D I M A C S

Center for Discrete Mathematics and Theoretical Computer Science

ANNUAL REPORT - December 1996

DIMACS Center • Rutgers University • P.O. Box 1179 • Piscataway, NJ 08855-1179 • USA
TEL: 908-445-5928 • FAX: 908-445-5932 • EMAIL: center@dimacs.rutgers.edu
WWW: http://dimacs.rutgers.edu/ • TELNET: telnet info.rutgers.edu 90

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs, Bellcore, and Bell Labs

Contents

1 Executive Summary

The period starting in September 1995 has been a period of dramatic change at DIMACS. There have been outstanding scientific developments and at the same time major changes in the center's view of itself and in the way it operates. In this report, we summarize the accomplishments at DIMACS in the period September 1, 1995 to January 30, 1997 under its National Science Foundation Science and Technology Center award.

The Purpose of DIMACS

During the past year, the center has adopted the following statement of purpose: DIMACS is a collaborative effort of Rutgers and Princeton Universities, AT&T Labs, Bell Labs and Bellcore and their scientific personnel, as well as their colleagues nationwide, working together to play a national leadership role in the development, application and dissemination of discrete mathematics (DM) and theoretical computer science (TCS) by

P1. promoting and supporting high quality research within these fields, especially research areas that are poised for rapid advancement and have a potential for high impact;

P2. promoting and supporting the application of the methods and results of these fields in other areas of science, the mathematical sciences, and industry;

P3. encouraging and facilitating efforts of researchers in these fields to draw more of their research inspiration from other areas of science, the mathematical sciences, and industry;

P4. informing the broader mathematical community about the research currently taking place in the fields of DM/TCS and involving them in these fields; and

P5. promoting and facilitating the incorporation of contemporary mathematical ideas and techniques in these fields into classrooms at all levels.

Special Years

The year beginning September 1995 saw a continuation of DIMACS' extremely successful Special Year in Mathematical Support for Molecular Biology, and it also featured an outstanding Special Year in Logic and Algorithms. The Molecular Biology year continued into its third year, starting in September 1996. That month also saw the beginning of a Special Year on Networks and a Special Focus on Discrete Probability, the latter in partnership with the Institute for Advanced Study, and a "prespecial year" in Massive Data Sets. These activities are described in Sections 4.1, 4.2, 4.3, 4.4, and 4.5. As have all previous special years, the 1995-96 special years led to major scientific advances, serious new partnerships and collaborations, and changes in scientific direction on the part of a variety of people. These results are described in Sections 5.1 and 5.2.

Educational Programs

In the area of K-12 education, we have continued our highly praised Leadership and Young Scholars Programs, and introduced some extremely promising innovations. We have taken the first steps to franchise the former by running satellite programs in Rhode Island and Virginia and have found a way to continue our programs for high school teachers even without specific funding for them. We have initiated a major effort aimed at identifying areas where DIMACS can make substantial contributions in undergraduate education, both locally and nationally. Meanwhile, we have continued our well-received REU program and our students during the summer of 1996 did excellent work. See Section 6.

DREI

We initiated a major new program in the summer of 1996, with the first year of our DIMACS Research and Education Institute, aimed at integrating research and education. This institute, funded by a $1.9M grant from NSF, and carried out in cooperation with the Institute for Advanced Study/Park City Mathematics Institute, was very successful in its first year. See Section 6.5.

Postdoctoral Programs

We once again had a terrific group of postdocs in 1995-96 and have another wonderful group in 1996-97. Our postdoctoral fellows have obtained many interesting and exciting scientific results, have contributed greatly to the life of the center, in particular to the success of its special years and also to its educational programs, and have gone on to find good positions, some in "non-standard" places. See Section 3.1 for detailed information about the work of our postdocs.

Other Scientific Programs

DIMACS continued its leadership role in experimental computing by continuing its series of innovative DIMACS "challenges." It also hosted the second Conference of African American Researchers in the Mathematical Sciences, an outstanding event. In addition, DIMACS sponsored a variety of workshops aimed at exploring hot topics, examining potential special focus themes, returning to past special focus topics, and establishing new partnerships. See Section 4.6.

Multi-disciplinary, Multi-institutional research projects

DIMACS has won awards for two major multi-disciplinary, multi-institutional projects on simulations of telecommunications systems, and continued its multi-institutional, NSF-supported LINK project, aimed at developing software for research and education in discrete mathematics. See Section 5.5.

Changes in Personnel

There have been major changes in personnel since September 1995. Fred Roberts became the Director of DIMACS in January 1996. Over the course of the year, Sarah Donnelly became the Center Administrator, Pat Pravato the Workshop Coordinator, Hangbiao Shi the Systems Administrator, and Robert Hochberg Educational Coordinator. These people were hired to replace people who had left. We also hired Elaine Foley as DREI administrator and to take charge of evaluation programs and publicity efforts at the center.

Strategic Planning

Beginning in October 1995, DIMACS entered a period of complete rethinking of what the center was about and what it was trying to accomplish and this led to the dramatic changes that we describe below. The intense period of self-examination and strategic planning resulted in a detailed statement of goals and purposes for the organization (of which the statement of purpose included above is just a very small piece), an examination of all DIMACS programs, and a connection of planned programs to the goals and purposes of the organization. The emphasis in the first stage of this planning activity was years 8-11 of our STC award, i.e., the period beginning February 1, 1996 until the end of the 11 year STC award on January 31, 2000. The DIMACS strategic plan that resulted has formed the basis for the changes that we have made.

Standards for Evaluation

Along with the strategic plan, we developed detailed standards for evaluation that give us specific ways of determining our progress toward our strategic goals. While perhaps overly ambitious, they are specific and we feel that we have made substantial progress on many of the criteria described in those standards.

Emphasis on Special Focus Programs

A central part of our strategic plan is to put the major emphasis of our research program on special focus activities. These are of durations varying from 6 months to 3 years, and are modeled after our special years. We will often run more than one such activity at a time - we are running three such programs during 1996-97. Our other major programs will be special educational activities.

Outreach

Another part of our strategic plan is that our programs should in almost all cases include a significant outreach component, involving outreach to another field of science, the mathematical sciences, or industry. Our molecular biology, logic and algorithms, and networks special years all had or are having significant industrial outreach components. Molecular biology and discrete probability involve major partnerships with other scientific communities, the latter with statistical mechanics. Discrete probability also has major outreach to another field of the mathematical sciences, namely probability theory.

Prespecial Years

We are running and will run programs that explore potentially exciting topics, prior to deciding upon them as special focus activities, or as preparation for the more intensive special focus activities. We call these prespecial years. For example, during 1996-97, we are planning activities in massive data sets, in particular workshops and tutorials, preliminary to our 1997-98 special year on that topic. (See Section 4.5.)

Continuing Activities

When a special focus or other program at DIMACS has been successful, we have decided that we will make sure that it leads to an activity that is self-standing, or return to it with occasional future programs, or develop it into a regular, continuing activity of the center. The continuation of molecular biology is a case in point. We returned to discrete and computational geometry, the topic of DIMACS' first special year, with an intense three-workshop program in the summer of 1996 as part of our DREI program, and will return to it again with a workshop in 1998. We will return to logic and algorithms with a followup workshop in 1997 and another in 1999.

National and International Scope

We are aiming to give DIMACS a more national and international scope and make the center more of a national resource. To make DIMACS more visible nationally, we have run a series of DIMACS sessions at national meetings of professional societies, run articles about DIMACS in professional science newspapers, solicited program suggestions from the community, produced a brochure about DIMACS that is appropriate for general audiences, and improved our WWW site. We have also participated in a program for state legislators in Trenton and at a display at the House Office Building, hosted Congressman Rodney Frelinghuysen and NSF Director Neal Lane at our center, and hosted many dignitaries at our DREI program, including Lane, two state assemblymen, and the head of the New Jersey Department of Education. (We did not use STC funds for political purposes.) We have initiated a variety of international collaborations with institutions in the Czech Republic, the Netherlands, and Israel, and in particular have entered into a very ambitious and promising partnership with the new DIMATIA center in the Czech Republic.

Integration of Research and Education

We are aiming to fully integrate our research and educational programs. We are asking special year organizers to plan on educational components (the logic and algorithms year, for example, sponsored a workshop on teaching of logic) and seeking to involve special year participants in educational activities. We have instituted a new policy of asking each of our postdocs to participate in either an educational or an outreach activity of the center. We have instituted the DREI program referred to above, and it has been very successful so far. More details about our integration efforts can be found in Section 6.4.

Graduate Student Involvement

DIMACS believes that graduate students are and should be a central part of a science and technology center. NSF's STC guidelines make this clear. Because our funds for local graduate student support were cut out by NSF effective in the Fall of 1996, we debated several alternatives for involving students. We are finding ways to involve local graduate students by supporting them on separate grants and by including them in programs that do not give them direct financial support. We feel that a program of strong involvement by local students, receiving financial support from the center so that they may be fully integrated into its programs such as its special focus programs, should be a major part of our game plan, and we are working on plans to make this happen. During the academic year 1996-97, we have expanded our program of visits by non-local students in connection with our special focus activities and expanded our budget for participation by non-local students in our workshops.

"Departmental Allocations"

DIMACS used to give to each academic unit (the Princeton CS Department, and the Rutgers CS, Math, and O.R. units) an allocation for visitors. New allocations have been discontinued. Where unexpended commitments exist, we have reallocated the funds to the DIMACS visitor committee, which has representatives from each DIMACS unit and makes centralized decisions about visitor requests that are intended to reflect the major goals and purposes of the center.

Organizational Members

Because of the split of AT&T, our organizational member, AT&T Bell Labs, has split into two organizational members, AT&T Labs and Bell Labs. We are now operating as a five-organization center, with Rutgers, Princeton, and Bellcore being the other three members. (The transition is still ongoing. The names keep changing, from AT&T Research to AT&T Labs, for instance, and from Bell Labs/Lucent Technologies to Bell Labs.) Support for DIMACS-related activities at our industrial partners has increased substantially over the past year, and the commitment of all of our partners toward DIMACS activities continues to be very strong in spite of down-sizing pressures and difficult budgets. More details about institutional support for DIMACS can be found in Section 2.

New Jersey Commission on Science and Technology

The Commission has been phasing out support of its centers. However, its support of DIMACS has continued at a level amount. More details about this can be found in Section 2.

Institute for Advanced Study Program

DIMACS and IAS are partners in a variety of joint programs, including a shared postdoctoral program, the special focus in Discrete Probability, a visitor program, and the DIMACS Research and Education Institute. IAS' computer science and combinatorics program was "created to be complementary to and supportive of the work of DIMACS."

Affiliated Organizations

In order to increase our interactions with other groups, to make DIMACS more of an outward-looking institution, and to add new excitement to our programs, we have been actively engaging new partners for our many programs. We have defined a new category of organizational affiliation, an affiliated organization, which is an organization that agrees to share in a joint program with DIMACS over a period of years. We have been talking to a variety of organizations about this kind of affiliation, and anticipate adding several formal affiliated organizations.

New Management Structures

We have streamlined the executive committee and made this smaller group the main decisionmaking body that advises the director; made the larger council an advisory body; completely redesigned our committee structure and implemented an active committee system, with each committee having representatives from all DIMACS units; and developed DIMACS' first set of bylaws, with an emphasis on a "strong director" model.

Planning for After Year 11 of the STC Award

After completing our major strategic plan for years 8-11, we began to plan for a DIMACS beyond year 11. This type of planning is very difficult, given the uncertainties about future funding of science, budgets at universities and companies, and about NSF plans for a recompetition for centers. The process of planning for a DIMACS beyond year 11 has led to a detailed draft of a long term plan. The member organizations are enthusiastic about having the center continue after the 11 year STC award ends. We would certainly like to compete if there is another STC competition. If there is not, or if we do not succeed in this competition, we are committed to seeking alternative ways to continue the center in the future, and we have obtained commitments from our five member organizations that should allow us to continue in operation at a relatively modest level at the end of year 11, even without external funds, and to establish a base from which we can seek significant external support for our programs. Given the strength of DIMACS, the active involvement of so many world-class researchers, and the dramatic scientific achievements that have come out of our center, our long term plan is to seek funding that will let us continue our operations at a high level. We feel that we should be able to present a strong case for continued long term funding at a substantial level, and we are both optimistic and ambitious with respect to this goal.

Overview of the Report

The rest of this report is organized as follows. In Section 2, we summarize the financial situation at DIMACS. Section 3 describes the work of the scientific personnel who were funded by DIMACS, in particular the postdoctoral and graduate student fellows. We also list all the visitors to the center in the period covered by this report. DIMACS has some 150 permanent members (see

Appendix D), almost all from its five organizational members plus NEC Research, but none of them received salaries from the NSF STC award. In Section 4, we describe our scientific programs. We give information about our special focus activities – their themes, their organizers, and special year workshops. We also describe other workshops, conferences, and scientific programs at DIMACS. In Section 5, we give highlights of scientific achievements at the center, starting with the recent special years. We also mention recent results that can be traced back to earlier special focus activities at DIMACS, and then give a selection of the exciting recent results of our postdocs, students, and visitors. Finally, Section 6 describes our educational programs and our efforts to integrate research and education.

ABBREVIATIONS USED THROUGHOUT:

ATTL = AT&T Labs

BC = Bellcore

BL = Bell Labs

DM = discrete mathematics

GS = graduate student

IAS = Institute for Advanced Study

PD = postdoc

PM = permanent member

PU = Princeton University

RU = Rutgers University

SF = special focus

STC = science and technology center

SY = special year

TCS = theoretical computer science

V = visitor

2 Financial Report

With the addition of Sarah Donnelly as the DIMACS Center Administrator, the center has added one of the most talented and experienced budget people at the university. Thanks to her expertise, DIMACS has made significant improvements in its budget management. She and the director have worked very closely to develop procedures to track commitments, project center expenditures and revenues, and develop budgetary flexibility. In this section, we summarize the center's finances, discussing its STC budget, funding from its institutional partners, the role of the New Jersey Commission on Science and Technology and the Institute for Advanced Study, and the center's external funding.

The STC Budget

The STC award from NSF was $2,445,000 in year 7 (2/1/95-1/31/96). This included a one-time increase for equipment upgrades and an REU supplement. Because of the phase-out of $300,000 for graduate student support over a two year period, the year 8 budget is $2,110,000 and the year 9 budget is anticipated to be $1,960,000. The year 10 budget is projected to be 83% of year 9, and year 11 to be 64% of year 9. This is part of the built-in winding down period for the STC's.

We also had available an allowed carryover of $674,000 from earlier years, which we are committed to spend in years 8 and 9. Our budget plan for year 8, including an anticipated $25,000 REU supplement and $406,500 from the carryover, called for spending $2,541,500. Year 9 calls for spending $2,227,500.

Institutional Support from Partner Organizations

DIMACS receives support from its five institutional partners. Rutgers provides the director's salary, salary for part of a secretary, space, and operating and maintenance expenses. Rutgers also returns to DIMACS a significant portion of the indirect costs on the main STC award and other grants, amounts between 40% and 50% of the overhead. Princeton provides half of a postdoc, graduate student tuitions in the years when we support students, secretarial support, and space. Our industrial partners, AT&T Labs, Bell Labs, and Bellcore, each provide workshop support, appoint DIMACS postdocs and contribute to the cost of DIMACS visitors. For the one year period beginning 9/1/96, the total amount of this support is estimated to be $1,690,649.

All five institutions strongly encourage volunteer work by their scientists, and the number of hours spent on DIMACS committee work and organizational activities by their scientists is very large. For example, over 40 permanent members of DIMACS, representing all of the five member organizations, are actively involved in the committees that run the center and its programs, the Executive Committee, Council, Strategic Planning Committee, Budget Committee, Projects Committee, Education Committee, Postdoc Committee, Equipment Committee, Graduate Student Selection Committee, REU Committee. More are involved in an organizational role in our many special focus programs, workshops, and educational projects.

New Jersey Commission on Science and Technology

Since the start of DIMACS, with the exception of one year, the center has been receiving an annual award of $100,000 from the New Jersey Commission on Science and Technology. Recently, the Commission made the decision to phase out its support of centers over a three year period, replacing the centers program by a program of competitive awards for research projects. Most of the centers supported by the state received major cuts for the fiscal year beginning July 1, 1996 (FY97), and some were cut off entirely. However, recognizing that the State had some obligation to continue to support DIMACS, the Commission did not cut our center and we once again received $100,000 for the present fiscal year. For FY98, the Commission's Budget Committee is recommending that the Commission's support of centers be cut to 50% of its FY97 level. However, it is recommending that the support for DIMACS be continued in full. While the amount of Commission support is not large, we consider it significant given the current situation.

Institute for Advanced Study Program

The Institute for Advanced Study is sharing in funding several joint projects, including our joint postdoc program and our joint special focus in discrete probability. The cost of the IAS computer science and combinatorics program, which is "created to be complementary to and supportive of the work of DIMACS," is approximately $270K per year, not including administration, secretarial support, Institute overhead, or Institute faculty salaries. IAS is also a partner with DIMACS in the DIMACS Institute for Research and Education (DREI), an externally-funded program that integrates research and education and is described in detail in Section 6.5.

External Funding

DIMACS has external funding for a variety of its special focus activities, group research activities, and educational activities. The largest sources of such funds in currently active accounts are as follows:

- (1). a grant of $1.9M over four years from NSF for our DIMACS Research and Education Institute;
- (2). a grant of approximately $1.5M over three years from NSF for a multi-disciplinary, multi-institutional project on Simulations of Integrated Communications Systems;
- (3). a grant of $3.1M over three years from DARPA for our project on Scalable Self-Organizing Simulations;
- (4). a multi-institutional grant of $350K over three years from NSF for the development of LINK, a software system for research and education in discrete mathematics;
- (5). a grant of $2.6M over four years from the Education and Human Resources Directorate of NSF for support of the Leadership Program in Discrete Mathematics;
- (6). a grant of $196K over two years from the Education and Human Resources Directorate of NSF for support of the Young Scholars Program in Discrete Mathematics;
- (7). grants from NSF, NSA, CDC, National Institute for Allergy and Infectious Diseases (NIAID), the National Center for Human Genome Research (NCHGR), and support from Roche Molecular Systems, SmithKline Beecham, BIOSYM Tech, Merck Research, Ciba-Geigy, and Sandia, totaling $625K over three years, for support of our molecular biology activities;
- (8). funds from ONR for approximately $105K over three years to support a senior visitor in a project on binary hierarchy (clustering);
- (9). funds from ONR for approximately $123K over three years to support a graduate student in a project on measurement and clustering.

3 Postdoctoral Fellows, Graduate Student Fellows, and Visitors

In this section, we describe the scientific work of the postdoctoral fellows and graduate student fellows who were supported by DIMACS, and we list the visitors to the Center. We include postdocs in both academic years covered by this report, 1995-96 and 1996-97, and the graduate students in academic year 1995-96. There were no graduate students supported on the STC grant in 1996-97. DIMACS has some 150 permanent members (see Appendix D), almost all from its five organizations, plus NEC Research. None of the permanent members were supported on the STC grant during the period covered by this report. However, many of them were actively involved in DIMACS as organizers of programs, serving on committees, and in other jobs. Moreover, many of them collaborated with our postdocs, students, and visitors. Some of the scientific results coming out of these collaborations are reported on in this section and others in Section 5, when we highlight the recent scientific accomplishments at the center.

3.1 Postdoctoral Fellows: September 1995 - January 1997

Richa Agarwala (1994-96) (SY in Molecular Biology)

Currently Systems Analyst, R. O. W. Sciences, Inc. (on a full-time subcontract to NIH)

Much of Agarwala's research has focused on the problem of devising efficient algorithms for inferring the evolutionary relationships of a set of species, a fundamental problem in computational biology. Rooted trees called *phylogenies* describe the evolution of a set of species S, with the leaves representing the species in S and the internal nodes representing hypothetical ancestral species. The underlying assumption is that species have evolved from a common ancestor, and that the diversity of life is a result of the change and divergence through time. The difficulty in carrying out the construction stems from the fact that we do not have the information about all the ancestral species. The problem of finding phylogenies, while interesting in its own right, is also of both theoretical and practical value in biology, geology, linguistics, etc. Phylogenies are used for taxonomic classifications of species and to evaluate alternative hypotheses of adaptation, evolutionary mechanisms, and ancient geographical relationships. In one of the standard models, the species are described by an $n \times m$ matrix M, where n is the number of species, m is the number of characters, and M_{ij} is the state exhibited by the jth character of the ith species. The problem of finding a *perfect phylogeny* is to define a phylogeny and character states for the internal nodes in such a way that the set of species having any one character state form a connected subgraph. Collaborating with David Fernandez-Baca (V), Agarwala has found an $O((r-n/m)^m rnm)$ algorithm for determining whether a set of n species has a perfect phylogeny, where m is the number of characters used to describe a species and r is the maximum number of states that a character can be in. The perfect phylogeny algorithm leads to an $O((2e/k)^k e^2 k)$ algorithm for triangulating a k-colored graph having e edges. Agarwala also worked on the problem of fitting trees to distance metrics. Vineet Bafna (PD), Martin Farach (PM, RU, and assistant chair of the special year in molecular biology), Babu Narayanan (PD), Mike Paterson (V), Mikkel Thorup (V) and Agarwala gave the first approximation algorithm for this problem. Later, Farach and Sampath Kannan (V) used this work as a building block in a clever algorithm for computing maximum likelihood trees, which is likely to have a great impact on the field.

Agarwala has also developed novel approaches to different aspects of optimization theory and its applications, and in particular efficient methods to solve what are known as Lagrangian relaxation problems. Relaxation approaches are used to estimate the value of the optimum solution to difficult problems. As the name suggests, the idea is to "relax" the problem constraints enough so that the resulting problem is easier to solve, but not so much that the result obtained from the relaxed problem is too far from the true optimum. Lagrangian relaxation has proved to be an effective approach to hard optimization problems, since it tends to produce better estimates than other relaxation methods. Working with David Fernandez-Baca (V), Agarwala has developed a new approach to an important class of relaxation problems that is significantly faster than the algorithms that have been proposed to date. In particular, she developed a weighted version of Megiddo's multidimensional search technique. This speeds up algorithms for certain convex optimization problems in the plane.

Papers

R. Agarwala and D. Fernandez-Baca. Fast and simple algorithms for perfect phylogeny and triangulating colored graphs. Int. J. of Foundations of Computer Science, 7 (1996), 11-21.

R. Agarwala and D. Fernandez-Baca. Weighted search in the plane. Information Processing Letters, in press.

R. Agarwala, V. Bafna, M. Farach, B. Naryanan, M. Paterson, M. Thorup. On the approximability of numerical taxonomy. To appear in SIAM J. on Computing.

Jeremy Avigad (Spring 1996) (SY in Logic and Algorithms)
Currently Assistant Professor, Carnegie Mellon University

Together with Richard Sommer, Avigad developed a model-theoretic approach to ordinal analysis via the finite combinatorial notion of an α-large set of natural numbers. Using these methods, they provided sharp upper bounds on the proof-theoretic strength of a number of important theories. Avigad also developed a simple propositional proof system which consists of a single axiom schema and a single rule, and used this system to construct a sequence of combinatorial tautologies that, when added to any Frege system, p-simulates extended-Frege systems. Solovay has shown that if O is an open subset of $P(\omega)$ with code S and no infinite set avoids O, then there is an infinite set hyperarithmetic in S that lands in O. Avigad provided a direct proof of this theorem that is easily formalizable in ATR_0

Papers

J. Avigad. Open sets are Ramsey. Submitted to Archive for Mathematical Logic.

J. Avigad. Plausibly hard combinatorial tautologies. Submitted to Proceedings of DIMACS Workshop on Proof Complexity.

J. Avigad and R. Sommer. The model-theoretic ordinal analysis of predicative theories. To appear in J. Symbolic Logic.

Talks

"Model-theoretic Ordinal Analysis." 1996 ASL Spring Meeting, Madison, Wisconsin.

"Plausibly Hard Tautologies." DIMACS Workshop on Proof Complexity.

"Proof Theory and Subsystems of Second-order Arithmetic." Logic seminar at University of Pennsylvania.

A number of talks at the Rutgers Logic seminar.

Vineet Bafna (1994-96) (SY in Molecular Biology)
Currently Senior Computational Biologist, SmithKline Beecham Pharmaceuticals

When Bafna began at DIMACS, he was mainly interested in two orthogonal areas, approximation algorithms and algorithms for molecular biology, particularly in molecular evolution. A fundamental problem in evolution is that of measuring the amount of evolutionary divergence between two species, often computed by looking at their molecular sequences. His thesis work was centered around looking at entire genomes, and estimating divergence by measuring global genome rearrangements, which has a natural formulation as a combinatorial optimization problem. Together with Pavel Pevzner (V), his advisor and a visitor in the special year in molecular biology, Bafna gave approximation algorithms for this problem. They continued this collaboration at DIMACS, working with different types of rearrangements.

At DIMACS, Bafna also started collaborating with other postdocs, R. Ravi (PD), Babu Narayanan (PD) and S. Muthukrishnan (PD). With Narayanan and Ravi, he considered the problem of extracting a set of high scoring local alignments in the presence of repeats and rearrangements. Combinatorially, the problem translates to finding independent sets in claw-free graphs, and they analyzed the use of local-change heuristics for this problem. Bafna also looked at the problem of computing similarity between RNA sequences with Ravi and Muthukrishnan. They looked at many different similarity measures between RNA strings, as well as the problem of predicting structure by homology search. Their algorithms for these problems were based on techniques from string matching and dynamic programming techniques.

Bafna's interests shifted during the special year, from computing evolutionary divergence between pairs of species to computing evolutionary trees for many different species. He worked with Martin Farach (PM, RU) who has, among other things, worked on many different combinatorial problems in evolutionary tree reconstruction. The problem they considered was a long-standing one, of fitting trees to distance metrics. Richa Agarwala (PD), Farach (PM, RU), Narayanan (PD), Mike Paterson (V), Mikkel Thorup (V) and Bafna gave the first approximation algorithm for this problem. Later, Farach and Sampath Kannan (V) used this work as a building block in a clever algorithm for computing maximum likelihood trees. It can be expected that this will have a great impact on the field.

In work not related to molecular biology, Bafna collaborated with Piotr Berman and Toshihiro Fujito on approximations of feedback vertex sets in weighted undirected graphs. A *feedback vertex set* of a graph is a subset of vertices that contains at least one vertex from every cycle in the graph. Bafna showed that a feedback vertex set approximating a minimum one within a constant factor can be efficiently found in undirected graphs. In fact, the derived approximation ratio matches the best constant ratio known today for the vertex cover problem, improving the previous best results for both weighted and unweighted cases. The existence of an approximation preserving reduction of the vertex cover problem to the feedback vertex set problem suggests that further improvement of the obtained ratio will require entirely new ideas. Bafna, Berman, and Fujito extended their approach to handle graphs of bounded vertex degree, and showed an improved performance ratio for this case. These approximation bounds were obtained using a nontrivial generalization of the classical local ratio principle of Bar-Yehuda and Even.

Papers

V. Bafna, B. Narayanan, and R. Ravi. Non-overlapping local alignments. To appear in Discrete Applied Mathematics.

V. Bafna, P. Berman, and T. Fujito. Constant ratio approximations of feedback vertex sets in weighted undirected graphs. Submitted.

V. Bafna, S. Muthukrishnan, and R. Ravi. Computing similarity between RNA strings. Submitted.

R. Agarwala, V. Bafna, M. Farach, B. Naryanan, M. Paterson, M. Thorup. On the approximability of numerical taxonomy. To appear in SIAM J. on Computing.

Talks

V. Bafna and P. Pevzner. "Sorting Permutations by Transpositions." Symposium on Discrete Algorithms (SODA95).

V. Bafna, B. Narayanan, and R. Ravi. "Non-overlapping Local Alignments." Workshop for Algorithms and Data Structures (WADS95).

V. Bafna, S. Muthukrishnan, and R. Ravi. "Computing Similarity between RNA Strings." Combinatorial Pattern Matching Conference (CPM95).

R. Agarwala, V. Bafna, M. Farach, B. Naryanan, M. Paterson, M. Thorup. "On the Approximability of Numerical Taxonomy." SODA96.

Robert Beals (1994-96) (NSF postdoc)
Currently Assistant Professor, University of Arizona

Together with L. Babai, J-y. Cai, G. Ivanyos, and E.M. Luks, Beals has considered the solvability of the equation

$$\prod_{i=1}^{k} A_i^{x_i} = B$$

and generalizations, where the A_i and B are given commuting matrices over an algebraic number field F. In the *semigroup membership problem*, the variables x_i are constrained to be nonnegative integers. While this problem is NP-complete for variable k, they gave a polynomial time algorithm if k is fixed. In the *group membership problem*, the matrices are assumed to be invertible, and the variables x_i may take on negative values. In this case they gave a polynomial time algorithm for variable k and give an explicit description of the set of all solutions (as an affine lattice). Tits has shown that a finitely generated linear group either contains a non-abelian free group or has a solvable subgroup of finite index. Beals gave a polynomial time algorithm for deciding which of these two conditions holds for a given finitely generated matrix group G over an algebraic number field. There exist matrix groups (even 4×4 integer matrix groups) for which the membership problem is undecidable. These groups all contain nonabelian free subgroups, so the Tits alternative seems to separate tractable groups from intractable groups. Beals showed that for groups with solvable subgroups of finite index, problems such as membership testing are decidable, and some important subcases have polynomial time algorithms.

Together with R. Chang, W. Gasarch, and J. Toran, Beals has investigated the enumerability of the function SGA, the number of automorphisms of an undirected graph, in relation to the computational complexity of the Graph Isomorphism problem.

Beals has also considered Fourier transforms over the symmetric groups in the context of the theory of quantum mechanical computers. Many recent developments in quantum complexity theory, including Shor's celebrated algorithms for factoring and discrete logs, have made use of Fourier transforms over abelian groups. That is, at some point in the computation, the machine is in a superposition of states corresponding to elements of a finite abelian group G, and in quantum polynomial time (i.e., polynomial in $\log |G|$), the machine is transformed according to the Fourier transform to a superposition of states corresponding to the irreducible representations of G. Beals gave a quantum polynomial time algorithm for the Fourier transform for the symmetric groups S_n.

E. Allender (PM, RU), Beals, and M. Ogihara have characterized the complexity of some natural and important problems in linear algebra. In particular, they have identified natural complexity classes for which the problems of (a) determining if a system of linear equations is feasible and (b) computing the rank of an integer matrix, (as well as other problems), are complete under logspace reductions.

In several pieces of work, alone and in collaborations with K. Tanaka and T. Nishino, Beals has investigated the effect of limiting the number of negation gates in a boolean circuit. A theorem of Markov precisely determines the number r of NEGATION gates necessary and sufficient to compute a system of boolean functions F. For a system of boolean functions on n variables, $r \leq \lceil \log_2(n+1) \rceil$. We call a circuit using the minimum number of NEGATION gates *negation-limited*. This work continues recent investigations into negation-limited circuit complexity, giving both upper and lower bounds. A circuit with inputs $x_1, \ldots, x_n$ and outputs $\neg x_1, \ldots, \neg x_n$ is called an *inverter*, for which $r = \lceil \log_2(n+1) \rceil$. Fischer has constructed negation-limited inverters of size $O(n^2 \log n)$ and depth $O(\log n)$. Recently, Tanaka and Nishino have reduced the circuit size to $O(n \log^2 n)$ at the expense of increasing the depth to $\log^2 n$. Beals, et al. have constructed negation-limited inverters of size $O(n \log n)$, with depth only $O(\log n)$, and they conjecture that this is optimal. Next, they have obtained a number of lower bounds on the size and depth of negation-limited circuits computing several symmetric functions. Furthermore, they showed nonlinear lower bounds on the size of certain kinds of negation-limited circuits computing symmetric functions.

Beals has considered the size of algebraic decision trees. Two models of algebraic decision trees have been studied. The *binary* model allows queries of the form "is $p(x_1, \ldots, x_n) \leq 0$?", while in the *ternary* model, the value of the query polynomial is compared to zero and one of three possible answers $(<, =, >)$ is returned, each answer leading to a distinct subtree. It is straightforward to simulate a ternary tree with a binary tree at a cost of doubling the size and depth of the tree. However, no simulation in the other direction was previously known. Indeed, Grigoriev, Karpinski, and Yao have recently shown that if the degree of the query polynomials is bounded by a constant, then there exist binary decision trees for which such a simulation necessarily entails an exponential increase in size. Beals has shown that, contrary to the conjecture of Grigoriev, Karpinski, and Yao, a binary algebraic decision tree T can be converted to a ternary algebraic decision tree T' at the expense of a quadratic increase in size and a linear increase in depth.

Papers

L. Babai, R. Beals, J-y. Cai, G. Ivanyos, and E. M. Luks. Multiplicative equations over commuting matrices. Proc. 7th ACM-SIAM SODA (1996), pp. 498–507.

R. Beals. Algorithms for matrix groups and the Tits alternative. Proc. 36th IEEE FOCS (1995), pp. 593–602.

R. Beals, R. Chang, W. Gasarch, J. Torán. On finding the number of graph automorphisms. Submitted for publication.

R. Beals. Quantum computation of Fourier transforms over the symmetric groups. In preparation.

E. Allender, R. Beals, M. Ogihara. The complexity of matrix rank and feasible systems of linear equations. To appear in STOC 1996.

R. Beals. Improved construction of negation-limited circuits. DIMACS Tech. Report 95-31.

R. Beals, K. Tanaka, T. Nishino. More on the complexity of negation-limited circuits. Proc. 27th ACM STOC (1995), pp. 585–595

T. Nishino, K. Tanaka, R. Beals. Negation limited circuit complexity of symmetric functions. Submitted for publication.

R. Beals. Equivalence of binary and ternary algebraic decision trees. Algorithmica. To appear.

Amos Beimel (1996-97) (Special Year in Networks)

The main subject of Beimel's research is cryptographic tools that are information theoretically secure (namely without any complexity theoretic assumptions). He has considered secret sharing schemes and key distribution schemes. Both schemes are carried on a multiparty system in which the parties can communicate via a communication network. In his work he has analyzed the effect of communication and the security assumption of the communication channels on the efficiency of the schemes. That is, he has considered secure private channels and broadcast channels. Key distribution schemes enable every subset of parties to generate a secret key (different subsets have different keys). These keys can be used in private key crypto systems. Secret sharing schemes enable a dealer to distribute a given secret such that only some predefined sets of parties can reconstruct the secret. These schemes allow the storing of secret information in a network, such that only "good" (for example, large enough) subsets can reconstruct the information. Furthermore, by using these schemes we can allow only "good" subsets to perform actions in a system (e.g., sign a check). One of the most important issues when designing schemes in the information theoretic model is the size of the shares - the pieces of information given to the parties. (The size of the shares is the size of the binary strings that represent them.) If the size of the shares is too big then the parties will not have enough memory to store their shares (leaving aside the question of secure storage). In his research, Beimel studies the space efficiency of key distribution schemes, secret sharing schemes and threshold schemes with additional properties. Beimel is working on communication in key distribution schemes; secret sharing with public reconstruction; computing functions of a shared secret; lower bounds for linear secret sharing schemes; and universally ideal secret sharing schemes. In the secret sharing schemes that have been considered so far, the security is guaranteed even if the "bad" parties have unlimited power. In the general secret sharing schemes, where an arbitrary collection of subsets of parties should be able to reconstruct the secret, the known information theoretic schemes are not practical since the size of the shares are too big. However, in modern cryptography the common assumption is that the "bad" parties are computationally limited, i.e. they run in (probabilistic) polynomial time. Beimel is seeking to construct efficient computational schemes for collections that do not have efficient information-theoretic secret sharing schemes. That is, he would like to construct practical schemes for more collections. For most collections of reconstructing sets, the size of the shares in the most efficient schemes known to date is exponential in the number of parties in the system. That is, even in small systems the parties will not have enough memory to store the shares. The question of whether there exist more efficient schemes, or if this exponential blowup is unavoidable, is open. Beimel conjectures that the latter is true, and is trying to prove this conjecture.

Jonathan Berry (1995-96) (LINK Project)
Currently Assistant Professor, Elon College

Berry worked on the LINK software, a combinatorics and graph theory workbench for applications and research. Before Berry arrived, LINK, which was developed through a multi-institutional collaboration and supported by NSF, had been developed by Nate Dean (PM, BC and BL), Steve Skiena (V, Stony Brook), Mark Goldberg (V, RPI), and Greg Shannon (V, Indiana and Los Alamos and Spanning Tree Technologies). Berry made major progress on getting LINK in position to make a substantial impact on education in discrete mathematics. LINK is a set of C++ class libraries that supports applications in discrete mathematics. The libraries include a commandline interpreter and a graphical user interface that allow access to basic data structures such as Sets and Lists, and a graph hierarchy that includes undirected, directed, and "mixed" graphs. Many standard data structures including arrays, lists, heaps, and binary search trees are within a container hierarchy. Sets and sequences are supported within a collection hierarchy. The data structure hierarchies enable the user to experiment with competing data structure implementations, and with more complex and sophisticated data structures. If an algorithm has several possible choices of a data structure to be used, a single object can be created that is templated with the particular data structure desired. LINK also contains a set of graph generators, layout algorithms for hypergraphs and binary graphs, and numerous graph algorithms. Interactive visualization of hypergraphs, graphs, and subgraphs is included in LINKGUI, the research tool application for combinatorics and graph theory built on top of the LINK libraries. LINKGUI is a collection of libraries that includes a Motif-based graphical user interface and a Tcl-based command-line interface. The ability to select, contract and expand subgraphs and nested subgraphs in either hypergraphs or binary graphs is included in LINKGUI. LINKGUI enables a user to perform unary operations, such as complement, on a particular graph or to combine graphs via binary operations such as graph union or intersection. The resulting graphs can be displayed in multiple windows. Algorithms can be run on graphs where subgraphs have been contracted into single vertices. Other LINK applications can be developed using various parts of the library code and adding code specific to the application. This new code can include an entirely different interface from that supplied by LINKGUI. LINK is now available for download by anonymous FTP from the DIMACS home page.

In other work, Berry and Mark Goldberg (V) have developed a new heuristic for graph partitioning called Path Optimization, and have made an extensive set of empirical comparisons of the new algorithm with two very well known algorithms for partitioning: The Kernighan-Lin algorithm and simulated annealing. The results for geometric graphs, which have become a frequently-used benchmark in the evaluation of partitioning algorithms, show that Path Optimization holds an advantage over the others.

Papers

J. Berry, N. Dean, P. Fasel, M. Goldberg, E. Johnson, J.MacCuish, G. Shannon, and S. Skiena. LINK: A combinatorics and graph theory workbench for applications and research. DIMACS Tech Report 95-15.

J. Berry and M. Goldberg. Path optimization for graph partitioning problems. DIMACS Tech Report 95-34.

Maria Bonet (First Half of 1995-96) (SY in Logic and Algorithms)

Currently Assistant Professor, Universitat Politecnica De Catalunya

Bonet has worked in logic and complexity, on the topic of length of proofs. Her work has been concerned with studying the strength of various deduction Frege systems in terms of counting the number of lines. A *deduction Frege system* is a Frege system that has one more rule, the deduction rule. The simplest form of this rule is: If a formula B has a proof from an additional, extra-logical hypothesis A, then there is a proof of $A \to B$. To have a deduction rule in a Frege system makes the proofs much easier to build. Bonet, working with Samuel Buss (V), has answered the question: Does some form of deduction also make proofs shorter? The notion of shorter can be measured in terms of the number of lines or number of symbols. She has proved that for nested deduction Frege systems, the nested deduction rule gives a speed up of at most the inverse of the Ackerman function. In the process of simulating nested deduction, Bonet and Buss defined a new type of transitive closure problem on trees. An instance of this problem is a tree with m edges and a list of n closure edges. They proved that a solution to this problem will require at most $O((n+m)\alpha(m))$ steps, where $\alpha(m)$ is the Ackerman function. Working with Buss (V) and Tonian Pitassi (V), Bonet investigated the problem of finding natural combinatorial tautologies which separate Frege and Extended Frege, i.e., that have short Extended Frege proofs but require long Frege proofs. The first tautology to be proposed to lead to such separations is the pigeon-hole formula, which is known to have short Extended Frege proofs. Buss showed that it also has polynomial size Frege proofs. Later on, Krajicek proposed Bondy's principle as a tautology that might require exponential size Frege proofs. Bondy's theorem states that, in any $n \times n$ matrix containing n distinct rows, there exists a column such that, if the column is deleted, the resulting $(n-1) \times n$ matrix still has n distinct rows. Bonet, Buss, and Pitassi showed that Bondy's principle has polynomial size Frege proofs.

Bonet has also started working in computational biology and has worked on constructing evolutionary trees in the presence of polymorphic characters. Determining the evolutionary history of a set S of objects (taxa or species) is a problem with applications in a number of domains, including biology, comparative linguistics, and literature. Primary data used to compare different taxa can be described using *characters*, where a character is a function $c : S \rightarrow Z$, where Z denotes the integers and thus represents the set of possible states of c. Together with Cynthia Phillips, Tandy Warnow (V), and Shibu Yooseph, Bonet has considered tree construction when characters are permitted to have more than one state on a given object. This is the *polymorphism problem.* A character which is permitted to have more than one state on a given object is called a polymorphic character and one which can have only one state for every object is called a monomorphic character. Bonet and her collaborators have found efficient algorithms for the construction of perfect phylogenies from polymorphic data. These methods have been used to help construct the evolutionary tree proposed by Warnow, Ringe, and Taylor for the Indo-European family of languages, which was presented by invitation at the National Academy of Sciences in November 1995.

Papers

M.L. Bonet, S.R. Buss, and T. Pitassi, Hard examples for Frege proof systems. Feasible Mathematics II, to appear.

M.L. Bonet and S.R. Buss, On the serial transitive closure problem. SIAM J. on Computing, to appear.

M.L. Bonet, C. Phillips, T. Warnow, and S. Yooseph. Constructing evolutionary trees in the presence of polymorphic characters. Proceedings of the ACM Symposium on the Theory of Computing (STOC), 1996.

Hui Chen (1996-97) (SF in Discrete Probability)

Chen's special emphasis is on probabilistic methods and on molecular biology. He is working on the threshold of broadcast trees in random graphs. Frieze and Molloy showed that a random graph $G_{n,p}$ with probability at least 1 - $o(1)$ has an optimal broadcast tree when $p \geq 18 \ln 2 \log_2 n/n$. A trivial lower bound is $\log_2 n/n$. It is conjectured that a threshold exists. Chen has estimates for lower and upper bounds.

In molecular biology, Chen is working in hybridization using unique probes. He is considering a probabilistic model due to Alizadeh, Karp, Weisser, and Zweig, for the physical mapping of chromosomes. Given measured data D, assuming there is no noise in the data, there is a polynomial time algorithm which will give all the possible orderings that comply with D. When we assume that the occurrence of error in the measured data is unavoidable, then this problem becomes NP-hard. Chen is working on the problem of deciding a threshold value for the probabilities of the error occurrence below which we can still reconstruct the correct ordering with high probability. Chen is also working on introns, extrons and genes. Dyer, Frieze, and Suen show a threshold for the existence of a unique ordering or clone library using Poisson probes. A more realistic model is that the clone lengths are variable and their results don't generalize to this general case. Assuming that the lengths are uniformly distributed within $[1, 1 + \Delta]$, Chen is studying the problem of determining for which clone densities there exist a unique ordering of the clones and the probes.

Another aspect of Chen's work deals with resolution of contention for wireless networks. He is studying protocols that resolve the contention issue in wireless networks. The main difference between a wireless network and a normal wirelinked computer network is that the geometric locations of senders and receivers also affect the protocol design. Thus conventional protocols fail in most cases. The models he is studying are based on those of Raghavan and Upful. Recently R. Ravi has developed $O(\log^3 n)$ approximation algorithms for broadcast trees in undirected networks with the topology being known. Chen is trying to see if the method can be extended to directed graphs, which has been a hard case for many approximation algorithms. On-line algorithms are also being studied.

Scott E. Decatur (1996-97) (SY in Molecular Biology)

The research that Decatur is pursuing is within the fields of molecular biology, electronic commerce, and machine learning. The aspects of molecular biology that he is working on include: Uses of computers for predicting and modeling biological/biochemical processes and algorithms for the support of molecular biology. A central prediction problem in molecular biology is the task of predicting the three-dimensional structure of proteins given their primary amino acid sequence. A more tractable version of this problem is that of motif recognition in which one simply attempts to detect whether a given fold, or motif, forms from the primary sequence of test proteins. He is working on motif recognition with an emphasis on supersecondary-structure motifs, i.e. those whose structure partially describes the full three-dimensional structure of the protein. Another modeling problem on which he has begun working is a computational and geometric investigation of the polymerization of actin. Actin is a protein which polymerizes within cells to form long filaments. These filaments then form either bundles or meshes and perform a wide variety of functions in the cells of humans and other organisms such as cell shape, motility and division. The polymerization of actin into large global structures is known to be governed by local binding events dependent on biochemical conditions and the presence of other proteins. Thus, it may be possible to characterize high level actin polymerization by relatively simple local rules in a manner similar to the method of assembly hypothesized for virus protein shells. The elucidation of such rules governing actin's polymerization could be used to better understand its formation and to possibly alter the process in well defined ways.

In the area of electronic commerce and Internet security, with digital data being produced and distributed on the scale as we are now seeing, new problems are arising which were either not present, or at least not as severe, in the analogous non-electronic settings. Two such areas of research in which Decatur is interested are: developing methods to electronically perform transactions which are currently done either "off-line" or on a more restrictive electronic network, and developing methods to make electronic information widely available in a convenient manner while retaining desired aspects of access control and anonymity. A central component of the first area is payment protocols for electronic transactions. Due to the variety of specifications and properties desirable in such transactions, there will be a need to develop a spectrum of protocols able to meet the needs of these electronic markets and current proposals have begun to cover the space of possibilities. In addition, there are numerous trade-offs between measures such as time, security, anonymity, and convenience; thus different protocols achieving varied measures of these qualities can be useful and in some cases, the trade-offs can even be parameterized.

In the area of information dissemination and control, Decatur is interested in broadcast encryption and electronic fingerprinting. When broadcasting encrypted information, by differentiating the decoding process among recipients, it becomes possible to enable selective access control as well as trace disclosed decryption keys. When providing copyrighted information, electronic fingerprinting allows unauthorized copies to be traced back to the user.

Decatur's research interests in machine learning are mainly concerned with bridging the gaps between theory and practice. Much of his previous research has characterized learnability in formal models that, due to their modeling of imperfect data, begin to approach settings found in practice. One direction for future work is the use of techniques, or even specific algorithms, with theoretical foundation to attack applied problems of prediction and detection in the presence of imperfect training data. Many applications in molecular biology provide opportunities to incorporate such theoretical lessons. One such example is work he is beginning in collaboration with the Genetics Institute biotech company in which they are looking at how to predict optimal oligos for certain hybridization experiments and how to detect relevant patterns in gene expression data.

Kousha Etessami (1995-96) (SY in Logic and Algorithms)

Currently Postdoctoral Researcher, University of Aarhus (BRICS)

A significant part of Etessami's work was focused on applying the tools developed in descriptive complexity to answering questions in temporal and other logics used in automated verification. In particular, in joint work with Thomas Wilke (PD), Etessami answered several questions in temporal logic: He established a strict hierarchy of expressive power for nesting depth of Until operators, based on very simple counting properties, and he provided decidable characterizations for several fragments of temporal logic. The main tool used for these results is a game-theoretic characterization of temporal logic, which is a method used extensively in finite model theory and descriptive complexity. Temporal logic is used widely today to specify and automatically verify properties of open systems. It is of interest to practitioners to know exactly what properties are expressive within the fragments of this logic that they use. Etessami came into contact with these issues because of both the workshops and tutorials held on verification during the special year on logic and algorithms, as well as through discussions with the various researchers present at DIMACS. In later joint work with Moshe Vardi (V) and Thomas Wilke (PD), Etessami continued the study of fragments of temporal logic and their relationship to first-order logic, and provided new upper and lower bounds on the complexity of the satisfiability problem for these logics, again using techniques refined in finite model theory. Understanding the complexity of problems like satisfiability and model-checking is of great importance to the automated verification community, where such algorithms are a central feature of most automated verification tools.

Papers

K. Etessami, M. Vardi, and T. Wilke. First-order logic with two variables and temporal logic with unary predicates. To appear.

K. Etessami and T. Wilke. An until hierarchy for temporal logic. In Proc. 11th Symp. on Logic in Computer Science, 1996, pp. 108-117.

Anna Gal (1995-97) (DIMACS/IAS postdoc)

A large part of Gal's work last year was related to the study of span programs. The model of span programs was introduced by Karchmer and Widgerson in 1993. It is a computational model based on linear algebra with applications to lower bounds in other models, including Boolean formulae. The monotone version of this model is strongly related to the cryptographic problem of secret sharing. Proving lower bounds for monotone span programs has been a challenging open problem; the known techniques for proving lower bounds for monotone circuits (Razborov and Haken's methods) do not help here. In a paper with Amos Beimel and Mike Paterson, Gal developed a new lower bound technique that allows one to prove lower bounds for monotone span programs by considering a problem in extremal set theory. As a first application of the technique they proved a lower bound of $\Omega(n^{2.5})$ for the 6-clique function. Subsequently, in joint work with Laszlo Babai and Avi Widgerson (V), she showed that this technique can yield superpolynomial lower bounds for explicit functions. Her construction of an explicit function family that requires superpolynomial size monotone span programs is based on set systems obtained from Paley-type bipartite graphs. Her results give the first superpolynomial lower bounds for linear secret sharing schemes. Gal showed that this model can be much more powerful than monotone span programs, but require superpolynomial size monotone circuits and exponential size monotone formulae. In particular, this implies that lower bound methods for the circuit model cannot be directly applied to span programs. She also showed that the perfect matching function can be computed by polynomial size (non-monotone) span programs over arbitrary fields.

Papers

A. Biemel, A. Gal and M. Paterson. Lower bounds for monotone span programs. To appear in Computational Complexity.

L. Babai, A. Gal, J. Kollar, L. Ronyai, T. Szabo, A. Widgerson. Extremal bipartite graphs and superpolynomial lower bounds for monotone span programs. In Proc. 28th ACM STOC, 1996, pp. 603-611.

L. Babai, A. Gal, A Widgerson. Superpolynomial lower bounds for monotone span programs. Submitted to Combinatorica.

A. Gal. A simple function that requires exponential size read-once branching programs. Submitted to Information Processing Letters.

Anders Johansson (1995-97) (DIMACS/IAS postdoc)

Johansson has worked on chromatic numbers and list chromatic numbers of graphs. A *list coloring* of a graph arises when we are given a list of allowable colors at each vertex, and we seek an ordinary graph coloring in which each vertex receives a color chosen from its list. List colorings have a variety of practical applications in channel assignment problems, traffic phasing problems, and elsewhere. The *list chromatic number* of a graph is the smallest k so that for every assignment of lists of size k to each vertex, there is always a list coloring. Johansson has obtained asymptotic upper bounds on the chromatic number or list chromatic number of a graph for several classes of graphs with no r -clique. The bound is in terms of the maximum vertex degree – a Brooks'-type bound. The methods for obtaining these results use randomized arguments that should help shed light on the behavior of "simulated annealing," a much-used heuristic for practical applications. Johansson has also worked on applying the same methods to other types of coloring and partition problems. In other work, Johansson has studied *equitable edge-colorings* of eulerian graphs, edge-colorings where the degree of each chromatic factor differs by at most one at all vertices in the graph. He has investigated the relation of this notion to a result of Lovasz and Cherkasskij concerning the possibility of linking the vertices of a set T by edge-disjoint paths, given that the vertices outside T have even degree.

Dominic Mayers (1996-97)

In classical cryptography, it is taken for granted that an eavesdropper has complete access to the messages exchanged between legitimate participants. One purpose of cryptography is to scramble the information so as to make it incomprehensible to the eavesdropper. Quantum cryptography, which was pioneered by Bennett and Brassard in the early 1980's, provides an alternative approach to the ancient problem of secure communication. When information is encoded as non-orthogonal quantum states, it becomes impossible for an eavesdropper to gain information on the transmission without running a risk of spoiling the information irreversibly, which would alert the legitimate users. The more information the eavesdropper attempts to obtain, the bigger her risk of detection. When this is used for the transmission of a random cryptographic key, it allows two parties to communicate in absolute secrecy under the nose of an eavesdropper regardless of her computing power. Quantum techniques may also assist in the achievement of subtler cryptographic goals, such as protecting private information while it is being used to reach public decisions. Quantum protocols that accomplish important cryptographic tasks were known in the late 1980's and some work on their security had been accomplished. Central to all these protocols is the transmission of a large number of photons from one participant called Alice to another called Bob. Except for the quantum bit commitment protocol, security proofs for these protocols considered only a restricted type of measurement on the photons, known as complete orthogonal quantum measurements. Mayers' first contribution in this area was a security proof of a quantum oblivious transfer protocol against the most general measurement that Bob may execute individually on each photon. In this protocol, a test is executed by Alice to guarantee that Bob measures the photons before Alice announces classical information about their quantum state. It was the first time that a security proof considered the possibility that Bob could execute an incomplete measurement on the photons, just enough to pass the test, and then, later on, illegally complete the measurement in view of the classical information announced by Alice. The security of this quantum oblivious transfer protocol against measurements executed on all the photons together was subsequently established by Yao. However, despite the great power of oblivious transfer, the security of quantum protocols that accomplish other cryptographic tasks such as key distribution had not yet been established. The techniques used in Yao's security proof for quantum oblivious transfer could not be applied to quantum key distributions. Mayers' second contribution was a reduction of the security of a quantum key distribution protocol to the security of a quantum oblivious transfer protocol. It follows that secure quantum key distribution is also possible. Mayers has also found a flaw in a quantum bit commitment scheme in the literature and is working on a full security proof of an alternative quantum bit commitment protocol. Mayers designed a quantum oblivious transfer that requires no quantum bit commitment. The problem with this protocol is that it requires the entanglement of n photons. Nevertheless, this approach is of theoretical interest and he would like to develop it further. Using a similar approach, he is also working on a quantum bit commitment protocol with the possibility of providing a commitment for the exclusive-or of two commitments.

Erez Petrank (1996-97) (SY in Networks)

Petrank's research has been focused mostly in Security and Cryptography, but it also contains work on the hardness of approximations, on randomized distributed protocols, and on coding theory. His research in cryptography includes issues in knowledge complexity, non-interactive zero knowledge protocols, and secure storage of files. One of the basic issues in the foundation of Cryptography is the control of the amount of information that leaks during an interaction. The word "information" is misleading since the theory known as "Information Theory" does not provide the appropriate approach to this study. Goldwasser, Micali, and Rackoff initiated the study of how much knowledge leaks in an interaction. They defined knowledge to be the computational advantage gained through an interaction. Their study concentrated on the zero level which is well known as *zero knowledge*. Petrank has suggested ways to measure knowledge complexity of protocols when it is not zero. He has also studied what can be done in low knowledge complexity. That is, what languages can be proven in low knowledge complexity? Petrank has also studied noninteractive zero-knowledge proofs in the shared random string model proposed by Blum, Feldman and Micali and developed a non-interactive proof system for NP with general assumptions (the existence of one way permutations) which is highly efficient in terms of the number of shared random bits used by the prover and verifier.

Petrank is also interested in a natural problem in secure systems, namely storing classified files on-line. A system of classified files contains a set of files (or documents), a set of users, and an authorization structure which defines the subset of files each user is authorized to see. We want a user in the system to be able to see every file in his "authorized list" but not any other file. Petrank has developed definitions of secure classified file systems and studied various aspects of such systems, such as static versus dynamic models, resilience, the role of a central trusted entity, etc.

In his work on hardness of approximations, Petrank has suggested the refinement of the complexity analysis of approximation problems by relating it to a new parameter called *gap location*. He has developed definitions and hardness results of new approximation versions for some known NP-Complete optimization problems. Petrank is also interested in distributed agreement protocols in the presence of malicious failures. The problem is that using a deterministic algorithm yields a protocol with an inefficient running time, whereas using coin tosses induces the possibility of (rare) infinite runs.

Sridhar Rajagopolan (1995-96)
Currently Research Associate, IBM Almaden Research Lab

In joint work with Leonard Schulman, Rajagopolan proved a coding theorem for distributed network protocols. More precisely, he showed that any distributed protocol which runs on a noiseless network in time T, can be simulated on an identical but noisy network with a slow-down factor proportional to $\log(d+1)$, where d is the maximum degree in the network, and with exponentially small probability of error. The theorem is a network analogue of the classical coding theorem of Shannon (1948), which addresses the issue of the transmission of data over a noisy link. Although Shannon's original theorem is non-constructive – it proves the existence of codes that transfer data at asymptotically optimal rates – the subsequent development of coding theory makes this result constructive. Shannon's theory, however, does not address issues that arise when noise occurs in the communications of a wide area computational network with a number of links available for simultaneous use. This question of what happens when computation is interleaved with noisy communications was proposed and first addressed by Schulman in the special case of a pair of computers connected by a bidirectional communication link. In this case, Schulman obtained an analogue of Shannon's theorem. Like those of Shannon and Schulman, the results of Rajagopolan are nonconstructive. However, in his case, it is possible to obtain constructive simulations by compromising the reliability with which they run. Shannon's theorem and these new results both show that, in principle, it is possible to simulate a noise-free environment on a noisy one, but neither provides an explicit simulation technology. One of the main motives of coding theory is to find a constructive alternative to Shannon's code, a goal toward which science has made remarkable progress. The analogous question in the context of distributed computing, the efficient construction of *tree codes*, still remains substantially open. All that is known is that there exist small circuits which compute them. In further joint work with Schulman, Rajagopolan has described a simple randomized mechanism, which falls short of the goal of constructing tree codes, but nevertheless suffices to simulate noiseless networks on noisy ones.

In work with Vijay Vazirani (V), Rajagopolan built on the classical greedy sequential set cover algorithm, using the primal-dual schema, to obtain simple parallel approximation algorithms for the set cover problem and its generalizations. These algorithms run in RNC^3 in the case of set cover and RNC^4 for its generalizations. In the most general case, they address covering integer programs. These are integer programs of the form $\min \underline{c} \cdot \underline{x}$ subject to $A\underline{x} \geq \underline{b}$ and where $\underline{x}$ is restricted to be integral and entries of the vectors $\underline{b}$,$\underline{c}$ and matrix A are all non-negative. As can be seen by their general form, covering integer programs model a large number of applications. Recent work of Lund and Yannakakis has shown that it is NP complete to approximate covering integer programs better than $c \log n$. Rajagopolan's work establishes that there is a constant $C > 0$ such that approximating such programs to within $C \log n$ is in RNC. Fast parallel approximation algorithms were known before for set cover though not for all of its generalizations.

The following question is equivalent to a number of interesting combinatorial optimization problems: Given an undirected graph $G = (V, E)$ with non-negative edge lengths and a positive integer k, find a set $K \subseteq V$ of size k such that the length of the minimum spanning tree on K is minimized. In joint work with Vazirani, Rajagopolan has presented an $O(\log k)$ factor approximation algorithm for this problem, improving on the earlier $O(\log^2 k)$ factor result of Awerbuch, Azar, Blum and Vempala.

Rajagopolan has also provided an $\tilde{O}(n^2)$ time randomized algorithm to check whether a given operation $f : S \times S \to S$ is associative (letting $n = |S|$). He proved this performance is optimal (up to polylogarithmic factors) even in the case where the operation is "cancellative". No sub-n^3

algorithm was previously known for this task. More generally he gave an $O(n^c)$ time randomized algorithm to check whether a collection of c-ary operations satisfy any given "read-once" identity.

Papers

S. Rajagopolan and L. Schulman. Coding theory for distributed computation. IEEE International Symposium on Information Theory, Vancouver, December 1995.

S. Rajagopolan and L. Schulman. Constructive simulation on noisy networks. In preparation.

S. Rajagopolan and V. Vazirani. Primal-dual RNC approximate algorithms for Multi-set, Multi-cover and covering integer programs. SIAM J. Computation, to appear.

S. Rajagopolan and V. Vazirani. Logarithmic approximation of minimum weight k trees. In preparation.

S. Rajagopolan. Verifying Identities. FOCS 96.

Talks

"Coding Theory for Distributed Computation." At DIMACS, Bell Labs, Princeton, IBM-Almaden, Brown, Maryland.

"Primal-dual RNC Approximate Algorithms for Multi-set, Multi-cover and Covering Integer Programs." At RUTCOR and INFORMS.

"Logarithmic Approximation of Minimum Weight k Trees." At RUTCOR, Princeton, Berkeley.

"Verifying Identities." At FOCS, DIMACS.

Edgar Ramos (1995-96)

Currently Postdoctoral Fellow, Institute fur Informatik in Stadtwald

Ramos' main area of research is computational geometry, and he recently has concentrated on the theory and applications of geometric sampling, primarily in its deterministic version. During his year at DIMACS he also started to work in computational metrology, an area of emerging interest in computational geometry, because of the potential for contributing solutions to a real problem (determining how well a manufactured object conforms to the intended shape). Also, as a result of interaction and discussions while at DIMACS, he has started to investigate problems in the area of geometric algorithms for large data sets, an area also relevant to real problems (an often cited example is Geographical Information Systems). In many problems in computational geometry, the input consists of a number n of geometric objects in the plane, three dimensional space, or higher dimensional spaces, and the output is some structure determined by these objects. For many geometric objects, a small random sample of the input provides important information that can be used to divide the original problem into smaller subproblems thus leading to a *divide-and-conquer* strategy to solving the problem. This approach has found many applications starting with those included in the work of Clarkson and Shor, where they established the basic theory and gave several applications. Along with the success of this random geometric sampling, there has been an effort to match its achievements with deterministic algorithms. Here, the properties of a random sample that are useful in algorithmic applications are identified, and deterministic algorithms are found to construct such samples. The divide-and-conquer approach also leads to efficient parallel algorithms. Approximations and nets of range spaces are essential tools in the derandomization of geometric algorithms. M.T. Goodrich in an ACM-SIAM SODA 1996 paper, gave fast constructions of approximations and nets in the CRCW and EREW PRAM models. Goodrich and Ramos discovered a technical problem in one of the arguments, but were able to correct the situation through new algorithms so that the same algorithmic results still hold. Furthermore, they found that for an important and large family of geometric objects (basically all the situations that appear in applications) all the results in the original paper still apply.

In other work, Ramos and Amato considered a collection S of n *sites* in the plane, each one with an associated distance function. The *Voronoi cell* of a site is the set of points closer to that site than to any other, and the (planar) *Voronoi diagram* of S is the decomposition of the plane into the Voronoi cells of the sites in S. Ramos and Amato considered a large class of Voronoi diagrams in which the bisector between two sites is a curve (open or closed) that is *piecewise algebraic*, that is, it consists of a constant number of pieces that are algebraic of bounded degree. They have described a deterministic (sequential) algorithm for this problem that runs in time asymptotically optimal $O(n \log n)$. No such general algorithm was previously known. More recently, using the results described below on computing lower envelopes of functions of one variable, the algorithm was implemented so that its running time depends on the output size f of the Voronoi digram: it runs in time $O(n \log f)$. Again such result was previously known only for a few cases, and not even using randomization in the general case. Furthermore, using results of Goodrich and Ramos on constructing $1/r$-approximations in the CRCW PRAM model, Ramos has obtained optimal work deterministic algorithms that run in time $O(\log n \log\log n)$.

In other work, Amato and Ramos have considered the construction of weighted Euclidean Voronoi diagrams of point sets in R^d. Equivalently, this corresponds to the intersection of halfspaces in R^{d+1} bounded from below by non-vertical hyperplanes. Let n be the number of points and let f be the number of faces in the output. For $d = 4$, their algorithm runs in time $O((n + f)\log^3 n)$ which is close to the known lower bound $\Omega(n \log f + f)$. For $d = 3$, Amato and Ramos observed that a previous algorithm by Chan, Snoeyink and Yap works for degenerate input. The $d = 4$ case, however, requires non-degenerate input (which can be simulated using standard techniques).

In other work, Ramos considered n functions of one variable $f_i : R \to R$, $i = 1, \ldots, n$ each of which consist of a constant number of pieces that are algebraic (the solution of a polynomial equation). The goal is to compute the function $f(x) = \min_i f_i(x)$. Using the divide-and-conquer approach based on geometric sampling has led to a deterministic algorithm running in time $O(\lambda(n) \log n)$, and that also leads to a new output sensitive algorithm that runs in time $O(\lambda(n) \log f)$. Here, $\lambda(n)$ is an upper bound on the combinatorial complexity of f.

Ramos has also worked on a problem concerning the diameter of a point set in 3-space. Given a set of n points in 3-dimensional space, the *diameter* is the furthest distance between any pair of points. Clarkson and Shor gave a randomized algorithm with optimal running time $O(n \log n)$. Since then, the question of matching that running time with a deterministic algorithm has been considered. Ramos is now within a $\log n$ factor of achieving this.

Recently within computational geometry attention has been directed to a class of problems motivated in the area of manufacturing that is concerned with measuring how well a manufactured object fits the intended shape. In particular, the round and planar shapes have drawn great attention (the *roundness* and *flatness* problems). However, the theoretically efficient algorithms obtained are very complicated, and not as efficient as desirable. In joint work with M.T. Goodrich and his student C.A. Duncan at Johns Hopkins University, Ramos gave some simple and efficient approximation algorithms for these two problems, and performed some experimentation with some implemented heuristics.

Papers

N. M. Amato, M. T. Goodrich, and E. A. Ramos. Parallel algorithms for higher-dimensional convex hulls. In Proc. 35th Annual IEEE Sympos. Found. Comput. Sci., pages 683–694, 1994.

N.M. Amato and E.A. Ramos. On computing Voronoi diagrams by divide-prune- and-conquer. Proceedings of the ACM Symposium on Computational Geometry, 1996.

C.A. Duncan, M.T. Goodrich and E.A. Ramos. Efficient approximation and optimization algorithms for computational metrology. To appear in the Proceedings of the SIAM Symposium on Discrete Algorithms, 1997.

M.T. Goodrich and E.A. Ramos. Bounded-independence derandomization of geometric partitioning with applications to parallel fixed-dimensional linear programming. Submitted to Discrete and Computational Geometry.

Talks

ACM Symposium on Computational Geometry in Philadelphia in May.
DIMACS Research and Education Institute in June.
Geometry seminar, Courant Institute

Dana Randall (1994-96) (DIMACS/IAS postdoc)
Currently Assistant Professor, Georgia Institute of Technology

During her two year postdoc at DIMACS and the Institute for Advanced Study, Randall studied uses for Markov chains, particularly focusing on applications in Statistical Physics and Computer Science. Markov chains are used in these disciplines to study features of a combinatorial set by allowing us to sample configurations and estimate the cardinality of the set. In Statistical Physics such sets arise as models of simple physical systems, where the configurations typically represent arrangements of molecules in the system. Sampling provides insight into the thermodynamic properties of the system, such as the specific heat and free energy. In work with Michael Luby and Alistair Sinclair, Randall developed rapidly mixing Markov chains for efficiently sampling planar lattice configurations such as domino or lozenge tilings and Eulerian orientations. These problems arise as dimer systems modeling diatomic molecules being adsorbed onto the surface of a crystal and as a model for ice. For each of these problems there is a natural Markov chain used in practice which starts with an arbitrary configuration and repeatedly performs local updates in order to move to a random configuration. The efficiency of these algorithms remains open. They showed how to add some additional moves in order to turn this into an efficient algorithm with rigorous performance guarantees. The analysis is based on a simple coupling argument and the rich geometric structure underlying the sets of configurations. A second paper extends these results to handle not only planar lattice regions with fixed boundary conditions, where we only consider configurations which agree on the boundary, but also free boundary conditions where we take the union over all possible boundary constraints.

In work with Neal Madras, Randall developed theorems for decomposing Markov chains to use bounds on the mixing rates of various pieces in order to bound the original Markov chain. The state decomposition theorem factors the state space into overlapping pieces and shows that the original Markov chain is mixing if it is when restricted to each of the pieces and if the pieces overlap enough. They applied this to Markov chains for generating three-colorings of lattice regions to present the first provably polynomial time algorithm for this problem. Three-coloring arises in the Ice model and the 3-state Potts model (for antiferromagnetism). A second type of decomposition called the "temperature decomposition theorem" shows how to bound the mixing rate of a type of Metropolis Markov chain by viewing it as an average over several temperatures. This result has been used by Madras and Piccioni to analyze an algorithm for simulated tempering.

Papers

M. Luby, D. Randall, A. Sinclair. Markov chain algorithms for planar lattice structure. FOCS 95.

D. Randall. Generating tilings with free boundary conducts using Pfaffians. In preparation.

N. Madras and D. Randall. Factoring graphs to bound mixing rates. FOCS 96.

Kathleen Romanik (1994-96) (Hoechst-Celanese Postdoc)
Currently Research Associate, National Institute of Standards and Technology

Romanik continued doing research in computational geometry, particularly in the areas of geometric probing, robot navigation and graph visualization, and developed probing strategies for identifying objects in a computer image. These strategies are useful for model-based object recognition. She continued research on finding strategies for localizing a robot in a known environment using minimal travel, with colleagues from McGill University, as well as with a colleague from University of Freiburg in Germany, whom she met at DIMACS. In one scenario, she showed that the localization problem is NP-hard when the environment is modeled by a simple polygon, and she developed a polynomial time strategy that travels close to the optimal distance. In the other work, she designed a competitive strategy for localizing a robot in a geometric tree that improves upon the best previous strategy. She continued research begun at McGill University on visibility representations of graphs. A visibility representation for a graph maps vertices to geometric objects and expresses edges as visibility relations between these objects. She studied a class of rectangle visibility graphs in the plane and proved that the class of directed acyclic graphs that can be represented in this way has unbounded dimension. She finished a research project with Jeffrey Vitter that applied concepts from learning theory to develop better methods for determining the testing complexity of software. They examined the complexity of testing different control flow statements of standard programming languages by defining a complexity measure that is similar to a measure used in computational learning theory and applying it to different control flow statements and combinations of statements. This measure gives probable bounds on the number of random test points needed to determine that a program is approximately correct and gives insight into the relative difficulty of testing different syntactic constructions.

Papers

K. Romanik. Directed rectangle-visibility graphs have unbounded dimension. Discrete Applied Mathematics, 31 (1996).

G. Dudek, K. Romanik, S. Whitesides. Localizing a robot with minimum travel. SIAM Journal on Computing, to appear.

K. Romanik and J.S. Vitter. Using Vapnik-Chervonenkis dimension to analyze the testing complexity of program segments. Information and Computation, 128 (1996), pp. 87-108.

K. Romanik and S. Schuierer. Optimal robot localization in trees. Twelfth ACM Symposium on Computational Geometry, May 24-26, 1996.

K. Romanik. Directed rectangle-visibility graphs have unbounded dimension. Seventh Canadian Conference on Computational Geometry. Quebec City, Canada, Aug 10-13, 1995.

K. Romanik. Geometric probing and testing - A survey. VI Congreso Espanol de Geometria Computacional. Barcelona, Spain, July 5-7, 1995.

R. Freimer, S. Khuller, J.S.B. Mitchell, C. Pitako, K. Romanik, and D. Souvaine. Localizing an object with finger probes. Vision Geometry III, part of SPEI's Intl. Symp. on Photonics for Industrial Applications. Boston, MA, Oct. 31 - Nov. 4, 1994.

K. Romanik. Directed VR-representable graphs have unbounded dimension. Graph Drawing '94. Princeton, NJ, October 10-12, 1994.

Talks

"Rectangle Visibility Representation for Graphs." American Mathematical Society, Northeastern University, October 7-8, 1995.

"Localizing a Robot with Minimum Travel." Hogskolan i Lulea, Institutionen for Systemteknik. Lules, Sweden, September 7, 1995.

"Computational Geometry Applied to Robot Navigation and Automatic Target Recognition." Oak Ridge National Laboratory, Oak Ridge, TN, August 23, 1995.

"Geometric Probing and Testing - A Survey." VI Congreso Espanol de Geometria Computacional. Barcelona, Spain, July 7, 1995.

"A Three Dimensional Visibility Representation for Graphs." Pace University Mathematics Seminar, April 19, 1995.

"Localizing a Robot with Minimum Travel." UMIACS Seminar on Algorithms, University of Maryland, February 20, 1995.

"A Three Dimensional Visibility Representation for Graphs." Combinatorists of New England (CONE) one day conference on Combinatorics and Graph Theory, Smith College, February 18, 1995.

"Localizing a Robot with Minimum Travel." McGill University Computational Geometry Seminar, November 24, 1994.

"Localizing a Robot with Minimum Travel." Fourth MSI Workshop on Computational Geometry, Cornell University, October 14-15, 1994.

"Using Vapnik-Chervonekis Dimension to Analyze the Testing Complexity of Program Segments." AT&T Bell Laboratories, Murray Hill, NJ, September 21, 1994.

Mona Singh (1995-97) (SY in Molecular Biology)

During the past year, Singh has been working primarily on computational methods for protein structure prediction. Biologists are interested in this problem because the structure of a protein provides insight into its biological function. Experimental approaches to determining a protein's structure are difficult and time-consuming. As a result, much recent research has focused on applying mathematical and computational techniques to predicting protein structure from one dimensional sequence data. A first step to tackling this complex problem is a solution to the structural motif recognition problem: Given a known commonly occurring three-dimensional substructure, or motif, determine whether this motif occurs in a given amino acid sequence. With Bonnie Berger (V) from the Math Department at MIT, Singh has worked on an iterative algorithm that uses randomness and statistical techniques to improve existing methods for recognizing protein structural motifs. Their algorithm is particularly effective in situations where large numbers of sufficiently diverse examples of the motif are not known. These are precisely the situations that pose significant difficulties for previously known methods. They have implemented their algorithm, and show that it is able to substantially improve performance on the coiled coil motif. Using their program, they, with Peter Kim of MIT and the Whitehead Institute, have identified one and occasionally two coiled-coil-like regions in the envelope proteins of many retroviruses, including Human Immunodeficiency Virus (HIV) and Human T-cell Lymphotropic Viruses (HTLV). In addition, they have been exploring an application of the method to signaling proteins (with Andrea Cochran of Genetech, Peter Kim and Bonnie Berger). Singh is also working on a project for predicting the full three dimensional structure of a protein. With Teresa Przytycka (V), she has been working on a "coarse grain threading" approach. Traditional threading methods try to predict protein structure by aligning protein sequences onto three dimensional structures, using a pairwise energy potential. Singh and Przytycka's method augments the usual threading algorithm with predictions of the secondary structure of the sequence. They have formally defined the coarse grain threading approach in a graph theoretic model. It turns out that this problem is NP-complete, and they are now working on approximation algorithms for the problem, as well as exact solutions for important special cases. After they complete the theoretical study of this problem, they will implement their algorithms. Singh has very recently started working with Martin Farach (PM, RU) and Scott Decatur (PD) on a project studying simple models of protein folding. This project is still in the developmental stages, but will be a computational study of folding on the H/P lattice model. The work will not focus on structure prediction but rather on elucidating properties of protein folding. Biophysicists have studied this model for some time, but until recently computer scientists have not.

Paper

B. Boyn and M. Singh, An iterative method for improved structural motif recognition. First Annual Conference on Computational Molecular Biology (RECOMB 97).

Lubos Thoma (1996-97)

Thoma's research falls into the areas of probabilistic methods in combinatorics, extremal combinatorics, partially ordered sets, and 0-1 laws for random structures. Consider a random linear ordering of the k-element subsets of an m-set, with all permutations being equally likely. Recently, Kuperberg asked if the greedy algorithm yields a sufficiently large family of subsets. The answer to this can be obtained using r-uniform hypergraphs and a "nibble" approach: The random greedy algorithm yields almost surely a packing of "sufficiently" large size. Thoma is also interested in cover graphs of finite lattices. An old problem of Ore is to characterize graphs that are cover graphs of finite posets. Nesetril, Rodl, and Brightwell finally showed that this was NP-complete. Thoma has extended this result to finite lattices. Thoma has also worked on convergence laws for random mappings. A *random mapping* on the set $[n]$ is a function chosen uniformly at random from the family of all functions which map $[n]$ into itself. Thoma has considered properties of random mappings u_n which can be expressed in the second order monadic language known as $L^{\pi}_{\rightarrow}$. This language, besides logical symbols, contains two countable sets of variables, one interpreted as elements and the other as sets of elements, and it also contains a membership predicate and a predicate $\rightarrow$, where $x \rightarrow y$ is interpreted as $u_n(x) = y$. With T. Luczak, Thoma proved that for every sentence from this language, the probability that the sentence holds for a random mapping u_n converges as n gets large. Thoma is working on applying this result to random graphs. Specifically, if $G_{n,p}$ is the random graph with edge probability equal to p, Shelah and Spencer have shown that for probability p such that $pn \rightarrow c, 0 < c < \infty$, the limit of the probability that a given sentence holds in $G_{n,p}$ exists provided $c < 1$ and does not provided $c > 1$. Thoma is working on the case $c = 1$. Finally, Thoma's research is concerned with the use of linear algebra methods to solve extremal problems in finite set theory. A typical question is: Is there an $\epsilon > 0$ such that for each family of subsets of $[n]$ satisfying the condition that the cardinality of the intersection of any two distinct sets in the family is different from $\lfloor n/4 \rfloor$, one concludes that the size of the family is at most $(2-\epsilon)^n$? This question was answered affirmatively by Frankl and Rodl, but Thoma is investigating alternative proofs and generalizations.

Pavel Valtr (1996-97)

Valtr's research interests include discrete geometry, classical combinatorics, probabilistic methods in combinatorics, Ramsey theory, stochastic geometry, and computational geometry. Valtr is especially interested in the class of finite planar point configurations. A large part of his work concerns so-called dense sets. *Dense sets* are finite planar point sets for which the ratio between the maximum and the minimum distance spanned by pairs of points of the given set is essentially smallest possible, i.e. of size $O(\sqrt{n})$, where n is the number of points of the set. Motivations for the study of dense sets come both from discrete and from computational geometry. He is working on algorithms for computing the convex hull of a dense set, the size of a convex polygon that can be guaranteed to exist in a dense set of n points in the plane, and the maximum size of two equal-sized mutually avoiding sets in any dense set of n points. Valtr is also interested in stochastic geometry. He has shown the exact formula for the probability that n random points chosen uniformly and independently from a square or from a triangle are vertices of a convex n-gon. This gives a partial solution to a natural generalization of an old problem of Sylvester. He hopes to find such a formula when a square or triangle is replaced by another convex body (planar or more-dimensional, e.g., disk, cube). He is also working on proving the conjecture that, with n fixed, the above probability is maximum for a disk (ball) and minimum for a triangle (simplex). Probabilistic methods play an important role in his research in combinatorics as does Ramsey Theory.

Thomas Wilke (1995-96) (SY in Logic and Algorithms)

Currently Assistant Professor, Christian-Albrechts-Universitat zu Kiel

Wilke's work has been concerned with the investigation of propositional linear temporal logic, henceforth abbreviated by temporal logic. *Temporal logic* is a language for expressing relationships between the order of events occurring over time. It arises in various guises in several areas of computer science as different as computer-aided verification (temporal logic as a specification language for reactive and concurrent systems) and databases (temporal logic to represent knowledge involving time and to deduct on time-dependent data). Temporal logic is viewed as a fundamental logical language, as it is known to be exactly as expressive as first-order logic over linear orders with unary predications. Collaboration with Kousha Etessami (PD), another special year on logic and algorithms postdoc, and Denis Therien (V), who visited DIMACS from Montreal, has led to the following major contributions to the study of temporal logic during the special year: A combinatorial tool for analyzing the expressive power of temporal logic, a so-called Ehrenfeucht-Fraisse game for temporal logic; a game-theoretic method for characterizing fragments of temporal logic in terms of the structure of the corresponding finite automata; a general principle that connects finite semigroups and temporal logic, the so-called "semidirect-product/substitution principle". There are two other separate sets of special year results Wilke has been involved in. One is joint work with Doron Peled from Bell Labs and Pierre Wolper (V) from Liege, who was one of the lecturers for the introductory special year tutorial. The knowledge that in a given reactive system certain actions can be executed concurrently without influencing each other can be used to decrease the complexity of checking the system against specifications in temporal logic. For this to work, it is, however, necessary that the specified property does not distinguish between two interleavings of the same concurrent run of the system. Wilke, Peled and Wolper's results imply that one can verify in polynomial space whether a temporal property satisfies this condition. The other set of results is concerned with the comparison of temporal logic and two-variable first-order logic. That is joint work with Kousha Etessami (PD) and Moshe Vardi (V and co-chair of the special year). They show that the fragment of temporal logic that uses "next," "previously," "eventually," and "sometime in the past" as the only temporal operators is expressive as two-variable first-order logic in the language with successor and order relation, and, similarly, that the fragment of temporal logic that uses "eventually" and "sometime in the past" as the only temporal operators is as expressive as two-variable first-order logic in the language with order relation only. Moreover, they determined the differences in succinctness between first-order logic and temporal logic in the two cases and the complexity of the satisfiability problem for each of the four logical languages involved.

Papers

K. Etessami and T. Wilke. An until hierarchy for temporal logic. In 11th Annual IEEE Symposium on Logic in Computer Science, pages 108-117, New Brunswick, NJ, 1996. IEEE Computer Society Press.

D. Peled, T. Wilke, and P. Wolper. An algorithmic approach for checking closure properties of ω-regular languages. In Concur '96: Seventh International Conference on Concurrency Theory, Pisa, Italy. Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1996. To appear.

D. Therien and T. Wilke. Temporal logic and semidirect products: An effective characterization of the until hierarchy. Technical Report 96-28, DIMACS, Piscataway, NJ, 1996.

D. Therien and T. Wilke. Temporal logic and semidirect products: An effective characterization of the until hierarchy. In Proceedings of the 37th Annual Symposium on Foundations of Computer Science. Burlington, Vermont, 1996. IEEE Computer Society Press. To appear.

3.2 Graduate Student Fellows: September 1995 - January 1997

3.2.1 1995–96 DIMACS-supported Graduate Students at Rutgers University

Tamas Badics

Badics was a student of Endre Boros (PM) at RUTCOR, the Rutgers Center for Operations Research. The title of his thesis was "Approximations of some nonlinear binary optimization problems." Badics' work was in two areas: Computational study of maximum flow algorithms (part of DIMACS implementation challenge) and approximation algorithms for pseudo-Boolean minimization (3/4 guarantee for a new simple algorithm for MAX-SAT). He spoke on this topic at the DIMACS challenge workshop mentioned above, and wrote a paper on the topic with his advisor. He is now working as a "Sr. Software Engineer" at Parametric Technology Corp.

paper

T. Badics and E. Boros. Implementing a maximum flow algorithm: Experiments with dynamic trees. In D. Johnson and C. McGeoch (eds.), Network Flows and Matching, First DIMACS Implementation Challenge. AMS-DIMACS Series, Providence, RI.

talk

"Implementing a Maximum Flow Algorithm: Experiments with Dynamic Trees." DIMACS Challenge Workshop.

Tom Bohman

Bohman was a student of Jeff Kahn (PM) in Mathematics. He defended in the spring of 1996 and is now a postdoc at MSRI in Berkeley, CA. The title of his thesis is "On two combinatorial problems: The Conway-Guy sequence and discrete threshold growth." His research involved a model of crystal growth and an analysis of the Conway-Guy sequence, in an effort to solve a 60-year-old problem of Erdős. Erdős' problem was to find out how large the largest element of a set with n elements needs to be in order for that set to have distinct subset sums. Conway and Guy proposed a sequence of sets which would give a small upper bound if they could show that the sets had distinct subset sums, and 28 years later, Bohman showed that this sequence had the desired properties. He discussed related sequences as well, and now has the best known upper bounds for Erdős' question.

Tor Gunston

Gunston is a graduate student in the Mathematics department. His advisor is Wolmer Vasconcelos (PM), and the focus of his research is how algebraic structures can arise from combinatorial structures. There are several ways in which one can derive from a combinatorial object an object in commutative algebra, usually a standard graded algebra over a field. For example, one can start with a graph and derive a monomial algebra by considering a polynomial ring (over some field) whose variables correspond to the vertices of the graph and taking the subalgebra generated (over the field) by those squarefree, degree two monomials that correspond to edges in the graph. Similarly, one may construct an algebra from a simplicial complex by taking the quotient of the polynomial ring whose variables correspond to the vertices of the complex by the ideal generated by those squarefree monomials corresponding to the minimal non-faces of the complex (this is called a Stanley-Reisner ring). By constructing a dictionary between the properties of the combinatorial object and the algebraic object, one hopes to gain insight into algebra. The general aim is to understand what sorts of algebras are out there. For example, it is unknown what are the possible Hilbert functions of a Gorenstein standard graded algebra which is a domain. Investigations were made into the properties of the monomial algebra associated to a graph in the case where the graph is bipartite and planar. In particular, it was shown how to read the Hilbert series from the graph. Also investigated was the combinatorial "meaning" of the homological degree of a Stanley-Reisner ring. Gunston's expected graduation date is May, 1998.

paper

Doering and Gunston. Algebras arising from bipartite planar graphs. Communications in Algebra, 24 (1996), 3589-3598.

Vincenzo Liberatore

Liberatore has been working on a problem in algorithmic information theory under the direction of Mike Fredman (PM) in the Computer Science department. The problem he is working on is as follows: "Given n subsets of a totally ordered universe of size n, find the maximum of each subset with a small number of comparisons". The best lower bound for this problem is $\Omega(n)$ comparisons and no matching upper bound is known except for special cases. A generalization of the rank-sequence algorithm had been conjectured to solve this problem with $O(n)$ comparisons, but he disproved it. He then generalized an algebraic approach to yield a linear algorithm for a large class of instances. He is currently writing up a paper on the topic.

Alexander Ng

Ng is working with Joszef Beck (PM) in the Math department. His research problems include discrete geometry and mathematical finance. In discrete geometry, he obtained an algorithm to compute the minimum width strip containing 3 among N given points in a unit square in $O(N^2)$ time and $O(N)$ space. The algorithm relies on a duality transformation which transforms the point configuration to a line arrangement. The topological line sweep is used to compute the best triple. The algorithm can be extended to find the minimum width strip containing k among N points. The duality approach is also promising for computing asymptotic upper bounds on the smallest width guaranteed to contain some 3 points in any N point configuration since it allows the use of recent results on line arrangements. This open problem is related to Heilbronn's famous minimum area triangle conjecture. In mathematical finance, he is investigating the correct pricing of a zero coupon bond under the assumption of a stochastic inflation rate, nonzero constant tax rate and required rate of return. When taxation is continuous, the initial price P is implicitly defined by a stochastic differential equation. This problem naturally leads to optimal stopping consideration depending on whether early exercise of the bond is allowed or not. Both the continuous time and discrete N period cases are not well understood. Previous researchers have typically assumed no inflation and tax considerations. Those resulting pricing formulas are much simpler.

Ming Ouyang

Ouyang is working with Vasek Chvatal (PM) in the Computer Science department on the SAT problem: Given a boolean formula in the CNF form, decide if it is satisfiable. He gave a talk about his work at *Optimization Days,* May 1996, Université de Montreal. His publications include DIMACS Technical Report 96-38 about the Davis-Putnam-Logemann-Loveland algorithm, one of the most popular algorithms for solving the satisfiability problem. Its efficiency depends on its choice of a branching rule. In this report, Ouyang constructed a sequence of instances of the satisfiability problem that fools a variety of "sensible" branching rules in the following sense: When the instance has n variables, each of the "sensible" branching rules brings about $\Omega(2^{(n/5)})$ recursive calls of the Davis-Putnam-Logemann-Loveland algorithm, even though only $O(1)$ such calls are necessary.

paper

M. Ouyang. How good are branching rules in DPLL? DIMACS Technical Report 96-38.

talk

"Davis-Putnam-Loveland versus Resolution Search." Optimization Days. Université de Montreal. May 1996.

Aleksandar Pekec

Pekec was a student in Mathematics, working under the direction of Fred Roberts (PM). His main area of research fell into the interface among combinatorial optimization, combinatorial algorithms, and the theory of scales of measurement. He was concerned with the problems that might arise if we claim that something is an optimal solution to a problem where certain parameters are measured using scales of measurement. If admissible transformations of scale (such as change of unit or of zero point) are made, it is possible that a solution which was optimal might no longer be optimal. The simplest example of this might be in the shortest path problem, where if we transform each edge-weight by a function of the form $d(w) = aw + b$, $a > 0$, i.e., we change both units and zero points, we can change the shortest path from one vertex to another. Pekec got interested in this issue in connection with a practical problem of scheduling in which the penalty for late arrival of a passenger or package depended upon the priority of the passenger or package. He showed situations where the optimal solution to such a scheduling problem can change under admissible change of scale used to measure priority, and described situations when this problem does not occur. This led him to obtain results about the limitations of some well known combinatorial optimization algorithms, and specifically about the objective functions for which the algorithms still lead to the same optimal solution after admissible change of scale. It also led him to similar results about different solution concepts in the theory of multiperson cooperative games. In addition, this kind of analysis led Pekec to some fundamental results connecting ordinal scales and greedy algorithms, results that can be stated in a surprisingly general, abstract setting.

Pekec worked with Mike Rothkopf (PM, RU) to solve some problems involving simultaneous auctions allowing combined bids. The theory of auctions and bidding has enormous practical significance, and is used widely in industry, for instance in the recent bids by telecommunications companies for pieces of the FCC radio spectrum and in oil lease bids. Yet, many fundamental problems of auctions and bidding remain unresolved because of their combinatorial difficulty.

There is a long literature on optimal orientations of graphs, in connection with designs of transportation networks. Following on numerous efforts devoted to finding optimal strongly connected orientations of rectangular grid graphs, Pekec, together with J.C. Bermond, J. Bond, C. Martin, and F. Roberts, has found optimal strongly connected orientations of circular grids with a central vertex. He has also, on his own, gone further and analyzed the case of circular grids without a central vertex. His ideas have helped to organize a variety of methods into a coherent attack on these problems.

There has been a great deal of interest recently in the study of combinatorial games, and in particular in their applications to finding derandomizations of probabilistic proofs. With this as motivation, Pekec has solved a difficult problem of Joszef Beck (PM, RU) concerning winning strategies in Ramsey games.

Because of his interest in applications ranging from social influence models to neural networks, Pekec got interested in majority processes on graphs as models of consensus methods. These processes involve iterative coloring processes in which vertices are colored in one of two colors and a vertex changes color if a majority of its neighbors have opposite color. He gave a necessary condition for a graph to end up monochromatic in the color which is in the majority at the start of the process. The problem of determining when this might happen has interesting potential applications to consensus methods and has been intriguing discrete mathematicians and theoretical computer scientists.

Pekec is currently on a postdoctoral fellowship at BRICS in Aarhus, Denmark.

papers

J-C. Bermond, J. Bond, C. Martin, A. Pekec, F.S. Roberts. Orientations of annular cities. In preparation.

N.V.R. Mahadev, A. Pekec, and F.S. Roberts. On the scheduling problem with priorities and lateness/earliness penalties: When is an optimal solution not optimal? Operations Research, to appear.

N.V.R. Mahadev, A. Pekec, and F.S. Roberts. On the scheduling problem with priorities and lateness/earliness penalties: Some meaningful conclusions with nonconstant arrival times. Submitted.

A. Pekec. A winning strategy for the Ramsey graph game. Combinatorics, Probability, and Computing, to appear.

A. Pekec. Optimization under ordinal scales: When is a greedy solution optimal? Mathematics of Operations Research, to appear.

M.H. Rothkopf, A. Pekec, and R. Harstad. Computationally manageable combinatorial auctions. Submitted.

talks

Second International Colloquium on Graphs and Optimization, Loeche-les-Bains, Switzerland, August 1994.

ORSA/TIMS joint national meeting, Detroit, MI, October 1994.

Croatian Operations Research Society, January 1995.

Southeastern Conference on Combinatorics, Graph Theory, and Computing, Boca Raton, FL, March 1995.

INFORMS National meeting, Los Angeles, CA, April 1995.

Southeastern Conference on Combinatorics, Graph Theory, and Computing, Baton Rouge, LA, February 1996.

Reuben Settergren

Settergren is working at RUTCOR, the Rutgers Center for Operations Research, under Farid Alizadeh (PM). His research involves algorithms for, and probabilistic modeling of, physical mapping of DNA. The most common technologies for physical mapping involve clone fingerprinting. A clone is a subsequence of the target DNA, created by random fragmentation. After a library of overlapping clones is obtained that covers the genome to be mapped, fingerprinting experiments are performed, which give evidence of each clone's overlap with the others, or their containment of the landmarks which will eventually form the physical map. From this fingerprinting evidence, the correct order of the clones on the genome must be determined. After the clones are ordered, there might also be more work to build the actual physical map. He has dealt with various aspects of all three of these issues in the past year. He has given two talks at the RUTCOR brown-bag seminar.

papers

1) Alizadeh, Chen, and Settergren. Thresholds for ordering libraries of variable length clones. RUTCOR Research Report RRR # 5-96, 1996.

2) Fasulo, Jiang, Karp, Settergren, and Thayer, An algorithmic approach to multiple complete digest mapping. To appear, RECOMB, Jan 1997.

3) Settergren. 'What we have here, is a failure to cooperate'. To appear in Discrete Mathematics in the Schools: Making an Impact. D. Franzblau, F. Roberts, J. Rosenstein, eds. DIMACS-AMS, 1997.

talks

"Analyzing the Asymptotic Threshold Behavior of the Clone Densities Necessary to Build Clone Libraries which cannot be Reconstructed Incorrectly." RUTCOR brown-bag seminar.

"Resolving a Physical Map from the Ordered Clones, with Fingerprint Data from Restriction Fragmentation." RUTCOR brown-bag seminar.

Wenan Zang

Zang defended his dissertation "Three topics in graph theory" in the Fall of 1995. He worked under the direction of Vasek Chvatal (PM) in Chvatal's role as a member of the graduate program in Operations Research. A paper with F.K. Hwang (PM), based on his work at DIMACS, is in review. For data consistency in distributed database systems, it is necessary to compare remotely located copies of a file. The cost of such comparison is in the amount of communication required to identify the erroneous pages in the different copies. In this paper he assumes that the sites at which the copies reside can communicate only with a coordinator and that for each page the majority of copies are correct. The minimum amount of communication necessary to identify any given number of erroneous pages is estimated and a technique to achieve their bound is presented. Zang is currently a lecturer in Mathematics at the University of Hong Kong.

paper

F.K. Hwang and W. Zang. Detecting corrupted pages in M replicated files. Submitted.

3.2.2 1995–96 DIMACS-supported Graduate Students at Princeton University

Dan Boneh

Boneh does research in cryptography, network security, computational number theory and DNA computing. His research in cryptography includes proving the security of various cryptographic primitives as well as designing new protocols. In network security he designs and implements various systems. These results usually involve a fair amount of computational number theory. The research he does in DNA computing includes finding applications for DNA computers as well as designing "algorithms" that take advantage of the computational power available from manipulation of DNA strands. He defended his thesis in June of 1996 under the direction of Richard Lipton (PM), and is currently a research scientist at the cryptography research group at Bellcore.

papers

Boneh and Rubin. Revocation of unread E-mail in an untrusted network. Submitted to 1997 IEEE Symposium on Security and Privacy, Oakland, California.

Boneh, DeMillo and Lipton. On the importance of checking computations. Submitted to Eurocrypt 97, Konstanz, Germany.

Boneh and Venkatesan. Rounding in lattices and its cryptographic applications. To appear in SODA 97.

Boneh and Lipton. A revocable backup system. Proceedings 6th USENIX Security conference, 1996, San-Jose, California.

Boneh and Venkatesan. Hardness of computing the most significant bits of secret keys in Diffie-Hellman and related schemes. Proceedings Crypto96, Santa Barbara, California,

Boneh and Lipton. Algorithms for black box fields and their application to cryptography. Proceedings Crypto96, Santa Barbara, California.

Boneh, Dunworth, Lipton and Sgall. On the computational power of DNA. Discrete Applied Mathematics, Special Issue on Computational Molecular Biology. (Princeton CS Tech-Report TR-499-95).

Boneh and Lipton. Making DNA computers error resistant. Proc. of Second Annual DIMACS Conference on DNA Computing, 1996. (Princeton CS Tech-Report TR-491-95).

Baum and Boneh. Running dynamic programming algorithms on a DNA computer. Proc. Second Annual DIMACS Conference on DNA Computing, 1996.

Annalisa DeBonis

DeBonis was in the Computer Science department at Princeton, working under Robert Tarjan (PM), when she was a DIMACS-supported student, and is now working at Rutgers under Diane Souvaine (PM). Her research has been concerned with problems related to the implementation and complexity of data structures. In particular, it has been focused on problems related to the splay trees and priority queue data structures. The splay tree is a very efficient type of binary search tree which allows one to perform all the standard search tree operations in $O(\log n)$ amortized time per operation. The idea which allows one to achieve this bound is to restructure the tree every time a node is accessed in a way such that future operations can be performed more efficiently. DeBonis has also considered the problem of implementing efficient priority queues. This topic has been extensively investigated by many authors. The Fibonacci heaps of Fredman and Tarjan achieve the best amortized time since they require an amortized time of $O(\log n)$ for the Delete and Delete-Min operations and amortized constant time for the other operations. The relaxed heaps of Discroll *et al* achieve worst case $O(\log n)$ time for the Meld operation and for the Delete and Delete-Min operations and constant worst case time for the other operations. Brodal has recently described a data structure which requires worst case constant time for all operations except for the Delete and Delete-Min operation. Two main issues concern the implementation of this data structure. One issue concerns the construction of the data structure which involves many complicated details. Another problem connected to the implementation of Brodal's data structure is that it makes use of arrays and therefore is not implementable on a pointer machine. Her research has been concerned with the problem of making Brodal's data structure practically applicable by simplifying its construction. At the same time she explored the possibility of giving an alternative implementation which does not make use of arrays.

Christopher Dunworth

Dunworth is working in the field of DNA computing under Richard Lipton (PM) of the Computer Science department. Presently, he is writing a biological simulator that will allow him and his colleagues to more rapidly design and test experiments on a small scale before heading into the laboratory. They want the design to be extendible, the interface to be intuitive and informative, and, of course, the model to be accurate. He intends to finish this project by January of 1998.

Alejo Hausner

Hausner is a graduate student in the Computer Science department under David Dobkin (PM). He is currently developing a system for animating and designing geometric algorithms over the web. This involves writing a large system in Java within which geometric algorithms can display their activity visually. In other words, this system eases the creation of animations of geometric algorithms. The project has two goals: The first goal is pedagogical, since seeing an animation of an algorithm often conveys the ideas behind the algorithm better than a written description. Embedding the animations in a web page allows us to mix animations with written and verbal descriptions. The second goal is to ease the development of new geometric algorithms. Geometers often struggle to translate the numerical contents of their programs into the naturally visual geometric objects which these numerical contents represent. The project aims to provide a robust system for developing and visualizing new algorithms. He expects to complete his Ph.D. thesis around December, 1997.

talks

"Geometric Algorithm Animation," Princeton Univ. CS Dept., Nov. 1996.

"Demonstration of a Geometric Algorithm Animation," DREI '96 workshop, July 1996, Princeton.

Liviu Iftode

Iftode is a student in the Computer Science department working under Kai Li and Jaswinder Pal Singh (PM). His current research involves improving the performance of shared virtual memory. Ten years after the shared virtual memory (SVM) concept was introduced, there is still debate on its efficiency and limitations compared with expensive, hardware cache-coherent systems. A new trend in developing SVM systems has recently emerged: These systems use various degree of custom hardware support to improve the performance of SVM running on a network of workstations, as long as it doesn't significantly alter the commodity nature of the system as a whole. One such mechanism is Automatic Update. This is a simple memory-mapped communication mechanism supported by the SHRIMP network interface, which forwards local writes to a remote memory in a transparent manner. The goal of the research is to further improve the performance of shared virtual memory by enhancing the hardware support, protocols and applications. Iftode and his colleagues have demonstrated that the automatic update support and the home-based software coherence protocols can achieve substantially better performance compared with previous shared virtual memory systems.

papers

Iftode, Li and Zhou. Performance evaluation of two home-based lazy release consistency protocols for shared virtual memory systems. Proceedings of the 2nd Symposium on Operating Systems Design and Implementation (OSDI 96), October 1996.

Iftode, Li and Pal. Scope consistency: A bridge between release consistency and entry consistency. Proceedings of the 8th Annual ACM Symposium on Parallel Algorithms and Architectures, June 1996.

Iftode, Li and Singh. Understanding application performance on shared virtual memory. Proceedings of the 23rd Annual International Symposium on Computer Architecture, May 1996.

Dubnicki, Felten, Iftode and Li. Improving release-consistent shared virtual memory using automatic update. Proceedings of the 2nd International Symposium on High-Performance Computer Architecture, February 1996.

Alpert, Bilas, Blumrich, Clark, Damianakis, Dubnicki, Iftode and Li. Early experience with message-passing on the shrimp multicomputer. Proceedings of the 23rd Annual Symposium on Computer Architecture, May 1996.

Dubnicki, Felten, Iftode and Li. Software support for virtual memory-mapped communication. Proceedings of the 10th International Parallel Processing Symposium, April 1996.

Caceres and Iftode. Improving the performance of reliable transport protocols in mobile computing environments. IEEE Journal on Selected Areas in Communications, June 1995.

Haim Kaplan

Kaplan defended his thesis "Purely Functional Lists" in November, 1996 working under Bob Tarjan (PM). He has recently joined AT&T Labs. The thesis contains some of his results on data structures (joint work with Bob Tarjan). Some of these results and others were presented in STOC'95 and STOC'96. Among other things, Kaplan has worked on purely functional programming. He developed methods for purely functional representations of sorted lists, implemented as finger search trees. The most efficient of three implementations he has developed achieves logarithmic access, insertion, and deletion time, and double-logarithmic catenation time.

papers

Kaplan and Tarjan. Persistent lists with catenation via recursive slow-down. 27th Annual ACM Symposium on the Theory of Computing, Las Vegas, Nevada. 1995, 93–102.

Kaplan and Tarjan. Purely functional representations of catenable sorted lists. To appear in 28th Annual ACM Symposium on the Theory of Computing, Philadelphia PA. 1996.

Kaplan, Shamir and Tarjan. Tractability of parameterized completion problems on chordal and interval graphs: Minimum fill-in and physical mapping (Extended abstract). Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS), November 1994, IEEE Computer Science Press, 780–791.

Kaplan, Shamir and Tarjan. Faster and simpler algorithm for sorting signed permutations by reversals. To appear in The Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'97).

Arora, Frieze and Kaplan. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. Proceedings of the 37th Annual Symposium on Foundations of Computer Science (FOCS'96), IEEE Computer Science Press, 21–30.

talks

STOC '95
STOC '96

3.3 Center Visitors: September 1995 - January 1997

- **Foto Afrati**; National Technical University of Athens;
  Research Area: Logic & Algorithms; Dates: 5/3/96 - 6/4/96

- **Noga Alon**; Tel-Aviv University ;

Research Area: Discrete Probability; Dates: 9/1/96 - 12/31/96

- **Andris Ambainis**; Institute of Mathematics and Computer Science, University of Latvia;
Research Area: Molecular Biology; Dates: 10/28/96 - 12/4/96

- **Martin Anthony**; London School of Economics;
Research Area: General; Dates: 10/16/95 - 11/2/95

- **Roy Armoni**; Hebrew University;
Research Area: General; Dates: 12/15/95 - 2/29/96

- **Richard Arratia**; University of Southern California;
Research Area: Discrete Probability; Dates: 12/8/96 - 12/16/96

- **Imre Barany**; Hungarian Academy of Science;
Research Area: General; Dates: 5/14/96 - 5/28/96

- **Stephen Bellantoni**; Analytica & Co, Inc.;
Research Area: Logic & Algorithms; Dates: 12/1/95 - 1/12/96

- **Steve Bloch**; Adelphi University;
Research Area: Logic & Algorithms; Dates: 6/3/96 - 6/30/96

- **Egon Boerger**; University of Pisa;
Research Area: Logic & Algorithms; Dates: 10/2/95 - 11/25/95

- **Bela Bollobas**; University of Cambridge;
Research Area: Discrete Probability; Dates: 9/1/96 - 12/15/96

- **Eleanora Bounimova**; Middle East Technical University;
Research Area: Logic & Algorithms; Dates: 7/16/96 - 8/6/96

- **Andreas Brandstaedt**; University of Rostok, Germany;
Research Area: General; Dates: 3/1/96 - 5/30/96

- **Graham Brightwell**; London School of Economics;
Research Area: Discrete Probability; Dates: 8/22/96 - 9/17/96

- **Oleg Burdakov**; Center for Research and Advanced Training in Scientific Computation (CERFACS);
Research Area: General; Dates: 6/2/96 - 6/8/96

- **Alessandra Carbone**; Universitaire Scientifique de Jussiev;
Research Area: Logic & Algorithms; Dates: 4/8/96 - 4/29/96

- **Shiva Chaudhuri**; MPI Informatik;
Research Area: Logic & Algorithms; Dates: 5/28/96 - 6/7/96

- **Kiran Chilakamarri**; Central State University;
Research Area: General; Dates: 7/29/96 - 8/19/96

- **Dennis Dams**; Einhoven University of Technology;
Research Area: Logic & Algorithms; Dates: 7/4/96 - 7/30/96

- **Bhaskar DasGupta**; Rutgers,Camden-Department of Computer Science;
Research Area: General; Dates: 10/1/96 - 6/30/97

- **Anuj Dawar**; University of Wales Swansea;
Research Area: Logic & Algorithms; Dates: 10/30/95 - 11/30/95

- **Bernard Derrida**; Laboratoire de Physique Statique;
Research Area: Discrete Probability; Dates: 10/23/96 - 11/8/96

- **Shlomi Dolev**; Ben Gurrion University;
Research Area: General; Dates: 7/9/96 - 7/24/96

- **Dannie Durand**; The University of Pennsylvania;
Research Area: Molecular Biology; Dates: 9/1/95 - 8/31/96

- **Weiwu Fang**; Chinese Academy of Science;
Research Area: General; Dates: 11/15/96 - 11/15/97

- **Paul Flavell**; University of Birmingham;
Research Area: General; Dates: 4/15/96 - 5/22/96

- **Biano Gal**; University of Minnesota;
Research Area: General; Dates: 1/15/96 - 3/15/96

- **Giorgio Gallo**; Universita di Pisa Dipartimento di Informatica;
Research Area: General; Dates: 5/6/96 - 5/17/96

- **Martin Grohe**; Albert-Ludwigs-Universitat Freiburg;
Research Area: Logic & Algorithms; Dates: 10/16/95 - 11/17/95

- **Orna Grumberg**; Technion;
Research Area: Logic & Algorithms; Dates: 7/7/96 - 8/3/96

- **Vladimir Gurvich**; Technion;
Research Area: General; Dates: 7/8/96 - 8/31/96

- **Leonid Gurvits**; NEC Research Institute;
Research Area: General; Dates: 10/24/96 - 6/30/97

- **Armin Haken**; Czech Academy of Science;
Research Area: General; Dates: 8/1/96 - 7/31/97

- **Peter Hamburger**; Indiana-Purdue University at Fort Wayne;
Research Area: General; Dates: 9/23/96 - 11/23/96

- **Neil Immerman**; Cornell University;
Research Area: Logic & Algorithms; Dates: 10/1/95 - 10/13/95

- **Guoping Jin**; University of Cambridge, England;
Research Area: Discrete Probability; Dates: 12/7/96 - 12/21/96

- **Istvan Juhasz**; Miami University;
Research Area: General; Dates: 5/14/96 - 6/15/96

- **Gil Kalai**; Hebrew University;
Research Area: General; Dates: 7/1/95 - 9/15/95

- **Bruce Kapron**; University of Victoria;
Research Area: Logic & Algorithms; Dates: 7/14/96 - 7/31/96

- **Marek Karpinski**; University of Bonn;
  Research Area: General; Dates: 3/16/96 - 3/24/96
- **Alexander Kelmans**; University of Puerto Rico;
  Research Area: General; Dates: 6/8/96 - 8/8/96
- **Pascal Koiran**; Laboratoire de l'Informatique du Parallelisme, Ecole Normale Superieure de Lyon;
  Research Area: General; Dates: 4/22/96 - 6/22/96
- **Phokion Kolaitis**; University of California, Santa Cruz;
  Research Area: Logic & Algorithms; Dates: 7/15/96 - 7/26/96
- **Alexandr Kostochka**; Laboratory of Discrete Extremal Problems Institute of Mathematics, RAS;
  Research Area: General; Dates: 10/3/96 - 11/3/96
- **Jan Krajicek**; Academy of Sciences, Prague;
  Research Area: Logic & Algorithms; Dates: 4/4/96 - 4/25/96
- **Brenda Latka**; Lafayette College;
  Research Area: Logic & Algorithms; Dates: 9/1/95 - 8/31/96
- **Vladimir Levin**; Middle East Technical University;
  Research Area: Logic & Algorithms; Dates: 7/17/96 - 8/6/96
- **Nathan Linial**; Hebrew University;
  Research Area: General; Dates: 10/24/95 - 11/16/95
- **James Lynch**; Clarkson University;
  Research Area: Logic & Algorithms; Dates: 1/2/96 - 6/30/96
- **Oded Maler**; VERIMAG;
  Research Area: Logic & Algorithms; Dates: 11/6/95 - 11/16/95
- **Dominic Mayers**; University of Montreal;
  Research Area: General; Dates: 4/29/96 - 5/9/96
- **Monica McArthur**; Indiana University;
  Research Area: Logic & Algorithms; Dates: 7/1/96 - 7/26/96
- **Angus McIntyre**; Oxford University;
  Research Area: General; Dates: 9/1/95 - 9/30/95
- **Fred R. McMorris**; University of Louisville;
  Research Area: Molecular Biology; Dates: 11/11/96 - 11/16/96
- **Roy Meshulam**; Technion;
  Research Area: General; Dates: 9/1/96 - 10/15/96
- **S. Muthukrishnan**; University of Warwick;
  Research Area: Molecular Biology; Dates: 7/15/96 - 9/30/96
- **Leszek Pacholski**; Institute of Computer Science;
  Research Area: Logic & Algorithms; Dates: 6/3/96 - 8/2/96

- **Roderick Page**; University of Glasgow;
  Research Area: Molecular Biology; Dates: 11/8/96 - 11/17/96
- **Panos Pardalos**; University of Florida;
  Research Area: Logic & Algorithms; Dates: 3/9/96 - 3/15/96
- **Rohit Parikh**; City University of New York;
  Research Area: Logic & Algorithms; Dates: 7/8/96 - 8/2/96
- **Peter Perkins**; College of the Holy Cross;
  Research Area: General; Dates: 9/1/94 - 8/31/95
- **Stephen Ponzio**; Hebrew University;
  Research Area: Logic & Algorithms; Dates: 9/25/95 - 10/20/95
- **Pavel Pudlak**; Academy of Sciences, Prague;
  Research Area: Logic & Algorithms; Dates: 4/4/96 - 5/2/96
- **A.J. Radcliffe**; University of Nebraska at Lincoln;
  Research Area: Discrete Probability; Dates: 12/1/96 - 12/8/96
- **Alexander Razborov**; Steklov Mathematical Institute;
  Research Area: Logic & Algorithms; Dates: 4/18/96 - 5/18/96
- **Soren Riis**; Leeds University, UK;
  Research Area: Logic & Algorithms; Dates: 4/8/96 - 4/27/96
- **Yosi Rinott**; University of California at San Diego;
  Research Area: General; Dates: 9/10/96 - 9/16/96
- **Neil Robertson**; Ohio State University;
  Research Area: General; Dates: 6/20/96 - 9/13/96
- **Eric Rosen**; University of Pennsylvania;
  Research Area: Logic & Algorithms; Dates: 1/12/96 - 2/12/96
- **Miklos Ruszinko**; Szamitastechnikai Intezet (Institute for Computer Technology), Budapest, Hungary;
  Research Area: Discrete Probability; Dates: 12/1/96 - 12/10/96
- **Vladimir Sazonov**; Artificial Intelligence Research Center PSI RAS;
  Research Area: Logic & Algorithms; Dates: 1/16/96 - 4/30/96
- **Richard Schelp**; University of Memphis;
  Research Area: Discrete Probability; Dates: 12/5/96 - 12/8/96
- **Holger Schlinghoff**; Institut fur Informatik der Tu;
  Research Area: Logic & Algorithms; Dates: 8/23/95 - 9/5/95
- **A.D. Scott**; University College, London;
  Research Area: Discrete Probability; Dates: 12/12/96 - 12/20/96
- **Anil Seth**; The Institute of Mathematical Sciences of India;
  Research Area: Logic & Algorithms; Dates: 4/15/96 - 7/30/96

- **Jiri Sgall**; Mathematical Institute, AVCR;
Research Area: Logic & Algorithms; Dates: 4/15/96 - 4/25/96

- **Thomas Shermer**; Simon Fraser University;
Research Area: General; Dates: 2/12/96 - 6/15/96

- **Miklos Simonovits**; Mathematical Institute of the Hungarian Academy of Sciences;
Research Area: Discrete Probability; Dates: 4/16/96 - 5/16/96; 11/16/96 - 11/26/96

- **Tamas Szonyi**; Attila Jozsef University;
Research Area: General; Dates: 3/26/96 - 4/5/96

- **P.S. Thiagarajan**; SPIC Science Foundation, India;
Research Area: Logic & Algorithms; Dates: 7/1/96 - 7/31/96

- **Helmut Thiele**; University of Dortmund;
Research Area: General; Dates: 10/1/95 - 10/30/95

- **Robin Thomas**; Georgia Tech;
Research Area: General; Dates: 6/20/96 - 9/13/96

- **Mikkel Thorup**; University of Copenhagen;
Research Area: Molecular Biology; Dates: 6/11/96 - 6/16/96

- **Ann Trenk**; Wellesley College;
Research Area: General; Dates: 9/1/95 - 8/30/96

- **Jerzy Tyszkiewicz**; Mathematische Grundlagen der Informatik;
Research Area: Logic & Algorithms; Dates: 10/30/95 - 11/28/95

- **Juoko Vaananen**; University of Helsinki;
Research Area: Logic & Algorithms; Dates: 12/12/95 - 1/20/96

- **Jacob van den Berg**; CWI;
Research Area: Discrete Probability; Dates: 1/10/97 - 6/2/97

- **Ron van der Meyden**; Weizmann;
Research Area: Logic & Algorithms; Dates: 6/26/96 - 8/2/96

- **Moshe Vardi**; Rice University;
Research Area: Logic & Algorithms;
Dates: 8/27/96 - 9/1/95; 12/2/95 - 12/5/95; 1/11/96 - 1/17/96; 3/26/96 - 3/28/96; 5/13/96 - 5/17/96; 5/28/96 - 8/17/96

- **Vijay Vazirani**; Georgia Tech;
Research Area: General; Dates: 9/1/96 - 10/15/96

- **Rineke Verbrugge**; Vrije University, Amsterdam;
Research Area: Logic & Algorithms; Dates: 8/13/95 - 8/19/95

- **Igor Walukiewicz**; University of Aarhus;
Research Area: Logic & Algorithms; Dates: 3/3/96 - 3/31/96

- **Farn Wang**; Institute of Information Science Academia Sinica;
Research Area: Logic & Algorithms; Dates: 6/13/96 - 7/5/96

- **Jinchang Wang**; Missouri Western University;
Research Area: General; Dates: 7/1/96 - 7/31/96

- **Rafe Wenger**; Ohio State University;
Research Area: General; Dates: 3/19/96 - 3/21/96

- **Avi Wigderson**; Hebrew University;
Research Area: General; Dates: 9/1/95 - 8/31/96

- **Amiram Yehudi**; Tel-Aviv University;
Research Area: General; Dates: 8/1/96 - 9/30/96

- **Young-jin Yoon**; Kunsan National University;
Research Area: General; Dates: 6/15/96 - 8/20/96

- **Sergio Yovine**; VERIMAG, France;
Research Area: Logic & Algorithms; Dates: 10/20/95 - 11/5/95; 7/12/96 - 7/26/96

- **David Zuckerman**; University of Texas at Austin;
Research Area: General; Dates: 6/3/96 - 6/14/96

4 Special Year, Focus and Workshop Programs

4.1 Special Year on Mathematical Support for Molecular Biology, 1994-97

Overview: This program began in August 1994 as a special year intended to last 25 months. With significant additional external funding of some $525K plus additional funds from Rutgers, the center was able to plan an extensive series of tutorials and workshops, to support 4 postdoctoral fellows for two year appointments, host several long term visitors, and host a distinguished lecturer series.

The program is now extending into a third year of activity, aided by an additional grant of $100K from the National Center for Human Genome Research. The current year has two new postdoctoral fellows, several workshops, and an active seminar series. The first two years built up several collaborative relationships with other centers and departments in the biological sciences, and this year is building on those relationships.

Organizers:

- Fred Roberts (Chair), Rutgers
- Joachim Messing(Chair), Waksman Institute - Rutgers
- Lawrence Shepp (Co-Chair), AT&T Bell Labs
- Michael Waterman (Co-Chair), University of Southern California
- Martin Farach (Assistant Chair), Rutgers
- Sampath Kannan (Publicity Chair), University of Pennsylvania

Steering Committee: David Axelrod, Rutgers; Michael Bulmer, Rutgers; Charles Cantor, Boston University; Daniel Gusfield, UC - Davis; Leroy Hood, Center for Molecular Biotechnology, Washington; Richard Karp, Washington; Eric Lander, MIT - Whitehead Institute; Eugene Lawler, UC - Berkeley; Joshua Lederberg, Rockefeller; Joseph Naus, Rutgers: Michiel Noordewier, Rutgers: Wilma Olson, Rutgers: David B. Searls, University of Pennsylvania; William Sofer, Rutgers; Tandy Warnow, University of Pennsylvania; Peter Winkler, AT&T Bell Laboratories; and Andrew Yao, Princeton.

Collaborating Centers and Institutes:

- Ecology & Evolutionary Biology Department, Princeton University
- Genome Center, University of Pennsylvania
- Molecular Biology Department, Princeton University
- Center for Molecular Biophysics and Biophysical Chemistry, Rutgers University
- Center for Molecular Biotechnology, University of Washington
- Program in Mathematics and Molecular Biology, UC - Berkeley
- Center for Theoretical and Applied Genetics, Rutgers University
- Mt. Sinai School of Medicine
- Waksman Institute for Molecular Genetics, Rutgers University

Theme of program:

Discrete structures have been used to formulate some of the most fundamental concepts of molecular biology. It is now well-known that information storage within a cell is by means of long nucleic molecules which can be thought of as long strings of smaller units called nucleotides. Some of the most important problems of modern molecular biology involve combinatorial and algorithmic questions about these strings. Models of proteins, DNA, and RNA molecules as sequences form a basic tool of molecular biology.

Throughout its history, molecular biology has interacted with methods of discrete mathematics and theoretical computer science. The first nucleic acid sequence was determined in 1965 by R.W. Holley and his co-workers at Cornell University (it contained 77 bases). The method used was the fragmentation stratagem, which involved the combinatorial problem of reconstructing an unknown sequence from fragments obtained by certain enzyme decompositions. Graph-theoretical methods, using eulerian chains, were developed by Hutchinson to give algorithms for implementing the fragmentation stratagem. Of course, these methods are not used any more, and indeed were used only for a short time before other, more efficient methods were adopted. However, they played an important role in the development of molecular biology. In this special year, we are investigating similar uses of DM/TCS in molecular biology.

The special year has a number of specific goals. These are to expose a large community of discrete mathematicians/theoretical computer scientists to problems of molecular biology that are fundamentally DM/CS problems; provide an opportunity for intensive collaboration to discrete mathematicians/theoretical computer scientists who have already discovered these problems; introduce some outstanding young people to the field in a very concentrated way; forge lasting partnerships between mathematical scientists and biological scientists; involve biological scientists in a primary role in defining the agenda of computational biology; and introduce molecular biologists to a community of mathematical scientists who are interested in helping them with their problems.

Workshops:

We provide detailed information here about workshops held since September 1995. During the period August 1994 to August 1995, we held a tutorial series on mathematical support for molecular biology, workshops on mapping and sequencing, sequence alignment, and phylogeny, an international conference on HIV sequence variation and statistical methods, and mini-workshops on combinatorial structures in molecular biology, DNA topology and structure, global minimization of nonconvex energy functions, sequence-based methods for protein folding, DNA based computers, and geometrical methods for conformational modeling. Details about these events can be found in the DIMACS web pages and in earlier reports.

- **4th DIMACS Algorithm Implementation Challenge: Two Problems in Computational Biology: Fragment Assembly and Genome Rearrangements**
  - Dates: Sept 14 - 15, 1995
  - Location: Rutgers University
  - Attendance: 55
  - Organizers: Michael Waterman, University of Southern California; Martin Vingron, German National Research Center for Computer Science (chair); Ellson Chen, Applied Biosystems; Sorin Istrail, Sandia National Laboratory; David Johnson, AT&T Bell Laboratories; John Kececioglu, UC - Davis; Joachim Messing, Rutgers; Joseph Nadeau, Jackson Laboratory; Pavel Pevzner, Pennsylvania State University; Peter Rice, Sanger Center

There still is a large gap in Computational Molecular Biology between biological application and algorithmic work aimed at answering biological questions. The goal of this implementation challenge was to help bridge this gap. *Sequence assembly* is the problem of reconstructing a DNA sequence knowing only the sequences of many overlapping fragments of it. *Genome rearrangements* is the problem of comparing two orders of genes under the assumption that one order is transformed into another by a series of re-arrangements. For both problems there exist algorithms and software that in the case of fragment assembly are widely used. The Challenge was a year long program that began with these problems; participants developed a variety of projects to test algorithmic approaches, gather useful data sets, and build libraries of data structures to use experimental methods to study these problems.

Ellson Chen, a "cutting edge biological sequencer," defined real data sets and provided intermediate biological feedback. Participants are highly motivated to continue coordinated work on testing their algorithms on common benchmmark data. The addition of further, large, and difficult data sets was agreed upon. A major result was that scientists from industry not only participated actively, but helped with the benchmark data sets by giving away their raw data at a very early stage – something rather unusual. We received data from Applied Biosystems, a Division of Perkin-Elmer, and Genome Therapeutics and anticipate more from the Institute for Genomic Research. These are among the leading genome sequencing companies. Since participants from both academia and industry tested their programs in the same way, there was a lively discussion. It has become apparent that current experimental techniques lead to a need for extremely powerful software and this observation is now strongly influencing ongoing research.

- **Mini-workshop on Mathematics of Drug Discovery**
  - Dates: Sept. 21-22, 1995
  - Location: Rutgers University
  - Attendance: 75
  - Organizers: Tim Havel, Harvard and I. G. Kuntz, UC - San Francisco

The search for new drugs to combat disease is an endeavor of growing importance, especially as organisms once thought conquered begin to develop immunity to existing compounds. For many years, the chief strategy involved finding compounds by screening large numbers of naturally occurring substances for activity. The active compounds were then identified, and synthesized in hundreds or thousands of variations on their basic structure until something acceptable was found. This is often described as the "shotgun" approach.

Attempts to predict biological activity from chemical structure have led to a veritable alchemy of correlations, which continue today under the name of "quantitative structure activity relations", or QSAR. With the advent of reliable methods of determining spatial molecular structure and inexpensive high-speed computers, the field of 3-D structure-based drug design was born. To date, however, the lack of a simple and reliable method of predicting the binding affinity of a "ligand" to a "receptor" protein of known structure (molecular complementarity) has limited its success. More recently, methods based on "combinatorial chemistry" and "test-tube evolution" have been developed. These promise to make it possible to screen large numbers of compounds simultaneously and/or systematically search for those with improved activity. In each of these diverse approaches, significant computational and mathematical problems exist which limit their applicability and effectiveness. This mini-workshop explored recent developments in such fields as discrete and computational geometry, statistics, data analysis and systems theory, which may prove applicable to any or all of the above approaches, while at the same time helping mathematicians and computer scientists to identify the relevant problems.

The majority of speakers and participants were from industry. There was financial sponsorship from Biosym Tech., Ciba Geigy, Merck Research, and SmithKline Beecham. This mini-workshop was designed to acquaint pharmaceutical and computational chemists with the latest developments in such fields as discrete and computational geometry and to introduce discrete mathematicians and theoretical computer scientists to the problems of the field coming to be known as "combinatorial chemistry," a field that so far consists largely of heuristic algorithms based upon chemical intuition. The mini-workshop led to lots of discussion and to some promising beginning contacts, for example one between a Chiron Corp. chemist and a DIMACS mathematician who have common interests and ideas about the use of large-scale cluster analysis in drug design.

- **Mini-workshop on Gene Finding & Gene Prediction**
  - Dates: Oct. 14-15, 1995
  - Location: University of Pennsylvania
  - Attendance: 100
  - Organizers: David Searls, Genome Center, University of Pennsylvania; Jim Fickett, Los Alamos; Michiel Noordewier, Waksman Institute, Rutgers; and Gary Stormo, Colorado

The Workshop on Gene-Finding and Gene Structure Prediction was concerned with the increasingly important activity in computational biology of discovering protein-encoding genes in otherwise uncharacterized primary sequence data. This has traditionally been done in genomic sequences by discriminating likely coding regions based on a variety of statistical analyses and by detection of landmark sequences such as splice junctions. Recent approaches have involved combination of such evidence using rule-based and/or connectionist architectures, and have dealt in a variety of ways with the combinatorial problem of exon assembly (dynamic programming, clustering, etc.) The recent profusion of expressed-sequence data and related techniques has also raised new issues and opportunities. This workshop explored topics such as compositional measures of exonic tendency (including approaches founded in statistics, information theory, and signal processing), the effects of genome heterogeneity, the role of models of biological signals and processes, dealing with incomplete and error-prone sequence data, algorithmic and probabilistic techniques, and similarity-based gene prediction. Problems of interest include detecting coding sequences and assembling gene models from large-scale genomic sequence, collections of expressed sequence fragments, and sets of putative exons from a region. Practical issues of interest included dataset and performance metric standardization, annotation of genome databases, and software interoperability. Major funding for this mini-workshop was provided by SmithKline Beecham.

- **Second Sandia National Laboratories Workshop on Computational Molecular Biology**
  - Dates: March 4-6, 1996
  - Location: Albuquerque, NM
  - Attendance: 135
  - Organizers: Sarina Bromberg, David Greenberg, William E. Hart, Sorin Istrail, and Cynthia Phillips (all of Sandia National Laboratories)

The workshop, jointly sponsored by DIMACS, focused on open problems and research directions in Computational Molecular Biology. In addition to talks given by active researchers in this field, several plenary lectures were presented by speakers of highest distinction in computational molecular biology. Plenary speakers were Ken Dill, Richard Karp, Martin Karplus, Jonathon King, Irwin Kuntz, and Michael Waterman.

- **Workshop on Computational Biology** organized in collaboration with the University of Pennsylvania as part of the 50th Anniversary of ENIAC
  - Dates: May 17-19, 1996
  - Location: Princeton University
  - Attendance: 90
  - Organizers: Tandy Warnow, University of Pennsylvania; Craig Benham, Mt. Sinai; Sue Davidson, University of Pennsylvania; Warren Ewens, University of Pennsylvania; Martin Farach, Rutgers; Sampath Kannan, University of Pennsylvania

  The conference was based around sessions focusing on important research areas within the field of computational biology. The special research topics included: Structure Prediction, Tandem Repeats, Databases, Sequencing and Mapping, and Genome Rearrangements. There was also a special session on interdisciplinary education in computational biology, in which workshop organizers and students in various training programs throughout the country participated.

- **2nd DIMACS Workshop on DNA Based Computers**
  - Dates: June 10-12, 1996
  - Location: Princeton University
  - Attendance: 110
  - Organizers: Eric Baum, NEC; Dan Boneh, Princeton; Peter Kaplan, Princeton; Richard Lipton, Princeton; John Reif, Duke; Nadrian Seeman, New York University

  The workshop hosted papers in all areas that relate directly to computing with DNA, including: Algorithms, applications, techniques, architectures, practical obstacles to DNA based computers, proposed solutions to such obstacles, computational processes in vivo, and relevant ideas regarding biological evolution.

- **Workshop on Mathematical Hierarchy and Biology**
  - Dates: Nov. 13-15, 1996
  - Location: Rutgers University
  - Attendance: 75
  - Organizers: Martin Farach, Rutgers; Boris Mirkin, DIMACS; F.R. McMorris, Louisville; Fred Roberts, Rutgers; Andrey Rzhetsky, Pennsylvania State University

  Viewing hierarchies as combinatorial objects underlies many important areas of current scientific investigation. Examples of areas where this is the case are biology, evolutionary studies and taxonomy, general classification theory, data/knowledge bases, cognitive models and linguistical structures. The mathematical study of hierarchies is being pursued in many disciplines, such as computer science, statistics, biology, psychology, and, of course, mathematics. Topics of presentations included the representation and enumeration of various types of hierarchies, the comparison of hierarchies, constructing hierarchies from various kinds of data, modeling of hierarchical behavior, etc.

- **Computational and Physical Aspects of Protein Folding and Related Problems** as part of the semi-annual Statistical Mechanics Conference at Rutgers.
  - Dates: Dec. 15-17, 1996
  - Location: Rutgers University
  - Attendance: TBD
  - Organizers: Joel L. Lebowitz, Rutgers; Shura Grosberg, MIT; Mona Singh, DIMACS; Scott Decatur, DIMACS

  This special session reviewed physical aspects of biological problems, with an emphasis on protein folding. There were talks on random walks on microtubules, DNA and recognition, why proteins look like proteins, the statistical mechanics of protein folding, cellular protein folding, combinatorial methods for protein design, thermodynamics vs. kinematics in protein folding, correlations and heterogeneity of model protein and heteropolymer energy landscapes, conformations of charged heteropolymers, freezing transition of compact polyampholytes, spectral analysis of biological sequences, specificity and affinity of biomolecular interactions, modeling cellular guts with colloidal soups, cell dynamics of model proteins, and physics from jelly fish.

Workshops Planned for the Rest of 1996-97:

- **Mini-Workshop on Protein Structure**
  - Winter 1997: Dates to be determined
  - Location: To be announced
  - Organizers: Mona Singh, DIMACS; Scott Decatur, DIMACS
- **Workshop on DNA Topology II**
  - Date: April 3-4, 1997
  - Location: Rutgers University
  - Organizers: Wilma Olson, Rutgers; Bernard Coleman, Rutgers; Tamar Schlick, New York University; Irwin Tobias, Rutgers
- **Second DIMACS Workshop on Phylogeny**
  - Spring/Summer 1997: Dates to be determined
  - Location: To be announced
  - Organizer: Martin Farach, Rutgers
- **Third DIMACS Workshop on DNA Based Computers**
  - June 1997
  - Location: University of Pennsylvania
  - Organizers: Harvey Rubin, University of Pennsylvania (chair); Len Adleman, University of Southern California; Carter Bancroft, Mt. Sinai; Anne Condon, Wisconsin; Richard Lipton, Princeton; David Wood, Delaware

4.2 Special Year on Logic and Algorithms, 1995-96

Overview:

The special year on logic and algorithms started in August 1995 with a Summer School on Applied Logic with international participation and ended with the center hosting FLOC'96 - four international logic related conferences federating together in August 1996. During the year there were 12 workshops on topics ranging from CAV to proof complexity. Over 60 visitors, several hundred workshop attendees, and many local scientists participated in the research programs. In summer 1997, the center is planning a workshop to follow up on the results that have been continuing to come during the current academic year.

Organizers:

- **Eric Allender (Co-Chair), Rutgers**
- **Bob Kurshan, (Co-Chair), AT&T Bell Labs**
- **Moshe Vardi (Co-Chair), Rice**
- **Stephen Bloch (Publicity Chair), Adelphi University**

Steering Committee: Paul Beame, Washington; Sam Buss, UC - San Diego; Gregory Cherlin, Rutgers; Ed Clarke, Carnegie-Mellon; Steve Cook, Toronto; Allen Emerson, Texas; Joan Feigenbaum, AT&T Bell Labs; Orna Grumberg, Technion; Phokion Kolaitis, UC - Santa Cruz; Daniel Leivant, Indiana; Richard Lipton, Princeton; Amir Pnueli, Weizmann Institute; and Peter Winkler, AT&T Bell Labs.

Industrial Advisory Board: Hao Nham, AT&T QUEST; Paul Loewenstein, Sun Microsystems; Carl Pixley, Motorola; Zeev Shtadler, Intel; and Kurt Keutzer, Synopsys.

Theme of program: A dichotomy in theoretical computer science is best demonstrated by looking at the 1994 Handbook of Theoretical Computer Science. Volume A discusses algorithms and complexity, while Volume B treats formal models and semantics. Theoretical computer science in the United States is largely "Vol. A"-ish, while European theoretical computer science is largely "Vol. B"-ish. The goal of this Special Year was to bridge the gap between the two branches, focusing on three bridge areas: Computer-Aided Verification, Finite-Model Theory, and Proof Complexity. All three are emerging research areas that fit naturally between Vol. A and Vol. B.

Below is a brief description of each topic. We started the year with a series of week-long tutorials, one in each topic, intended to introduce students, recent graduates, and professionals from other areas to the topic.

Computer-Aided Verification: Computer-Aided Verification studies algorithms and structures for verifying properties of programs. It draws upon techniques from graph theory, combinatorics, automata theory, complexity theory, Boolean functions and algebras, logic, Ramsey theory and linear programming. Since the DIMACS CAV workshop in 1990, worldwide interest in CAV has grown enormously, not only in academia but in companies like Intel, DEC, SGI, Motorola, Sun and AT&T. This creates an unusual and rewarding opportunity to see theory put directly into practice.

Finite-Model Theory: Model theory is the study of mathematical structures which satisfy sets of axioms. Recent work on the finite models of a set of axioms has yielded elegant connections with theoretical computer science, including model-theoretic characterizations of complexity classes. Further, when a class of finite mathematical structures (e.g. graphs) is equipped with probability measures, one can often develop powerful meta-theorems called zero-one laws, which give conditions under which probabilities must approach zero or one as the structure size goes to infinity.

Proof Complexity: Two related notions of "proof complexity" currently motivate research at the interface between computer science and logic. One notion centers on the length of a proof, and the other on the complexity of the inference steps within the proof. It is well known that NP=co-NP iff all propositional tautologies have short proofs. But the connection between proof length and complexity theory goes much deeper. Lower bounds on circuits are closely tied to those on proof length in restricted systems, and advances on one front often lead quickly to progress on the other. By restricting the complexity of inference steps within a proof, one obtains a fragment of Peano Arithmetic called Bounded Arithmetic, which defines exactly the predicates in the polynomial hierarchy. Exciting recent work has shown that if certain theories of bounded arithmetic can prove lower bounds in complexity theory, then corresponding cryptographic systems cannot be secure.

Workshops:

- **Summer School in Applied Logic:**
  - Dates: August 2 - Sept. 1, 1995
  - Location: Rutgers University
  - Attendance: 55 for each of three weeks
  - Organizers: Special Year Committee

  This three week program consisted of one week each of finite model theory, proof complexity, and computer aided verification. The program was designed to allow scientists with knowledge of one area to quickly grasp the principal models, problems and approaches of other areas. The program attracted world-wide attendance.

- **Workshop on Verification and Control of Hybrid Systems**
  - Dates: Oct. 22-25, 1995
  - Location: Rutgers University
  - Attendance: 130
  - Organizers: Rajeev Alur, AT&T; Tom Henzinger, Cornell; Eduardo Sontag, Rutgers

  A *hybrid system* consists of a discrete program embedded in a continuously changing environment. While the continuous dynamics has been studied in control theory and the discrete aspects are well understood in computer science, the combination poses new problems. Reliable design of many real-world systems, such as automated highways, demands solutions to these problems. The goal of the workshop was to bring together two communities, computer scientists and control theorists, who usually have little interaction. The response to this marriage between communities was, in the words of the workshop organizers, "overwhelmingly positive." There were 130 participants from computer science, electrical engineering, mathematics, and industry. The workshop "was a significant step in finding plausible solutions by bridging the gap between computer-science and control-theory approaches." It was also realized that hybrid systems is an emerging area of increased importance and consequently a steering committee was formed to ensure that an annual workshop on hybrid systems will be organized in the future. The organizers "believe that the hybrid systems workshop will become a significant annual event, and this year's DIMACS workshop will be remembered as its birthplace."

- **Workshop on Logic and Random Structures**
  - Dates: Nov. 5-7, 1995
  - Location: Rutgers University

- Attendance: 70
- Organizers: Ravi Boppana, NYU; James Lynch, Clarkson; Kevin Compton, University of Michigan

The central theme was the relationship between logic and probabilistic techniques in the study of finite structures. In one direction, this topic includes recent research on limit laws and zero-one laws. A limit law holds in a class of finite structures if all properties describable in some logical language have limiting probabilities as structure size grows; a zero-one law holds (as in the classic work of Glebskii et. al. and Fagin) if the limiting probabilities are always 0 or 1. This work has been applied to areas such as analysis of algorithms for database query optimization and polynomial time approximation of NP optimization problems. In the other direction, this topic covers use of probabilistic methods to establish lower bounds in circuit complexity. Particular emphasis was placed on the connections between first-order logic and constant depth circuits with unbounded fan-in.

- **Finite Models and Descriptive Complexity**
  - Dates: Jan. 14-17, 1996
  - Location: Princeton University
  - Attendance: 96
  - Organizers: Neil Immerman, University of Massachusetts; Phokion Kolaitis, UC - Santa Cruz

The main aim of descriptive complexity is to classify algorithmic problems according to the resources needed to express such problems in various logics on finite models. This area has been developing through a continuous interaction between computational complexity, finite model theory, database theory, and combinatorics. Research in descriptive complexity has amplified the close relationship between logic and computation, as evidenced by natural characterizations of major complexity classes in terms of logical expressibility on finite models. The workshop brought together contributors to all of these fields.

- **Workshop on the Satisfiability Problem: Theory and Applications**
  - Dates: March 11-13, 1996
  - Location: Rutgers University
  - Attendance: 65
  - Organizers: Ding-Zhu Du, University of Minnesota; Jun Gu, University of Calgary; Panos Pardalos, University of Florida

The satisfiability (SAT) problem is central in mathematical logic, computing theory, and many industrial application problems. There has been a strong relationship between the theory, algorithms, and the applications of the SAT problem. The main focus of this workshop was to bring together the best theorists, algorithmists, and practitioners working on the SAT problem and on the industrial applications involving the SAT problem, enhancing the interaction among the three research groups. This workshop featured the application of theoretical/algorithmic results to practical problems as well as the presentation of practical problems for theoretical/algorithmic study. Major topics covered in the workshop included practical and industrial SAT problems and benchmarks, significant case studies and practical applications of the SAT problem and SAT algorithms, new algorithms and improved techniques for satisfiability testing, specific data structures and implementation details of the SAT algorithms, and the theoretical study of the SAT problem and SAT algorithms.

- **Workshop on Computational and Complexity Issues in Automated Verification**
  - Dates: March 25-28, 1996
  - Location: Rutgers University
  - Attendance: 55
  - Organizers: Bob Brayton, UC - Berkeley; Allen Emerson, University of Texas; Joan Feigenbaum, AT&T Bell Labs

  The correctness of computer hardware and software is an area of growing theoretical interest as well as practical importance. It is now widely acknowledged that the best way to reason about program correctness hinges on two criteria: (i) the use of appropriate formalisms, such as temporal logic or automata, for rigorously specifying correct behavior; (ii) the use of mechanical reasoning algorithms, such as model checkers, to permit proofs of correctness to be constructed automatically that could not reasonably be constructed by hand owing to the intractable amount of tedious detail.

  A principal limiting factor to automated verification is the computational complexity of various associated mechanical reasoning problems, which might be prohibitively high due to the combinatorial state explosion problem in various guises. The workshop focused on such topics as the theoretical complexity as well as the practical efficiency of model checking algorithms and of testing satisfiability of logics, both in general and in applications-oriented special cases. Other methods for overcoming state explosion such as abstraction through the use of simulations and bisimulation relations were also presented. Other talks examined the use of compositionality and modularity to simplify the construction of a correct program from correct constituent subprograms, program checking, testing and correcting, and program synthesis. A portion of the workshop was devoted to tools for verification and testing, evaluation of their strengths versus weaknesses, and reports on real experiences with industrial applications.

  An "Industrial and Systems Day" at the workshop featured a variety of talks by industrial speakers, in particular an interesting presentation on the "Formalcheck/Cospan" verification tool by G. de Palma of AT&T (now Lucent), who provided a genuinely industrial perspective to the audience. Other speakers represented such companies as Intel, Motorola, and IBM-Haiffa.

- **Workshop on Feasible Arithmetics and Length of Proofs**
  - Dates: April 21-23, 1996
  - Location: Rutgers University
  - Attendance: 50
  - Organizers: Paul Beame, University of Washington; Sam Buss, UC - San Diego

  This workshop focused on two complementary approaches to the complexity of proving a mathematical fact (with emphasis on their connections to circuit complexity and computational complexity). Feasible theories of arithmetic are weak first- or second-order systems whose theorems correspond to problems in computational classes such as the levels of the polynomial time hierarchy or the NC hierarchy. The "length of proofs" approach, on the other hand, studies the growth rate (in number of symbols, number of lines, logical depth, etc.) of proofs of instances of some theorem within a fixed, quantifier-free proof system. Proof complexity and feasible arithmetics are closely related to each other and to fundamental problems from circuit complexity and computational complexity, such as the NP versus coNP problem. Over the last decade there has been significant progress in the study of these areas and many deep connections between them have been discovered. One example is the application of methods from circuit complexity lower bound proofs to obtain new lower bounds on proof complexity. Another example is the recent work on natural proofs, which has shown that if certain theories of bounded arithmetic can prove lower bounds in complexity theory, then certain levels of cryptographic security cannot be achieved.

Topics addressed included lower bounds for resolution, natural proofs, descriptive complexity, hashing on AC^0 RAMS, monotone circuits, cutting-plane arguments, model theory, and group theoretic methods.

- **Workshop on Controllers for Manufacturing and Automation: Specification, Synthesis, and Verification Issues**
  - Dates: May 13-15, 1996
  - Location: Rutgers University
  - Attendance: 68
  - Organizers: Mohsen A. Jafari, Rutgers (Industrial Engineering); Bud Mishra, NYU

The workshop on Controllers for Manufacturing and Automation brought together researchers working on various aspects of the specification, synthesis, and verification of supervisory control of manufacturing and automation systems. The workshop brought together engineers and quality managers from industry (pharmaceutical, mechanical, chemical, and electrical) and academicians of diverse backgrounds (computer scientists, mathematicians, logicians, control theorists, and engineers). Companies represented included Rockwell, Warner Lambert, Global Pharmaceuticals, Uniview, Siemens, Merck, Smith-Kline-Beecham, Carter-Wallace, Verimag, and Intel. Specific applications discussed by both academic and industrial speakers included intelligent highway systems, chemical processes, food and pharmaceutical applications, robotics, air-traffic control, smart sensors, and operating system design. There was also considerable discussion of the practical aspects of enforcing validation requirements, in particular when they are framed by regulatory agencies such as FDA and cannot be unambiguously interpreted. Other important issues of interest included the difficulty of model building for chemical and mechanical processes, diagnostics and recovery, and interface with legacy controllers. One result of the workshop was a consortium to develop software for controllers that will integrate the available theory to be used to solve practical problems of industrial importance. The interaction among the different types of participants at the workshop was particularly valuable. This workshop received major support from the Office of Naval Research.

- **Workshop on Partial Order Methods in Verification (POMIV'96)**
  - Dates: July 24-26, 1996
  - Location: Princeton University
  - Attendance: 74
  - Organizers: Doron Peled, AT&T; Gerard Holzmann, AT&T; Vaughan Pratt, Stanford University

Research in the last few decades on the connection between partial order semantics and interleaving semantics has resulted in a class of new verification techniques for distributed systems that may prove themselves to be of great practical value in routine industrial applications of formal verification techniques. Much of the new work in this area is being pursued by people from quite distinct fields, who rarely have an opportunity to meet and exchange ideas directly. These groups include, for instance, verification tool builders, logicians, semantic researchers, and process algebra specialists.

The main topics of workshop talks were:

- Experience with the application of verification methods and tools based on partial order semantics and partial order reduction techniques.
- Formal description techniques, logics and algebras, based on partial orders.
- Formal specification methods where partial order semantics are more natural than interleaving semantics.

Automated verification tools based on automata theory and partial order reduction techniques are beginning to gain acceptance as effective tools for distributed system design, and have been integrated into tools such as SPIN and VFSM-valid, which are both used at AT&T and Lucent for software verification. The workshop led to a lively discussion and interaction among theoreticians and practitioners about all aspects of this software verification method. This workshop received major support from the Office of Naval Research.

- **Workshop on Teaching Logic and Reasoning in an Illogical World**
  - Dates: July 25-26, 1996
  - Location: Rutgers University
  - Attendance: 65
  - Organizers: Susanna Epp, DePaul; David Gries, Cornell; Peter Henderson, SUNY - Stony Brook; Ann Yasuhara, Rutgers

  Logic and logical thinking are central to all disciplines and are critical in the mathematical and computer sciences. This symposium explored the teaching of introductory logic and logical thinking, with a primary focus on the college level and a secondary focus on the high school level. The symposium was interdisciplinary, emphasizing and contrasting approaches used in mathematics, computer science, natural sciences, and engineering. Topics included:

  - Pedagogical Approaches
  - Cognitive Models of Logical Reasoning
  - Empirical Studies
  - Exemplary Course Material
  - Innovative Approaches
  - Courseware for Teaching Logic

  One offshoot of the workshop was a new DIMACS "working group" on curriculum issues in logic and computer science, under the direction of Moshe Vardi of Rice University.

- **Workshop on Computational Complexity and Programming Languages**
  - Dates: July 25-26, 1996
  - Location: Rutgers University
  - Attendance: 50
  - Organizers: Bruce Kapron, University of Victoria; Jim Royer, Syracuse

This workshop addressed problems arising from higher type programming languages that have not been adequately addressed in computational complexity. The theory of algorithms is basically oriented to the RAM model, while new languages such as ML are introducing higher-type objects for which there is only a rudimentary theory of complexity. Object oriented languages which are increasingly used also lack a precise way to talk about the complexity of a module or object. This workshop invited a group of researchers who have been working in relative isolation to discuss progress on these problems and their different approaches. Speakers included Neil Jones (Copenhagen), Guy Blelloch (CMU), Dan Leivant (Indiana), and Samson Bramsky (Edinburgh).

- **Workshop on SPIN96 – 2nd International SPIN Verification Workshop: Algorithms, Applications, Tool Use, Theory**
  - Dates: August 5, 1996
  - Location: Rutgers University
  - Attendance: 55
  - Organizers: Jean-Charles Gregoire, INRS Telecommunications (Montreal), Gerard J. Holzmann, AT&T, and Doron Peled, AT&T

  SPIN is a reachability analysis tool designed for the verification of distributed systems. The tool was first made available publicly in 1991, and has quickly become a standard in both academia and in industry. In academia the tool is used for both teaching verification techniques, and to support research in, e.g., search algorithms, reduction methods, and tool construction. In industry the tool is used routinely on large scale verification problems. Presentations topics included collision avoidance, real-time SPIN, improving memory efficiency in SPIN, analysis of SA/RT models with SPIN, self-stabilizing protocols, and closure/stuttering analysis.

- **FLoC'96: the 1996 Federated Logic Conference**
  - Dates: July 27-August 3, 1996
  - Location: Rutgers University
  - Attendance: 600
  - Organizers: Jon Riecke, AT&T

  The federated Logic Conference consisted of four conferences on the applied logic area that met together in New Brunswick over an 8 day period. DIMACS was a co-sponsor. The conferences were:

  - CADE (Conference on Automated Deduction)
  - CAV (Conference on Computer-Aided Verification)
  - IEEE LICS (Logic in Computer Science)
  - RTA (Conference on Rewriting Techniques and Applications)

4.3 Special Year on Networks, 1996-97

Overview:

The Special Year on Networks began in July 1996 and is expected to last well into the 1997-98 year. The program is stressing postdoctoral training and workshops that will reach a broad segment of the DM and TCS research community as well as engage practitioners.

Organizers:

- **Stuart Haber, Bellcore / Surety Technologies**
- **David Johnson, AT&T Labs**
- **Mihalis Yannakakis, Bell Labs**

Scientific Advisory Board: Baruch Awerbuch, Johns Hopkins; Br. Badrinath, Rutgers; Endre Boros, Rutgers; Ding-Zhu Du, Minnesota; David Goodman, Rutgers; Jack Holtzman, Rutgers; Frank Hsu, Fordham; Srinivasan Keshav, Cornell; Arjen Lenstra, Citibank; Udi Manber, Arizona; Ray Miller, Maryland; Debasis Mitra, Bell Labs; Clyde Monma, Bellcore; Panos Pardalos, Florida; K.K. Ramakrishnan, AT&T Labs; Greg Shannon, Spanning Tree Technologies; Scott Shenker, Xerox; Subhash Suri, Washington U. in St. Louis; Eva Tardos, Cornell; Peter Winkler, Bell Labs; and Yacov Yacobi, Microsoft

Theme of program:

This special year is intended to bring together researchers and practitioners in a wide variety of fields connected to the common theme of "Networks" and ranging from questions of infrastructure to ones of applications. Special emphasis is placed on Internet-related topics, as this is a young and growing area where theoretical computer science and discrete mathematics have a chance to make a real impact.

At present planning, the Year's topics are grouping themselves into four overlapping areas, which we list below along with examples of planned or possible workshop topics.

- **Network Security:** protocols for electronic commerce, public key infrastructure, network security and threat models.
- **Network Applications:** information retrieval, databases and the Web, multimedia transmission, and coding distributed and collaborative computing.
- **Network Control:** routing, congestion and admission control, network protocols, mobile computing.
- **Network Design:** topology and connectivity, switch design, facilities location.

The goal is to engage theoreticians with these problems and foster communication between them and real-world practitioners. Practitioners have much to tell theorists about what the key issues are, and theorists may themselves be able to provide crucial insights and solutions to practitioners.

Workshops:

- **Workshop on Databases and the WWW**
  - Dates: July 30, 1996
  - Location: Bell Labs, Murray Hill
  - Attendance: 40
  - Organizer: Steven Levy

  The emergence of the World Wide Web as a framework for the global exchange of data is continuing to lead to new applications. The applications already available address electronic commerce, consortial data exchange and market research, to name a few. The fledgling nature of these applications has led to several technical problems, including some that have traditionally been associated with database technology.

The aim of the workshop was to bring together people who are interested in database issues associated with the World Wide Web. The format was flexible, and the setting informal. The focus was on new work in the area.

- **Networks Day**
  - Dates: Sept. 27, 1996
  - Location: Rutgers University
  - Attendance: 100
  - Organizers: Special Year Committee

  This one-day program to launch the special year featured 4 distinguished lecturers:

  - Scott Shenker, Xerox PARC *Incentive Issues in Networks*;
  - Baruch Awerbuch, Johns Hopkins University *Online Algorithms for Maximal Dense Trees and Selective Multicast*;
  - Dan Bienstock, Columbia University *Computational Experience with Network Design Algorithms*; and
  - Matthew Franklin, AT&T Labs *Verifiable Secret Sharing in Electronic Commerce.*

- **Workshop on Trust Management in Networks**
  - Dates: Sept. 30 - Oct. 2, 1996
  - Location: Rutgers University
  - Attendance: 160
  - Organizers: Ernie Brickell, Bankers Trust; Joan Feigenbaum, AT&T; David Maher, AT&T

  This workshop focused on a major problem in implementing public key systems and any of the applications such as banking, commerce, or electronic intellectual property, based on them: How does one know a public key is authentic? The workshop brought together designers and advocates of four major approaches to certificate authentication for talks that explored the strengths and limitations of each. The attendance reflected the broad interest in the topic - the program attracted several people from law, banking, companies involved in building digital commerce applications and other sectors.

- **Tutorials on Networking Technology**
  - Dates: Nov. 5, 1996
  - Location: Rutgers University
  - Attendance: 60
  - Organizers: David Johnson and K. K. Ramakrishnan both of AT&T Labs

  This one day program had two tutorial lectures to explain workings of networking. Henning Schulzrinne of Columbia University described Internet protocols to introduce the subject to theorists. He described architecture, addressing and routing, IP, TCP and UDP transport protocols, TCP congestion control issues. He discussed applications such as real-time audio and their implementation on the Internet. Anujan Varma of UC - Santa Cruz introduced ATM networking. He described the history, key protocols and control mechanisms, and challenges for ATM.

- **Workshop on Performance of Real-time Applications on the Internet**
  - Dates: Nov. 6-8, 1996
  - Location: Rutgers University
  - Attendance: 140
  - Organizers: Ramon Caceres, Lucent Technologies; Nick Duffield, AT&T; Mario Garzia, AT&T; Christoph Lindemann, Technical University of Berlin; Mike Luby, DEC/SRC and ICSI/UC Berkeley; Kishor Trivedi, Duke University

  This workshop examined underlying issues raised by Internet applications with real-time requirements such as telephony, video, commerce, and web-searching. With the premise that the success of these applications depends on performance, quality of service, cost, security, and other criteria, the workshop invited presentations addressing a broad range of issues. Talks focused on such aspects as: Performance requirements for electronic commerce, multi-media coordination requirements, service discrimination, traffic characterization, measurements of existing applications, stochastic analysis of real-time protocols, among others.

- **Workshop on Network Threats**
  - Dates: Dec. 4-6, 1996
  - Location: Rutgers University
  - Attendance: 96
  - Organizers: Rebecca Wright, AT&T; Peter Neumann, SRI; Steve Bellovin, AT&T

  This workshop focused on security related threats that are emerging in the Internet: Netscape, Java, Internet protocols, and electronic commerce all are demonstrating weaknesses. The workshop addressed ways to manage that risk. Topics included attacks on network security, prevention/detection of attacks, threat models, threats to individual privacy, risk management, formal methods of security analysis, and political, legal, and social aspects of network security.

Workshops Planned for the Rest of 1996-97

- **Workshop on Architecture and Algorithmic Aspects of Communication Networks**
  - Date: January 26-28, 1997
  - Location: Rutgers University
  - Organizers: Eva Tardos, Cornell University; Serge Plotkin, Stanford University; Israel Cidon, Technion and Sun Microsystems
- **Workshop on Selected Topics Related to the Electronic Marketplace**
  - Date: Spring, 1997 - Dates to be announced
  - Location: TBA
  - Organizers: Matt Franklin, AT&T; Avi Rubin, Bell Labs
- **Workshop on Economics, Game Theory, and the Internet**
  - Date: April 18-19 1997
  - Location: Rutgers University

  - Organizers: Eric Friedman, Rutgers; Scott Shenker, Xerox PARC

- **Workshop on Network Design: Connectivity and Facilities Location**
  - Date: April 28-30, 1997
  - Location: Nassau Inn, Princeton, NJ
  - Organizers: Ding-Zhu Du, University of Minnesota; Panos Pardalos, University of Florida

- **Workshop on Non-blocking and Almost Non-blocking Switching Networks**
  - Date: July 7-9, 1997
  - Location: Princeton University
  - Organizers: Frank Hwang, AT&T; Ding-Zhu Du, University of Minnesota

- **Workshop on Traffic and Network Modelling and Analysis**
  - Date: Spring-Summer, 1997 - Dates to be announced
  - Location: TBA
  - Organizers: Sandeep Bhatt, Bellcore; Andrew Ogielski, Bellcore; Debasis Mitra, Bell Labs (Advisor)

4.4 Special Focus on Discrete Probability, 1996-97

Overview:
This program is scheduled for the 1996 - 1997 academic year. It is jointly organized and sponsored with the Institute for Advanced Study, which is cohosting visitors and four workshops with DIMACS.

Organizers:

- **Enrico Bombieri, Institute for Advanced Study**
- **Jeff Kahn, Rutgers**
- **Peter Winkler, Bell Labs**
- **Anders Johanson (Publicity Chair), Rutgers**

Collaborating Centers and Institutes: This program is jointly organized and funded with the Institute for Advanced Study of Princeton, NJ.

Theme of program:
Probabilistic methods and ideas have by now become central to large parts of discrete mathematics and theoretical computer science, too much so to allow a convenient summary here. At the same time, recent years have seen a number of major developments in parts of probability and statistical mechanics dealing with topics of an essentially discrete nature (e.g. percolation, random walks and Markov chains, particle systems, branching processes).

A major goal of this program is to foster interaction between these communities. This is already happening to some extent, with both the application of increasingly sophisticated tools from probability to DM and TCS, and the solution of problems from probability by combinatorial means. Nonetheless, such interactions, and even awareness of what these communities have to offer each other, are still far more exception than rule. So the program is intended to give a substantial push to a broader and more rapid development of these connections.

Workshops:

- **Two Distinguished Lecturers on Discrete Probability**
  - Dates: October 4, 1996
  - Location: Rutgers University
  - Attendance: 70
  - Organizers: Special Focus Committee

  This program had two presentations:

  - Christian Borgs, Leipzig, *Birth of a Discontinuity: Finite-Size Scaling in Magnetic Systems and Lattice Gases* and
  - Jennifer Chayes, UCLA, *Birth of the Infinite Cluster: Finite-Size Scaling in Percolation.*

- **Workshop on Probabilistic Methods in Discrete Mathematics**
  - Dates: Oct. 14-18, 1996
  - Location: Rutgers University
  - Attendance: 80
  - Organizers: Noga Alon, Tel Aviv; Joel Spencer, NYU

  The Probabilistic Method was developed by Paul Erdos as a technique for proving the existence of a combinatorial object, coloring, tournament, graph or whatever, by proving that an appropriately defined random object has the desired properties with positive probability.

  Workshop presentations covered topics such as the development of powerful methods to deal with graph coloring problems, expander graph constructions and applications to communication and sorting networks, design of error correcting codes, and derandomization methods, rapidly mixing Markov chains, and random walks. Because of the passing of Paul Erdos, a special lecture by Bela Bollobas, describing Erdos' scientific contributions, was added to the program.

Workshops Planned for the Rest of 1996-97

- **Workshop on Statistical Physics Methods in Discrete Probability, Combinatorics and Theoretical Computer Science**
  - Date: March 23-27, 1997
  - Location: Rutgers University
  - Organizers: Jennifer T. Chayes, UCLA; Dana Randall, Georgia Tech

- **Workshop on Probabilistic Analysis of Algorithms**
  - Date: May 11-14, 1997
  - Location: Princeton University
  - Organizers: Alan Frieze, Carnegie-Mellon; Michael Molloy, Toronto

- **Workshop on Microsurveys in Discrete Probability**
  - Date: June 2-6, 1997
  - Location: Institute for Advanced Study, Princeton, New Jersey
  - Organizers: David Aldous, UC - Berkeley; Jim Propp, MIT

4.5 Prespecial Year in Massive Data Sets, 1996-97

Overview:

DIMACS is planning a 1997-98 special year on massive data sets, with prespecial year activities beginning in 1996-97. This will be a complex year. In planning for it, we have identified about 20 faculty members at Rutgers and Princeton from departments outside the DIMACS departments, and we are starting to engage them in a dialogue about their massive data set problems to help to identify common issues. We will be running "get acquainted" events with "users" and are exploring the idea of a tutorial for them on modern techniques from computer science for organizing and accessing data, as well as a workshop. We are also organizing, together with the Division of Mathematical Sciences at NSF, a session on mathematical problems of massive data sets at the American Mathematical Society meeting in San Diego in January 1997.

Organizers:

- **Robert Sedgewick, Princeton (chair)**
- **Jon Bentley, Bell Labs**
- **Joan Feigenbaum, AT&T Labs**
- **Laney Kulsrud, IDA-CCR, Princeton**

Theme of Program:

The emergence of the computer as an essential tool in scientific research and an essential ingredient in commercial systems has led to the emergence of massive amounts of data: Data from astrophysical models, genetic sequencing, geographic information systems, world-wide ecological monitoring, weather prediction, telecommunications applications, biological models for medical applications, government information systems. Researchers in all of these applied areas currently face daunting computational problems in organizing and extracting useful information from these massive data sets. DIMACS' special year will seek to stimulate new research to address these problems. We will seek to learn common properties among diverse applications, identify fundamental computational problems that transcend particular applications, develop and analyze new practical algorithms, establish partnerships with researchers in other areas, with the goal of developing solutions for specific scientific and commercial problems.

4.6 Other Research Workshops, Conferences, and Institutes

Overview:

In addition to workshops that are part of the concentrated activity of special years, DIMACS has other workshops and conferences that are often used to visit topics of special current interest, that might develop into special focus topics, that "revisit" former special focus topics, or that involve new collaborations with other units.

Workshops:

- **Mathematical Association of America, New Jersey Section**
  - Date: November 18, 1995
  - Location: Rutgers University
  - Attendance: 75
  - Hosted by DIMACS

- Organizers: Lawrence D'Antonio, Kay Gura, Ramapo College; Theresa C. Michnowicz, Jersey City State College; Sandeep Mahewhwari, William Paterson College; Judith Lenk, Ocean County College

Welcome for the MAA meeting was given by Fred Roberts. The speakers were Michael Beals, Rutgers University; Sean Mahon, Trenton State College; Charles Peskin, Courant Institute, NYU; Ann Trenk, Wellesley College and DIMACS; and Peter Winkler, AT&T Bell Laboratories. Trenk and Winkler were "DIMACS" speakers.

- **Workshop on Combinatorial Number Theory**
  - Dates: Feb. 5 - 9, 1996
  - Location: Rutgers University
  - Attendance: 85
  - Organizers: Carl Pomerance, University of Georgia; Jeff Lagarias, AT&T; Ronald Graham, AT&T; Mel Nathanson, Rutgers; J.-L. Nicolas, France; A. Sarkozy, Hungary; Endre Szemeredi, Rutgers

This workshop focused on research that has emerged in recent years that explores connections between combinatorics and number theory. Principal topics were: Multiplicative problems, divisors and applications, linear additive problems, probabilistic methods, pseudorandom sequences, and combinatorial aspects of algorithmic number theory and vice versa.

- **The DIMACS Research and Education Institute (DREI): Geometry, Visualization and Computing**
  - Dates: June 23 - July 13, 1996
  - Location: Princeton University
  - Attendance: 30 teachers; 70 researchers.
  - Organizers: Bernard Chazelle, Princeton; David Dobkin, Princeton; Scot Drysdale, Dartmouth; Diane Souvaine, Rutgers

Details about DREI can be found in Section 6.5. This 3-week summer institute featured three research workshops:

- **Software and Mathematical Visualization**
- Dates: June 24-28, 1996
- Organizer: Ayellet Tal, The Weizmann Institute of Science

A wide variety of research topics in software and mathematical visualization were discussed, as well as the use of visualization for education in all levels. The research areas included, among others, the visualization of geometry, knot theory, calculus, dynamical systems, tiling, minimal surfaces, 3-manifolds, statistics, the visualization of very large software systems, combinatorial structures, and algorithm animation.

- **Computational Geometry Problems in Aerodynamics**
- Dates: July 1-3, 1996
- Organizer: Tim Baker, Princeton

The aim of this workshop was to bring together two groups of researchers, computer scientists specializing in computational geometry and computational scientists whose research deals with problems of an essentially geometric nature. In particular, the workshop addressed the issues of mesh generation, surface representation and solid modeling. The emphasis was on identifying the key problem areas (e.g., mesh quality and optimal triangulations for the solution of flow problems in 2-D and 3-D, the blocking problem of decomposing arbitrary regions into a collection of hexahedral domains, surface meshing to achieve optimal approximation of curved surfaces, etc.). There were a series of presentations by leading researchers in both the numerical methods and the computer science communities, from academe and industry, as well as a number of panel discussions to identify areas that might benefit from a closer collaboration between the two groups.

- **Hot Topics in Computational Geometry**
- Dates: July 8-12, 1996
- Organizer: Bernard Chazelle, Princeton

The workshop on Hot Topics in Computational Geometry presented an opportunity to hear the latest results in the computational, discrete, and combinatorial aspects of geometry.

- **2nd Conference for African-American Researchers in the Mathematical Sciences (CAARMS)**
  - Dates: June 26-28, 1996
  - Location: Rutgers University, Bell Labs (Lucent) and IAS
  - Attendance: 85
  - Organizers: Nathaniel Dean, AT&T Research; William A. Massey, Bell Laboratories

This meeting highlighted current research by African-American researchers and students in mathematics. Its goal was to strengthen the mathematical sciences by encouraging the increased participation of African-Americans and underrepresented groups, facilitate working relationships between them and helping to cultivate their careers. The meeting was jointly sponsored by the Institute for Advanced Study and Bell Labs. Invited speakers were Stella Ashford, Curtis Clark, Jonathan Farley, Arthur Grainger, Isom Herron, Lee Lorch, Carolyn Mahoney, Walter Miller, Richard Tapia, Nathaniel Whitaker, Floyd Williams, Scott Williams, and Leon Woodson. Topics ranged over the entire breadth of mathematics. The conference received major funding from the Sloan Foundation.

- **Theory and Hypermedia for Big Science in Musicology**
  - Dates: Oct. 26, 1996
  - Location: Rutgers University
  - Attendance: 30
  - Organizer: Michael Leyton, Rutgers Center for Cognitive Science

This workshop, co-sponsored by the Rutgers Center for Cognitive Science, was a one day course offered by Guerino Mazzola of the University of Zurich and author of *The Geometry of Tones* (Birkhauser). Mazzola discussed the theoretical basis of his work, based on modern information technology, mathematics and semiotics, which has given precise models and language to describe musicological facts. It is based on three factors: (1) symmetry,

(2) local-global gestalts and (3) the Yoneda classification by "variation of perspectives". He discussed his hypermedia implementation of the theory and his views that the representation and control of music's inhomogeneous and semantically open data, make music a prototypical research subject asking for Big Science effort and funding in the Humanities. The software RUBATO(R) developed at Zurich was described in detail.

- **The Fifth DIMACS Implementation Challenge: 1995-1996**
  Priority Queues, Dictionaries, and Multi-Dimensional Point Sets
  - Dates: Oct. 28-30, 1996
  - Location: Rutgers University
  - Attendance: 65
  - Organizers: David Johnson, AT&T; Cathy McGeoch, Amherst

  Implementation challenges, part of DIMACS' emphasis on experimental computing, have become a tradition at the Center. This implementation challenge was a year long effort to develop algorithm implementations, data structure support, test suites and experimental methods on the priority queues and dictionaries data structures. (There were no multi-dimensional point set projects.) Two days of the workshop reported on the projects; a third day focused on methodology for algorithm experimentation with talks that addressed both major errors in techniques as well as successes.

Workshops Planned for the Rest of 1996-97

- **DIMACS Workshop on Combinatorial Optimization and Disordered Materials: Recent Progress and Algorithmic Challenges**
  - Date: May 15-17, 1997
  - Location: Rutgers University
  - Organizers: Phil Duxbury, Michigan State; Deborah Franzblau, City University of New York, College of Staten Island; M.F. Thorpe, Michigan State; Joel Lebowitz, Rutgers; Paul Leath, Rutgers
- **DIMATIA Inaugural Conference, The Future of Discrete Mathematics**
  - Date: May 19-25, 1997
  - Location: Prague, Czech Republic
  - Organizers: Ronald Graham, AT&T Labs; Jan Kratochvil, DIMATIA; Jarik Nesetril, DIMATIA; Fred Roberts, Rutgers

  Conference to inaugurate DIMACS' sister center, the Center for Discrete Mathematics and Theoretical Informatics in Prague.
- **DIMACS Quantum Computing Tutorial and Workshop**
  - Date: August 11-15, 1997
  - Location: Princeton University
  - Organizers: Andre Berthiaume, AT&T Labs; Peter Shor, AT&T Labs; Andy Yao, Princeton
- **Workshop: Special Year on Logic and Algorithms - One Year Later**

- Date: To be announced
- Location: To be announced
- Organizers: Eric Allender, Rutgers; Robert Kurshan, Bell Labs; Moshe Vardi, Rice University

5 Research Highlights

5.1 Special Year in Mathematical Support for Molecular Biology

The Special Year led to many significant research accomplishments. Perhaps the most important are the results of or beginnings of significant interdisciplinary collaborations, which are illustrated in our first few examples.

Computer Modeling of Ultraselfish Genes in Mice. The *t-haplotype* is a mutant chromosomal region in mice that is passed on to progeny more often than the "normal" version of this region. This "ultraselfish" behavior is quite surprising because under Mendelian genetics, offspring inherit paired genetic segments with equal probability. Males carrying two t-haplotypes are sterile and most forms of the t-haplotype also carry a recessive lethal mutation. The evolution and persistence in nature of such a deleterious gene is therefore equally surprising and a number of explanations have been proposed in the literature. Since the complex social structure of mice makes it difficult to develop mathematical models of the population genetics of the t-haplotype, Dannie Durand (PM, BC, and V) developed a Monte Carlo simulator to test these hypotheses in collaboration with Lee Silver of the Department of Molecular Biology at Princeton. The results of these simulations upset an explanation of t-haplotype persistence which had been widely accepted for thirty years and have suggested new hypotheses to be tested through laboratory and field work. This collaboration is continuing.

DNA Computing. Princeton biologist Laura Landweber and Dick Lipton (PM, PU), who met through the special year, have started a serious collaboration on DNA computing. The collaboration has led to the idea that DNA computation can solve certain problems fundamental to DNA itself – an idea also espoused by others such as Pevzner. These problems include, for example, the analysis of DNA sequence composition, DNA diagnostics, and mutation detection. All use a simple encoding principle in which DNA molecules anneal to their target sequences, with each DNA probe encoding a specific tagged sequence. The idea is to use molecular selection with DNA probes to interpret existing DNA sequences. This is similar to how RNA enzymes and functional RNA molecules act in cells on target RNA substrates, through use of standard base-pairing. Up until now, the field of DNA computers has focused on only a small number of "classical" problems or calculations which can be more easily performed on a conventional computer or on paper. By expanding the role of DNA computers to solve molecular biological problems, Landweber and Lipton hope to provide a bridge for this field into the main stream of technology. More importantly, they expect that their initial ideas will define a class of problems for which DNA computers are ideally suited, and even better suited than a digital computer: Problems inherently involving DNA. This work has progressed to the stage where Landweber and Lipton are ready to test it by experiment in Landweber's wet lab.

Models of Duplication. For the last 20 years, an important idea in molecular evolutionary biology has been that at least some of the differences in patterns of divergence in single-gene-family based phylogenies must be explained by duplications of the genes involved (paralogous evolution). In this framework, a likely evolutionary species' phylogeny is supposed to be most parsimonious, that is, involving a minimum number of the duplications and corresponding loss-of-a-copy events. (This is very much different from the parsimonious trees currently of wide interest: The latter minimize the number of between-symbol differences in the molecular sequences.) The duplication idea has previously had no explicit expression except for some very technical constructions based on mapping the nodes (leaf clusters) of one tree into their least common ancestor (lca) nodes in another tree, constructions having no explicit biological meaning. An explicit graph-theoretic model of the gene duplication in a species tree (involving duplicate-copy labeling of the descent species) was suggested by Boris Mirkin (V), a long-term DIMACS visitor. His model led to an explicit definition of corresponding loss events which allowed him to develop a set theoretic framework for defining a collection of between-tree difference measures as counts of various duplication-based events needed for explanation of the difference between two trees. In the framework of this model, the technical suggestions about counting duplication based events made in the earlier literature were partly proved correct and partly disproved. A conjecture of Mirkin's about the equivalence of a model-based measure of difference with a combinatorial one (counting the total number of the intermediate nodes between the lca images of all child-parent pairs in the gene tree) was proved correct independently by Martin Vingron (V) and Oliver Eulenstein and by L. Zhang. This also led to an invitation to Mirkin to visit the Department of Theoretical Informatics of the National Cancer Center of Germany in Heidelberg, where he continued this work in collaboration with Vingron. Mirkin and Vingron are investigating the relationship between Mirkin's model and yet another cladistic-minded approach being developed in Glasgow, UK by biologist Robert Page (V) and his associates, with an aim toward a larger area of application (host-parasite co-evolution, biogeographical evolution). The conjecture is that there is an equivalence between Mirkin's model and that approach. This conjecture was investigated during a recent DIMACS visit by Page. Mirkin and Vingron are also developing a model based algorithm for combining different gene-trees into a most parsimonious species tree, one needing the least number of the duplication-based events explaining all differences.

Reconstructing Evolutionary Trees and the Evolution of Modern Languages. Inferring the evolutionary relationships of a set of species is a fundamental problem in computational biology and one to which we have devoted considerable attention during the special year. Tandy Warnow and Sampath Kannan developed an algorithm for reconstructing evolutionary trees while Kannan was a postdoc at DIMACS. During the special year, when both Warnow and Kannan were DIMACS visitors, the algorithm was used to provide a new picture of the evolution of modern languages, and this work was highlighted in the New York Times.

Approximating Numerical Taxonomy. A common problem in phylogeny is to fit an $n \times n$ distance matrix D by a tree metric T. Richa Agarwala (PD), Vineet Bafna (PD), Martin Farach (PM, RU), Babu Narayanan (PD), Mike Paterson, and Mikkel Thorup (V) presented an $O(n^2)$ algorithm for finding an additive tree T such that the difference between the matrix of tree distances and the distance matrix D is at most three times the optimal. This is the first algorithm for the problem with a performance guarantee. They also showed that it is NP-hard to find such a tree T with the difference at most $9/8$ times the optimal. Later, Farach and Sampath Kannan (V) used this work as a building block in a clever algorithm for computing maximum likelihood trees, which is likely to have a great impact on the field. Experiments undertaken by Rutgers graduate student Jaime Cohen and Farach have shown that these distance methods construct better trees in practice than do the most popular heuristics used by biologists.

Phylogenetic Tree Reconstruction Under Stochastic Models of Nucleotide Substitution. Biologists seek to reconstruct evolutionary trees from aligned genetic sequences for an increasing number n of species. How fast must the sequence length N (a measure of the cost of getting input data) grow, as a function of n, in order to accurately recover the underlying tree with probability $1-\epsilon$, if the sequences evolve according to simple stochastic models of nucleotide substitution? This was the problem considered by Mike Steel (V), Laszlo Szekely (V), and Peter Erdos (V). They showed that for a certain model, a reconstruction method exists for which the sequence length N can grow surprisingly slowly with n (sublinearly for a wide range of parameters), and even as a power of $\log n$ in a narrow range (which roughly meets the lower bound from information theory). In contrast, a more traditional technique (maximum compatibility) provably requires N to grow faster than linearly in n. Their results were based on a new, and computationally efficient approach, for reconstructing phylogenetic trees from aligned DNA sequences.

Phylogenetic Tree Reconstruction from Quartet Splits. A significant problem in phylogeny is to reconstruct a slb tree from few valid quartet splits of it. It is well known that every slb tree is determined by its set of all valid quartet splits. Peter Erdos (V), Mike Steel (V), Laszlo Szekely (V), and Tandy Warnow (V) strengthened this result by showing that its local (i.e., small diameter) quartet splits infer by a dyadic inference rule all valid quartet splits, and hence determine the tree. They also presented a polynomial time algorithm to recover the tree.

Sequencing by Hybridization. Sequencing by hybridization (SBH) is a promising alternative to the classical DNA sequencing approaches. However, the resolving power of SBH is rather low; with 64 Kb sequencing chips, unknown DNA fragments only as long as 200 bp can be reconstructed in a single SBH experiment. To improve the resolving power of SBH, Broude and others have recently suggested positional SBH (PSBH), allowing one to measure (with additional experimental work) approximate positions of every l-tuple in a target DNA fragment. Sridhar Hannenhalli (V), Herbert Lewis, Pavel Pevzner (V), and Steve Skiena (V) have studied the positional eulerian path problem motivated by PSBH. In this problem, the input is an eulerian graph in which every edge has an associated range of integers and the goal is to find an eulerian path such that the range of the ith edge in the path contains i. They showed that the positional eulerian path problem is NP-complete even when the maximum outdegree of any vertex in the graph is 2. However, they found polynomial algorithms to solve a special case of PSBH where the range of the allowed positions for any edge is bounded by a constant.

Computing the Similarity between RNA Strings. RNA strings are strings over a four-letter alphabet A, C, G, U with a secondary structure of base-pairing between $A - U$ and $C - G$ pairs in the string. Edges are drawn between two bases that are paired in the secondary structure and these edges have traditionally been assumed to be noncrossing. The noncrossing base-pairing naturally leads to a tree-like representation of the secondary structure of RNA strings. Vineet Bafna (PD), S. Muthukrishnan (PD), and R. Ravi (PD) have addressed several notions of similarity between two RNA strings that take into account both the primary sequence and secondary base-pairing structure of the strings. They have found efficient algorithms for exact matching and approximate matching between two RNA strings. They defined a notion of alignment between two RNA strings and devised algorithms based on dynamic programming. They discovered a method for optimally aligning a given RNA sequence with unknown secondary structure to one with known sequence and structure, thus attacking the structure prediction problem in the case when the structure of a closely related sequence is known.

5.2 Special Year in Logic and Algorithms

An Until Hierarchy for Temporal Logic. As a dramatic result of DIMACS' efforts to bring together two distinct research communities, ideas from finite model theory were used to solve a problem of computer-aided verification (CAV) concerning the expressibility of temporal logic statements. Kousha Etessami (PD), Denis Therien (V), and Thomas Wilke (PD) designed a new Ehrenfeucht-Fraisse game to capture the power of temporal logic, one of the workhorses of CAV. Ehrenfeucht-Fraisse games are a standard tool in finite model theory, but have not found much use previously in the study of temporal logic. As a result, Etessami, Therien, and Wilke are also now able to provide simple and elegant proofs of theorems concerning temporal logic, whose previous proofs involved long and intricate algebraic constructions.

Connections between Bounded Variable Logic and Temporal Logic. The work of Etessami, Therien, and Wilke described in the previous example drew the attention of Moshe Vardi (V and special year co-chair) to the connection between temporal logic and bounded-variable logics, and it led to his collaboration with F. Gradel (V) and P. Kolaitis (V) in which it was proved that the decision problem for FO^2 (first-order logic with two variables) is NEXPTIME-complete. This collaboration among two computer scientists and a logician solved a classical open problem in logic. Motivated by the special year workshop on verification, Etessami (PD), Wilke (PD), and Vardi also obtained results on the connection between bounded-variable logic and temporal logic from a complexity-theory perspective.

Model Checking of Open Systems. Orna Kupferman (PD, BL) and Moshe Vardi (V and special year co-chair) have obtained beautiful results on the problem of model checking of open systems. In computer system design, a *closed* system is one whose behavior is completely determined by the state of the system; an *open* system is one that interacts with its environment and whose behavior depends on this interaction. Current model-checking algorithms and tools, which are based on temporal logic, and which are designed for closed systems, are not able to take into account uncertainty regarding the environment. Kupferman and Vardi proved both a positive and negative result. They showed that (assuming specifications in CTL or CTL*), although model checking of closed and open systems coincide for the *linear-time paradigm*, model checking of open systems is much harder than model checking of closed systems for the *branching-time paradigm.*

Verification of Fair Transition Systems. Orna Kupferman (PD, BL) and Moshe Vardi (V) have also obtained interesting results on verification of fair transition systems. In program verification, we check that an implementation meets its specification. Both the specification and the implementation describe the possible behaviors of the program, though at different levels of abstraction. Kupferman and Vardi distinguished between two approaches to implementation of specifications. The first approach is *trace-based implementation*, where we require that every computation of the implementation must correlate to some computation of the specification. The second approach is *tree-based implementation*, where we require that every computation tree embodied in the implementation must correlate to some computation tree embodied in the specification. The two approaches to implementation are strongly related to the linear-time versus branching-time dichotomy in temporal logic. Kupferman and Vardi have examined the trace-based and tree-based approaches from a complexity-theoretic point of view, comparing the complexity of verification of fair transition systems, modeling both the implementation and the specification, in the two approaches. For the trace-based approach, the corresponding problem is *fair containment* and for the tree-based approach, it is *fair simulation.* They show that while both problems are PSPACE-complete, their complexities in terms of the size of the implementation do not coincide and the trace-based approach is easier. As the implementation is normally much bigger than the specification, they see this as an advantage of the trace-based approach. The results are surprising because they are at variance with the known results for the case of transition systems with no fairness, where no approach is evidently advantageous.

Progress on the Berman-Hartmanis Conjecture. The Berman-Hartmanis conjecture, an old and well-known conjecture, states that all NP-complete sets are *isomorphic* via isomorphisms that are polynomial-time computable and invertible. However, most complexity theorists currently believe that there exist "very strong" one-way functions (functions that are easy to compute but hard to invert), leading to strong doubts about the truth of the conjecture. Manindra Agrawal (V) and Eric Allender (PM, RU, and special year co-chair) have found evidence for the conjecture, by proving a restricted version: They show that the sets in NP which are complete under AC^0 reductions (reductions computed by families of constant-depth circuits of polynomial size) – rather than those complete under polynomial-time reductions – are also isomorphic via AC^0 reductions. Reductions of this sort arise in a natural way from logical characterizations of NP and other complexity classes, which were explored in the special year. It is important to note that one-way functions exist in AC^0 if and only if they exist in polynomial time. Thus the "intuition" against the isomorphism conjecture mentioned above seems to hold also for AC^0 reductions. Agrawal and Allender also showed that the sets complete for the complexity class NC^1 are isomorphic via AC^0 reductions.

Solution to a 25-Year-Old Problem of Erdos and Renyi. Saharon Shelah (PM, RU) is a part-time faculty member at Rutgers. Funds from the special year enabled him to spend a substantial additional amount of time at DIMACS during the year. During this stay, he solved an important combinatorial problem of Erdos and Renyi, open for over twenty-five years. He proved that a Ramsey graph has exponentially many pairwise nonisomorphic subgraphs, i.e., for every $c > 0$, there is a $d > 0$ such that every graph of n vertices not containing homogeneous subsets of size $c \log n$ has $2^{d \log n}$ pairwise nonisomorphic subgraphs.

5.3 Continuing Impacts of Earlier Special Years

The success of a special year or even a single workshop cannot be measured immediately upon its conclusion, nor even a year or two later. Participants in a DIMACS special year or workshop may work on a problem during an event, and then continue to work or collaborate afterwards. Results, methods, or collaborations found or initiated during a special year or workshop may lead to questions that influence the field for many years. A variety of examples continue to illustrate this point.

An Optimal Convex Hull Algorithm. For example, the first DIMACS special year, in 1989-90, was on Discrete and Computational Geometry. The work of Jiri Matousek (V) proved to be tremendously influential. Some of the best people in the field spent the next few years extending and generalizing it, and finding more applications. That work still goes on. Bernard Chazelle (PM, PU) has recently solved an 18-year-old open problem on finding an optimal convex-hull algorithm for every dimension, using Matousek's technique of nonrandom geometric sampling in an essential way.

Experimental Results on the Traveling Salesman Problem Experimental computing is a key continuing theme at DIMACS. DIMACS' emphasis on this subject led to the visit in 1990-91 by Lyle McGeoch of Amherst College, and to a continuing collaboration between McGeoch and David Johnson (PM, ATTL), which has recently led to major results on the traveling salesman problem, together with Ed Rothberg. These results are extensive experimental results analyzing the gap between the Held-Karp traveling salesman bound and the optimal value for a variety of instance classes, real and probabilistic, and also its asymptotics for standard classes of randomly generated instances. For randomly generated Euclidean instances, the commonly-cited value for the limiting constant was shown to be off by over 7%.

Improved Computer Proof of the Four Color Theorem. In a collaboration that goes back to the DIMACS special year on graph theory and algorithms, Neil Robertson (V) and Robin Thomas (V) have been returning to collaborate with Paul Seymour (PM, PU). This collaboration has led to an improved computer proof of the four-color theorem.

Solution to the Even Cycle Problem. The collaboration among Neil Robertson (V), Robin Thomas (V), and Paul Seymour (PM, PU), mentioned in the previous example and supported by DIMACS and also by Bellcore, has led to the solution of a long-standing open problem of finding a polynomial time algorithm which does the following: Given a directed graph, decide if there is a directed cycle of even length in it. This problem originated from the famous mathematician George Polya in connection with evaluating the permanent of a matrix. It is also known to be equivalent to many other problems, including matrix sign-stability (a very important problem in economics that goes back to Nobel Prize winner Paul Samuelson), Pfaffian orientation of bipartite graphs, and variants of hypergraph 2-coloring. Essentially, Robertson, Seymour, and Thomas explicitly found all digraphs which have no even directed cycle. They were all built from certain planar digraphs and one exceptional one by glueing all digraphs together on small cutsets. Then to test if a digraph has no even cycle, one tests if it has this structure.

5.4 Work of Postdocs, Students, Visitors

Many of the exciting results of our postdocs, students, and visitors is not connected to a particular special year. Here, we give some examples of such results, many obtained in collaboration with our permanent members or with other postdocs, students, and visitors.

Simulations of Random Algorithms. The role played by randomness in computation is an area of intense interest and research activity. One of the most outstanding results to come out of DIMACS recently came from a collaboration between Michael Saks (PM, RU) and Shiyu Zhou (GS, RU). They showed that any function computable by a randomized algorithm that uses space S (and is guaranteed to halt) can be simulated by a *deterministic* algorithm that runs in space $O(S^{3/2})$. For the special case of randomized space computations (with 1 or 2 sided error), this improves the $O(S^2)$ bound for simulating a non-deterministic algorithm by a deterministic one from Savitch's 1971 theorem.

Dispersers and Randomized Algorithms. Michael Saks (PM, RU), Shiyu Zhou (GS, RU), and Aravind Srinivasin (PD) also collaborated on other work strengthening results on randomized algorithms. An area of considerable research interest in recent years concerns how dependent randomized algorithms are on the quality of the random source: Given an algorithm that uses random coins, can it be made robust in the sense that it works correctly even if the coins deviate from randomness to some extent? This question has been shown to be related to the explicit construction of graphs called *dispersers*, whose edges are in some sense nicely distributed. Saks, Zhou, and Srinivasin have constructed the best known family of dispersers, to show that every randomized algorithm can be converted to one that works correctly even with a very weak source of random bits.

Monotone Span Programs Another area in which DIMACS researchers have made considerable progress is in proving lower bounds within a new model of computation called a span program. A major goal of complexity theory is to prove lower bounds on the computational complexity of functions (in terms of time, space, or circuit size), and one of the major obstacles blocking progress toward this goal is the "sequential" nature of computation, as it evolves over time. Proving that no circuit (or algorithm) of a certain kind can exist means that, in some sense, the tree of all possible execution sequences must be examined, and that it must be shown that none of these sequences can perform a certain task. The span program is of great interest because the model is general enough to simulate rather powerful classes of circuits and it is not sequential. For example, the problem of proving lower bounds for certain classes of circuits reduces to proving lower bounds on span program size, which in turn is a purely combinatorial question, with none of the additional complications caused by the evolving nature of computation. Work of L. Babai, A. Gal (PD), J. Kollar, L. Ronyai, T. Szabo, and A. Wigderson (V) gave lower bounds for monotone span programs, complementing existing lower bounds for monotone circuits, which are also of great interest in complexity theory. This work in turn implied explicit lower bounds for monotone Boolean formulae and certain kinds of circuits. Later work by Babai, Gal, and Wigderson gave the first superpolynomial lower bounds for monotone span programs computing explicit functions. They exhibited the surprising power of monotone span programs by exhibiting a function computable in this model in linear time while requiring superpolynomial size monotone circuits and exponential size monotone formulae.

Uniform Span Programs over the Rationals. The work of Gal, et al. referred to above concerns span programs over finite fields. Until now, not much was known about span programs over infinite fields, such as the rationals. A new result by Eric Allender (PM, RU), together with Robert Beals (PD) and M. Ogihara, gave an exact characterization of the computational power of uniform span programs over the rationals.

Purely Functional Programming. The power of purely functional programming in the construction of data structures has received much attention, not only because functional languages have many desirable properties, but because structures built purely functionally are automatically *fully persistent*: Any and all versions of a structure can coexist indefinitely. Recent results illustrate the surprising power of pure functionality. One such result was the development of a representation of double-ended queues with catenation that supports all operations, including catenation, in worst-case constant time. Haim Kaplan (GS, PU) and Robert Tarjan (PM, PU) have developed methods for purely functional representations of sorted lists, implemented as finger search trees. The most efficient of three implementations they have developed achieves logarithmic access, insertion, and deletion time, and double-logarithmic catenation time. The bounds to find, insert, and delete were the same as the best known bounds for an ephemeral implementation of these operations using finger search trees. The representations developed are the first that address the issues of persistence and pure functionality, and the first for which fast implementations of catenation and split were presented. They are simple to implement.

Packet Waiting Times in Processor-ring Communication. Ed Coffman (PM, BL), Nabil Kahale (V), and F.T. Leighton have obtained a tight asymptotic bound on packet waiting times in processor-ring communication. They consider N processors communicating unidirectionally over a closed transmission channel, or ring. Each message is assembled into a fixed-length packet. Packets to be sent are generated at random times by the processors, and the transit times spent by packets on the ring are also random. Packets being forwarded, i.e., packets already on the ring, have priority over waiting packets. The objective of their research was to analyze packet waiting times under a greedy policy, within a discrete Markov model that retains over-all structure of a practical system, but is simple enough so that explicit results can be proved. Independent, identical Bernoulli processes model message generation at the processors, and i.i.d. geometric random variables model the transit times. The emphasis in the research was on asymptotic behavior for large ring sizes, N, when the respective rate parameters have the scaling λ/N and μ/N. The main result is that if the traffic intensity is fixed at $\rho = \lambda/\mu < 1$, then as $N \to \infty$ the expected time a message waits to be put on the ring is bounded by a constant. This result verified that the expected waiting time under the greedy policy is within a constant factor of that under an optimal policy.

Duchet's Conjecture and Kernel-Solvability. Work at DIMACS has led to important breakthroughs on several graph-theoretical questions. In 1980, Pierre Duchet conjectured that odd directed cycles are the only edge minimal kernel-less connected digraphs, i.e., in which after the removal of any edge a kernel appears. This conjecture was disproved by Apartsin, Ferapontova, and Vladimir Gurvich (V). Endre Boros (PM, RU) and Gurvich proved that a modification of Duchet's conjecture still holds, by showing that odd holes (odd non-directed chordless cycles of length 5 or more) are the only connected graphs which are not kernel-solvable but which, after removal of any edge, are kernel-solvable.

Asymptotic Upper Bounds on Chromatic Number. A *list coloring* of a graph is an ordinary graph coloring in which the color assigned to a vertex is chosen from a list of acceptable colors associated with that vertex. List colorings arise in many practical applications of graph coloring, including channel assignments in communications, traffic phasing, and fleet assignment problems. The *list chromatic number* of a graph is the smallest k so that the graph always has a list coloring whenever each vertex is given a list of size k. Anders Johansson (PD) has obtained asymptotic upper bounds on the chromatic number or list chromatic number of a graph for several classes of graphs with no r-clique. The bound is in terms of the maximum vertex degree – a Brooks'-type bound. The methods Johansson used to obtain these results use randomized arguments that should help shed light on the behavior of "simulated annealing," a much-used heuristic for practical applications.

Dependence of Optimal Solutions on Scales of Measurement. Work at DIMACS has related to a variety of interesting practical problems. For instance, mathematical formulation of an optimization problem often depends upon data which can be measured in more than one acceptable way. Motivated by practical problems of scheduling, Aleksandar Pekec (GS, RU), Fred Roberts (PM, RU), and N.V.R. Mahadev (V) have shown examples where an admissible change of scale of measurement can change an optimal solution to a problem into one that is no longer optimal and have identified conditions under which this anomaly is sure to be avoided. Pekec has also shown that under certain general conditions on a combinatorial optimization problem, if a solution remains optimal under all possible admissible changes of scale and the scale is admissible up to ordinal transformations, then the greedy algorithm will give an optimal solution.

The Russian Option. Not long ago, Larry Shepp (PM, ATTL) and Albert Shiryaev (V) proved a theorem in pure probability that gives the stopping time which maximizes the expected time-discounted supremum of an exponential Brownian motion and an explicit solution for this optimal value. The exponential Brownian motion is the widely accepted Black-Scholes model for asset fluctuations, so the theorem can be interpreted as an explicitly valued stock market hedge option. In the past year, this option, known as the Russian option, has become widely discussed on Wall Street and is beginning to be sold as a standard long buy hedge option.

An Application to Auctions and Bidding. The theory of auctions and bidding has enormous practical significance, and is used widely in industry, for instance in the recent bids by telecommunications companies for pieces of the FCC radio spectrum and in oil lease bids. Yet, many fundamental problems of auctions and bidding remain unresolved because of their combinatorial difficulty. Aleksandar Pekec (GS, RU), Michael Rothkopf (PM, RU), and Ronald Harstad have combined to study simultaneous auctions for situations in which the value of assets to a bidder depends upon which other assets he or she wins. In such cases, bidders may well wish to submit bids for combinations of assets. When this is allowed, the problem of determining the revenue maximizing set of nonconflicting bids can be a difficult one. Pekec, Rothkopf, and Harstad analyzed this problem, identifying several different structures of combinatorial bids for which computational tractability is constructively demonstrated and some structures for which computational tractability cannot be guaranteed.

5.5 Multi-Disciplinary, Multi-Institutional Projects

DIMACS has three multi-disciplinary, multi-institutional research projects, each funded separately from its main STC award. Our three projects of this kind are all scientifically exciting and have taken us into new and exciting research areas. Two of these, on simulations of integrated communications systems and scalable self-organizing simulations, can be directly traced to collaborations initiated during our special year on massively parallel computing. The third, the LINK project, is an offshoot of our continuing focus on experimental computing. These projects are described below.

In such a project, DIMACS brings scientific expertise and experience with large, complex projects that involve a variety of institutions and disciplines, and such projects provide opportunities for students, postdocs, faculty members, and industrial researchers to share in research experiences and learn more about each others' "cultures" and ways of doing science. Also, DIMACS provides a neutral site for collaboration. Companies find it much easier to collaborate through a neutral third party when direct collaboration would be difficult. DIMACS provides such a neutral third party through which files can be shared and other collaborations take place.

Simulations of Integrated Communications Systems

Project Summary: The project goals are to (1) develop a network description methodology, similar in spirit to VHDL (very high speed integrated circuit hardware design language), suitable for the description of current and future telecommunication networks (2) provide validated high-performance simulation software prototypes based on this methodology, and (3) demonstrate the use of the tools for network construction and operation problems that cannot be answered otherwise. The project is producing public domain software tools and is encouraging further development, adoption and applications of the simulation software.

Participating Scientists: Sandeep Bhatt (PM, BC and RU), Richard Fujimoto (Georgia Tech), Albert Greenberg (PM, ATTL), Jim Kurose (University of Massachusetts), David Nicol (William and Mary / Dartmouth), Andrew Ogielski (PM, BC), Diane Souvaine (PM, RU), Don Towsley (University of Massachusetts), and Paul Wright (PM, ATTL).

Collaborations: This project directly brings together the expertise of two of the nation's leading telecommunications companies in partnership with leading academic sites. We expect the collaboration to influence research directions toward developing solutions that will impact industry practice in designing new communications networks. The project also includes graduate and postdoctoral training opportunities that will benefit participants through exposure to industry problems and scientists. It overlaps with DIMACS' Special Year on Networks (1996 - 1997), which involves a broader audience of mathematical and computer scientists in research problems motivated by networking and computer security issues. The Simulations project is funded by the National Science Foundation for a three year period over 1995 - 1998, for a total of $1.5M.

Project Description: The rapidly developing synthesis of communications and computing technologies, the ongoing construction of immense digital data networks, and the global deregulation of telecommunications industries will have a tremendous impact on the national economy and infrastructure. This project brings together leading experts in network modeling, network design and analysis, and high performance discrete-event simulation to focus on the development and validation of high-performance modeling and simulation tools for broadband networks and wireless communication services. The project is applying these tools to cutting edge problems in network design and operations that cannot be solved otherwise. Despite the recent publicity in the popular media, research in this area has not kept pace with industrial demands. The primary obstacle has been the size and immense complexity of large telecommunications systems, considerably exceeding the complexity of the largest integrated circuits and computers.

This research has the potential for significant impact. The tools are being used by team members for network system design and analysis research. In order to stimulate further research and initiate technology transfer, the software tools and a repository of validated descriptions will be placed in the public domain. Success with this project will provide a proof of concept and software prototypes for future industrial development of modeling tools and proprietary applications. Our results will facilitate the deployment of telecommunications networks and services that are more reliable, more capable, more flexible, and less expensive than they would otherwise be.

Scalable Self-Organizing Simulations (S3)

Project Summary: Modeling the global-scale, heterogeneous communication domains that are being built in military and civilian contexts requires revolutionary advances in the state-of-the-art in telecommunications modeling. The goal of this project is to achieve a generational transition in performance and manageability of simulations of telecommunication systems. The design and management of heterogeneous, distributed simulations with millions of complex communicating objects requires new modeling techniques and new software architecture for the simulation framework. Existing discrete event simulation tools do not scale to systems of such complexity.

Participating Scientists: Sandeep Bhatt (PM, BC and RU), Marina Chen (Boston University), James Cowie (Cooperating Systems Corporation), Richard Fujimoto (Georgia Institute of Technology), Don Leskiw (Ultra Corporation), David Nicol (Dartmouth), Andrew Ogielski (PM, BC), Paul Wright (PM, ATTL), Ellen W. Zegura (Georgia Institute of Technology).

Collaborations: The daunting complexity of modern communications technologies on the one hand, and of the large software systems required on the other, demands that a project of this sort be carried out by an interdisciplinary team of experts in parallel simulations, telecommunications, applied mathematics, software engineering, and defense applications. We have put together a team of scientists from seven different locations to work on the project, including participants from two small companies. Current plans are to recruit one postdoctoral fellow and two graduate research assistants to work on the project. This project is funded by DARPA for a three year period over 1996-1999, for a total of $3.1M.

Project Description: The project will develop the modeling techniques and software for the scalable self-organizing simulation (S3) framework. An S3 framework involves (1) dynamic hierarchical modeling techniques, supporting aggregation and deaggregation of subsystems at runtime, and (2) adaptive refinement of the space/time simulation partition to efficiently handle computational load. The design and implementation of the S3 simulation framework software meeting these requirements requires an extensible communications description language, the parallel simulation runtime system, and the compiler that maps one to the other. Concurrently, we are building a prototype simulation of a system of systems communications architecture to verify the performance in defense-specific applications.

Defense plans envision global telecommunications services employing wireless, wireline and hybrid networks of networks interconnecting millions of diverse network and user nodes. Technical requirements on such systems place them one or two generations ahead of those currently deployed by commercial providers. Verification of design, operations, and management architectures relies on fast and accurate simulations.

We introduce a set of Self-Organizing, Scalable Simulation techniques (S3) to meet the challenge of simulating global-scale telecommunications architectures. The project will proceed along two fronts: (a) core systems research (foundational modeling, design of description language and the S3 framework, and the parallel and distributed simulation kernels), and (b) defense applications tests: validation of our new technologies by testing on a realistic specification of an extended BM/C3I communication architecture which targets a "one million agent" mark.

The LINK Project

Project Summary: LINK is a software system designed to be a general-purpose environment for experimentation with discrete mathematics. The environment is friendly enough so that non-technical users will be able to visualize problems easily, yet powerful enough for computer scientists and mathematicians to express complicated computations. The system is designed to support educational pursuits, to be used as a research tool, and to be useful in solving real-life problems. LINK's release date was July 1996. The system is freely available for teaching and research via the DIMACS web site.

Participating Scientists: Jonathan Berry (Elon College), Nathaniel Dean (PM, BL), Mark Goldberg (Rensselaer Polytechnic Institute), Greg Shannon (Spanning Tree Technologies), Steve Skiena (SUNY Stony Brook).

Collaborations: This project brings together the expertise of researchers who had each had previous experience in the design of software packages for discrete mathematics. They gathered together at DIMACS in July, 1991, and proposed LINK. The success of some DIMACS workshops on such software packages gave their idea impetus by clearly indicating that there is a strong and growing interest in the use of computers to make discrete mathematical techniques readily available for experimental research and for industrial and educational applications. LINK has been funded by an NSF grant to DIMACS for a three year period over 1993 to 1996, for a total of $350K, and it has also received support from the Office of Naval Research and Los Alamos National Laboratory.

Project Description: Details about the LINK project can be found in Section 3.1, in the description of the research of Jonathan Berry.

An Interesting Application: LINK was used at Los Alamos National Laboratory to assist the Internal Revenue Service in detecting fraud in "faked" tax returns. Individual returns are represented as points in a graph and close similarity between two returns is represented by an edge link. Clusters of similar returns can be detected and selected returns are then examined for patterns of fraud involving fake income reports that show taxes paid. This system is in use by tax examiners at all five regional IRS service centers and has been successful at uncovering organized fraud. Pat Fasel of Los Alamos, who directed the IRS project, is also exploring new uses of LINK for modeling nuclear material proliferation, and to assist in detecting Medicare fraud.

6 Education Programs

DIMACS' educational programs divide into two kinds, K-12 programs and programs for undergraduates. (Our graduate student and postdoc programs are also educational in many aspects, but are presented in this report as part of our description of the research activities of the center.)

DIMACS' two most visible K-12 programs, the Leadership Program in Discrete Mathematics and the Young Scholars Program, are described in Section 6.1. Other K-12 programs for teachers include the "Workshops in Your District Program," through which trained participants in the Leadership Program have gone to districts to provide workshops to teachers, and we have reached thousands of teachers in this way. We also have a newsletter for high school teachers, "In Discrete Math," that is circulated free twice a year to approximately 3500 teachers.

DIMACS runs a variety of undergraduate programs and is planning major new initiatives in the undergraduate arena. DIMACS' programs are described in Section 6.2, and in particular its successful REU program is described in Section 6.3.

The center has made major strides toward integration of its research and education programs. Our efforts at integration are described in Section 6.4, and in particular our major new program, the DIMACS Research and Education Institute, which has integration as its primary theme, is described in Section 6.5.

6.1 K-12 Programs

6.1.1 Leadership Program in Discrete Mathematics

The Leadership Program in Discrete Mathematics (**LP**) has evolved through three phases: Phase I for high school teachers (1989-1991), Phase II for high school and middle school teachers (1992-1994), and now Phase III for K-8 teachers (1995-1998). In the present phase, teachers attend a two-week program during the summer, attend four follow-up days (eight are available) during the year, and then return for a week the following summer.

The content material for the first summer (two weeks) includes applications of graphs, and patterns in numbers and geometry. Participants begin by coloring maps and end by generating spirolaterals and fractals. These topics are presented during morning workshops lasting two hours, and consist of instruction together with hands-on activities and group work designed to engage the teachers in the material and in the problem-solving process. Also mentioned during these workshops, and emphasized in other sessions later in the day, are ways that the material can be brought into the classroom.

The **LP** is successful if the teachers return to their schools empowered to become "discrete math leaders." During the program it is emphasized that they have become experts of a sort and that they are encouraged and expected to try this material in their classrooms; as well as to give workshops for other teachers in their school or district, telling them about discrete math. Thus the **LP** includes special sessions on problem solving, assessment of mathematical learning, an overview of discrete mathematics (so that they can successfully answer the common question "what *is* discrete math?") and leadership issues. It is fair to say that the **LP** is as leadership-oriented as content-oriented. As of October 1996, 72 of the 116 participants in the 1995 K-8 **LP** had given 238 workshops attended by about 5,000 teachers.

Study group sessions follow the morning workshop. During this time, the teachers are broken into grade-heterogeneous groups to work on problems related to the morning workshops. These problems will be discussed further during the afternoon consolidation sessions and the following morning, before the next workshop. Afternoons at the **LP** find the teachers in grade-homogeneous groups exchanging ideas for implementing the material in the classroom and incorporating it into their already full syllabi. Teachers are also exposed to e-mail (and everyone is given a DIMACS account), the Internet and LOGO–Turtle Math.

The 1996 **LP** was quite a bit larger in scope than the 1995 program. In 1996 there were two two-week residential programs and one commuter program at Rutgers (same as 1995), as well as two off-site commuter programs (there were none in 1995) in Warwick, RI and Virginia Beach, VA. Those programs represent 161 teachers participating in the **LP** for the first time. Additionally, there were three one-week second-summer follow-up programs at Rutgers for the 1995 participants, two residential and one commuter, for another 79 teachers (of course, there were no follow-ups in 1995). We also held a one-week program for the high school and middle school "veterans" from the 1989-1994 programs and a new two-day "crash course" in discrete mathematics, each of which involved 45 high school teachers. Thus, we were able to continue our high school teacher programs even though we did not have a specific grant to fund these activities. Both of these programs will continue in 1997.

The experiment with off-site programs was successful because we were able to create comparable high-quality materials which were used at all different sites by a number of different workshop leaders. Three workshop instructors this past summer were working the **LP** for the first time, and were able to do very well with the materials given them (including slides, instructor's notes, handouts and participant materials) after some training from the **LP**.

The 1997 schedule is much like the 1996 schedule except that there will be many more participants returning for follow-up, second-summer weeks. There will be two residential programs and one commuter program at Rutgers, and two off-site programs, one in Warwick, RI and one in Scottsdale, AZ for new, first-summer participants; we expect about 180 teachers in these programs. Additionally, we will have the returning teachers from the summer of 1996, and veterans from all programs before that. There will also be another two-day "crash course" for high school teachers.

There have been a number of sites which have asked to host a **LP**, and will be considered for 1998. These include, for example, Wilmington, DE, Staten Island, NY, Stafford, VA, San Diego, CA, Carlisle, PA, Greenfield, MA and Louisville, KY.

The materials that have been developed for the leadership program workshops, including slides, instructor's notes, handouts and participant materials, are sufficiently detailed and, by the end of the summer of 1997, will have been sufficiently field-tested and evaluated, that it should be possible for other mathematicians to use them as the basis for independently-conducted ("franchised") parallel high-quality institutes in 1998, after suitable training provided by the leadership program.

6.1.2 Young Scholars Program in Discrete Mathematics

This is a four-week residential program for 40 mathematically talented students entering grades 11-12, with four follow-up sessions during the year. Week-long workshops are given by college and university faculty on topics ranging from number theory to fractals, designed to excite and challenge the students, motivate them to continue their study of mathematics, and consider careers in the mathematical sciences. Rutgers undergraduates and graduate students serve as group leaders and mentors. Maintaining diversity is a priority: each year the program has maintained a 50/50 balance between females and males, and approximately 20% of the students are minorities.

Beginning in 1994 we added a new research component, organized by Deborah Franzblau and now by Robert Hochberg, in which DIMACS-affiliated researchers develop projects and work on them with groups of 4-6 students during the last three weeks of the program. 15 researchers have now participated, with project topics including secret communication, scheduling, polyomino tilings, and text compression. Both students and researchers have found the projects valuable. The researchers for the 1996 program were Adam Buchsbaum (PM, ATTL), "Paying the Piper—Balancing the Costs of Different Tasks," Matt Franklin (PM, ATTL), "Search Games on Graphs," Deborah Franzblau (PM, RU), "Sticks and Stones or Knucklebones: How do You Play that Game?," Joseph Gerver (Rutgers Camden), "Soap Bubbles and Minimal Surfaces," Robert Hochberg (PM, RU), "Ominous Ominos," Stephen Penrice (PD), "Can you Design a Fair Draft Lottery?," Nick Reingold (PM, ATTL), "Dotland," and Nina Shteingold (RU), "Trinkets and More."

For the past two years there has been a significant technology component to the Young Scholars Program. Each year an entire week was devoted to doing mathematics with computers and graphing calculators, together with an exciting "LEGO LOGO" project. Janice Kowalczyk (an educator from Rhode Island, now working with the Leadership Program full-time) runs three days of workshops on programming in the LOGO language and how the interface works with LEGO motors and products, at the end of which the students have a competition to design and build a robot to perform some specified task. This project integrates mathematics, programming, design and construction.

During the summer of 1996, we added a connection to the center's molecular biology activities through the DIMACS-Mt. Sinai Computational Biology program, and the participation in YSP of Professor Gary Benson from Mt. Sinai.

The program has run since 1990 at Rutgers University, is funded by the NSF, and is co-directed by Robert B. Davis (RU), Michael O'Nan (PM,RU) and Joseph G. Rosenstein (PM,RU).

6.2 Undergraduate Education Initiatives

From the outset, DIMACS made a substantial commitment of time and resources to K-12 education because we had a talented and energetic leader who was willing to sustain and initiate programs in the area; because of the success of our activities, which led to their expansion; because that seemed to be a priority at NSF; and because all of these activities evoked a favorable response from the members, the participants, the community and the funding agencies. However, we have not put comparable emphasis on involvement of undergraduates, or into playing a role in change of the undergraduate curriculum, or into programs for undergraduate teachers. This was partly a matter of believing that we had something more unique to offer in K-12 education, partly that our two undergraduate institutions already had strong programs in DM and TCS, so it was hard to see what additional things the center could do, and partly a matter of not having the resources to do more.

Starting in the Fall of 1995, we totally reassessed the goals and purposes of the center (see Executive Summary), and in February 1996 decided that DIMACS should play much more of a national leadership role relative to the inclusion of DM/TCS in the undergraduate curriculum and in introducing undergraduates and college teachers to the topics of DM/TCS and their applications. We began a serious assessment of alternative new programs. While we have been carrying out this assessment, we continued a number of programs we have been running and instituted a few small new ones.

- We continued our successful REU program and made some changes in it. See Section 6.3.
- We continued to host visitors from four year colleges. In 1995-96, we had year-long visitors from Wellesley, Lafayette, and Holy Cross.
- We have run programs for the NJ Chapter of the Mathematical Association of America and the Mathematical Association of Two Year Colleges of New Jersey, most recently hosting the joint Fall meeting of these two organizations in November 1995 and organizing the program for them.
- We ran a workshop on teaching logic during our Special Year on Logic and Algorithms, in the summer of 1996.

- We have organized a "working group" on curriculum issues in logic and CS under the direction of Moshe Vardi of Rice University.
- We began a DIMACS-Mt. Sinai Computational Biology Program with the Mt. Sinai School of Medicine, and took our REU students on a field trip to Mt. Sinai. We plan expansion of this collaboration. (We also collaborated on the Young Scholars Program for high school students and plan a collaboration on postdoc training.)

To explore ways that DIMACS might have much more of an impact, we initiated, in September 1996, an electronic dialogue under the direction of Steve Maurer at Swarthmore. We asked for suggestions from the community about things DIMACS might do in the future. We are planning a meeting of the DIMACS Education Committee on December 11, 1996, to which we have invited members of the community, to discuss alternative plans for new DIMACS undergraduate initiatives, and we expect to implement several major new initiatives early in 1997.

6.3 REU

DIMACS has run a successful NSF Research Experiences for Undergraduates (REU) program since 1993. The program introduces mathematically talented students to research in DIMACS-related fields under the guidance of excellent research mentors, both to discover new results and provide a basis for making career and graduate school decisions. Students come to work one-on-one with senior DIMACS researchers for eight weeks during the summer, and enjoy the stimulating and supportive environment at DIMACS. Students report on their projects at weekly student-faculty lunches, and write a short report at the end of the summer. Most of the students discover original results, and some are impressive. For example, one of this year's students, Todd Mateer, working with Charles Sims (PM, RU), used computer algebra programs to obtain a much improved upper bound on the diameter of some Cayley graphs of the restricted Burnside group R(2,5), which may be useful for determining whether this group is infinite or finite—a big open problem in computational group theory. Another student, Chris Burrows, developed animations of graph algorithms for distribution with LINK, a graph manipulation software package being developed at DIMACS.

During the summer of 1996, there were several innovations in our REU program. We initiated a series of "popular" evening lectures, aimed not just at the REU students but also at summer students attending programs run by our industrial partners and housed at Rutgers. The highlight of this summer's program was a wonderful lecture by Avi Wigderson (V) entitled "The wonders of the digital envelope - A crash course in modern cryptography." We also introduced a connection to the center's thrust in molecular biology as part of the DIMACS-Mt. Sinai computational biology program. We took our REU students on a field trip to the Mt. Sinai School of Medicine in New York, where they were introduced to problems in computational biology by Mt. Sinai faculty. We plan to expand this program next summer.

In 1996, the Rutgers REU selected eight students, our biggest program to date, including three women and five men:

Student	Mentor	Project area	Home Institution
Chris Burrows	Jon Berry (PD)	Animations in LINK	Trenton State U.
James Ezick	Sven Dickinson (RU, CS)	Robotics	SUNY Buffalo
Todd Mateer	Charles Sims (PM,RU)	Permutation Groups	Grove City College
Priya Sangar-neswaran	Peter Hammer (PM,RU)	Chinese Productivity Patterns for Prediction	Oberlin
Jodi Schnieder	Wilma Olson (RU,Chemistry)	DNA Structure	St. John's College
Bhavin Shukla	Diane Souvaine (PM,RU)	Visualizations for Wireless Communication	Rutgers
Adam Wolfe	Richard Bumby (PM,RU)	Number Theory	Rutgers
Lily Yau	Adi Ben-Israel (PM,RU)	Iterative Methods and Complex Dynamics	Columbia

The eight students came from seven different schools in five different states, making it our most national REU so far. Six of the eight students are continuing their research during the academic year, one has submitted a paper with her advisor and two others are in preparation. One student will be presenting his work at the January meetings of the AMS/MAA in San Diego. The DIMACS REU program is now coordinated by Robert Hochberg, with Michael Saks (PM, RU) as faculty advisor.

6.4 Integration of Research and Education

One of DIMACS' goals is to integrate our research and education programs. We believe that involving our permanent members, postdocs, and graduate students in the educational programs of DIMACS can help instill in them a career-long commitment to helping in the education of young people, and that involving teachers with our research programs can help make education in the mathematical sciences more relevant and alive, and create long-term partnerships between researchers and educators. Our education and research programs are linked through the incorporation of educational outreach activities in the planning and implementation of special years (we are now asking special year organizers to plan on educational components for their years if possible; for example, as noted in Section 4.2, the logic and algorithms year had a conference on the teaching of logic), and the direct participation of special year researchers in educational outreach activities. We are asking each of our postdocs to participate in either an outreach activity or an educational activity while at DIMACS. In the past, and in particular in the past year, DIMACS postdocs, visitors, permanent members, and graduate students have participated in educational activities by acting as mentors in our research experiences for undergraduate programs and in our Young Scholars Program for high school students, by giving lectures or teaching courses in our Leadership Program for precollege teachers, and by supervising undergraduate theses. Participants in these programs have also included some of the most senior DIMACS leaders from all five institutions. A small portion of this involvement is described in Sections 6.1 and 6.3.

In the summer of 1996, we inaugurated a four year DIMACS Research and Education Institute (DREI), under a \$1.9M grant from the NSF program of Regional Institutes in the Mathematical Sciences, RIMS. The program is aimed at "integrating" researchers, postdocs, students, and teachers. Our program, in collaboration with the Institute for Advanced Study/Park City Mathematics Institute, is described in Section 6.5.

6.5 DIMACS Research and Education Institute

The DIMACS Research and Education Institute (DREI) is aimed at integrating research and education. It is broadly conceived as "The Special Year at DIMACS Revisited." Special years at DIMACS focus on cutting edge research issues in the areas of discrete mathematics and theoretical computer science. Whole new streams of research result. After a few years, however, it is imperative that the research community regather to assess the directions taken by the field, to explore new areas of interest, and to forge new collaborations. The perspectives gained in the intervening time, however, also make it timely to expose a far broader community to the area and to stimulate a dialogue on the potential impact of the new science and mathematics on curriculum at all levels, as well as on how best to develop the mathematicians and scientists of tomorrow. In the process, the special year revisited theme is ideal for integration of research and education.

DREI includes a research program for "researchers," an education program for high school teachers, and a program aimed at "integrating" the researchers and the teachers. Diane Souvaine (PM, RU) is the "convenor" of DREI and has overall responsibility for the program. Funded for four years, DREI was initiated in the summer of 1996 with a 3-week program at Princeton University whose overall theme revisited that of the first DIMACS Special Year: Discrete and Computational Geometry. The research program was chaired by Bernard Chazelle (PM, PU, and chair of the first DIMACS special year), and co-chaired by David Dobkin (PM, PU) and Scot Drysdale (Dartmouth). The Research Program incorporated university and industrial researchers, postdoctoral researchers and graduate students, and each of the three weeks had a separate focus: Software and Mathematical Visualization (Ayellet Tal, Organizer), Computational Geometry Problems in Fluid Dynamics (Tim Baker, Organizer), and Hot Topics in Computational Geometry (Bernard Chazelle, Organizer). The Education Program was organized along the same themes. Scot Drysdale led a program on Voronoi diagrams, convex hulls, and Delaunay triangulations during weeks I and II which segued into Iliana Bjorling-Sachs' program on visibility and art gallery theorems during week III. David Dobkin and Ken Steiglitz (PM, PU), assisted by Paul Burchard, as part of the education program, led a computing course with an Internet component, a web component, and a geometry visualization component. Discussion sessions provided opportunities for interaction between those in the research and education programs and for those in the education program (4 high school lead teachers, 28 high school teachers, 3 college/university teachers and 4 undergraduates) to integrate what they were learning.

The key to DREI is the plenary sessions. The education program is synchronized in such a way that the participants have the vocabulary and background necessary to profit from a mathematical science plenary speaker's presentation and to engage in lunchtime discussions thereafter with both researchers and other educators. Other plenary speakers treat mathematical sciences education, or uses of mathematics and science, or policy related to mathematics and science–these sessions are accessible to all participants without additional preparation and stimulate vigorous discussions.

DREI '96 provided researchers in geometric areas the opportunity to enhance their own research by sharing their state-of-the-art results with other active researchers, and gather feedback and new ideas. It also exposed them to new ideas about pedagogy and to important issues in education. The summer portion of DREI '96 provided education program participants with stimulating exposure to new mathematical ideas. The first two followup sessions for the high school teachers (September 28, 1996 and November 9, 1996) brought these new mathematical concepts down-to-earth with presentations that focused on where these concepts fit in the high school curriculum. By December 1, these teachers had submitted abstracts for lesson plans that they will develop. By March 1, 1997, the final versions will be submitted, and presentations will be made at the March 22, 1997 and the June 30, 1997 followup sessions. Through the plenary sessions, the joint meals, and the joint social activities, DREI '96 provided all participants with an expanded academic community, with an understanding of the issues facing teachers in high schools, of the issues facing teachers in universities, of the nature of state-of-the-art research, and of enhancements that can take place in classrooms both at the high school and university level. Email exchanges across this expanded academic community continue, as does producing web pages with K-12 mathematical material.

Plans are well underway for DREI '97, which will revolve around the theme Cryptography and Network Security, and will be chaired by Joan Feigenbaum (PM, ATTL) and Stephen Rudich (Carnegie-Mellon). In an experimental program such as DREI, the need for revisions is inevitable. But we are pleased with the successes of DREI '96, and anticipate even greater success with DREI '97.

A Management

- Director: Fred S. Roberts, Rutgers University
- Co-Director: Bernard Chazelle, Princeton University
- Associate Director: Stephen Mahaney, Rutgers University
- Associate Director for Education: Joseph Rosenstein, Rutgers University

B Executive Committee

- Sandeep Bhatt, Bellcore
- Bernard Chazelle, Princeton University
- Peter Hammer, Rutgers University
- David Johnson, AT&T Labs
- Stephen Mahaney, Rutgers University
- Fred S. Roberts, Rutgers University
- Joseph Rosenstein, Rutgers University
- Michael Saks, Rutgers University
- William Steiger, Rutgers University
- Robert Tarjan, Princeton University
- Mihalis Yannakakis, Bell Labs

C External Advisory Board

- Louis Auslander, CUNY
- Cynthia Dwork, IBM Almaden Research Center
- Avner Friedman, Univ. of Minnesota (Chair)
- Richard Karp, University of Washington
- Daniel Kleitman, M.I.T.
- Laszlo Lovasz, Yale University
- Jill Mesirov, Boston University
- Christos Papadimitriou, Univ. of CA/Berkeley
- William Pulleyblank, IBM Watson Research
- Dana Scott, Carnegie Mellon University
- Eva Tardos, Cornell University
- Fred Wan, Univ. of CA/Irvine

D Permanent Members of DIMACS and their Research Interests

James Abello, AT&T Labs
Research Areas: Discrete and computational geometry and applications; very large graph theory; visualization and animation.

Alfred Aho, Columbia University
Research Areas: Pattern matching algorithms; digital libraries; software engineering

Bill Aiello, Bellcore
Research Areas: Cryptography and distributed computing

Farid Alizadeh, Rutgers University
Research Areas: Combinatorial optimization, convex optimization, computational biology

Eric Allender, Rutgers University
Research Areas: Computational complexity theory

Rajeev Alur, Bell Labs, Lucent Technologies
Research Areas: Formal methods, verification, logic and concurrency

Sanjeev Arora, Princeton University
Research Areas: Complexity theory, probabilistically checkable proofs, approximation algorithms, computational uses of randomness.

Brenda Baker, Bell Labs, Lucent Technologies
Research Areas: Analysis of algorithms and software tools

Eric Baum, NEC
Research Areas: Cognition, artificial intelligence, neural nets, learning theory, DNA computing.

Jozsef Beck, Rutgers University
Research Areas: Combinatorics, combinatorial games, number theory

Adi Ben-Israel, Rutgers University
Research Areas: Matrix theory; convexity and optimization; mathematical programming; mathematical economics

Jon Bentley, Bell Labs, Lucent Technologies
Research Areas: Experimental algorithms

Sandeep Bhatt, Bellcore
Research Areas: High performance parallel and distributed computing; simulations of communications systems; algorithms and data structures

Dan Bienstock, Bellcore
Research Areas: Optimization and parallel computing

Matt Blaze, AT&T Labs
Research Areas: Security

Enrico Bombieri, Institute for Advanced Study
Research Areas: Probability

Endre Boros, Rutgers University
Research Areas: Discrete and combinatorial optimization; finite geometries

Adam L. Buchsbaum, AT&T Labs
Research Areas: algorithms and data structures, especially with applications in speech recognition and translation; graph problems

Richard Bumby, Rutgers University
Research Areas: Number theory

Robert Calderbank, AT&T Labs
Research Areas: Coding Theory

Tamra Carpenter, Bellcore
Research Areas: Telecommunications applications; interior point methods; stochastic optimization

Bernard Chazelle, Princeton University
Research Areas: Computational geometry, data structures

Gregory Cherlin, Rutgers University
Research Areas: Logic, model theory; model theoretic algebra

Fan Chung, Univ. of Pennsylvania
Research Areas: Combinatorics

Vasek Chvatal, Rutgers University
Research Areas: Algorithms; combinatorics; graph theory; operations research.

Kenneth Clarkson, Bell Labs
Research Areas: Discrete and computational geometry; applications

Edward G. Coffman, Jr., Bell Labs, Lucent Technologies
Research Areas: Combinatorial scheduling and packing theory, probabilistic analysis of algorithms, stochastic modeling, stochastic optimization

Edith Cohen, AT&T Labs
Research Areas: Algorithms design and analysis

John Conway, Princeton University
Research Areas: Combinatorics, games

Bill Coughran, Bell Labs, Lucent Technologies
Research Areas: Scientific computing

Ingemar Cox, NEC
Research Areas: Computer vision

Valerie DeBellis, Rutgers University
Research Areas: Mathematics education; interactions between affect and cognition during mathematical problem solving

Nathaniel Dean, Bell Labs
Research Areas: Graph algorithms

Richard DeMillo, Bellcore
Research Areas: Software reliability; security

David Dobkin, Princeton University
Research Areas: Graphics, analysis of algorithms, geometry

Nick Duffield, AT&T Labs
Research Areas: Applied probability, queueing, telecommunications, statistical mechanics

Dannie Durand, On leave from Bellcore
Research Areas: Computational biology; parallel computation

Martin Farach, Rutgers University
Research Areas: Algorithms; computational molecular biology

Joan Feigenbaum, AT&T Labs
Research Areas: Security

Peter Fishburn, AT&T Labs
Research Areas: Utility and decisionmaking; social choice; graph theory; foundations of probability

Steven Fortune, Bell Labs, Lucent Technologies
Research Areas: Computational geometry

Matt Franklin, AT&T Labs
Research Areas: Security

Michael Fredman, Rutgers University
Research Areas: Data structures and algorithms; computational complexity

Emden Gansner, AT&T Labs
Research Areas: Graph drawing and algorithms; programming languages

Michael Garey, Bell Labs, Lucent Technologies
Research Areas: Algorithms and complexity; discrete math; scheduling theory

Israel Gelfand, Rutgers University
Research Areas: Artificial intelligence; mathematics; neuroanatomy; cell biology

Phil Gibbons, Bell Labs, Lucent Technologies
Research Areas: Parallel computation; algorithms; databases

Lee C. Giles, NEC Research Institute
Research Areas: Artificial intelligence, neural networks, GAs, fuzzy systems, intelligent agents; applications

Simon Gindikin, Rutgers University
Research Areas: Theory of representations; integral geometry; several complex variables; mathematical physics; molecular biology problems

Ram(anathan) Gnanadesikan, Rutgers University
Research Areas: Statistical data analysis, graphics, pattern recognition

Gerald Goldin, Rutgers University
Research Areas: Mathematical physics, lie theory, mathematics education, psychology of problem solving

Ronald Graham, AT&T Labs
Research Areas: Combinatorics; Graph theory; number theory, theoretical computer science

Albert Greenberg, AT&T Labs
Research Areas: Networks, performance analysis

Michael Grigoriadis, Rutgers University
Research Areas: Algorithms for network optimization

Stuart Haber, Surety Technologies
Research Areas: Security and cryptography

Andras Hajnal, Rutgers University
Research Areas: Set theory, discrete mathematics

Peter. L. Hammer, Rutgers University
Research Areas: Logical analysis of data and discrete applied mathematics

Haym Hirsh, Rutgers University
Research Areas: Artificial intelligence; machine learning

Rob Hochberg, Rutgers University
Research Areas: Combinatorics and graph theory

Alan J. Hoffman, Rutgers University
Research Areas: Combinatorics and optimization; linear algebra

Frank Hwang, Chiaotung University
Research Areas: discrete math; optimization

Guy Jacobson, AT&T Research
Research Areas: Algorithms; data compression; games

Suresh Jagannathan, NEC Research Institute
Research Areas: Programming languages; distributed and parallel systems

Ravi Jain, Bellcore
Research Areas: theoretical computer science

David S. Johnson, AT&T Labs
Research Areas: Combinatorial optimization; complexity theory

Jeffry Kahn, Rutgers University
Research Areas: Extremal problems; probabilistic methods; geometry

Bahman Kalantari, Rutgers University
Research Areas: Mathematical programming; global and discrete optimization

Narendra Karmarkar, Bell Labs, Lucent Technologies
Research Areas: Linear programming

Michael Kearns, AT&T Labs
Research Areas: Computational learning theory

Leonid Khachiyan, Rutgers University
Research Areas: Discrete optimization; mathematical programming; complexity

Joe Kilian, NEC
Research Areas: Computational complexity; cryptography

Jeong Han Kim, AT&T Labs
Research Areas: Combinatorics

Alexander Kogan, Rutgers University
Research Areas: Boolean functions, discrete optimization, artificial intelligence

Janos Komlos, Rutgers University
Research Areas: Combinatorics; probability; theoretical computer science

Harold Kuhn, Princeton University
Research areas: Game theory

Robert Kurshan, Bell Labs, Lucent Technologies
Research Areas: Computer aided verification

Jeffrey C. Lagarias, AT&T Labs
Research Areas: Number theory; optimization; discrete geometry; computational complexity

Kevin Lang, NEC
Research Areas: Neural networks; learning theory; parallel computation

Andrea Lapaugh, Princeton University
Research Areas: Combinatorial algorithms, VLSI design

Joel Lebowitz, Rutgers University
Research Areas: Statistical mechanics of equilibrium and nonequilibrium processes; discrete and continuous time stochastic particle systems; computer simulation of statistical mechanical model systems; foundations of statistical mechanics.

David Lee, Bell Laboratories, Lucent Technologies
Research Areas: Network protocols

Richard Lipton, Princeton University
Research Areas: Architecture; Complexity; VLSI

Nicholas Littlestone, NEC
Research Areas: Computational learning theory

Carsten Lund, AT&T Labs
Research Areas: Complexity theory; interactive proofs

Richard Lyons, Rutgers University
Research Areas: Finite simple groups

Stephen R. Mahaney, Rutgers University
Research Areas: Computational complexity

Yossi Matias, Bell Labs, Lucent Technologies
Research Areas: Parallel algorithms

James McKenna, Bellcore
Research Areas: Math education

Michael Merritt, AT&T Labs
Research Areas: Distributed algorithms; cryptographic protocols; formal methods

Milena Mihail, Bellcore
Research Areas: Network algorithms; matroids

Clyde Monma, Bellcore
Research Areas: Combinatorial optimization; applications to telecommunications

Arun Netravali, Bell Labs, Lucent Technologies
Research Areas: Digital image processing

Michael O'Nan, Rutgers University
Research Areas: Permutation groups; simple groups

Andrew Odlyzko, AT&T Labs
Research Areas: Cryptography; number theory; combinatorics; coding theory; analysis; probability theory

Andrew Ogielski, Bellcore
Research Areas: Simulation

Alon Orlitsky, Bell Labs, Lucent Technologies
Research Areas: Information theory; communication and computation complexity

Rafail Ostrovsky, Bellcore
Research Areas: Computational complexity; interactive proofs; average case complexity

Steven Phillips, AT&T Labs
Research Areas: Algorithms in telecommunications

Andras Prekopa, Rutgers University
Research Areas: Stochastic processes; stochastic optimization; linear and nonlinear programming

Satish Rao, NEC
Research Areas: Parallel computing and combinatorial optimization.

Sivaram Rajagopalan, Bellcore
Research Areas: Cryptography; average-case completeness

Jim Reeds, AT&T Labs
Research Areas: Security; cryptography; geometry; probability theory; cellular telephony; history of cryptography

Marty Reiman, Bell Labs, Lucent Technologies
Research Areas: Applied probability: queueing theory, stochastic control, diffusion processes; applied to broadband networking and other computer/communication systems

Nick Reingold, AT&T Labs
Research Areas: Algorithms; network design

Mauricio G. C. Resende, AT&T Labs
Research Areas: Combinatorial optimization; interior point methods

Eric Ristad, Princeton University
Research Areas: Computational linguistics; statistical modeling

Fred Roberts, Rutgers University
Research Areas: mathematical models in the social, behavioral, biological and environmental sciences and of problems of communications and transportation; graph theory; measurement theory; utility, decisionmaking, and social choice.

Joseph G. Rosenstein, Rutgers University
Research Areas: Mathematics education; linear orderings; recursive model theory

Michael H. Rothkopf, Rutgers University
Research Areas: Practice of operations research; models of bidding; energy economics

Avi Rubin, Bellcore
Research Areas: Computational complexity; algorithms; extremal set theory; partially ordered sets

Michael Saks, Rutgers University
Research Areas: Computational complexity; algorithms; extremal set theory; partially ordered sets

Iraj Saniee, Bellcore
Research Areas: Network and combinatorial optimization; SONET and ATM network design

Yatin Saraiya, Bellcore
Research Areas: Databases

Peter Sarnak, Princeton University
Research Areas: Data structures; graph algorithms

Serap Savari, Bell Labs, Lucent Technologies
Research Areas: Data compression; communication; algorithms; operations research

Rob Schapire, AT&T Labs
Research Areas: Learning theory and applications

Robert Sedgewick, Princeton University
Research Areas: Analysis of algorithms; algorithm visualization

Ravi Sethi, Bell Labs
Research Areas: Programming languages

Paul Seymour, Princeton University
Research Areas: Graph theory

David Shallcross, Bellcore
Research Areas: Discrete optimization

David F. Shanno, Rutgers University
Research Areas: Linear and nonlinear programming; numerical analysis; parallel computing

Saharon Shelah, Rutgers University
Research Areas: Logic; model theory

Lawrence Shepp, AT&T Labs
Research Areas: Probability; tomography; combinatorics

Peter Shor, AT&T Labs
Research Areas: Quantum computing; algorithms; computational geometry

Victor Shoup, Bellcore
Research Areas: Algebraic algorithms; number theory

Charles Sims, Rutgers University
Research Areas: Computational group theory and algebraic algorithms

Jaswinder Singh, Princeton University
Research Areas: Parallel computing applications and systems; computational biology

Neil J. A. Sloane, AT&T Labs
Research Areas: Mathematics; statistics

Warren Smith, NEC
Research Areas: Fundamental limits on computation set by laws of physics; algorithms; chess programming; optimal shapes

Eduardo Sontag, Rutgers University
Research Areas: System and control theory; learning and neural networks

Diane Souvaine, Rutgers University
Research Areas: Design and analysis of geometric and graph-theoretic algorithms

William Steiger, Rutgers University
Research Areas: research interests: Algorithms; computational geometry; probability; parallel computation

Ken Steiglitz, Princeton University
Research Areas: Computing with particles; economic simulation; and computer music.

Don X. Sun, Bell Labs, Lucent Technologies
Research Areas: Industrial experimental designs; quality improvement through experimentations and data analysis; combinatorial design of experiments; signal processing and speech recognition; pattern classification and artificial intelligence; and graphical methods for large data analysis.

Mario Szegedy, AT&T Labs
Research Areas: Algorithms and complexity; approximations

Endre Szemeredi, Rutgers University
Research Areas: Number theory; extremal graphs; theoretical computer science

Robert Tarjan, Princeton University
Research Areas: Data structures; graph algorithms; complexity

Simon Thomas, Rutgers University
Research Areas: Model theory; infinite groups

Robert J. Vanderbei, Princeton University
Research Areas: Optimization

Wolmer Vasconcelos, Rutgers University
Research Areas: Commutative algebra and computational algebra

Venkie Venketesan, Bellcore
Research Areas: Cryptography

Kiem-Phong Vo, AT&T Labs
Research Areas: Discrete/graph data structures and algorithms; graph visualization; pattern matching; data compression; software architecture and design, reliability, libraries

Gordon Wilfong, Bell Labs, Lucent Technologies
Research Areas: Computational geometry; algorithms

Albert C. Williams, Rutgers University
Research Areas: Linear, nonlinear, and integer programming; applications of operations research

Peter Winkler, Bell Labs, Lucent Technologies
Research Areas: Discrete math; probability; theory of computing

Avishai Wool, Bell Labs, Lucent Technologies
Research Areas: Quorum systems; fast communication networks; and combinatorial optimization

Paul Wright, AT&T Labs
Research Areas: Probabalistic methods; stochastic processes

Rebecca Wright, AT&T Labs
Research Areas: Security

Mihalis Yannakakis, Bell Labs, Lucent Technologies
Research Areas: Algorithms; complexity; databases; testing

Andrew Yao, Princeton University
Research Areas: Computational complexity; analysis of algorithms

Ann Yasuhara, Rutgers University
Research Areas: Recursive functions; logic; theoretical computer science