Ambient Intelligence for Scientific Discovery

Yang Cai
School of Computer Science,
Carnegie Mellon University,
5000 Forbes Ave. Pittsburgh, PA 15213, USA
ycai@cmu.edu

Judith Klein-Seetharaman
Forschungszentrum Jülich, IBI-2: Institut für Biologische Strukturforschung,
D-52425, Jülich, Germany
j.klein@fz-juelich.de

Author Keywords
Ambient intelligence, discovery, interfaces, visualization, game, interaction design

ACM Classification Keywords
Information interfaces and presentation

OBJECTIVES

Human-Computer Interaction in scientific discovery has been gaining a momentum since the last decade due to its potential payoffs in applications, such as drug discovery, nano-materials or telecommunication. In HCI history, many advanced human interfaces were invented by members of the scientific community, such as the World Wide Web from CERN and the "Glass Cockpit" from NASA. With the growth of data streams and complexity of discovery tasks, there are demands for ambient intelligence such as peripheral vision, soft agents, data mining, and semantic web.

The goal of this one-day workshop is to gather interdisciplinary scientists and HCI professionals under one roof and explore ambient intelligence interfaces for the new needs in discovery. The expected results include a proceeding of selected papers, a position summary that states the current technologies and philosophy, and workshop for coming years.

The workshop is focused on three themes: 1) information reduction, 2) novel representations, and 3) discovery networks.

INFORMATION REDUCTION

Modern science and data communication technologies produce more data than can be analyzed by all humans together on Earth. For example, with the growing number of satellites and improving spatial and spectral resolutions, the data down-link rate will be 1 gigabit per second in 2010 and up to 10 gigabit per second by 2020. Much of this data is redundant and irrelevant to specific purposes, and can be classified as noise. Scientists spend up to 70% of their time on preprocessing the imperfect, redundant and noisy data. In many instances, it is not yet known how to distinguish noise from real signals. This is also the case in Bioinformatics, where protein sequences need to be interpreted with respect to their structure, dynamics and function, but to distinguish an important building block in the protein sequence from simple filling material is an unsolved mystery in biology. Furthermore, medical doctors are faced with new drugs on a regular basis and the amount of biomedical knowledge that could be useful for diagnosis and treatment is increasing, but to link the important pieces of information together that would result in actual improvement in human health is not yet possible. Delays during the analysis and lack of sufficient insight into the nature of a problem hinder the discovery prediction of disastrous situations or significant physical features in all of these and other examples. In addition, the overflow of information distributed over the Internet might be the cause for its end if the problem is not efficiently addressed.

We aim to reduce the information at the source side. Thus, we will share the experience from developers of ambient interfaces for data representation, semi-autonomous discovery agents, semantic data compression, feature indexing, and navigation tools for browsing large databases, knowledge or model based information understanding and information reduction.

NOVEL REPRESENTATIONS

Scientific discovery requires multidisciplinary insight and serendipity from routine work, i.e. the capability to make non-obvious connections between the complex interactions of the components of these systems. Such insightful solutions can often be found in an interactive and visual problem solving environment, as demonstrated for example by the fact that despite the modern numerical computing technologies, scientists today still use Gedanken experiments for concept development. [1] It is striking that simple intuitive simulation is still one of the most powerful approaches to creative problem solving. New ways of creating such interactive and visual problem solving environments are needed.

Since the early days of artificial intelligence, issues of interfacing human experts with computers have been explored. Especially, scientific reasoning and its representation, in particular for those connected with everyday knowledge of the behavior of the physical world, have been studied. At least three aspects have been studied: a) multiple representation, e.g. CaMeRa (Computation with Multiple Representations) that represents words with pictures [5], b) qualitative reasoning, e.g. Gardin's two-dimensional diagrams physical objects and their interactions [6] and Forbus' "qualitative physics" [7] and c) Ishii's "Tangible Bits" that represent data in tactile displays [8].

Recently, novel scientific discovery interfaces have being explored, for example the storytelling system [4] and a biological game [1] for fertilizing multidisciplinary biomedical problem solving, and tactile display for visualization of molecular structures [2]. In addition, analogies have been explored to allow the application of tools and expertise developed for one application carried over to a new application [2]. One of examples is the biology-language analogy [3] where scientists use the analogy between mapping of words to meaning and mapping protein sequences to their structure, dynamics and function (www.cs.cmu.edu/~blmt). Since humans intuitively "understand" how language works, this analogy provides them with a new way of thinking about the scientific biological question that enables them to generate creative ideas that would not have been possible without stimulating their thinking using the language analogy.

We are exploring under this theme new representation devices (e.g. neural network chips), analogies, new languages, tangible interfaces, ambient data displays, human-like referring expressions, game-based interaction design, semantic feature indexing, soft agent protocols, and others.

DISCOVERY NETWORKS

Contemporary discovery networks have been changing the landscape of scientific research. First, the ambient information environment, such as media, web, and telecommunication, build bridges among individuals who are in difference fields. For example, a web site about a computerized Traditional Chinese Medicine (TCM) has been visited by online readers at the rate of 167 per week. The readers were from five countries and four disciplines, which formed an *ad hoc* research community online. In the workshop, we will discuss how these scientific communities conduct day-to-day work and how they network with each other. Peter Jones' field study for Elsevier Scientific will serve as an example. Second, scientific discovery networks can be also applied to other areas, such as distributed health intelligence, healthcare for elderly people or people with special needs, homeland security and scientific education. For example, Julio Abascal explored the elderly care system that takes advantages of available network systems at home such as cable televisions, or wearable devices. Third, privacy protocols in scientific data will be explored. This is particularly urgent in medical practice, where patients' identities of DNA samples, diagnostic images and results could become available in public databases by accident. The consequences might cause lawsuits or criminal acts. Many publicly available databases, such as web browsers, web cams and satellite image archives can be used for data mining that can be invasive to privacy. Also, we need to balance privacy issues and effectiveness of scientific discovery. For example, when we encrypt a patient's identity, we would also ensure that the data can be retrieved properly.

FUTURE DIRECTIONS

Scientific discovery today has powerful vehicles such as the Internet, WiFi, SMS, System-On-Chip (SoC), data mining and Semantic Web. These contemporary information technologies enable ambient interfaces that can fuse and reduce information and, provide novel representations. We believe that these new technologies are creating a niche for Ambient Intelligence as a modal of scientific discovery. The Ambient Intelligence could be invisible but effective. The goal is to help scientists to focus in on important aspects of a problem, to make non-obvious connections and thus to aid them in the creative discovery process.

REFERECES

1. Cai, Y., Snel, I., Bharathi, B.S., Klein, C. and Klein-Seetharaman, J. (2003) Towards Biomedical Problem Solving in a Game Environment. In: *Lecture Notes in Computer Science* 2659, 1005-1014.

2. Roco, M.C. and W.S. Bainbridge, Editors (2002) Converging Technologies for Improving Human Performance: Nanotechnology, Biotechnology, Information Technology, and Cognitive Science, National Science Foundation: Arlington, USA.

3. Klein-Seetharaman, J. and Reddy, R. (2002) Biological Language Modeling: Convergence of Computational Linguistics and Biological Chemistry. In: Converging Technolgoy for Improving Human Performance – Nanotechnology, Biotechnology, Information Technology and Cognitive Science. Eds. Roco, M.C. and Bainbridge, W.S., National Science Foundation, 378-385.

4. Kuchinsky, A., K. Graham, D. Moh, A. Adler, K. Babaria, and M.L. Creech: Biological storytelling: A software tool for biological information organization based upon narrative structure. AVI (2002)

5. Simon, H.: Models of Thoughts. Vol. II. Yale Press (1989)

6. Gardin, F. and B. Meltzer: Analogical representations of naive physics. Artificial Intelli-gence 38 (1989) 139-159

7. Forbus, K.D.: Qualitative Process Theory. Artificial Intelligence 24 (1994) 85-168

8. Ishii, www.media.cmu.edu

9. Cai, Y. A Novel Imaging System for Tongue Inspection, IEEE IMTC'02, Alaska, 2002