Newsjournal of the Society for Industrial and Applied Mathematics

siam news

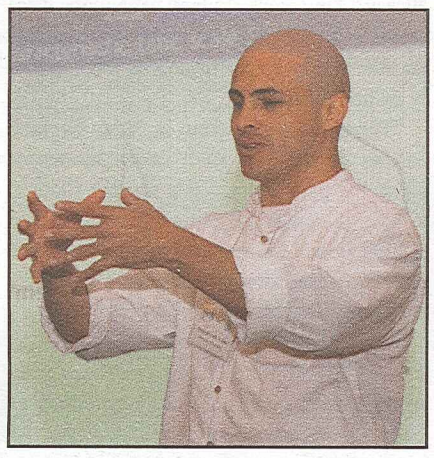
Volume 41/ Number 10
December 2008

11th SIAM Diversity Day Held in San Diego, July 2008

July was a good month for Jimena Davis, one of seven graduate students who spoke at Diversity Day, held in San Diego on July 9, during the 2008 SIAM Annual Meeting. Davis's talk—on the population dynamics of mosquitofish, a mosquito-control agent used as an alternative to pesticides in many rice-growing countries—was based on her dissertation, "Uncertainty Quantification in the Estimation of Probability Distributions on Parameters in Size-Structured Population Models," which she completed under the supervision of H.T. Banks at North Carolina State University. She received her PhD in July and is now a postdoc in the National Center for Computational Toxicology at the U.S. Environmental Protection Agency.

*The Diversity Day audience heard about the mosquitofish population model on which Davis worked at NC State, developing computationally efficient approximation methods for estimating the distribution of population growth rates. Davis, whose graduate study was supported by a Computational Science Graduate Fellowship from the Department of Energy, worked during the summers at Sandia National Laboratories—on the biochemistry of bacteria. Her experiences as a DOE Computational Science Graduate Fellow are described in the annual report of the program (*DEIXIS*, 2007–08, available from http://www2.krellinst.org/csgf/deixis/).*

Assessing Diversity Day as a valuable "opportunity to share my research with others as well as meet other minority graduate students and hear about their research," Davis encourages others to get involved in this "truly valuable experience for minority mathematicians."



In 1997, as a beginning graduate student at the National Autonomous University of Mexico, Marco Herrera went to the University of Arizona as part of a graduate exchange program between the universities. Not long afterward, he enrolled in a PhD program at Arizona to study dynamical systems, specifically the dynamics of molecular flux across membranes.

Realizing early on that his "fascination with dynamical systems extended beyond the models to the systems themselves," he set out to acquire the necessary background in the biological sciences. In parallel, he continued to study mathematics.

In July 2008, Herrera gave a talk at Diversity Day in San Diego titled "A Simple, Generic, and Physiologically Significant Model for the Pacemaker Dynamics of Cardiac Cells." A month later, he successfully defended his dissertation for a PhD in physiology; his research was on the relation between nearly coincident spiking and common excitatory synaptic input in motor neurons. He expects to receive a second PhD, in mathematics, in the spring of 2009.

Herrera's plans for the near future include postdoctoral work at the Mathematical, Computational and Modeling Sciences Center at Arizona State University. Not surpisingly, the career he envisions will be centered on "research combining mathematics and physiology."

"I think it is very important for graduate students to talk at meetings like the SIAM conference held last summer," Herrera says. "It helps the students realize how their work fits in a bigger picture concerning mathematics and science."

Additional coverage of Diversity Day 2008 can be found on page 3 in this issue.

Geometry Analysis and Signal Processing on Digital Data, Emergent Structures, and Knowledge Building

By Ronald R. Coifman and Mauro Maggioni

This article (which is based on the invited talk of the first author at the 2008 SIAM Conference on Data Mining) discusses "diffusion geometry," which, by generalizing classic tools of harmonic analysis, provides a synthesis of different approaches to data analysis and processing.

In the last few years exciting developments in data mining and machine learning have been applied to the analysis of large data sets arising in a wide variety of disciplines. With millions of text documents being converted to digital format, for example, many users would benefit from automatic ways to organize and extract information from large collections of documents, automatic recommendations of interesting documents based on their reading history, and so on.

Many of the problems that arise in this area fall broadly into two classes. The first class encompasses problems related to the geometry of the data: low-dimensional, low-distortion embeddings of large data sets in high-dimensional space and graphs, permitting visualization, human interaction and information extraction, denoising of data, outlier detection, and other capabilities. The second class includes problems about the approximation/fitting/learning of functions on the data from a few samples, with the goal of predicting the values of the functions at new data points. Of particular importance have been methods based on the assumption that the intrinsic geometry of the data plays an important role, and that the smoothness of relevant functions on the data should be defined in a way that is adapted to the geometry.

Diffusion Geometry

Ideas from harmonic analysis and spectral graph theory have played a fundamental role in recent advances in this area. Diffusion geometry starts from the premise that a similarity measure $A(x,y)$ between any pair (x,y) of nearby data points can be meaningfully defined. A typical choice for data points lying in R^D is $A(x,y) = \exp(-\|x-y\|^2/t)$, where t is a fixed scale parameter. In general, the choice of A is both data- and goal-dependent. If N is the number of data points, A is an $N \times N$ matrix, which we think of as sparse because only nearby data points are connected by an edge with a weight above some threshold.

We can renormalize A to obtain a Markov matrix P, which represents a random walk on the data points—that is, $P(x,y)$ is the probability of jumping from x to y in one step. In diffusion geometry, P and its powers are used to gain insight into the geometry of the data, e.g., by finding coordinate systems, as well as to construct dictionaries of functions, à la Fourier or wavelet analysis, for learning functions on the data.

As an example (see Figure 1), we consider a body of 1000 articles from *Science News* (kindly provided by J. Solka). We can represent each document as a high-dimensional vector by fixing a vocabulary of d words and letting the kth coordinate of a document be the frequency of the kth word in the dictionary for that document. In our case we selected 10,000 common English words and then retained the top 1000 with respect to a score for significance in the data set, based on mutual information. We define similarity between documents as the correlation between their word vectors when larger than 0.95, and as 0 otherwise. The

See **Diffusion Geometry** *on page 4*

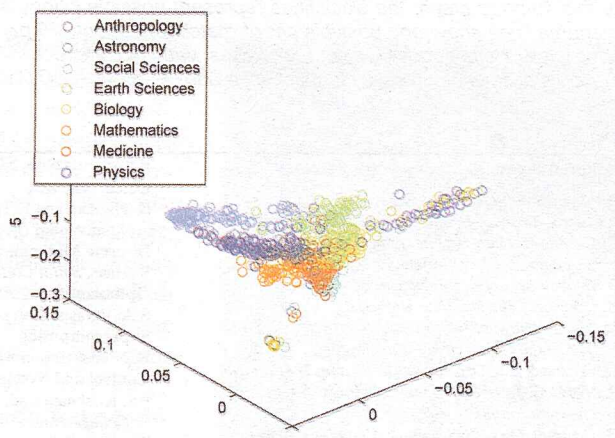


Figure 1. *Low-dimensional diffusion map of a body of documents from* Science News. *Data set courtesy of Jeff Solka.*

Matchmaking for Kidneys

In a 2004 article about mathematicians' spare-time pursuits, Dana Mackenzie led off with MIT graduate student Sommer Gentry, an accomplished swing dancer who had incorporated "the wordless communication used by dance partners" into her research on robotics. This fall, SIAM News *became aware of another of Gentry's talents: writing. Having been supported at MIT by a DOE Computational Science Graduate Fellowship, she was eligible to participate in the annual essay-writing contest sponsored by DOE for program alumni. Her essay—on matching donors and recipients for kidney transplants—took the top prize for 2007 (awarded by a trio of judges that included SIAM vice-president-at-large David Keyes). In the article that follows, Mackenzie brings readers up to date on Gentry, now a member of the faculty at the U.S. Naval Academy, her husband (her long-time dance partner and now a transplant surgeon), and results of their recent collaboration on kidney exchanges.*

By Dana Mackenzie

In April 2007, a 45-year-old African–American woman in Maryland made transplant history by becoming the first recipient of a kidney taken from a living donor and then flown all the way across the country. Anna (not her real name) had been waiting for a donated kidney for two years. Like many patients with kidney failure, Anna had a willing donor—her husband—but their blood types were a mismatch. She had blood type O and could receive an organ only from another person with blood type O. Her husband had blood type A.

When patients cannot find a compatible living donor, they often wait for what is called a "cadaver kidney"—a transplant from someone who has just died. But the odds against Anna were long. As a result of five pregnancies and numerous blood transfusions, her immune system was hypersensitive. Blood tests indicated that her body would reject organs from 96% of prospective donors, even those who matched her blood type.

Today, however, Anna is alive and has a normally functioning kidney, donated by a 50-year-old white man she had never met, who lived across the country in California. He had wanted to donate a kidney to his Asian–American wife, but also was denied because of a blood-type mismatch. Not only that, a third couple was involved, a Korean–American couple (both 48 years old) who likewise had incompatible blood types. By a series of three simultaneous exchanges—the 50-year-old in California donating to Anna, Anna's husband donating to the Korean woman, and the Korean man donating to the white man's wife—all three patients received the organs they needed, and nobody had to die.

For patients with renal failure, such games of "musical kidneys" are becoming more and more common, and the ending is almost always a happy one. Johns Hopkins

See **Matchmaking** *on page 2*

siam

SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

Nonprofit Org
U.S. Postage
PAID
Permit No 360
Bellmawr, NJ

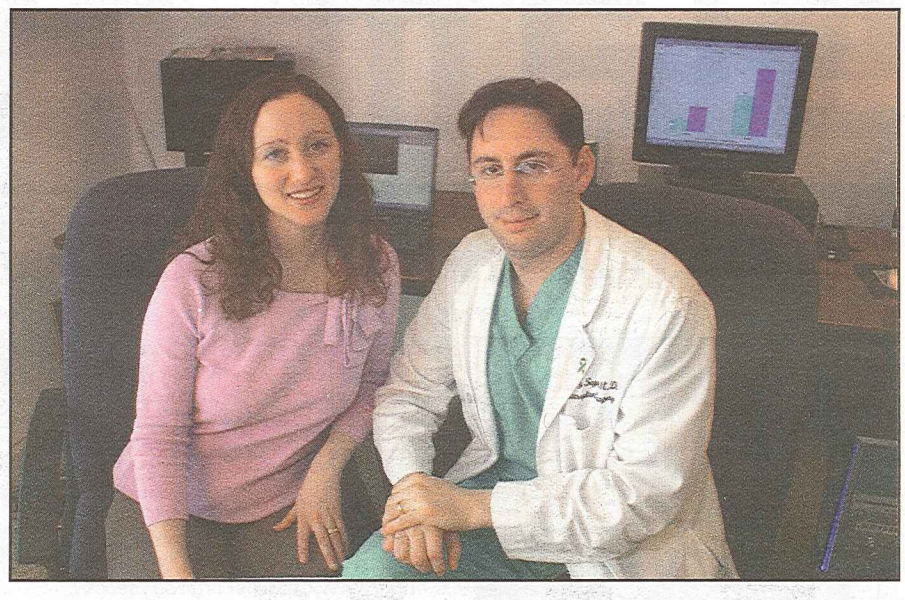
siam news
Volume 41/ Number 10/ December 2008

Matchmaking

continued from page 1

University (where Anna received her kidney) and the California Pacific Medical Center in San Francisco (where the other two transplants were performed) have been leaders in organizing simultaneous exchanges of pairs or triples of kidneys.

Although they have been highly publicized (a six-way exchange even appeared recently in the popular TV drama *Grey's Anatomy*), these multiple exchanges are no publicity stunt. They represent a very real way to make more kidneys available to more patients who need them. Sommer Gentry, an applied mathematician at the U.S. Naval Academy, has calculated that with a national database of incompatible donor–recipient pairs and a computerized matchmaking system, approximately 1000 to 2000 more kidney transplants could be performed every year—a major increase over the 15,000 transplants currently performed in a year. And the biggest beneficiaries would be people who, like Anna, are difficult to match. A nationwide pool would give them a better shot at finding a rare compatible donor.



Applied mathematician Sommer Gentry teamed up with her transplant surgeon husband Dorry Segev to optimize the decisions made by surgeons in matching donor–recipient pairs. Photo by Dorry Segev.

Multi-patient Exchanges: Too Ambitious

The idea of kidney exchanges was first proposed by Felix Rapaport, a former president of the Transplantation Society, in 1986. He pointed out that a donor who was ready to give a kidney to a relative, but was stymied by a blood-type mismatch, might have enough of an emotional investment to agree to a quid pro quo arrangement. But according to Alvin Roth, an economist at Harvard University, kidney exchanges were extremely rare until the early years of this decade, and were performed on a completely ad hoc basis.

"There had been four in New England before we started," Roth says. "In one of them, the patients met in a waiting room at a dialysis center." Presumably, the idea of trading donor kidneys came up as the two patients were comparing notes about their incompatible donors.

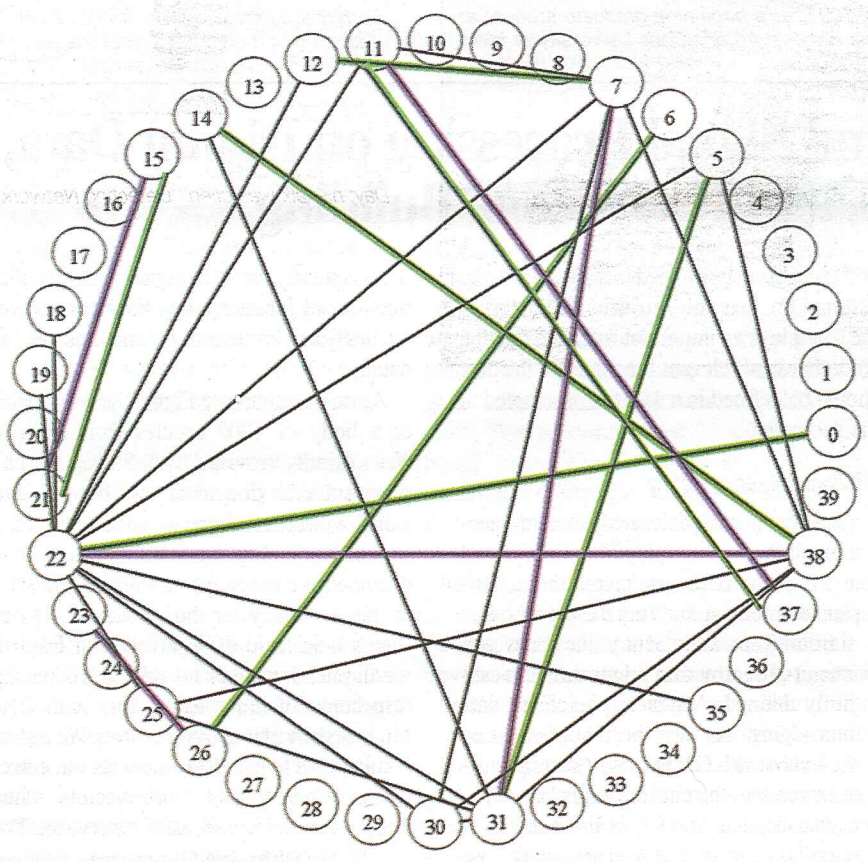
In early 2002, Roth and two colleagues, Tayfun Sonmez and Utku Unver, started to wonder if there wasn't a more rational way to approach kidney exchanges. "This is where economists come in," Roth says. "We shouldn't just rely on people bumping into each other." They proposed a kidney-matching algorithm, called TTCC, or Top Trading Cycles and Chains (see *SIAM News*, June 2004, http://www.siam.org/news/news.php?id=230), based on a similar method designed by the late mathematical economist David Gale to facilitate exchanges in a housing market.

"We sent [the paper] to every surgeon we could find," Roth says. "Only Frank Delmonico [a surgeon at Harvard University] came to lunch with us and said, let's start a program." Six years later, the New England Program for Kidney Exchange now links 17 transplant centers in a regional network.

The TTCC algorithm has one drawback, however: It produces large multi-patient exchanges. "Don't kid yourself, we'll never be that ambitious," Roth recalls Delmonico saying. Large exchanges are logistically more difficult. Every transplant requires two separate operating rooms (one for the donor, one for the recipient) and two teams of surgeons. And lots of things can go wrong between the identification of a likely match and the actual operation.

Even among people who have matching blood types and pass the antibody tests, there is roughly a 10% chance of an immune reaction, which can be detected only when the patient's blood is directly exposed to the donor's in a lab. Once a likely match has been identified, in other words, there is still a 10% chance that the operation will have to be called off. With three pairs, the probability climbs to 27%; with six pairs, it escalates to 47%. A larger number of patients also means a greater risk that one of the donors will back out—as dramatized in the *Grey's Anatomy* episode—or that a bumbling surgeon will drop one of the kidneys on the floor. (Oops, maybe that happens only on TV.)

For all these reasons, Roth says, Delmonico advised him to concentrate on pairwise exchanges at first. Ironically, that was the same idea that Gentry had. Her husband, Dorry Segev, is a transplant surgeon at Johns Hopkins. One Friday afternoon, as she picked him up from work, he told her about the problem of maximizing the total number of transplants. At the time, surgeons were using index cards or a magnetic board to match patients with donors by hand. "Does the kind of math you do say anything about this?" Segev asked.



In this 40-node graph, the black lines represent possible kidney paired donation matches. The magenta lines show one possible set of matches, but only 10 pairs get transplants. The green lines show the best possible set of matches, with 14 pairs receiving transplants. From Sommer Gentry's prize-winning essay (published in Compose, *the DOE CSGF annual essay contest journal).*

Pairwise Only: Not Ambitious Enough

By the next Monday, Gentry had an answer. She adapted an algorithm that had been discovered in 1965 by Jack Edmonds, a pioneer in the field of computational complexity. Edmonds's algorithm was designed to find the maximum number of matchings on a graph—in other words, the largest possible set of edges such that no two edges have a common vertex. In the kidney exchange problem, each edge would connect a mismatched donor–recipient pair to another pair who could exchange kidneys with

See **Matchmaking** *on page 3*

siam news

ISSN 1557–9573. To be published ten times in 2008: January/February, March, April, May, June, July/August, September, October, November, and December.

Advertisers: For display advertising rates and information, contact Kristin O'Neill at marketing@siam.org.
One-year subscription: $55.00 U.S., Canada, and Mexico, $65.00 elsewhere. SIAM members and subscribers should allow 8 weeks for an address change to be effected. Change of address notice should include old and new addresses with zip codes. Please request address change only if it will last 6 months or more.

Printed in the USA.

Diversity Day 2008

Diversity Day organizer Ricardo Cortez of Tulane University (far left), shown in San Diego with the day's seven graduate student speakers: Bassidy Dembele, Jimena Davis, Silvia Jiménez Bolaños, James Gatewood, Rosalyn Rael, Luis de la Torre, and Marco Herrera.

Silvia Jiménez Bolaños, now finishing a PhD at Louisiana State University, titled her Diversity Day talk "Corrector Theory for the Homogenization of Nonlinear Composite Materials." The workshop "gave me a great opportunity to present my work at a national meeting, with the advantage that it happens in a supportive and friendly environment," she says. "It was my first talk at a conference, and I cannot imagine how it could have been a better experience."
Under the supervision of Robert Lipton, Jiménez is studying the behavior of local fields in heterogeneous media, working to develop new multiscale tools to bound the singularity strength inside microstructured media in terms of the macroscopic applied fields. Future plans, along with application for postdoctoral positions, include generalization of her research results to the case of random materials and the search for novel properties that can be derived from nonlinear composite materials.
Students, especially from underrepresented minorities, should take advantage of opportunities like Diversity Day, she says. "Many students are unaware of the importance of getting involved in these activities. This workshop was a great place to find mentoring and networking that will help me in my future career."

Now an assistant professor in the Department of Mathematics and Computer Science at Grambling State University, in Grambling, Louisiana, Bassidy Dembele had just received his PhD when he spoke at Diversity Day. His talk, "Malaria Model in Periodic Environments," was based on his dissertation research, as is a paper that is slated to appear in the Journal of Biological Dynamics. *Dembele is now working on some of the "many open questions that can be generated from my work."*
"It is always a good thing to attend SIAM meetings," he says, citing the friendly atmosphere and the rewarding collaborations that often ensue.

University of Arizona graduate student Rosalyn Rael applies evolutionary game theory to systems of interacting organisms; she described her research at Diversity Day in a talk titled "Emergence of a Coexistence from Competition in an Evolutionary Game." She is also studying "how evolution can affect the outcomes of other systems, such as predator–prey systems." The goal, she says, is "to explore how evolution can change the outcomes of various ecosystem interactions."
Rael expects to receive her PhD in May 2009 and is now looking into postdoctoral fellowships. She found the Diversity Day program "a wonderful opportunity to present my research at a major conference. I was also able to develop some very valuable connections with people at the lunch"—including SIAM president Cleve Moler.
Like other Diversity Day participants, Rael considers the SACNAS (Society for Advancement of Chicanos and Native Americans in Science) conference, held this year in Salt Lake City, October 9–12, a valuable experience for minority students; her own participation dates back to her undergraduate years. She also credits two other programs—the Mathematical and Theoretical Biology Institute, then at Cornell, and the Los Alamos National Laboratory summer program for students—for paving her way to graduate school.

Luis de la Torre, currently a beginning graduate student at Northwestern University, was featured in SIAM News *as a 2007 participant in MSRI-UP, a summer research program for undergraduates. Shown here in San Diego with MSRI associate director Kathleen O'Hara, he was one of the seven graduate student speakers at Diversity Day; his talk was titled "Detecting Network Vulnerabilities Through Discrete Graph Methods."*

Lunchtime conversation with SIAM board member Petter Bjørstad was part of the Diversity Day experience for (standing, left to right) Donavion Huskey (Harvey Mudd College), Carmen Smith (Spelman College), and James Gatewood (Rensselaer Polytechnic Institute), and (seated) Gilbert Ymbert and Russell Carden, both of Rice University. Gatewood, one of the graduate student speakers, gave a talk titled "Traffic Flow Model of the Internet."

Matchmaking

continued from page 2

them. The algorithm could be tweaked to include real-world constraints (for example, giving priority to hard-to-match patients, avoiding long-distance matches for patients who couldn't travel). Best of all, Edmonds's algorithm was *fast*—it was, in fact, one of the first nontrivial examples of a polynomial-time algorithm.

In 2005, Gentry and Segev performed a computer simulation of the effects of implementing Edmonds's algorithm on a national database of incompatible donor–recipient pairs. (Only a simulation was possible, because no such database existed.) The simulation showed that 1000 to 2000 additional kidney transplants could be performed every year. Their article presenting the result appeared in the *Journal of the American Medical Association*. "The JAMA editors told me that it was the first article they had ever published that was a computer simulation, and they published it because it was so important," says Robert Montgomery, the chief of transplantation at Johns Hopkins, who was a co-author of the paper.

Unfortunately, Gentry's approach had a weakness that was just the opposite of the one holding back Roth's TTCC algorithm: Because it considered *only* pairwise exchanges, it wasn't ambitious enough. "If you limit the problem to two-way exchanges, the problem is simple, polynomial time," Roth says. "Or if you place no limit on the sizes, it's also simple. But the hard problem, which is NP-complete, is to compute the solution with a limitation on size." That turns out to be the most relevant problem for kidney exchanges. Three-way exchanges, such as the one involving Anna's kidney, significantly increase the number of transplant opportunities, but with more than three donor–recipient pairs, the practical problems start to outweigh the benefits.

Both Roth and Gentry have developed integer-programming algorithms to solve the problem for three-way exchanges, and their algorithms have been successful on a regional scale. Gentry's software identified the three-way match that Anna participated in, because one of the California donors had actually gone to Johns Hopkins and been entered into its computer system.

Toward a Truly National Registry

For a truly national registry with as many as 10,000 patients, however, better matchmaking algorithms would be needed. Both Gentry and Roth relied on the integer-programming package CPLEX, which just couldn't scale up to such a large network. Last year, a team led by computer scientist Tuomas Sandholm of Carnegie Mellon University developed special-purpose software that chops the problem into smaller parts that can be fed into CPLEX.

Now that a national matchmaking program seems technically feasible, will it actually happen? Last year, the major legal hurdle was removed. In 1984, Congress had banned the sale or transfer of human organs "for valuable consideration." Did another person's kidney constitute "valuable consideration"? This legal ambiguity prevented the United Network for Organ Sharing (UNOS), the U.S. government's watchdog agency for organ transplantation, from moving ahead with proposals to implement a national kidney-sharing network.

In 2007, Congress passed the Charlie Norwood Living Donation Act, which made it clear that organ exchanges are legal. "This proposal came up before three different Congresses, but then Charlie Norwood [a congressman from Georgia] died of transplant complications, and his colleagues passed it in honor of him," Roth says. With UNOS now legally able to support the plan, and fully committed to it, Montgomery says that the national registry of incompatible donor–recipient pairs will probably "go live" in 2010.

It probably would not have happened without the mathematicians. "At first, surgeons had not seen the problem as being interconnected to the degree that I was talking about," Gentry says. According to Montgomery, it was Gentry and Segev's unique combination of talents that made surgeons pay attention. Says Montgomery, "It's been a wonderful marriage of someone who was capable of solving the problem with someone who had the clinical ability to identify the problem and evaluate how well the solution fits."

Dana Mackenzie writes from Santa Cruz, California.