doi:10.1093/comjnl/bxh000

Editorial

Fionn Murtagh

Department of Computer Science, Royal Holloway, University of London, Egham TW20 0EX, England; and Science Foundation Ireland, Wilton Place, Dublin 2, Ireland
Email: fmurtagh@acm.org

Received 0; revised 0

I am demitting as Editor-in-Chief of the Computer Journal at the start of 2008. I am passing on the Editor-in-Chief role to an outstanding successor. Having been Editor-in-Chief for the past 7 years, I have to thank in particular Florence Leroy for providing so much administrative support. I would also like to thank Elaine Boyes, who heads Publishing, British Computer Society, and the team at Oxford University Press, especially Maxine Smith, Catherine Morgan and Clive Hemingway. A big word of thanks has to go to the authors and to the Journal's referees.

Particular High Points

In an ever more online world, the Computer Journal is doing very well indeed. Quality of research is the primary goal, over and above the form taken by research output. The Computer Journal continues to do great service to quality in our discipline.

From 2006, pursuing a proposal from Alex Gammerman, the Computer Journal Lectures were initiated and have established themselves with great presentations by leaders of our field, – Robin Milner, Tony Hoare, Alex Gammerman and Volodya Vovk, Keith van Rijsbergen, Samson Abramsky, Erol Gelenbe, and in early 2008, Steve Furber. It is with sadness that I note that someone who had intended giving a Computer Journal Lecture, Karen Spärck Jones, passed away in 2007.

The Wilkes Medal for the best paper in the previous volume was restarted, and a medal and prize launched.

The range of special focus issues testifies to tremendous vitality. Quite a large number are in the pipeline and will appear in print soon.

Some of the themes of the upcoming and planned special issues are hugely exciting. Below, I look at content coverage in the Journal. In my very initial faculty job in computer science, colleagues (and sometimes, I) used to sit at coffee/tea and debate what computer science really is. With the student recruitment crisis in the sector, after 2001 and up to 2006 or so, the same debates were still taking place. The Computer Journal points to what computer science is. And it should be even wider, embracing e.g. fields like complex systems, bioinformatics, computational neuroscience, and others.

With Fairouz Kamareddine's great help, summaries are produced of all articles. Thematic areas are now used to categorize articles.

A high point of my role as Editor-in-Chief was to see the 50th year of the Computer Journal. In the editorial on the occasion of Volume 50 (see issue 1 of that volume), some looking back occurred, by Maurice Wilkes, Peter Hammersley, Keith van Rijsbergen, and myself. As I recounted there, the Computer Journal's role is so similar to aims described by Henry Oldenburg, editor of the Philosophical Transactions, in its first issue of Monday 6 March 1665. It is an honour to pursue such aims.

Recent Content of the Computer Journal

The Computer Journal scores very highly as a scholarly journal of reference. With over 50 years of publishing it provides a unique vantage point on our discipline. This is ever more the case, since the Computer Journal is a generalist journal, with a very broad range of thematic areas represented.

Some thematic areas, and submissions from various parts of the world, have ebbed and flowed over my time as Editor-in-Chief. This has been noticeable, and interesting, especially to the extent that it represents underlying evolution and progress in interests. The viewpoint of variation over time in thematic content holds for submitted papers. Let me however take a closer look at thematic content of what has appeared in print in the Journal. Some years ago, Stella Mills collated computing topics from the first ten years of the Journal and the resulting paper appears now on the Journal's web site.

My look at thematic content fits in well with the categorizing of articles which was introduced two years ago. To allow for evolution of subject material, and indeed to be pragmatic in the labels used, I bootstrapped a set of subdiscipline labels. My aim

was to have an evolving, and flexible, set of subject categories.

For all articles from issue 1 of 2000, through to issue 5 of 2007, I looked at 377 papers. Editorials, book reviews, and obituaries were not taken into account. Then I looked at category labels actually used, or I retrofitted categories to older articles. What I came up with is the concept hierarchy or ontology seen in Table 1. Survey articles and history articles are included also as themes. There is no subdivision of bioinformatics, and the subdivisions used for software engineering are admittedly crude. Plagiarism is considered under machine learning. This collection of subject categories has some hierarchical structure, shown by the indentation, which often allowed use of narrower terms accompanying a broader term. The 377 articles were hand-classified, leading mostly to 2 of the terms in Table 1 being used for a given article. Each article had a maximum of 4 terms used and a minimum of 1. The articles × terms table is sparse.

Correspondence Analysis provides a way of studying even very faint structure in numerical, text, or semi-structured data. I use it here as a data and information representation technique, rather than a data modelling technique. For visualization, a Euclidean embedding is needed. Figure 1 is based on the primary thematic headings of *Systems – Physical*, *Information and Data*, and *Systems – Logical*. Respectively, out of 377 published articles, the (overlapping) numbers of articles in each of these primary thematic areas were 167, 149 and 161. All terms in Table 1 were used to characterize the published articles, and then the primary thematic areas were aggregated from terms. The planar projection shown in Figure 1 accounts for practically all information in the data since it is based on the published articles crossed by frequencies of occurrence of terms within primary thematic areas. That is to say, a 3-dimensional semantic space is projected into the 2-dimensional display. Shown as dots, the published articles are projected into this display. So also are the years of publication, 0 = 2000, 1 = 2001, ..., 7 = 2007. The years play an inactive role and are projected in after determination of the factors.

The triangular apex locations of the thematic areas follow from the ambient geometry of the semantic space. This configuration shows the polarity of Information and Data versus Systems (Physical or Logical) as being the most important distinction to be made. A large but less important distinction can then be made between the Physical and Logical Systems themes, with Data and Information being neutral for that polarity.

There is some "pull" from year to year in the direction of one or other of the thematic areas. This would be very interesting to follow up on over a considerably longer period of time. An immediate application of this analysis was that I initially took the theme of "Semantic Web" as a sub-theme of "Distributed Systems" only. But, having verified that it really was closely positioned

Systems – Physical

- Architecture, Hardware
  - Networks, Mobile
- Distributed Systems
  - System Modelling
  - Networks, Mobile
  - Grid, P2P
  - DS Algorithms
  - Semantic Web
  - Sensor Networks
- Networks, Mobile
  - Mobile Computing
  - Networks
  - Search, Retrieval

Data and Information

- Storage
  - Databases
  - Graphics
  - Imaging, Video
  - Memory
  - Non-Memory Storage
  - Network Storage
- Knowledge Engineering
  - Data Mining
  - Machine Learning
  - Search, Retrieval
- Data Mining
  - Imaging, Video
  - Semantic Web
- Machine Learning
  - Databases
  - ML Algorithms
  - Plagiarism
  - Reasoning
  - Representation
- QuantumProcessing
- Algorithms
  - Coding
  - Compression
  - Graphs
  - Strings, Lists
  - Trees
- Bioinformatics
- Computation Modelling

Systems – Logical

- Information Security
  - Networks, Mobile
- Software Engineering
  - Design
  - Education
  - Programming Languages
- System Modelling
  - Software Engineering
  - Testing
  - Ubiquitous Computing
  - Workflow
- Games
- Human Factors

TABLE 1. Concept hierarchy used, representing a view of appropriate subject headings for articles published over the past seven years in the Computer Journal.

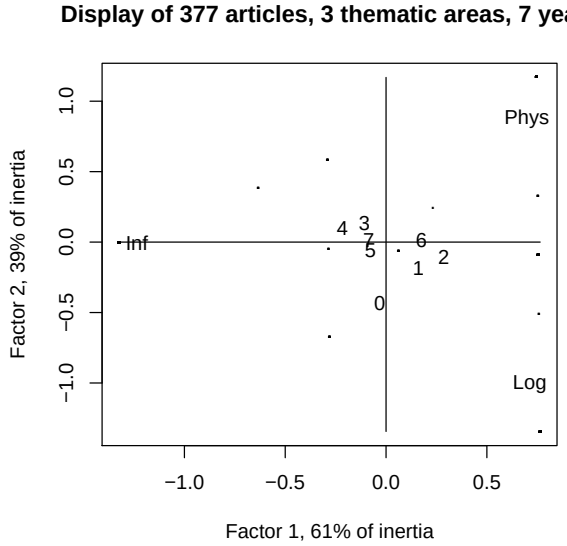

FIGURE 1. Nearly all information is presented in this semantic display, based on 377 published articles (positions shown with dots), crossed by 3 primary thematic areas. The latter are: Information and Data (Inf), Systems – Physical (Phys), and Systems – Logical (Log). The years of publication are shown (0 = 2000, 1 = 2001, etc.), used as supplementary elements in the analysis.

in semantic terms to the Information and Data thematic area, I responded appropriately and considered it also as being a sub-theme of "Data Mining".

The characterization of article content gives food for thought on the evolution of our field, and on the breadth and richness of themes that are represented in the Computer Journal. To the Journal under its new Editor-in-Chief, *Ad multos annos!*