Parabola Volume 50, Issue 1 (2014)

Editorial

Dear Readers

One of the most fashionable research pursuits of our time, in the mathematical sciences, is Big Data Science. Fashionable pursuits attract a great deal of interest and funding, which creates more interest and funding, so that fashionable pursuits can quickly dominate research effort. There's a nice article by Freeman Dyson (*The Mathematical Intelligencer* September 1983, Volume 5, Issue 3, pp 47-54) about fashionable and unfashionable research. Dyson points out that, "It has always been true, and it is true now more than ever, that the path of wisdom for a young scientist of mediocre talent is to follow the prevailing fashion". Dyson goes on to advocate that "we should set aside a certain fraction of our resources, perhaps a tenth or perhaps a quarter, for the support of unfashionable people doing unfashionable things". Over the past two decades I think it is fair to say that in the applied mathematical sciences, the fashionable pursuit was Complex Systems Science. In the decade before that it might have been Chaos and Fractals, and before that Solitons and Integrability, and before that Perturbation Methods. When Dyson suggests that we should set aside resources for unfashionable pursuits he is not suggesting a return to pursuits that were once fashionable. He is thinking more of things that may one day become fashionable.

One of the curious characteristics of fashionable science over the past several decades has been that the research areas have themselves been hard to define. Even now after a few decades of research most scientists would find it hard to give a clear definition for Complex Systems Science. Similar problems, although not to the same extent, were troubling in the areas of Chaos and Fractals. I recall attending a conference where Martin Kruskal (one of the discoverers of Solitons and Integrability) quipped that "You may not be able to tell me what Chaos is but I can tell you what Integrability is: It is the opposite of Chaos". The Wikipedia article on Big Data doesn't give a very precise definition of Big Data Science but it does give a sense of why it might be fashionable. If I could summarize in a few lines, Big Data is everywhere, it's a challenge to deal with it, and there are tremendous opportunities. This summary could equally apply to Complex Systems Science, and to a lesser extent, Chaos, whose assured ubiquitousness and the end of Newtonian Determinism must have led some to wonder how we ever got to the moon, and back.

A little over a decade ago, Big Data Science was characterized by the so-called three Vs; Volume, Velocity, and Variety. This characterization was first introduced by Doug Laney in an article entitled *3D Data Management: Controlling Data Volume, Velocity, and Variety*, published online by the Meta Group as file 949. The three Vs concept has gained some acceptance but you wouldn't have to search hard on the internet to find ten Vs or even three Cs. However it is defined, Big Data Science is a playground for mathematicians and statisticians. Mathematical and statistical methods are required to extract meaningful knowledge from data. Some of the areas of focus in the mathematical sciences are methods for managing high dimensional data, inference methods for making sense of noisy data, computational methods that are scalable for large data, nonparametric statistical methods, statistical machine learning, large scale optimization, and network analysis. The potential for major advances in these areas is enormous. There is already a major advance that has been made in Big Data Science to which most people could quickly relate to. The internet, as we know it, started some time in the 1980s. Millions upon millions of pages were being set up with no library archival system for management, and the prospect of searching through this ever expanding collection of pages for useful information on a particular topic seemed to be an intractable problem. Then in 1998 Sergey Brin and Lawrence Page came up with a mathematical solution for ranking the importance of internet pages (http://infolab.stanford.edu/pub/papers/google.pdf), and this led to the development of Google, which currently has a market value capitalisation of some $360 billion. There is no doubt that there are many other Big Data problems that need solutions. Suppose that every car on our roads had its position and velocity sent to a central computer every second. How could we use this big data to change traffic light settings in real time to reduce traffic congestion?

Big Data Science is now on centre stage. It is fashionable, and along with fashion comes hype. But there are important real mathematical challenges to be met. Now, more than ever, we need new mathematical and statistical tools to make sense of big data. It's a great time to study mathematics and statistics at university, if you don't mind being fashionable. But if being unfashionable is your thing then mathematics and statistics can take care of you there too. Why not study something abstract like surreal differential equations? The only danger there is if you could come up with a consistent method for solving surreal differential equations, it may become the most fashionable pursuit since Big Data Science.

Editor

Bruce Henry