□ 특별기고 □

Some Thoughts on Computer Science and Engineering Research

IBM T.J. Watson Research Center Se June Hong[†]

I have often been asked what specific area I work on. My stock answer has been computer engineering. "Is that software or hardware?" "Neither or both," I usually answer. So just what is this area called computer science and engineering? It has numerous sub-fields, and there are many researchers with very diverse backgrounds. To me there is one unifying theme for the entire area. It is the study of how to automate or to mechanize things effectively and efficiently. Earlier in the history of computers, the scope of things was limited to calculations. Soon afterwards, we used the word "computation" to include things that require a wider variety of algorithmic processing. Not only in theory but in practice, computers have become universal machines. Thus computer science and engineering now covers things like how to make good machines, things that have to do with how to interface the machine including programming and compiling for efficiency, things that relate to building and managing the flow of information through a network, things about how to handle large volumes of data, things about complex systems and things that reach into various other fields-- for each application area presents some new challenge to the task of mechanization. The scope of things that we consider mechanizable or that needs to be mechanized is ever increasing. Artificial intelligence, in this light, is also about how to mechanize complex behaviour that we regard as being smart.

Achieving effectiveness and efficiency are the two key goals of all these efforts. Therefore, the whole area is inherently "engineering" in nature. And considering the fact that the basic building blocks of computers are electro-magnetic devices, it is only natural that we see trends to form ECE or ECSE departments in the universities. This phenomenon also generates an additional benefit of teaching us the fundamental notion of engineering trade-off, something that was rather neglected by many computer scientists.

Whether something is done in hardware, software, middleware, firmware, etc., it all has to do with proper engineering trade-offs so that the task that is performed can be effective and efficient. Study of what can be mechanized, how fast certain generic tasks can be mechanized, and abstract views of machines and processes that run on them constitute what we call computer science.

[†]E-mail sjhong@us.ibm.com

Even here the science is relevant only if it has to do with being effective and efficient.

Thus, a relatively short prolog program that can be correctly written to make winning moves for a chess game is no accomplishment because, though this may be an executable spec in principle, the programmer lost sight of effectiveness and efficiency. (Such a prolog program would run for millions of years before determining a single move.) On the other hand, the success of Deep Blue is an engineering feat.

Computer science and engineering is in the core of a broader realm, information technology, which includes areas closer to the materials and devices level as well as information contents and delivery related areas on the other end. A successful researcher in computer science and engineering (by now the readers would have noticed that I have been using the term, computer science and engineering, as a singular concept) must have a general perspective on this full spectrum. Information technology, in turn, is the hub of most of the progress we will see in the 21st century, for virtually everything we do will be created, utilized, preserved, and appreciated through some use of information technology.

So, computer science and engineering is indeed a very important area.

A researcher in this area needs to have a deep appreciation of engineering principles and a solid background in mathematics. I think both of these aspects have not been emphasized enough in recent years. One should also question just what the result of his/her research is going to enable or enhance before expending the effort. What task is being mechanized, or if the research is on some mechanism (hardware or software), what kind of tasks will it facilitate?

Focusing one's research work on what is needed or strongly perceived as needed will make the results relevant and also much more appreciated. Even for those who do mostly theoretical work, a meaningful contribution cannot be made if there is no perspective on who will benefit, at the end of the food chain, from the theory or the new algorithm that will be developed. Challenging applications are often the most fruitful sources of such needs for both theoretical and application oriented research (e.g., electronic commerce inspires progress in communication, protocols, security, OLTP and database, data mining, retrieval, natural language processing, real-time processing, virtual reality, group-ware, agents, etc.).

Collaboration between industry and academia brings real problems to researchers in the academia and useful results to industry. The notion of technology transfer is outdated and ineffective. A new technology "transferred" (i.e. thrown over the transom) to the developers seldom gets to the market. A more effective paradigm is for researchers and problem originators (the users, the customers of the solutions) to work together. This helps researchers to truly understand the problems and the developers to fully absorb the details of the solutions. The IT industry would do well to encourage academic researchers to prowl through their establishments in short stints to develop mutually satisfying research goals. By the same token, academia will benefit by utilizing technical professionals in industry as visitors, adjunct faculty, and Ph.D. committee members. Joint research projects need not be just for tomorrow's products or a blanket support for some area just in case it may (usually not) pan out. With some planning and mutual trust some very rewarding and relevant work can be done together. There is a growing need for academia and industry to come to an understanding for fairly sharing the ensuing intellectual property rights.

When an advance is grounded in real needs, it is likely add to intellectual property. Patents and copyrights are valued output of a researcher along with published technical papers and demonstrated useful systems. The decision to apply for a patent must be made strictly as a business decision, however.

Researchers in industry are accustomed to this process, but in academia, there is a great need for business and legal advice to this end. Elegance of the method or apparatus invented notwithstanding, it will be a waste of effort and expense if there is no substantial business need now or in the near future.

There are many more patents that never get utilized than do, just as there are many more papers that are sooner forgotten than are meaningfully referenced by others. While one can not help producing some of this low impact output during one's career, one should be mindful of one's hit ratio. Producing results that are of no significant impact most of the time may continue to generate high points from some evaluators who know only to count the items, but wastes valuable resources and eventually tarnishes the credibility and reputation of both oneself and the profession. For patents, many industrial institutions publish what is called an invention (or patent) disclosure bulletin, in defence and for later protection, when there is no immediate business justification for incurring the cost of filing.

While I am on the subject of research, I must add what I believe to be a very effective way to proceed as a researcher. I am a firm believer in the fact that we detect flaws in someone else's ideas a lot quicker than in our own. It must be human nature. In fact, I have long learned not to trust proof-reading my own writing for I invariably miss many mistakes. Thus, close collaboration with a colleague can produce more useful results much faster than the analogy of parallel processing would suggest. The synergy really works. To do this, one should engage in technical discussions with a colleague on a routine basis.

Whenever I see a couple of technical people arguing over some points on the blackboard, I consider it a healthy and enviable sight. A problem is not to be owned and worked at by one researcher alone. A problem begs to be solved in the most expedient way even if the key idea may not be the poser's own. The efficiency and effectiveness principle also applies to the act of doing the research!

My thoughts on computer science and engineering research ran through the points I personally consider desirable. I wish to practice these myself and I share them with you hoping that the same will be useful for you, my fellow researchers in this area. Your contributions that are connected with all the things we will be automating in the 21st century are eagerly expected.

May the joy of discovery and creation accompany all your endeavors!