DIMACS Center
Rutgers University

DIMACS Research Experiences for Undergraduates (REU) Program

Annual Report

January 2005

Ia. Participants from the program

Participants:

DIMACS REU Domestic Program:

Jason Burrowes-Jones, Howard University (supported by other funds)
Zdenek Dvorak, Charles University, Prague, Czech Republic
Bryson Finklea, St. John's College
Daniel Halperin, Harvey Mudd College
Andrew Hodges, Manchester College
Vitek Jelinek, Charles University, Prague, Czech Republic
Jan Kyncl, Charles University, Prague, Czech Republic
Eva Ondrackova, Charles University, Prague, Czech Republic
Robert Renaud, Rutgers University
Ross Sowell, Sewanee College
Tomas Valla, Charles University, Prague, Czech Republic
Yinmeng Zhang, Carnegie Mellon University

DIMACS REU/RISE Program:

Susan Hope, East Carolina University

DIMACS REU Prague Program:

Shiri Azenkot, Pomona College
Logan Everett, Binghamton University
Tracy Grauman, Rutgers University
Diana Michalek, University of California, Berkeley
Bianca Viray, University of Maryland

Department of Mathematics REU Program:

Michael Alfare, Rutgers University
Yuriy Choliy, Rutgers University
Stephen Curran, Rutgers University
Andrew Dudzik, University of Chicago
Jana Gevertz, Rutgers University
Michael Grabchak, Rutgers University
Matthew Kohut, Rutgers University
Christopher Ross, Rutgers University
Vidya Venkateswaran, Rutgers University

Organizers:

Fred Roberts, Director, DIMACS, Rutgers University
József Beck, Dept. of Mathematics, Rutgers University
Sarah Genoway, Graduate Coordinator, Rutgers University
Janos Komlos, Dept. of Mathematics, Rutgers University
Daniel Kral, Graduate Coordinator, Charles University, Prague, Czech Republic
Jaroslav Nesetril, Director, DIMATIA, Charles University, Prague, Czech Republic

Seminar Speakers:

Eric Brown, Department of Mathematics, Princeton University
Amy Cohen, Department of Mathematics, Rutgers University
Paul Ellis, Department of Mathematics, Rutgers University

Dinesh Pai, Department of Computer Science, Rutgers University
Lara Pudwell, Department of Mathematics, Rutgers University
Michael Saks, Department of Mathematics, Rutgers University
Andrew Sills, Department of Mathematics, Rutgers University
Alexander Soifer, DIMACS, Princeton University, University of Colorado

Visiting Lafayette REU Program Speakers:

Ethan Berkov (mentor), Department of Mathematics, Lafayette College
- Jacob Carson, Lafayette College
- Jenna Hammang, Valparaiso
- Benjamin Hummon, Vassar
- Joy Kogut, Simmons

Gary Gordon (mentor), Department of Mathematics, Lafayette College
- David Eisenstat, University of Rochester
- Jennifer Feder, Washington University of St. Louis
- Greg Francos, Lafayette College
- Amanda Redlich, University of Chicago

John Meier (mentor), Department of Mathematics, Lafayette College
- David Glasser, MIT
- Lori McDonnell, University of Akron
- Kevin Penderghest, Lafayette College
- Mary Vacha, University of Nebraska

US Faculty Mentors:

Endre Boros, RUTCOR, Rutgers University
Ovidiu Costin, Department of Mathematics, Rutgers University
Graham Cormode, DIMACS, Rutgers University
Patrick DeLeenheer, DIMACS, Rutgers University
Stanley Dunn, Biomedical Engineering, Rutgers University
Martin Farach-Colton, Department of Computer Science, Rutgers University
Vladimir Gurvich, RUTCOR, Rutgers University
Paul Kantor, SCILS, Rutgers University
S. Muthu Muthukrishnan, Department of Computer Science, Rutgers University
Wilma Olson, Department of Chemistry, Rutgers University
Manish Parashar, CAIP, Rutgers University
Andrew Sills, Department of Mathematics, Rutgers University
Alexander Soifer, DIMACS, Princeton University, University of Colorado
Avy Soffer, Department of Mathematics, Rutgers University
William Steiger, Department of Computer Science, Rutgers University
Christopher Woodward, Department of Mathematics, Rutgers University

Faculty Mentors in Prague:

Martin Klazar, Charles University
Jan Kratochvil, Charles University
Martin Loebl, Charles University
Jiri Matousek, Charles University
Jaroslav Nesetril, Charles University and Director of DIMATIA
Pavel Valtr, Charles University

Ib. Participating Organizations

Telcordia Technologies: Facilities; Personnel Exchanges
Partner organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series.

AT&T Labs - Research: Facilities; Personnel Exchanges
Partner organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series. AT&T Labs hosted a field trip.

NEC Research Institute Inc: Facilities; Personnel Exchanges
Partner organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series.

Lucent Technologies, Bell Labs: Facilities; Personnel Exchanges
Partner organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series.

IBM Research: Facilities; Personnel Exchanges
Affiliated organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series. IBM Research hosted a field trip.

Princeton University: Facilities; Personnel Exchanges
Partner organization of DIMACS. Individuals from the organization are participating in the program planning and seminar series.

Charles University: Financial Support; Facilities; Collaborative Research; Personnel Exchanges
DIMATIA is a major research center at Charles University in Prague, Czech Republic, and our partner in the program.

Lafayette College; Personnel Exchanges
Mentors and students from the Lafayette College REU program visited DIMACS.

II. Project Activities

The DIMACS/DIMATIA REU program had the following three parts: (i) A group of students from all across the U.S. participated in an 8-week REU program headquartered at DIMACS; (ii) A second group of five students from all across the U.S. participated in the 8-week domestic REU program and then spent 3 additional weeks at DIMATIA; (iii) a group of students from the Czech Republic participated in the 8-week domestic REU program and then acted as hosts at DIMATIA when the U.S. students went to Prague. We collaborated with the Rutgers Department of Mathematics REU program, with participants from the two programs sharing office and living space, attending the same seminars, and participating in shared mentoring activities.

The overall goal of the program was to provide the participants with an exciting research experience that would help them decide on future educational and career paths. All students, including those participating only in the domestic program, got a taste of the international scientific enterprise. The U.S. students going to Prague got a more direct international experience and benefited from the scientific atmosphere at an international center of research, DIMATIA. The Czech students benefited similarly from exposure to their U.S. counterparts and, moreover, contributed to providing the global perspective that we sought. Our

REU program is unique because it is run in the context of two major research centers with many scientific activities. The richness of the intellectual community and the international flavor contributed by the many foreign scientists participating in DIMACS and DIMATIA activities at the same time as the REU students added to the overall atmosphere.

The key to our REU program was the one-on-one research experience under the direction of a mentor. The domestic part of the program officially began when the students arrived on the Rutgers campus in mid-June, moved into campus housing (where the students were housed as a group), and received offices and computer and library accounts at DIMACS. They were met by a graduate student coordinator, Sarah Genoway, who organized activities aimed at getting the REU students to meet each other and introduced them to their mentors to begin a program of directed study and research, including regular student/mentor meetings. There were regular lunches and teas, to which all the mentors and students were invited, as well as our weekly DIMACS REU Seminar Series hosting both local speakers and renowned outside speakers. The seminars were preceded by lunch and followed by opportunities to interact with the speaker. One of these seminars was devoted to a presentation about graduate school. Another was a workshop on how to give a presentation. The effectiveness of this workshop was evident from the high quality of the presentations made by the students during the course of the program.

We introduced the students to industrial research by making trips to AT&T Labs and to IBM Research, where our students got tours and there were technical presentations. Just as we feel it is important for the students to be exposed to multiple (and international) academic environments, we also found that the students were served well by the opportunity to explore two different corporate environments. Students were encouraged to take advantage of all of the activities at DIMACS. Many of them attended talks given by members of the DIMACS working group on Challenges for Cryptographers in Health Data Privacy. A few attended the BioMaPS/DIMACS/MBBC/PMMB Short Course: Transcriptional Regulation from Molecules to Systems and Beyond and the DIMACS Workshop on Usable Privacy and Security Software.

Students were asked to make several presentations about their projects during the course of the program. Near the beginning of the program, each student made a presentation describing his or her research problem. These short presentations, made before the entire REU group plus mentors, were aimed at introducing the whole group to the research topics of other members. This encouraged collaboration and discussion and provided the opportunity, as appropriate, to work on multiple projects or even switch projects. One day was spent hosting REU students from Lafayette College. Both the students in the Lafayette REU program and the DIMACS/DIMATIA students made presentations. Near the end of the domestic program, students made short presentations about their work to an audience consisting of REU students, mentors, and others in the DIMACS community. Students going to Prague repeated these presentations soon after their arrival at DIMATIA, in order to give the DIMATIA faculty and students an idea of the research activities of the visitors, and made further presentations of their work before they departed Prague.

In addition to oral presentations, students were asked to prepare personal websites. These websites describe their problem area and were filled in during the program and after the program ended with results that the students obtained.

The DIMACS part of the program did not end with the end of the DIMACS stay. Our U.S. and Czech REU students have been encouraged to stay in touch with their mentors by email, prepare web pages with their work, and keep those pages up to date after they left. Some of them ended up working on a research paper and/or preparing a presentation at a scientific meeting, still under the guidance of their mentor, after they left DIMACS. Many of our students planned or presented talks back at their home institutions and others planned to make their REU project a major piece of their senior thesis or even eventually of a master's thesis.

The goals of the experience in Prague were different from those for the experience in the U.S., and centered around introducing the participating students to a wealth of open problems and questions and appropriate problem-solving techniques and strategies.

We did not assign students going to Prague to an individual DIMATIA faculty mentor in advance. Rather, Jarosolav Nesetril, director of DIMATIA, served at first as acting mentor for each of these students. The DIMATIA faculty were involved in mentoring when the students arrived in Prague

Prior to leaving for Prague, students met several times with DIMACS staff to discuss the experience and the logistics. In addition, the Czech REU students and a Czech graduate student coordinator who accompanied the Czech students to DIMACS were available to offer pre-trip advice.

The Prague experience lasted three weeks. The U.S. graduate student coordinator accompanied the group and acted as the liaison person with the DIMATIA faculty and as a program coordinator while in the Czech Republic. The Czech students who participated in the DIMACS REU program in turn acted as hosts for the U.S. students in Prague. In the relatively short time the students were in Prague, they participated in several activities. They attended talks at the Eleventh Midsummer Combinatorial Workshop, a traditional problem oriented one-week international workshop on combinatorics. The speakers included distinguished professors and researchers from leading universities and institutes both in Europe (e.g., Renyi Institute, University in Bordeaux) and the U.S. (Princeton University, Emory University, University of Washington.) After the workshop, a series of small lectures, accompanied by informal discussions, by the members of the department at Charles University widened the knowledge of all the participants of the REU program on various fields of mathematics and computer science. The students also took part in various trips organized by Dan Kral, the DIMATIA graduate student coordinator, and the Czech undergraduates who had taken part in the first part of the program at Rutgers.

Students were met in Prague by the Prague student participants in our domestic program and moved into dormitory housing at Charles University that was provided by DIMATIA (through a companion Czech grant). Funds for meals in Prague were also provided by DIMATIA. During the first week, there were several days of special lectures by DIMATIA faculty Jan Kratochvil, Jiri Matousek, and Jaroslav Nesetril on combinatorics, graph theory, and combinatorial geometry to give students another introduction to the scientific interests of the local faculty, to prepare them for the topics of the Midsummer Workshop, and to present potential research problems. The students also took part in the Summer School in Combinatorics organized by Robert Babilon and Jiri Matousek.

The faculty members presenting tutorial lectures acted as mentors to the students, but we did not make formal assignments and we let the emphasis on problem solving lead to natural connections between students and mentors, often with more than one mentor per student. Students were encouraged to work on research problems as individuals or in groups or both. There were also presentations orienting the students to the culture and history of the city of Prague and the country, with special emphasis on the rich Czech mathematical tradition. The Czech REU students played a central role in this part of the program. The U.S. students also made informal presentations about their domestic REU research. Czech students were invited to attend the tutorials and presentations and joined the U.S. students for lunch and afternoon sightseeing. While at DIMATIA, the REU students participated in the Prague Midsummer Workshop, which has informal problem sessions each morning, aimed at identifying research topics. The Midsummer Workshop is attended by Czech students and faculty, and we involved the Czech students in afternoon sessions with the REU students to maximize the international exposure and contacts for the U.S. students. An REU student presentation was given by Zdeněk Dvořák. In the afternoon, the U.S. and Czech coordinators/mentors led discussions on some of these topics, helping the students to focus on research topics identified. The group meetings emphasized approaches to unsolved problems, problem-solving strategies, and group attacks on problems. Students were encouraged to explore, in collaboration with Czech mentors and students, the research questions that arose during the visit to Prague, either through faculty lectures, the Midsummer Conference, or informal discussions, and to pursue the research project begun at DIMACS with Czech mentors.

After the one week Midsummer Workshop, the REU program concluded the next week with more intensive one-on-one sessions between students and mentors and group meetings for problem-solving and for presentations of research results. While the short visit in Prague didn't give students as much time as they had at DIMACS to get deeply into research in pre-defined areas with faculty mentors, the experience taught the students a good deal about how research questions are formulated and pursued. Moreover, with the background gained during the DIMACS portion of the program, the students were able to make a good start on some of the open problems and research topics that they could pursue after leaving Prague. They were encouraged to remain in continuing email contact with their Czech mentors after leaving the Czech Republic.

III. Project Findings

Three Optimal Algorithms for Balls of Three Colors

Zdenek Dvorak (REU participant), Vit Jelinek (REU participant), Daniel Kral (REU graduate coordinator), Jan Kync (REU participant) and Michael Saks (REU mentor) studied a game played by two players, whom we will call Paul and Carol. Carol fixes a coloring of n balls with three colors. Paul cannot see the colors of the balls. At each step, he chooses a pair of balls and asks Carol whether the balls have the same color. Carol truthfully answers yes or no. In the Plurality problem, Paul wants to find a ball with the most common color. In the Partition problem, Paul wants to partition the balls according to their colors. Paul's goal is to minimize the number of questions he asks to reach his goal. Problems of this type are well motivated from the communication complexity point of view and their analysis usually requires clever combinatorial arguments. Dvorak, Jelinek, Kral, Kynel, and Saks found optimal deterministic and probabilistic strategies for the Partition problem and an asymptotically optimal probabilistic strategy for the Plurality problem. Results were presented at the Eleventh Midsummer Combinatorial Workshop organized by DIMATIA and the Department of Applied Mathematics of Charles University in Prague and the paper describing the results has also been accepted to the STACS'05 conference. The research has also been published as DIMACS Technical Report 2004-48. It was also published as Institut teoretické informatiky (ITI) Series, Universita Karlova v Praze 2004-214 and KAM-DIMATIA Series 2004-693. We anticipate that a journal version of this paper will appear as well.

An Infinite Ramsey-Like Game

At the 2003 REU, Vit Jelinek, (REU participant, now a PhD student at Charles University, Prague), Jan Kara (REU participant), Tomas Valla (REU participant) and Robert Samal (REU graduate coordinator) started working on a research project related to an infinite Ramsey-like game.

A complete graph is given as a game board. There are two players, each of them has a colored pencil (let's say that the colors are red and blue). The players take turns, and each player, on his/her turn, chooses an edge of the graph that has not been chosen before, and colors the edge with his/her pencil. The goal is to create a monochromatic clique on four vertices, i.e., a complete subgraph on four vertices whose edges are all colored with the same color (either red or blue). The first player to achieve this is the winner.

Jelinek, Kara, Valla, and Samal aimed to find a winning strategy for this game. Clearly, only the player who makes the first move of the game may have a winning strategy. If the game is played on a sufficiently large finite complete graph, the Ramsey theorem guarantees that the game cannot end in a draw, so there must exist a winning strategy for the first player. However, no explicit winning strategy was known. It was also unknown whether there is a winning strategy whose length is bounded by a constant independent of the size of the game board. If the game is played on an infinite complete graph, it is no longer obvious whether the first player can force a win, or whether under optimal play the game would continue forever.

Jelinek, Kara, Valla, and Samal found an explicit winning strategy for the first player. Its length is bounded by a constant and it works on an infinite game board, as well as on a sufficiently large finite game board. They have completed a partial preliminary draft of a paper "A winning strategy for the K_4 clique game."

Applying Space Filling Curves to an Information Discovery System

Dan Halperin (REU), working with his mentor Manish Parashar and Cristina Schmidt, interpreted the ideas of Discrete Space-Filling Curves in the context of the Filesharing/Text-based-query problem in peer-to-peer networks and came up with many optimizations and interesting results. In the Squid peer-to-peer framework, the peers (any machine on the network) are organized in a structured overlay that supports efficient lookup. The nodes are given IDs in a range of numbers, and the overlay lets any peer find another peer that corresponds to a given number quickly. Squid uses keyword searches, and it bounds both the length of the keywords and the number of keywords that can be used in a search. If every keyword is treated as a number, then a query, or set of keywords to make a search, can be considered as a point in a multi-dimensional keyword space. A space-filling curve (SFC) runs through the space so that it intersects every point in the space exactly once, and the points in the keyword space are then numbered sequentially based on where they occur on the SFC. A query corresponds to some subregion of the keyword space, and the parts of the SFC that it intersects correspond to the IDs of the peers that are relevant (e.g. store the results for) that search. This way of storing data is called a distributed hash table (DHT). There are a variety of SFCs. One that is used in many applications is the Hilbert SFC. It has been proven to preserve locality better than other known discrete SFCs. The idea is that if points are close in the multi-dimensional space, they are probably related to each other, and they should be as close as possible in the one-dimensional SFC so that their relationship is preserved. Halperin looked at text-based queries in file sharing, which is a specific application of Squid, and tried to optimize Squid's performance. Halperin investigated the effects of replacing the Hilbert SFC with a modified Z-Curve. There were a variety of reasons to consider this, but most notable is that the Z-Curve can be customized to the alphabet being used and hence is more storage efficient. He wrote a simulator for a P2P network and tested the performance of Squid implemented with the Z-Curve, ran several tests measuring performance in many areas, and came up with some promising results as well as deeper insight into the problem and areas for further research. He hopes to provide rigorous mathematical and algorithmic analysis and verification of his conclusions. He also hopes to investigate other issues such as the effects of load balancing on these algorithms and the importance of clustering on the efficiency of searching. He expects to submit a paper for publication.

Cache-Oblivious B-Trees

The run-time of a program is sometimes dominated not by the number of instructions it performs but by memory effects -- cache misses, the number of page swaps, etc. Some algorithmic research has focused on minimizing some particular memory effect -- say, minimizing the number of page faults. Can we minimize *all* memory effects simultaneously? Is this too much to ask for? Tracy Grauman (REU) considered both theoretical and experimental aspects of this question in her work with mentor Martin Farch-Colton. There are two algorithms Grauman implemented. The first produces a binary tree ignoring node sizes; the second produces a B-tree that is sensitive to block effects and block transfers and provides the optimal algorithm. She found that the first algorithm was reliable but slow. She is continuing her work on the implementation of the second, hypothetically faster, algorithm.

Author Identification: Who Wrote the Federalist Papers

Undergraduates Diana Michalek (REU) and Ross T. Sowell (REU) worked with mentor Paul Kantor exploring the following question: Given a list of possible authors, a sample of text written by each author, and a sample of text of unknown authorship, can it be automatically determined which author wrote the text?

Authorship identification techniques might be used to determine whether or not William Shakespeare is the actual author of the Shakespearean plays. Other applications are to determine whether a terrorist threat is authentic, or to validate the author of an electronic mail.

One classic example of an authorship identification problem is that of the disputed Federalist Papers. Eighty-five essays appeared in newspapers in 1787 and 1788 and were each signed with the pseudonym "Publius." The authorship of seventy-three of them is known. However, it is disputed as to whether Hamilton or Madison wrote the other twelve. In 1964, Frederick Mosteller and David Wallace used thirty function words and Bayesian inference to attribute all the disputed papers to Madison.

Michalek, Sowell, and Kantor developed a new approach to the author identification problem. Using the frequency of function words, and a much older approach based on the lengths of the words, as well as word length frequencies to classify the papers, they ranked these features according to their discriminating ability, and showed that by using the top ranked feature from each list, linear discriminant analysis can be used to classify all the disputed papers to Madison. The most significant aspect of their work is that they achieved success with simple methods, which ensure that *different kinds* of features all get to play a role in the classification. This suggests that when features are very distinct, balancing them in very simple ways may make it possible to use very simple statistical techniques to classify text based on these chosen features.

Optimal Protein Encoding

Logan Everett (REU) worked with mentor Endre Boros on a problem arising in mathematical biology. Their goal was to find the most efficient way to encode protein data into binary strings such that the Hamming distance between the binary strings is inversely proportional to the similarity metrics. Halperin, Buhler, Karp, Krauthgamer and Westover had recently proposed a method to construct metric embeddings of amino acids, in order to speed up searches in protein databases. This approach, based on semidefinite programming, generates very long binary strings and there is some distortion. Boros' and Everett's work focused on a method, based on linear programming, that gives the theoretical minimum distortion and much shorter binary strings. This project involved Logan in learning linear programming and writing a C++ program utilizing a professional LP-solver. Everett and Boros managed to solve several very large linear programs (~400 constraints and ~500,000 variables). Their report "Optimal Protein Encoding," DIMACS Technical Report, DTR 2004-55, was published in November 2004. They plan to work further on some additional issues arising from these results. They expect the final version to be published in a professional journal.

Mathematical Models for HIV Incorporating Mutation

Andrew Hodges, Manchester College, worked with DIMACS post-doc Patrick DeLeenheer on HIV models for describing the dynamics of the virus, its target cells (CD4+ T-cells), and infected T-cells in the blood stream of an infected individual. They assumed the existence of 2 strains of viruses and associated infected T-cells. By means of local stability analysis and through simulations they determined that the mutation rate has a significant effect on which strain will be dominant. (Mutation occurs often and on a rapid time-scale in HIV.) They found that a strain that may be dominant if there is no mutation might be cleared if the mutation rate is sufficiently high.

The McKay Correspondence

In work that was part of the associated Rutgers Department of Mathematics REU program, Andrew Dudzik (REU) and Stephen Curran (REU) worked with mentor Christopher Woodward and Sikimeti Mau, Rutgers graduate student, on the McKay correspondence, and related topics discussed in Kronheimer's paper "The construction of ALE spaces as hyperK'ahler quotients." The correspondence relates finite subgroups G of SU(2) with the simply-laced Lie algebras; for instance the icosahedral group corresponds to the Lie algebra E_8. One way of defining the correspondence is through representation theory (define the corresponding Dynkin diagram by taking the nodes to be irreducible representations and edges given by decomposing the tensor product with the standard representation).

They worked out the representation theory for finite subgroups of SU(2) and the resolution of singularities for the associated simple singularities C^2/G. The graph of components of the exceptional divisor (a node for each component and an edge for each intersection) gives another way of obtaining the Dynkin diagram. In fact, this process assigns a number to each node (how many blow-ups before it appears) and Dudzik worked out a formula for these numbers, which apparently was not previously known.

Next, they looked at an approach of Kronheimer to constructing deformations of the simple singularities via quiver representations. A quiver is an oriented graph; a quiver representation associates to each node a vector space and to each edge a linear transformation. The quiver in question has two edges in opposite directions for each edge of the Dynkin diagram. The moduli of representations can naturally be viewed as a symplectic quotient; by shifting the value of the quotient one deforms away the singularities. Curran and Dudzik worked out some examples, in particular D_4.

Finally they worked on a conjecture of M. Douglas, B. Fiol, and C. Romelsberger on the existence of certain quiver representations that arise in string theory. The problem is to show that certain moduli spaces, whose expected dimension is negative, actually do not exist. Curran and Dudzik managed to solve a few more cases than were solved in Douglas' paper.

A Mathematical Model of Real-time PCR Efficiency

Several real-time PCR (rtPCR) quantification techniques are currently employed to determine gene expression from rtPCR data (fluorescence intensity). In each of these quantification techniques, it is assumed that the efficiency of rtPCR is constant. In work that was part of the associated Rutgers Department of Mathematics REU program, Jana Gevertz (REU) worked with mentor Stanley Dunn and Charles Roth (Chemical and Biochemical Engineering) to analyze rtPCR data. They showed that even during the exponential phase of rtPCR, the efficiency of the reaction is not constant, but is instead a function of cycle number. Based on this observation, it seems reasonable that current quantification techniques can be improved upon if the behavior of rtPCR efficiency can be understood. For this reason, Gevertz, Dunn, and Roth have developed a mathematical model of the annealing and extension phases of the PCR process. By updating the quantities of each species (e.g., primers, templates, active DNA polymerase) after each cycle, the model has allowed them to predict the efficiency of rtPCR at any cycle number, given a set of initial conditions and parameter values, which can mostly be estimated from melting temperature and other biophysical data. The model predicts a precipitous decrease in cycle efficiency when the product concentration reaches a sufficient level for template-template reannealing to compete with primer-template annealing; this behavior is consistent with available experimental data. The rigorous understanding of rtPCR provided by this model allows the development of more accurate methods to quantify gene expression levels from rtPCR data.

Automated Proofs of Combinatorial Identities

In work that was part of the associated Rutgers Department of Mathematics REU program, mentor Andrew V. Sills worked with Yuriy Choliy (REU) to build a faster algorithm for finding proofs of combinatorial identities. A geometric series is a series $a_0 + a_1 + a_2 +\ldots$ in which the ratio a_{k+1}/a_k of consecutive terms is constant. A hypergeometric series (first studied by Gauss) is a series in which the ratio of consecutive terms is a rational function; thus the term "hypergeometric." Since the time of Gauss, many identities involving hypergeometric series have been discovered. In the course of studying such series, it became standard to look for, by ad hoc methods, recurrence relations satisfied by the series. Sister Mary Celine Fasenmyer's algorithm that finds a recurrence relation satisfied by a given hypergeometric term was a major advance in the 1940's. In the early 1990's Herbert Wilf (University of Pennsylvania) and Doron Zeilberger (Rutgers University), as a part of what is now known as the WZ theory (named in honor of two famous complex variables), found a much faster algorithm that accomplishes much the same goal as Sister Celine's algorithm. Furthermore, Zeilberger has implemented the algorithms in Maple, while Peter Paule and Axel Riese of the Research Institute for Symbolic Computation (RISC) in Linz, Austria have done the same in *Mathematica*. As a result of this work, it is possible to prove a large class of hypergeometric identities in a completely automatic fashion using a computer.

A crucial step in a WZ-type proof of a given identity involves finding a certain rational function R, which is unique for every identiity. To find this function, the WZ method uses an algorithm developed by R.J. Gosper. Because use of Gosper's algorithm can be very time- and memory-consuming, many important identities still cannot be proven using the WZ method. The goal of Sills and Choliy was to develop a new, faster and more efficient method for finding the function R. Sills and Choliy implemented their ideas in Maple, so they could do a direct comparison with Zeilberger's implementation. They tested their implementation versus Zeilberger's on the same machine on a collection of ten classical hypergeometric identities. Their implementation was 9.8 times faster, on average, than Zeilberger's. They plan on implementing their algorithm in *Mathematica* and testing it against the Paule-Schorn-Riese implementation. Meanwhile, Zeilberger and his student Mohamud Mohammed have an improved version of Zeilberger's algorithm and the WZ method that is theoretically much simpler than the original, which could potentially provide a serious challenge to the Sills-Choliy implementation, creating an exciting mathematical competition. Also, it is well known that the WZ theory extends quite naturally to q-analogs, with computations that are often much more computer intensive than ordinary symbolic algebra. It is quite plausible that if Sills' and Choliy's methods are employed in the q-case, the increase in speed could be very dramatic.

IV. Project Training/Development

In a very real sense, the entire project was about training and development, focusing outstanding undergraduates on opportunities in a research career in both academic and business environments and exposing them to the international nature of the research enterprise. The following comments we received from students in the program reflect the impact the program had on training them in this sense:

Comments from 2004 Participants

"I was always interested in research but this REU experience has solidified that. Even though my research was not in electrical engineering I still found computer science intriguing and I intend to continue this project. I have been exposed to the ups and downs of life as a researcher and I realized that results are not proportional to effort but results do not come without effort. I have also improved my communication skills due to the various presentations that were given during the program. If possible, I would definitely consider this experience once more." Jason Burrowes-Jones, 2004

"I am now in Budapest, learning more math and science in a 5th year to prepare myself for graduate school. Just finished a semester at the Budapest Semesters in Math. Next semester I'll be volunteering full-time in an experimental neuroscience lab here in Budapest. I met the researcher, Zoltan Nusser, through a U. Chicago professor that I met at a conference that I went to over my DIMACS REU summer. It was very nice to get support to go to this conference even though it was not directly related to my DNA project with Dr. Olson. The experience at the conference (Computational Neuroscience Society meeting in Baltimore, MD) has led to many connections and a decision to go to graduate school in computational/mathematical neuroscience. (The school is still undecided.) I am a good example of how allowing for some flexibility in the REU program for students to spend a part of their time following their own interests can be very helpful for the student. Similarly, I think I will prove a good example of how encouraging students who have slightly different backgrounds to come to the REU is a very good thing for the students. Although I did not write up my summer work for a journal, nor am I going to present it at a conference, the DIMACS REU still helped give me the time to learn about research and which field interested me, and I feel that I am now on the path to being a good researcher." Bryson Finklea, 2004

"I had met research before the REU, but never this close -- I had seen other people do research, but had not done it myself. It was quite interesting for me to see how you do mathematics, and feel the atmosphere of an American university as a whole. I appreciate the opportunity to meet some great mathematicians. Now I think about mathematics as about something that people do rather than something that can be found in books and never really understood." Eva Ondrackova, 2004

Recent Comments from Participants in Prior Years

"The REU was a great experience for me. I was at a liberal arts college (University of St Thomas in St Paul, MN) that didn't have any graduate programs in math and science, so the REU let me see what goes on at a major research university. After working on a research project and talking with graduate students and professors, I decided that graduate school in math was what I wanted to do. And here I am in graduate school for math." Sam Stechmann, 2002, (Sam is currently a graduate student in Mathematics at NYU.)

"The NSF REU's gave me a strong perception of the lifestyle of academics. The idea of doing 'original' research was profoundly introduced to me at these NSF REU's." Sabyasachi Guharay, 2003, (Sabyasachi is currently a graduate student at the Wharton School.)

"I enjoyed the DIMACS REU program. I think it was a great introduction on how to perform research. My enjoyment of the REU was a large factor in my decision to enter a Phd program." Steven Jaslar, 2003, (Steven is currently a graduate student in Applied Mathematics at Yale.)

"I was a leader of the Czech group two years ago. It was nice experience to have an opportunity to help organizing an event that has so much positive influence on everyone taking part. As for our result, we are still writing it down, as it turned out that the few missing cases in the proof are rather complex. But it will be finished no more than a year from now." Robert Samal, Charles University, Czech Republic, 2003

"REU reinforced my prior determination to pursue a scientific career. The REU program has vastly improved my ability to take part in a research project." Vit Jelinek, Charles University, Czech Republic, 2003

"I participated in the Research Experience for Undergraduates program 2004 (REU 2004) as a graduate student coordinator for the part of the program in Prague. In the past, I took part in the REU program as an undergraduate student in 2000 and 2001. The REU program was always an exciting event for me and after my first REU, I was always looking forward to the others. At the moment, it might be still too early to measure the exact impact of the program. I have realized the big impact of the program (from the time I took part in the program as an undergraduate student) on my future after several years: besides the first serious contact with research and new knowledge acquired during the program, I met an academic environment completely different from the one that I was used to. This broadened my view of the world of research. The participation in the program also confirmed my decision to continue in my studies to get my PhD. In addition to the academic impact, the program gave me a unique opportunity to learn something about the culture and the style of life in US, and I also met people that I would not have met otherwise. The program was one of the greatest events that I took part in during my life and I am sure that the undergraduate students who participated in the program this year will say the same. The big impact of the program on the professional career and the contact with multicultural environment have really been very important for me." Dan Kral, Ph.D.

V. Outreach Activities

A unique aspect of the program was its place within a vital and active research center with many other exciting programs, which were made available to our REU students. The 2004 program coincided with the DIMACS Bio-Math Connect Institute (BMCI), which provides research experiences for high school teachers and connects them with active researchers in bioinformatics and computational biology. The REU students interacted with both the high school teachers and researchers in BMCI.

The REU students also interacted with the DIMACS Reconnect Program, which is aimed at 2- and 4-year college faculty, and in particular aims to help them bring current research into the undergraduate program. Having a thriving REU program as an example was an important aspect of our summer Reconnect Program.

The students in the REU made connections with a variety of other DIMACS programs, including tutorials, workshops, and seminars, and we encouraged them to get involved with our industrial partners to expose them to industrial research as well as to our other partners such as the Institute for Advanced Study.

The REU students have been active in outreach activities following the summer program.

One of the Rutgers students, Tracy Grauman, spoke about her REU experience and her project at her FIGS (First-year Interest Group Seminar) that she taught Fall semester 2004 on the topic of Computer Science.

Tomas Valla (DIMATIA REU student) gave a talk to the participants of a computer science seminar for talented secondary school students, presenting his REU research results and some of the easier proofs. The seminar was held in Zaseka, Czech Republic, October 2004.

Shiri Azenkot (REU student from Pomona College) is hoping to give a talk at the Seventh Annual Nebraska Conference for Undergraduate Women in Mathematics, University of Nebraska, February 2005.

VI. Papers

A collection of realizable chains in the Weak Bruhat Order, Vishal Gupta (REU) and James Abello, in preparation.

A mathematical model of real-time PCR efficiency, Jana Gevertz (REU), Stanley Dunn, and Charles Roth, to be submitted to **Journal of Theoretical Biology**.

A winning strategy for the K_4 clique game, Vít Jelínek (REU), Jan Kára (REU), Tomas Valla (REU) and Robert Šámal (REU graduate coordinator), in preparation.

Geometric graphs with no three disjoint edges, Jakub Černý (REU), in preparation.

Improved upper bound on crossing number of polygons problem, Jakub Černý (REU), Zdeněk Dvořák (REU), Vít Jelínek (REU), and Pavel Podbrdský (REU), **Discrete Math.**, submitted.

Pancyclicity of strong products of graph, Daniel Král (REU), Jana Maxová, Pavel Podbrdský (REU), and Robert Sámal (REU), **Graphs and Combinatorics**, to appear.

On pattern coloring of cycle systems, Zdeněk Dvořák (REU), Jan Kára (REU), Daniel Král (REU), O. Pangrác, **Workshop on Graph Theoretical Concepts in Computer Science 2002 (WG'02)**, Lecture Notes Computer Science, Springer-Verlag, to appear. Also published as DIMACS Technical Report 2002-06.

On the number of intersections of polygons, Jakub Černý (REU), Jan Kára (REU), Daniel Král (REU), Pavel Podbrdský (REU), Robert Šámal (REU03 graduate student organizer), M. Sotakova, **Commentationes Mathematicae Universitatis Carolinae,** to appear.

Optimal protein encoding, E. Boros and L. Everett (REU), DIMACS Technical Report, DTR 2004-55, November 2004.

Stable matchings in three-sided systems with cyclic preferences, Endre Boros, Vladimir Gurvich, Steven Jaslar (REU) and Daniel Krasner (REU), **Discrete Math.,** vol. 289 (2004), pp. 1-10. Also published as DIMACS Technical Report 2002-34.

The Federalist Papers: A new approach to an old problem, Diana Michalek (REU), Ross T. Sowell (REU), and Paul B. Kantor, Rutgers University, in preparation and likely to be submitted to **Journal of American Society for Informatiton Science** or to **Information Retrieval**.

Three optimal algorithms for balls of three colors, Z. Dvořák (REU), V. Jelínek (REU), D. Král' (REU), J. Kynčl (REU), and M. Saks, DIMACS Technical Report: 2004-48. Also published as Institut teoretické informatiky (ITI) Series, Universita Karlova v Praze 2004-214 and KAM-DIMATIA Series 2004-693.

VII. Other Products

<u>Presentations</u>

One of the goals of our program was to emphasize student presentations of research. Students returning from our program made presentations in a variety of venues. For instance:

Zdeněk Dvořák presented "Three optimal algorithms for balls of three colors" at the Prague Eleventh Midsummer Combinatorics Conference in August 2004.

Zdeněk Dvořák, Vít Jelínek, Daniel Král', Jan Kynčl and Michael Saks will present "Three optimal algorithms for balls of three colors," at STACS'05 on February 24, 2005.

Dan Kral presented "Three optimal algorithms for balls of three colors" at the Combinatorial Optimization & Graph Algorithms, Seminar Technical University Berlin, Berlin, Germany on November 4, 2004.

Shiri Azenkot is hoping to give a talk at the Seventh Annual Nebraska Conference for Undergraduate Women in Mathematics, University of Nebraska, February 2005.

Tomas Valla gave a talk to the participants of a computer science seminar for talented secondary school students, presenting his REU research results and some of the easier proofs. The seminar was held in Zaseka, Czech Republic, October 2004.

Jana Gevertz gave a poster presentation at the Biomedical Engineering Society Annual Meeting, Philadelphia, Pennsylvania, October 2004.

Steven Jaslar gave a talk on his REU research on the Stable Marriage Problems at the Rutgers Research Day, sponsored by the General Honors Program, Spring, 2004.

Web Dissemination

http://dimacs.rutgers.edu/REU/

VIII. Contributions within Discipline

Most of the research results obtained were in the fields of discrete mathematics and theoretical computer science, broadly defined. These are areas in which DIMACS is very strong and has a world-wide reputation. Still, it is remarkable that undergraduates could, in at least some of the cases, do cutting edge research and solve some open problems.

The most important contribution to the disciplines of discrete math and theoretical computer science in particular, and to the discipline of the mathematical sciences broadly speaking (including all areas of mathematics, computer science, operations research, statistics, etc.), was the opening of new horizons to very bright undergraduates, some of whom we hope will become leaders in the field in the future.

IX. Contributions -- other Disciplines

Throughout the program, we emphasized applications of discrete mathematics and theoretical computer science. Probably the biggest impact in terms of applications came from topics at the interface between mathematics/computer science and biology. This includes the work of Logan Everett (REU) with mentor Endre Boros on optimal protein encoding, the work of Rutgers undergraduate Jana Gevertz (REU) with mentor Stanley Dunn (Biomedical Engineering) on gene expression levels, and the work of Andrew

Hodges (REU), with mentor Patrick DeLeenheer on HIV models. Bryson W. Finklea worked with mentor Wilma Olson. To understand how the genetic code is translated into three-dimensional structures used in biological processes, Professor Olson's group has been analyzing the known geometry of DNA bases in chemical structures and developing mathematical models to incorporate the sequence-dependent bending, twisting, and translation of known fragments in DNA molecules of varying lengths and chemical composition. The Finklea was directly involved in the group's research. In addition, there was outreach to other mathematical biology research at Rutgers through connections with DIMACS Special Focus on Computational and Mathematical Epidemiology and Special Focus on Computational Molecular Biology.

X. Contributions -- Human Resource Development

The impact on the careers of our students will take awhile to determine. However, we do have some self-assessment feedback from our recent survey of REU participants. Students who had only a general interest in research, now had more focused interests, paths they wanted to immediately pursue. Here are some of the comments we received:

“The REU was an amazing experience. I learned a lot in the specific area in which I researched, and about research in general. It motivated and inspired me to go to graduate school in computer science. The REU was also great fun. I especiallly enjoyed the trip to Prague--it was an amazing opportunity. Thank you for everything.” Shiri Azenkot, 2004

“Although I was considering graduate school before, this REU definitely cemented my plans and helped me pick a focus (Bioinformatics). My project definitely gave me a sense of how to apply novel mathematical solutions to real-world problems. Going to the conference in Prague, seeing the lectures at Rutgers, and just generally being around my peers that were Math majors instead of (or in addition to) CS majors has given me a much broader understanding of many topics in advanced mathematics.” Logan Everett, 2004

“I really enjoyed my REU experience and felt that it changed many of my perspectives on research and grad school. While I really enjoy school, I was previously convinced that I didn't really want to go on to do graduate work because I don't enjoy research all that much. I think my REU really changed that by being the first really positive experience I've had with research. It was a lot of fun. The colloquia we had also really broadened by understanding and interest. Many of the talks including the Knot Theory and the Graph Theory related talks really inspired my interest as well as the talks from the Lafayette students. I also had a wonderful time with the other students in the program as well as my advisor, the graduate students I worked with, and all of the DIMACS faculty and staff.” Dan Halperin, 2004

“The REU was a great experience! I don't think anyone can know that they want to go to graduate school and spend five years doing research until they've had some real research experience. Coming from a small liberal arts college where we have no graduate students except those in the seminary, it is difficult to know what life as a graduate student would be like. I had another summer research experience at Texas A&M University the previous summer, and that is when I first thought of going to graduate school and pursuing a Ph.D. in computer science. My summer experience at Rutgers certainly confirmed that initial intuition.” Ross T. Sowell, 2004

“My experience with DIMACS solidified my desires to go to graduate school. I had no clue what computer science research would entail, and although my summer project wasn't entirely fruitful, it gave me a lot of exposure to the field. Additionally, I was honored to be able to attend conferences both in Dagstuhl, Germany, and in Prague, Czech Republic, which gave me insight into various open problems in research today.” Tracy Grauman, 2004

XI. Contributions to Education and Human Resources

XII. Contributions Beyond Science and Engineering