Profile: Andrew R. Conn

Research Staff Member and Research Relationship Manager Chemicals and Petroleum, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, USA

BSc Honours, Mathematics, Imperial College, London University; MSc, Computer Science, University of Manitoba; Ph.D, Computer Science, University of Waterloo

MY EXPERIENCES AS AN INDUSTRIAL RESEARCH MATHEMATICIAN

The year was 1989. I had been at the University of Waterloo, except for a postdoctoral fellowship at the Hebrew University, Jerusalem, in 1971 and two sabbaticals in France, ever since I arrived as a PhD student in 1968. It is fair to say that I was becoming a little restless, in spite of how great the University of Waterloo was for me and my research, but it had never occurred to me to work anywhere other than at a university. I had always had reservations about moving to the U.S., but I was starting to feel that I should find out first-hand whether they were valid. In any case I was certainly considering appointments at U.S. universities, as it is not easy to leave Waterloo for another Canadian university. Then, one day at Oberwolfach, Ellis Johnson asked if I would consider working at the T. J. Watson Research Center. I said I would, went down for an interview, and accepted a position within a year.

Some points to consider: Working at T. J. Watson is not a typical industrial position. At the time of my recruitment, one of my main interests was large-scale nonlinear optimization (the LANCELOT book [2] was nearing completion), and I was anxious to work on solving problems in which people were extremely interested in the answer. Furthermore, it is really good for one's ego to be asked to apply for a position. The salary I was offered seemed enormous, and the working conditions sounded idyllic. Although, as can be expected, much did not turn out exactly as it appeared at the time, I have few regrets concerning the choice I made.

Since the most contentious issues are likely to be associated with my reservations about moving to the U.S., let me just say that many of my prejudices were unfounded, but many others gave way to new ones. Suffice it say that after twenty years I am still here. Moreover, crazy as it may seem to some, both my wife, Barbara, and I adore New York City, which we live adjacent to, for many reasons. It turned out that my move almost coincided with the beginning of a significant (albeit temporary) decline for IBM in the 90s—so much so that friends asked if it was my fault. Consequently, much that was promised (such as the freedom to travel as much as I wanted to, which seemed realistic given that in those days I saw Ellis everywhere except at IBM) were never realized. My interpretation of my salary was also naive. It turns out to be very expensive to live in New York. Many things that we had taken for granted, like universal health care at moderate cost, good public schools, and moderate property taxes and insurance costs, transpired to be very different in New York.

As to teaching, I had always claimed that it was great in theory but often a pain in practise. I still maintain that teaching (at the right level and to good students) is very useful for research, but it makes constant demands that may not always fit into an ideal agenda. I also suspect that I enjoyed my teaching much more than my students did. However, it is not unusual for researchers at IBM to teach in universities, and indeed, for three years I taught a course in the Operations Research Department at Yale, had excellent students, and was about to build up a group of doctoral students when the department was unfortunately closed down, at least in part as a cost-saving measure. Currently, some of my colleagues in mathematical sciences, whilst maintaining their full-time positions at IBM, teach at Yale University, Columbia University, and New York University.

When I left Canada, I had the largest NSERC (Natural Sciences and Engineering Research Council) grant in numerical analysis in the country ($165,374 for three years). Within a year of arriving at IBM, I had a DARPA (Defense Advanced Research Projects Agency) grant, also for three years, of $917,809. Even taking into account the different exchange rates and the fact that NSERC did not cover overhead, the difference was considerable.

The Environment at IBM Research

I strongly believe that an industrial research laboratory like IBM T. J. Watson has much to offer that is not available in a university environment, and vice versa. Thus, the relatively recent demise of industrial research laboratories, as opposed to development laboratories, is to be lamented. I do not think that this is good for science or the community. Basic research is an important luxury* and has to be acknowledged as such. If short-term gains are the single most important objective, basic research cannot thrive. The researchers also need extended periods of uninterrupted thought. I consider that I am reasonably well acquainted with a cross-section of the western world's leading research universities. I know of no place that offers a more industrious or brilliant group of colleagues in mathematical sciences than those I have at T. J. Watson. That fact is certainly one of the highlights of my

move. The pressure, real or perceived, to hurry to completion things that are adequate but could be better, whilst a realistic sign of the times, is one of the main "lowlights."

I am never in doubt that I work for an American company. As a university professor, I was essentially self-employed, and it was rare that anybody told me what to do. As an IBM employee, I may be unmanageable, but nevertheless I have a real and rather hierarchical management. Political correctness, intellectual property, and legal issues are real and can be a considerable obstacle at times, although I do have IBM's consent to publish articles like this one. As an employee of a public education system like Canada's, I was subject to the vagaries of the political system, but that seemed relatively stable and conservative, compared with the effects of the stock market and the economy on the morale of a private company.

In the long run, the two environments may be equally stable; the short-term variance seems much greater in private industry, however, which can be both good and bad. There are also conflicting agendas in industry that can often appear shortsighted. I often disagreed with education policy, but it rarely affected me much locally. Taking care of apparently superficial things can be an annoyance. I unquestionably had better computing support (both hardware and software) when I was at a university than I do at IBM, and it was significantly easier for me to travel and have exchange visits with colleagues. There are more challenges working in industry, but the rewards, both financial and intellectual, are considerable. There is no shortage of interesting applied projects to work on at IBM, and although it may sometimes appear that the customer is too important, when a project is successful the results really matter, the problems are challenging, and they can be a significant catalyst for innovations in basic research. This working atmosphere is a clear motivator for learning new things.

Projects at IBM Research: An Application
Early on in my career at IBM, I became involved in the design of circuits. The history of that involvement is, of itself, interesting. Essentially, an electrical engineering colleague at Watson wanted to do some minimax optimization and asked my advice. After some brief discussion, I gave him some pointers and he took care of the rest himself.

At a later date, one of his colleagues was organizing an internal meeting across IBM concerning how we handled circuit tuning. On the basis of the earlier encounter, it was suggested that I could be considered a relatively approachable mathematician, and I was invited to attend the workshop. I had time to take part in about half of the meeting, whereupon the organiser, Chandu Visweswariah, now a distinguished engineer and senior manager at IBM, asked me what I thought of it. I said that I found it very interesting but regretted that they were using 1960s algorithms in the 1990s. I was then asked to tell them about 1990s algorithms. That was the beginning of an extremely successful collaboration that ended up gaining IBM millions of dollars and resulted in an IBM award for those who worked on the project; in addition to financial rewards, my wife and I, along with other awardees, were cared for royally at a retreat in the spring of 2002.

You may well think that nothing could have been easier than persuading IBM that the use of modern optimization to design better circuits was a win-win situation—especially when you consider that the alternative was to purchase outside tools. To benefit sufficiently from those tools, IBM would have had to disclose aspects of its circuits and, in doing so, contribute to improved utility of the software that the vendors would then sell to our competitors.

In fact, after the initial success with what might be called a prototype, a considerable part of the next eighteen months was spent fighting for the survival of the project. We did survive, for a number of non-trivial reasons. Firstly, we did receive unfailing support from our immediate management. As is often the case, the conflicts were higher up—and in this case relatively easy to explain. Basic to the optimization was the underlying simulation. What we had was a good, fast (about seventy times faster than a more accurate SPICE-like simulator) and reliable simulator, based on piecewise constant approximations; equally important, it also provided reasonably accurate derivative approximations. What others promised did not exist, would be based on piecewise linear approximations, and would not provide derivatives. The trouble was the prevailing view among higher management that linear approximations are better than constant ones, along with a failure to appreciate the significance of derivatives to the optimization. One thing I learned early is that—amazingly to me—what you promise is often much more important than what you deliver!

Secondly, I also benefited enormously from having world-class circuit designers and electrical engineers who had the confidence and ability to know that what they were doing was right and eventually to persuade others that this was the case. It is good to realize that, given time, those who understood the details were able to convince those who didn't, and eventually the correct decision was made. So practitioners, from a healthy mistrust at the outset, eventually became our biggest supporters. The experience also impressed on me the importance of good software that works, in an environment in which those who have to use it are comfortable. Otherwise, algorithms twice as good would not succeed. I think it is fair to say that the project could not have been as successful without the state-of-the-art optimization; without the electrical engineers and the familiar interfaces, however, it would not even have got off the ground. Furthermore, this is a project that would not have happened in an academic environment. It needed the investment, level of commitment, and breadth of skills that would have been difficult to find in such

a cohesive way anywhere other than a place like T. J. Watson. It gave rise to a number of research papers in premier journals, several patents, and a best-paper award [1], as well as the IBM recognition.

So, having set the not-so-atypical scene for a major project, let me give an idea of what it entailed. In high-performance custom-designed chips, the sizes of individual transistors are tuned to maximize the performance of the circuit. The key parameters are transistor widths, which control the amount of current that flows through the transistors. In simple terms, wider transistors generally lead to faster circuits, but consume more power and area and increase the load on previous stages of the circuit. Designers tune their circuits by adjusting the widths of transistors to obtain optimal performance.

When we began the project, traditional methods were slow, tedious, manual, and error-prone. The computer-aided design tool we eventually came up with automated circuit tuning whilst permitting its use on far larger circuits than possible with previous techniques. It used what was then a state-of-the-art optimization technique (LANCELOT), which was subsequently replaced with an interior-point filter method (IPOPT)[5]. In addition to making IBM circuit designs better, it significantly improved our designers' productivity. Today, it is deployed as a standard tool within IBM and assists in the design of all our custom circuits.

A Few Concluding Remarks

I work in an area rich in problems. It is exciting because of the access we have to extremely knowledgeable people both inside and outside IBM. The mathematical tools include nonlinear optimization, mixed integer nonlinear programming, differential equations, modelling, risk assessment, high-performance computing, simulation, and statistics. Most of the applications at such a level result in archival articles in mathematical journals. Of course, as research professionals, my IBM colleagues and I are involved in refereeing, serving on professional boards and committees, and, last but not least, promoting the field and our colleagues in universities. So you see that there is plenty of opportunity to have a rewarding and challenging career in industrial mathematics, from the point of view of both basic research and applications. Indeed, after LANCELOT, I completed two additional books [3,4] (with co-authors inside and outside IBM) since joining IBM.

I cannot overemphasize the two most wonderful things about my job: superb colleagues and a plethora of extremely challenging problems. I encourage those of you so inclined to be a catalyst for strengthening the industrial research centers that exist today and perhaps in the future promoting new ones.

*One could argue that it is important enough to be a necessity, and not a luxury; unfortunately, when budgets are tight, basic research is often seen as dispensable.

References

[1] A. R. Conn, P. K. Coulman, R. A. Haring, G. L. Morrill, C. Visweswariah, and C. W. Wu, JiffyTune: Circuit optimization using time-domain sensitivities, IEEE Trans. Computer-Aided Design of ICs and Systems, 17 (1998), 1292–1309. Also appeared in The Best of ICCAD: 20 Years of Excellence in Computer Aided Design, Kluwer Academic Publishers, Dordrecht, 2003.

[2] A. R. Conn, N. I. M. Gould, and Ph.L. Toint, LANCELOT: A Fortran Package for Large-scale Nonlinear Optimization, Springer Verlag, Heidelberg and New York, 1992.

[3] A. R. Conn, N. I. M. Gould, and Ph.L. Toint, Trust-Region Methods, MPS–SIAM Series on Optimization, SIAM, Philadelphia, 2000.

[4] A. R. Conn, K. Scheinberg, and L. N. Vicente, Introduction to Derivative-Free Optimization, MPS–SIAM Series on Optimization, SIAM, Philadelphia, 2009.

[5] A. Wächter and L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, Math. Program., 106 (2006), 25–57.