

The Newsletter of the Statistical and Applied Mathematical Sciences Institute

## Algebraic Methods in Systems Biology and Statistics Finishes with a Bang, but Doesn't End

The Algebraic Methods in Systems Biology and Statistics Program has formally ended, but many of the working groups will continue researching together for many months to come. Research by the working groups has already produced preliminary results of considerable interest.

The year kicked off on September 14-17 where over 115 people gathered at the Radisson RTP to start a year of research in this emerging field. From the opening workshop, four working groups formed to work on more specific research areas. The program was organized by Reinhard Laubenbacher (Virginia Tech), Seth Sullivant (North Carolina State U.) and Ruriko Yoshida (U. of Kentucky).

“The program has catalyzed several collaborations across fields which promise to be very fruitful. It had a truly transformative effect on my research program and has led me to think in new directions. In addition, SAMSI has helped me and my collaborators to bring together a number of disparate algorithms in mathematical systems biology and package them in a cohesive software system, Polynome, which is now available to the public,” notes Laubenbacher.

Sullivant comments, “Research Fellows Elizabeth Allman, Sonja Petrovic, and John Rhodes and I proved the identifiability of phylogenetic mixture models for group based models as a result of the workshops and working groups that met during the year. The techniques developed should have applications to many other families of mixture models.”

The *Evolutionary Biology* working group has made several advances including introducing the idea of k-interval speciation to quantify the amount of co-evolution between two trees. The group proves that two trees satisfying 1-interval co-speciation are, equivalently, separated by one Nearest Neighbor Interchange operation, which has been well-studied in phylogenetics.

The Evolutionary Biology group also concluded “There is no caterpillar in a wicked forest.” This settles a conjecture of Degnan and Rosenberg. A (rooted) caterpillar is a type of tree in which there exists an interior node descended from all other nodes. A wicked forest is a set of trees with a particular nasty property. For any pair of tree topologies A and B in a wicked forest, an observation of a high proportion of gene trees with topology A is evidence that the species tree has topology B, and an observation of a high proportion of gene trees with topology B is evidence that the species tree has topology A. The result is that no topology in a wicked forest can be a caterpillar topology.

The group also presented a polynomial-time algorithm for finding the geodesic distance between two trees in tree space. It is



(Left-Right) Ruriko Yoshida (U. of Kentucky), Reinhard Laubenbacher (Virginia Tech) and Seth Sullivant (NC State U) were three of the major organizers of the program.

based on producing a sequence of paths, where each successive path is formed by “bending” edges of the previous path. These intermediate paths correspond to “sliding” the legs of the path through tree space to successively shorten the path until the geodesic is obtained.

The *Algebraic Statistics and Experimental Design* working group is investigating two major related projects. The first is polynomial representation of probabilities and the second is lifting cumulant theory from finite discrete distributions to continuous distributions using the concept of finite generation. This work would be a generalization of Morris’ classification.

The *Network Inference* working group has studied biochemical reaction networks with mass action kinetics defining a system of ODEs with polynomial nonlinear right hand side. They derived conditions for the existence of at least two positive distinct steady state solutions by introducing sufficient conditions for the existence of a transformation that reduces the polynomial system to a linear one.

Yoshida remarks, “This special year was for me a great place to meet new people. Through these workshops I have met Megan Owen, a postdoc at SAMSI, and we have written a paper “First steps toward the geometry of cophylogeny” (with P. Huggins) Also during the transition workshop I worked with my colleagues, H. Hara and P. Huggins, collaborating on new papers which we are currently writing for submission.”

Some of the working groups will continue to collaborate and more discoveries in this emerging field of work will result from the Algebraic program.

## From the director...

I'm writing this having just gotten back from the end of the Transition Workshop of the program on Algebraic Methods in Systems Biology and Statistics. It is always sad to see the formal end of such a successful program, but wonderful to see a review of the exciting research that has been happening in the program. Many people contributed to making the program such a success, but special thanks are due to Reinhard Laubenbacher – who proposed and directed the program – and to Seth Sullivant who, with Reinhard, did the bulk of the planning and organizational work.

While SAMSI has always emphasized opportunities for new researchers, this period has been an exceptional one on several fronts. First, the seven National Science Foundation (NSF) Mathematical and Statistical Sciences Research Institutes, in collaboration with the Division of Mathematical Sciences at the NSF, were able to create 45 new two-year postdoctoral positions. This was in response to the economic downturn – which saw a severe reduction in available academic and research jobs for new researchers – and is an exciting way to further develop new researchers, while keeping them “in the scientific pipeline” until job conditions improve.

At SAMSI, eight postdoctoral fellows from this initiative will be joining the existing eight postdoctoral fellows to form what is shaping up to be a truly exciting postdoctoral community that will be primarily focusing on the two year-long scientific programs for next year: Space-Time Analysis for Environmental Mapping, Epidemiology and Climate Change; and Stochastic Dynamics. Men-

tors for this large postdoctoral group will arise from next year's 7 senior full-year visitors, 30 senior semester-long visitors, and 11 local Faculty Fellows. We were initially worried that, because of the economic situation, universities would have to cut back on their support for all these senior visitors to SAMSI but, impressively, not a single visit was curtailed because of the economic situation.

Go to <http://www.samsi.info/SAMSI-ALLPostdoc2009-10.pdf> to meet all the SAMSI postdocs for next year!

There were also an unusual number of opportunities for graduate students at SAMSI, with three distinct events. The first was the 3rd Annual Graduate Student Conference in Probability, discussed in an article in this newsletter. The second, in July, will be our regular summer event, the Industrial Math/Stat Workshop for Graduate Students. The third, at the end of July, is the Summer School on Spatial Statistics, intended, in part, to be advance preparation for many new researchers who will be involved in the 2009-10 SAMSI program on the subject.

Finally, I'm delighted to announce that Pierre Gremaud will be becoming Deputy Director of SAMSI in January, 2010. Pierre will be bringing his considerable talents to the job of overseeing the ongoing scientific activity at SAMSI, in addition to continuing his significant participation in planning for future programs.

James O. Berger  
Director



## SAMSI Postdoc Profile: Megan Owen

How do we construct evolutionary trees and how do we compare them? SAMSI postdoc fellow Megan Owen has been examining these questions this year during the Algebraic Methods in Systems Biology and Statistics program.

Owen always had a natural affinity towards mathematics. She used to enjoy doing problems in math books when she was a child. Owen grew up in Ottawa, Canada and attended Queen's University receiving a Bachelor of Science degree in Mathematics and Engineering. She was thinking at the time that she would get her degree, and then look for a job in industry, but her program advisor suggested that she consider pursuing her Ph.D. in mathematics and becoming a professor. She decided that would be a great career path, so she applied and was accepted to Cornell University. She received her Ph.D. in Applied Mathematics in 2008.

While Owen was at Cornell, Bernd Sturmfels told Owen about the upcoming program on Algebraic Methods in Systems Biology and Statistics that was being held at SAMSI. Reinhard Laubenbacher, from Virginia Tech and one of the main organizers of the SAMSI program, also told her about the program and encouraged her to apply, so she did.

This year Owen has been collaborating with members of the *Evolutionary Biology* working group. The group has looked at how to reconstruct evolutionary trees from data and how algebraic techniques could be used to help. Owen has been looking at ways to quantitatively compare evolutionary trees. Her research looks at computing the length of a path between the trees in treespace.

“The trick, however, is to find the shortest path between the two trees,” explains Owen.

She has developed an efficient algorithm that answers this dilemma. Owen worked with Scott Provan from the University of North Carolina at Chapel Hill on this problem. The algorithm they developed computes the shortest distance between the two trees, by transforming part of the problem into a problem on graphs and repeatedly solving that.

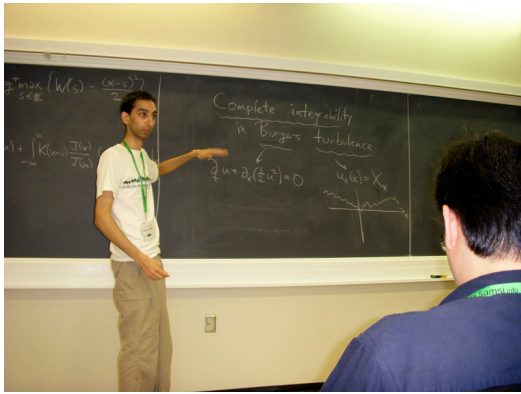
Owen presented her work in May at the 2<sup>nd</sup> Canadian Discrete and Algorithmic Mathematics Conference (CanaDAM). She also presented her work at the Society for Industrial and Applied Mathematics (SIAM) Annual Meeting in Denver.



Join Our Group! If you are on LinkedIn or Facebook please feel free to join SAMSI!

# Education and Outreach

## Probability Conference



The Third Annual Graduate Student Conference in Probability (GSCP) was held May 1-3, 2009 at the University of North Carolina at Chapel Hill. This conference was co-organized by a group of graduate students from the Mathematics Department at Duke University and the Statistics and Operations Research Department at the University of North Carolina at Chapel Hill. Approximately 100 graduate students and postdoctoral fellows from universities across the country attended this event. Planning is already underway for the 4th Annual GSCP to be held April 30- May 2, 2010 at Duke University!

## Undergraduate Workshop



The week long SAMSU interdisciplinary undergraduate workshop was held on May 18 - 22 both at SAMSU and on campus at NC State University in SAS Hall. Participants were selected from college juniors and seniors from across the country.

The main goal of the workshop is to provide the students with a hands-on experience of what mathematical and statistical modeling is about. Tutorials on mathematical and statistical methodology and on the physical experiments being used were given by SAMSU postdoctoral and graduate fellows. The students worked in small groups and presented their results at the end of the week.

## SAMSU Visitor Profile: Luis David Garcia-Puentes

Luis David Garcia-Puente, assistant professor of mathematics and statistics at Sam Houston State University, was very active this year as a visitor to SAMSU for the 2008-09 Program on Algebraic Methods in Systems Biology and Statistics.

Garcia-Puente grew up in Mexico City and attended Universidad Nacional Autónoma de México for his undergraduate work, earning his Bachelor's degree in 1999. He began his graduate studies in mathematics at New Mexico State University, where he met his Ph.D. advisor Reinhard Laubenbacher, one of the main organizers of the SAMSU 2008-09 Program on Algebraic Methods in Systems Biology and Statistics.

Over his career, he participated in programs and workshops in almost all of the NSF Mathematical Sciences Institutes.

"The work I have been doing at SAMSU is an interplay among statistics, algebraic geometry and geometric modeling. There is an underlying mathematical object, namely a toric variety, that appears in these different areas and facilitates the interchange of methods from one

field into the other," remarks Garcia-Puente about his work during his visit to SAMSU. Garcia-Puente spent the spring semester at SAMSU and also participated in the working groups throughout the year.

Several collaborations resulted from Garcia-Puente's visit to SAMSU. He is working with Seth Sullivant in identifiability problems arising in Gaussian Bayesian networks. This work also involves an undergraduate student from Sam Houston State University that participated in the SAMSU Two-Day Undergraduate Workshop held last February. Garcia-Puente also continued working together with the team led by Dr. Laubenbacher on parameter estimation for algebraic models in systems biology.

"One of the things that really impressed me was the WebEx meetings. It is great to be able to hold weekly seminars with people from around the world. That is something I don't think I've seen at other places yet," notes Garcia-Puente.



samsu.info

### Directorate:

James O. Berger | Director  
Duke University

Pierre Gremaud | Associate  
Director  
North Carolina State University

Michael Minion | Associate  
Director  
The University of North Carolina  
at Chapel Hill

Nell Sedransk | Associate  
Director  
National Institute of Statistical  
Sciences

### SAMSU Staff:

Rita Fortune | Administrative  
Assistant  
rita at samsu.info

Katherine Kantner | Webmaster  
kak at niss.org

Debbie Lesitkow | Program  
Assistant  
dcleisti at samsu.info

Cammy Cole Manning |  
Interdisciplinary Undergraduate  
Coordinator  
manningc at Meredith.edu

Sue McDonald | Senior  
Program Coordinator  
sue at samsu.info

Terri Nida | Workshop Specialist  
tnida at samsu.info

Jamie Nunnally | Communica-  
tions Director  
nunnally at niss.org

Debbie Smith | Program  
Assistant  
dsmith at samsu.info

James Thomas | Computational  
Systems  
help at samsu.info

# Calendar of Events for SAMSI

For more information about SAMSI programs and workshops, visit SAMSI's web site at <http://www.samsi.info>

2008-09 Education and Outreach Program  
**Industrial Math/Stat Workshop for Graduate Students**  
July 20-28, 2009

2009-10 Program on Space-time Analysis for Environmental Mapping, Epidemiology and Climate Change  
**Summer School on Spatial Statistics**  
July 28 - August 1, 2009

2009-10 Program on Stochastic Dynamics  
**Opening Workshop**  
August 30 - September 2, 2009

2009-10 Program on Space-time Analysis for Environmental Mapping, Epidemiology and Climate Change  
**Opening Workshop**  
September 13-16, 2009

2009-10 Education and Outreach Program  
**The Mathematics Institutes Modern Math Workshop at SACNAS**  
October 14-15, 2009

2009-10 Program on Stochastic Dynamics  
**Self-Organization and Multi-Scale Mathematical Modeling of Active Biological Systems**  
October 26-28, 2009

2009-10 Education and Outreach Program  
**Two-Day Undergraduate Workshop**  
October 30 - 31, 2009

2008-09 Program on Sequential Monte Carlo Methods  
**Transition Workshop**  
November 9-10, 2009

2009-10 Program on Space-time Analysis for Environmental Mapping, Epidemiology and Climate Change  
**GEOMED: Spatial Epidemiology 2009 Workshop**  
November 14-16, 2009  
Charleston, SC

