IMAges NZ

ISSUE THREE

INSIDE

New Zealand Institute of Mathematics & its Applications

The single letter that grew

When two statisticians met in a corridor at the University of Auckland back in 1990 they had no inkling that the initial of their first names would become the most commonly cited in mathematics. Jenny Rankine explains.

The encounter between Ross Ihaka, of Ngati Kahungunu and Pakeha ancestry, and Canadian Robert Gentleman led to their collaboration on the statistical software system R. Their original goal was to create something with which they could teach their first year introductory statistics course. They based R on existing software called S, and developed programming techniques to reduce memory demands and performance difficulties when dealing with large problems. They called it R after their names and because single letters can't be trademarked.

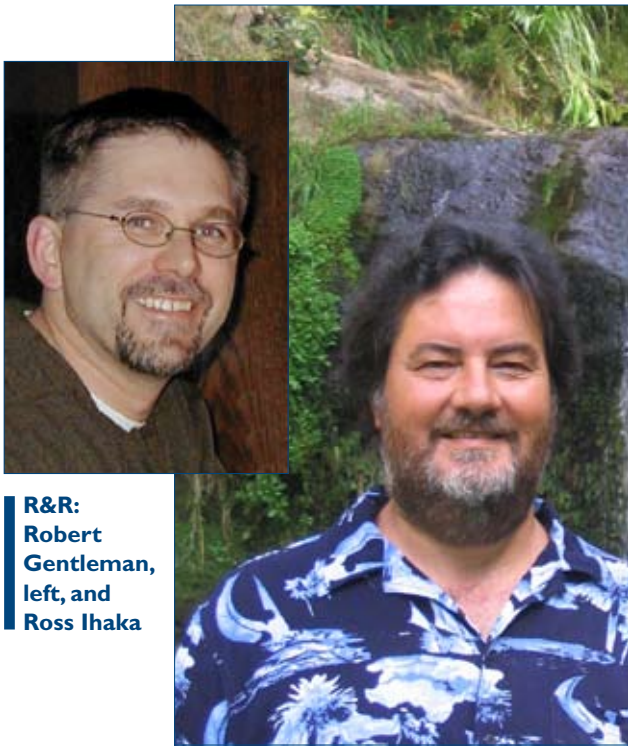
Rather than take the commercial route that S had, they took out a GNU public licence. "By making the software free," says Ihaka," we started to pick up a lot of collaborators. People aren't keen on working on commercial software, because their work gets taken over by the companies. We have the top researchers in the world involved in R, but we couldn't afford to pay them to work on it. They made R more and more useful, and more and more people started to use it. Now it's probably the most fully-featured piece of software in the world."

R can be used for linear and generalised linear models, nonlinear regression models, time series analysis, classical parametric and nonparametric tests, classification, clustering and smoothing. It also displays data in a range of well-designed charts, graphs and other diagrams.

R gets a life of its own

Ihaka and Gentleman's original paper about R has accumulated "a couple of thousand citations", as did another paper with the core collaborators, says Ihaka. The software took on a life of its own; the R Foundation formed by the R Development Core Team in 2002, with Ihaka and Gentleman as presidents, is based at the Vienna University of Technology in Austria. Mathematical publisher Springer told Ihaka it is preparing about 30 books for its series on applications of R.

Gentleman now heads the Program in Computational Biology at the Fred Hutchinson Cancer Research Center in Seattle, USA. Ihaka still "dabbles" on uses for R, but this self-taught programmer is now consumed by writing another programme. "The working name is L; I hope it will be better, a thousand times faster and



R&R: Robert Gentleman, left, and Ross Ihaka

able to handle much bigger problems than R."

His impetus is that "our ability to process and even store data is far exceeded by the rate at which we can collect it. Twenty years ago, we used to work with 50 to 100 numbers - now we work with billions. We might be comparing fertilisers using 50 plots of wheat in a field, or analysing high-resolution images from electronic telescopes, or analysing people's purchasing habits from their super-market transactions. We need better tools."

"I'm using Lisp, which dates back to the 1950s and is used in artificial intelligence and large-scale programming. R was an implementation of Lisp, but we didn't know much about it so there are ▶2

Welcome

We've had some great feedback on the first two issues of IMAges. We hope you enjoy this one just as much. Our new MathsReach initiative features, with a range of items about the work and interests of New Zealand mathematicians and statisticians.

Find out more from www.nzima.org.

Marston Conder and Vaughan Jones
Co-Directors

fundamental limits on in speed and problem size in R as a result. Then the machines you needed to run Lisp were bigger than we could afford, but now it's available on the smallest PC."

At the moment, the project occupies Ihaka and a Masters student. "People think you should get commercial sponsors, but we can't do it like that. We need hundreds of people working on it, and you can't get the investment to pay that many people."

Ihaka describes his work on R as "enormous fun. I get the occasional bottle of Scotch or free meal - a woman who teaches in a poor Black university in South Africa sent me a book they'd produced. A man researching tropical diseases in South America told me R was the only software they could afford. In a lot of developing countries it's the only one they use. Because it's so widely used, R has provided us with all sorts of contacts."

Ihaka is writing a book on R programming and one on statistical graphics and visualisation. He wanted to recommend a book to students about using colour and drawing graphs, but found few that were accurate about colour wavelengths and how our eyes perceive colour.

See also

www.r-project.org/foundation/main.html
www.gnu.org/

Remote but live

In March, audience members around the country were able to participate live in Professor Marcus du Sautoy's public lecture on the Music of the Primes.

The NZIMA and University of Auckland/ BeSTGRID presented the lecture to participants at the Universities of Auckland, Canterbury and Massey - Palmerston North and Albany campuses - and the Auckland University of Technology using portable AccessGRID nodes on the Kiwi Advanced Research and Education Network (KAREN) courtesy of BeSTGRID.

Two-way video and audio meant that another 100 people watching remotely were able to ask questions and participate in the discussion after the lecture with the 250 in one of the two University of Auckland lecture theatres needed for this event. In fact, the audience in other centres asked more questions than the one in Auckland.

Marston Conder, a NZIMA co-director, described the broadcast as a landmark event in New Zealand mathematics. Many remote participants had never been part of a lecture like this before, and many wanted to do similar broadcasts themselves.

Presenters can write their usual whiteboard notes on a tablet laptop, which is then projected up onto a wall for the audience.

See also

www.math.auckland.ac.nz/~bonning/video/marcus-du-sautoy.wmv for a sample video of the lecture on AccessGRID

www.bestgrid.org/index.php/Main_Page for information about the BeSTGRID and AccessGRID technologies

Associate Professor Paul Bonnington, Director of BeSTGRID, phone 09 373 7599 x 86623 or 021 623 206, email p.bonnington@auckland.ac.nz

ISSN: 1177-4819

Design:
Jenny Rankine,
Words and Pictures

New Zealand Institute of Mathematics and its Applications

Co-Directors
Marston Conder and Vaughan Jones

Research Manager
Margaret Woolgrove

c/o University of Auckland, Private Bag 92019, Auckland

P +64 (0)9 373 7599 x 82025

F +64 (0)9 373 7457

W www.nzima.org

E nzima-admin@ nzima.auckland.ac.nz

It is easier to square a circle than to get round a mathematician.
Augustus de Morgan

Southern Fields

The only two Fields Medallists from the southern hemisphere, Terry Tao (2006) and Vaughan Jones (1990) now work in the same part of the world. Professor Tao was brought up in Adelaide and is based at the University of California in Los Angeles. Professor Jones, born in Gisborne and a co-director of the NZIMA, is based at UC in Berkeley.