Special Contribution

Internet challenges for informatics research

Gérard HUET

French National Institute in Informatics and Control (INRIA)

Informatics is a young discipline indeed. After computers appeared 50 years ago, the field was broadly divided between designing hardware (electronics) and programming software (combinatorics, mathematical logic). There was a dichotomy between scientific computing (engineering and science) and data processing (administration, services) - to the extent that two families of computers were developed for the two application fields. Progressively the two lines merged, programming became the dominant paradigm (to the extent that nowadays hardware is macro-generated by software), and data processing was for a while relegated to the sub-field of databases. Mathematical logic evolved to adapt itself as the foundation of constructive mathematics needed for expressing algorithms, data structures and networking protocols. The computing centers vanished, while personal computers invaded people's homes and offices. Computers shrunk to get embedded in cars, cameras, telephones, televisions, games, home appliances. Most importantly, networks merged into Internet, a major new actor in our civilization, which opens unforeseen opportunities for societal mutations.

In this unprecedented technical revolution, the terminology "computer science" is becoming increasingly obsolete. Algorithms, logical specifications, process calculi and numerous relevant mathematical areas are being interlinked in an emerging major science, Informatics, with vigorous fundamental research. But in this major revolution brought about in engineering, medicine, commerce, transportation and most sectors of human activity, programs proper are somehow becoming secondary to the data they manipulate. Text, voice, images, physical measurements in a range of scales from astronomy to cell and atomic structures, are sensed, stored, computed upon daily in a worldwide computing maze. Already 20 years ago, Steve Jobs launched the NeXT cube, delivered to every customer with Shakespeare's complete works on disk. Today any PC owner has a complete GPS cartography of the planet at his disposal for navigation on the Web through Google™ Earth.

The works of art of passed centuries, the artifacts of ancient civilizations, the testimonies of passed and present cultures, are being digitalized, analyzed, digested in virtual museums open to everyone's perusal. Bamiyan's buddhas are resurrected from the dust of destruction, forever smiling in their pristine detachment from history's turmoil. World's cultural heritage is here and now at our fingertips. Archives, libraries and museums are opening up in everyone's home, ready to deliver their secrets and treasures. Knowledge is plied out of academic institutions, education is liberated from intermediation, significant information is exchanged out of the control of official channels. The maze of traditional knowledge is deciphered in the common Web, taken away from the cabinets of erudite scholars for the benefit of ordinary citizens.

Moreover, the traditional dichotomy between passive consumers and active producers of Art and Science is fading away. Everyone may reveal his talents as literary author, painter, musician, film director, multimedia creator. Anyone may edit, produce, and distribute his or her own creations. When Andy Warhol said that everyone would be able to afford an instant of celebrity, he was still thinking in terms of traditional media such as television. Today every kid on the block may expose his blog diary for the whole world to see, appreciate, comment upon, and interact. Everyone may put his own knowledge to the service of cooperative endeavors such as the Wikipedia, instantly reverberated to far away lands and people, forever stored away in as-good-as-eternal digital archives. The challenges of channeling this sea of knowledge in durable and dependable form, without hampering the free creativity of its authors, are formidable indeed.

The grand challenges of software development are

well documented - interoperability, legacy management, quality of service and reliability, safety and security, privacy protection, crime prevention, protection of intellectual and commercial property, are well understood as important concerns, and are the source of important R&D developments. But the grand challenges of data management are at least as important social concerns, while not having been articulated so far with the same degree of awareness. Furthermore the ever-changing evolution of ways to represent and manage information is a constant technological challenge. The whole language of man-machine interaction is evolving at such a pace that research must proceed at unprecedented speed to keep from obsolescence. SMS, smilies, cookies, forums, chats, blogs, wikies, such new forms of communication are mushrooming without deliberate planification. Skype™, unknown medium a year ago, has today more than 2 million users online at any moment. Surfing on this information tsunami in an intelligent manner will foster important new research and innovation. Data mining is only at its beginning, since intelligent search and retrieval in multimedia multilingual dynamic documents of gigantic size is still in its infancy.

The global problems that our children will face during their lifetime are awesome. Pollution and energy shortage will make tougher a world where unemployment and other social deceases are on the rise in an aging industrial society, while hunger and illness still plague the rest of humanity. Informatics will help, notably for health care, and more generally for mathematical modeling of various phenomena, leading to optimization and control of processes, better management of information and citizenship governance. But Informatics will also have the charter to organize the stocking and dissemination of worldwide knowledge. Commerce itself will be increasingly dominated by knowledge. Active access to information channels will give the potential to anyone to be a creative actor of society, irrelevantly of his or her physical location. New forms of learning, better tailored to individual talents, and affordable to the masses, will appear. The digital divide will have to be bridged in order to insure equal opportunities for a nation of minds. This revolutionary social transition has to be accommodated intelligently.

Ultimately, laws will have to be passed to better regulate the circulation and manipulation of information. Such laws, however, are not to be expected out of blue sky; their necessity is a direct consequence of technical evolutions. Here too the Informatics experts have the duty to explain to legislators and judges not only

the technical details of the feasibility of such or such operation, but also recommend their best practice, alert to possible attacks or misuse, work on standards and implementations facilitating their proper use. Research being at the forefront of technical innovations, it also has a responsibility in advising ethical rules consistent with future evolutions of the discipline.

This quick survey of Informatics research challenges in the face of Internet deployment advocates an active involvement of researchers in societal issues. This is unusual for a scientific discipline, where the gold standards are traditionally publications by peer review and industrial impact through patents and other instruments. But information and knowledge are immaterial goods of strategic importance for the whole of society, so this involvement must be faced with. Furthermore, tackling the proper preservation of the cultural heritage of national communities brings new and exciting research challenges, at the forefront of knowledge representation and man-machine interaction. We foresee an important development of these research areas in the years to come. Let us wish that the new generation of researchers will take up this challenge and lead us towards an Informatics renaissance.

Gérard HUET

Gérard Huet is a researcher at the French National Institute in Informatics and Control (INRIA). His main research contributions concern the logical foundations of computing, automated deduction, functional programming, and software architecture. He is the author of an algorithm for unification in higher order logic, and made pioneering contributions to rewriting theory, sequential algebraic computing, and constructive type theory. He is the author or co-author of the Mentor programming environment, the KB theorem prover, the Caml programming language, the Coq proof assistant, and the Zen computational morphology toolkit. He currently works in Computational Linguistics, developing a platform for Sanskrit processing. He is member of the French Academy of Sciences and Academia Europaea. He received the Herbrand award for his work in computational logic, and a Doctorate Honoris Causa in Technology from Chalmers University.