributions on the area of large-scale data

mining using metaheuristics. Potentially viable research themes are:

- Learning paradigms based on metaheuristics, evolutionary algorithms, learning classifier systems, particle swarm, ant colonies, tabu search, simulated annealing, etc

-

Hybridization with other kinds of machine learning techniques including exact and approximation algorithms

-

Knowledge representations for large-scale data mining

-

Advanced techniques for enhanced prediction (classification, regression/function approximation, clustering, etc.) when dealing with large data sets

-

Efficiency enhancement techniques

-

Parallelization techniques

-

Hardware acceleration techniques (vectorial instuctions, GPUs, etc.)

-

Theoretical models of the scalability limits of the learning paradigms/representations

-

Principled methodologies for experiment design (choosing methods, adjusting parameters, etc.)

-

Explanatory power and visualization of generated solutions

-

Data complexity analysis and measures

-

Ensemble methods

-

Online data mining and data streams

-

Examples of real-world successful applications

Instructions for authors

Papers should have approximately 20 pages (but certainly not more than 24 pages). The papers must follow the format of the Memetic Computing journal: http://www.springer.com/engineering/journal/12293?detailsPage=contentItemPage&CIPageCounter=151543

Papers should be submitted following the Memetic Computing journal guidelines and stating in the accompanying letter that the paper is being submitted to the special issue on Large Scale Data Mining.

Important dates

Manuscript submission: March 31st, 2009

Notification of acceptance: May 31st, 2009

Submission of camera-ready version: July 31st, 2009

--

---- Jaume Bacardit, PhD Lecturer in Bioinformatics University of Nottingham Automated Scheduling, Planning and Optimisation research group, School of

Computer Science, Jubilee Campus, Nottingham, NG8 1BB, UK Multidisciplinary Centre for Integrative Biology, School of Biosciences, Sutton Bonington, LE12 5RD, UK Tel: +441159516276 Fax: +44 1159516292 Email: jaume _dot_ bacardit _at_ nottingham _dot_ ac _dot_ uk Web: http://www.cs.nott.ac.uk/~jqb

---