D I M A C S

Center for Discrete Mathematics and Theoretical Computer Science

ANNUAL REPORT

December 1998

DIMACS Center • Rutgers University • 96 Frelinghuysen Road • Piscataway, NJ 08854-8018 • USA
TEL: 732-445-5928 • FAX: 732-445-5932 • 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 - Research, Bellcore, Bell Labs, and NEC Research

Contents

1 Summary

In this report, we summarize the accomplishments at DIMACS in the period February 1, 1998 to January 30, 1999, under the DIMACS National Science Foundation Science and Technology Center award as well as activities partially supported by other sources or tracing their roots to the DIMACS STC award. This year marks the tenth birthday of the center and we are very proud of its many accomplishments.

Special Focus Activities

The primary scientific activity at DIMACS remains the "special year" or special focus," a focused and coordinated collection of activities that span a period of time, usually at least six months and often more than one year. The focus concentrates on an area of discrete mathematics/theoretical computer science that is ripe for rapid development, contains problems appropriate for high quality research mathematicians and computer scientists, and, in most cases, has strong potential for connections to other areas of the sciences, the mathematical sciences, and/or industry. (We use the terms "special focus" and "special year" somewhat interchangeably. A special focus tends to be smaller in scope than a special year, but sometimes we rename something a special focus after it has been a special year and sometimes we continue a special year more than one year.) During the academic year 1997-98, we ran five special focus activities. We introduced one major new special focus, on Massive Data Sets, and continued our special foci on Networks, Discrete Probability, the latter joint with the Institute for Advanced Study, DNA Computing, and Mathematical Support for Molecular Biology. During 1997-98, we also geared up for the 1998-99 Special Year on Large Scale Discrete Optimization by running a "pre-special year" workshop on that topic. In 1998-99, in addition to launching the Large Scale Discrete Optimization special year, we are continuing the Molecular Biology, Networks, Massive Data Sets, and DNA Computing special foci and, in a modest way, our special focus on Discrete Probability. Our Networks special focus is being supported in part by the National Security Agency, allowing us to extend it until the year 2000. We are also planning a 1999-2000 Special Year on Computational Intractability. Every special focus activity has led to dramatic scientific achievements – solutions of major open problems, the change of direction of research by a large community, the stimulus of new kinds of collaborations. The scientific achievements of our special focus activities are summarized in Section 3.

Multi-disciplinary, Multi-institutional Research Projects

In addition to special years, DIMACS has continued several major multi-disciplinary, multi-institutional research projects on "Simulations of Integrated Communications Systems" under NSF support and on "Scalable Self-Organizing Simulations" under DARPA support and continued our LINK project on software for research and education in discrete mathematics. See Section 3.9.

Continuing and Exploratory Activities

It is the plan of DIMACS to return from time to time to successful past foci. Computational Geometry is a case in point. We returned to it in 1998 with a workshop on Combinatorial and Algorithmic Geometry as part of our international partnership with the DIMATIA Center in the Czech Republic. The 1995-96 Special Year in Logic and Algorithms is another case in point. We are planning a return to it with a joint DIMACS-DIMATIA workshop on Algebraic Methods and Arithmetic Circuits in June 1999. The Graph Theory and Algorithms Special Year was followed up by an emphasis on graph theory in the summer of 1998 at the DIMACS Research and Education Institute (DREI). One way to develop topics for possible future special foci is to hold exploratory workshops. We explored one such topic at the interface between the mathematical and the physical sciences with a workshop on discrete mathematical chemistry held in March 1998 and are planning to explore on-line decision making, cognitive science, financial modeling, and unusual applications of number theory in 1999.

Evaluation Reports on Special Years

We continued our program of evaluating the impact of earlier special years by commisioning an evaluation of DIMACS' second special year, the 1990-91 Special Year on Complexity Theory of Interactive Computing as seen from the perspective of eight years. We have now received five such reports, with more on the way. The report, which is discussed in Section 3.8, provides an exceptionally strong piece of evidence of the long-term impact of DIMACS' progams. A similar report on our 1994-99 Special Year on Mathematical Support for Molecular Biology says that "The special year has been enormously successful in meeting its goals" and is discussed in Section 3.4. This year, we have also commissioned a report on our 1992-93 Special Year on Combinatorial Optimization. The report is expected soon.

Regular Evaluation of our Programs

We have in place a program of regular evaluation by our workshop participants, visitors, postdocs, graduate students, and undergraduate students, run by one of our staff members. We also evaluate our special programs such as Reconnect and DREI, using professional evaluators. The data coming out of these evaluations offers powerful evidence that we are accomplishing our purposes. Evaluations by our visitors and workshop participants are summarized in Section 5.4. Evaluations by our postdocs, graduate students, REU students, and of our DREI and Reconnect programs are discussed in the corresponding sections of this report and we are really excited about the many positive results that we get as well as the useful feedback, especially for the newer programs.

Undergraduate Initiatives

This year brought to fruition the hard work of previous years to develop new and innovative programs in undergraduate education. We launched a new module series aimed at rapid, electronic publication of modules aimed at undergraduates and teachers of undergraduates and in some cases high school teachers and students. We initiated a "Reconnect" project aimed at reconnecting to the mathematical sciences enterprise faculty whose primary job is teaching undergraduates to the mathematical sciences enterprise. As part of this project, we ran a summer conference aimed at introducing such faculty to current research topics that are relevant to the classroom and at having them develop written materials usable in the classroom; and we ran a series of two-day and one-day conferences with a similar makeup aimed at two year college faculty. Some of the material developed by participants in our "reconnect" program is being developed in collaboration with editors in our Module Series. Our traditional undergraduate program, the research experiences for undergraduates (REU) program, continued with 12 students. We tried something new here as well, by, for the first time, involving a student who was just completing her freshman year. This was done in a new collaboration with the Douglass College Program for Women in Math, Science, and Engineering. One of our REU students, Carly Klivans from Cornell University, won the undergraduate research award (the Alice T. Schaeffer Prize) from the Association for Women in Mathematics, based on her work at DIMACS.

Precollege Programs

Our highly successful precollege programs have continued, with several major new developments. Our concentration on ways to expand the impact of our Leadership Program beyond the DIMACS site has led to a new emphasis on running off-site programs and on developing portable materials that can be used to "franchise" our programs. NSF has phased out support of Young Scholars programs like ours. We turned instead to a combination of tuition payments and gifts to run the program, and were lucky to receive a major gift from the AT&T Foundation to support it. In earlier years, we played a central role in the adoption of a new Curriculum Framework in Mathematics in New Jersey. The "Framework" and the 1997 DIMACS volume, *Discrete Mathematics in the Schools*, have begun to have a real impact on mathematics education in the schools.

Graduate Student Programs

Graduate students are a central part of a science and technology center. Our local graduate student program consists of small DIMACS fellowships to support the research of local students, small awards for summer research projects, and jobs for which graduate students are paid. The response to this program has been highly enthusiastic. We supported 27 students this way in the past year, and this support has led to rather impressive scientific payback, as demonstrated by some of the scientific results cited in Section 3. We have also continued our programs of sponsoring DIMACS visits by non-local graduate students and continued our long-standing policy of providing a supplemental allocation to each workshop for support of non-local graduate students.

Postdoc Programs

One of DIMACS' goals is to involve outstanding young people in the life of the center. Our postdoc program is designed to accomplish this. We have leveraged our funds for support of postdocs by sharing postdocs with the Institute for Advanced Study and with the University of Pennsylvania, obtaining grant support for postdocs, and including in our postdoc program postdocs appointed by our industrial partners. AT&T Labs has given us a gift for support of joint postdocs, another way we are now beginning to leverage our postdoc funds, and we have similar plans with NEC. We have instituted regular meetings with our postdocs to discuss the job market. Continuing procedures developed in previous years, each postdoc is assigned a faculty mentor. Moreover, because we want to train our postdocs to take an interest in educational activities and in outreach beyond the mathematical sciences, we ask each postdoc take part in such activities. Our postdocs meet this requirement in a variety of ways, as described in Section 4.3. Our postdoctoral fellows from past years have been remarkably successful in obtaining positions, both in academics and in industry. They have also rapidly risen to the top of their profession.

Integration of Research and Education

We have worked to integrate research and education by including educational components in our special foci, involving postdocs and graduate students in educational programs, and through our new undergraduate initiatives such as the module series and the "Reconnect" project. We also ran the second year of our DIMACS Research and Education Institute (DREI), under a $1.9M grant from NSF. This program is aimed at integrating researchers, postdocs, students, and teachers, through a series of research workshops, a high school teacher course, and plenary activities that link these two main groups. DREI was extensively evaluated by professional evaluator Dr. Norman Webb, who found extremely strong evidence of real interactions between teachers and researchers.

Institutional Membership

In December 1997, the NEC Research Institute, Princeton, NECI, became a member institution in DIMACS. The first full year of NECI's formal involvement with DIMACS has proven fruitful for the center, with the active involvement of NECI scientists in an organizational and advisory role in many DIMACS projects.

Partnerships

A key goal of DIMACS is outreach to other communities, both within and without the mathematical sciences. In the past year, we have made major progress in enhancing and expanding our many partnerships. The highly successful partnership with the Institute for Advanced Study has continued with joint workshops and shared postdocs. We ran a joint postdoc and workshop program with the University of Pennsylvania. We have initiated several joint activities with the Center for Molecular Biophysics and Biophysical Chemistry at Rutgers. We held a workshop sponsored jointly with Georgia Tech and Carnegie Mellon University and we laid plans for future joint workshops with Georgia Tech and with the Institute for Research in Cognitive Science (IRCS) at U. Penn and the Institute for Mathematics and its Applications (IMA) in Minnesota. On the international front, our contacts with other institutions have been formalized and have matured into joint programs. During this year, we have run two joint workshops with our partners in the DIMATIA Center in Prague (with NSF support), one with the new Erdős Institute in the Hungarian Academy of Sciences, and one with the Computer Science Department of the Hong Kong University of Science and Technology, and we have exchanged visitors with these partner institutions and others such as INRIA. Further international collaborations are in the process of being developed.

Programs for Underrepresented Groups

We initiated a collaboration with the Douglass College (Rutgers University) Program for Women in Mathematics, Science, and Engineering. Fred Roberts joined the Advisory Board for this program and, through this program, we invited a Douglass College freshman to participate in our REU, Reconnect, and DREI programs with the costs shared between DIMACS and Douglass. We continue to have very strong participation in the center by women. Several African Americans and Hispanics play prominent roles in DIMACS and the involvement by African Americans in at least some of our programs such as Young Scholars is strong. We opened discussions about joint programs with the University of Puerto Rico, continued discussions with Spelman College and Howard University about joint activities, and offered a postdoctoral fellowship to a Hispanic woman.

Changes in Personnel

Julia Abrahams will be completing her first year as Associate Director in January. Val Debellis, our education coordinator, left DIMACS in August for a position at East Carolina University. She continues to work on DIMACS programs as a consultant. Christine Spassione has joined us the new Visitor Coordinator, with Christine Houck taking on the role of publicity coordinator. Later, Christine Houck also left for another position, leaving us with the need to replace her as publicity coordinator.

DIMACS Book Series

The DIMACS series, published with high quality hard cover volumes, reached a milestone with the publication of the 40th book in the series. There are now 45 volumes.

2 Purposes, Goals, and Programs of DIMACS

This section summarizes the purpose of DIMACS and its goals and discusses its major programs relative to its purpose and goals.

2.1 DIMACS Purpose

The Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) is a collaborative effort of Rutgers and Princeton Universities, AT&T Labs-Research, Bell Labs, Bellcore, NEC Research Institute (NECI) 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.

2.2 DIMACS Goals

Motivated by its purpose, DIMACS has established a number of specific goals.

- Goal 1. High Quality Research

  All of the research programs carried out under DIMACS' auspices should be of the highest quality.

- Goal 2. Choose Appropriate Areas of Research Emphasis
  - Goal 2a. Choose areas of research emphasis that are poised for rapid advancement and have the potential for high impact.
  - Goal 2b. In choosing areas of research, assure that a large percentage of them have a strong potential for applications in other areas of science, the mathematical sciences, and/or industry.
  - Goal 2c. In choosing areas of research, assure that a large percentage of them have a strong potential for introducing researchers in dm/tcs to problems that arise in other subject areas.
- Goal 3. Take Advantage of the Center Mode of Operation
  - Goal 3a. Provide opportunities for concentrated efforts that explore an important field in ways that could not be achieved by individual researchers working alone or in small groups.

- Goal 3b. Utilize DIMACS permanent members' experience and reputation to organize large, complex programs to achieve scientific goals beyond what can be done with individual PI's or in traditional academic settings.
- Goal 3c. Provide mechanisms for individuals at the local institutions to cooperate, leveraging the strong talent that exists and its eagerness to organize and lead activities.
- Goal 3d. Provide a dynamic and intellectually exciting environment that attracts visitors, thus enhancing the possibilities for scientific collaborations.

- Goal 4. Facilitate Interactions and Collaborations within dm/tcs
  - Goal 4a. Make it easy for researchers both locally and from all over the country (and the world) to interact, thus facilitating rapid development of an area.
  - Goal 4b. Connect researchers in dm/tcs in differing communities or with differing areas of emphasis.
- Goal 5. Widespread and Rapid Dissemination of Plans, Results and Materials
  - Goal 5a. Disseminate widely and rapidly the latest results, ideas, and problems from dm/tcs and their applications, so that as large a community of researchers as possible can use or be involved in the development of an area.
  - Goal 5b. Disseminate or promote widely outstanding educational materials in dm/tcs that are either produced at DIMACS or elsewhere, so as to maximize their impact on education.
- Goal 6. Involve Outstanding Young Researchers

  Involve outstanding young people in all DIMACS programs, so as to maximize the long-term impact of these programs.
- Goal 7. Develop a National and International Scope and Make the Center a National Resource
  - Goal 7a. Involve scientists in dm/tcs from all parts of the nation and the world so as to maximize its impact on the field of dm/tcs.
  - Goal 7b. Make DIMACS a resource for the dm/tcs community nation-wide.
  - Goal 7c. Further national goals by involving as large a number of under-represented groups and women in our programs as possible.
- Goal 8. Long-Term Impact

  Have a lasting impact on the directions of the field of dm/tcs.
- Goal 9. Involve Scientists from Other Communities
  - Goal 9a. Expose leading researchers in other areas of science, the mathematical sciences, and industry to the methods of dm/tcs and their potential for usefulness in these other areas.
  - Goal 9b. Assist researchers in other areas of science, the mathematical sciences, and industry in translating their problems into forms for which dm/tcs can be helpful.
  - Goal 9c. Establish partnerships between researchers or practitioners from other fields of science, the mathematical sciences, and industry with leading researchers from dm/tcs, so that the researchers from other fields might pose problems that will help to stimulate the development of dm/tcs.

- Goal 9d. Involve outstanding young researchers in partnerships with researchers or practitioners from other fields of science, the mathematical sciences, and industry, aiming to create a career-long interest in achieving stimulation from outside the field of dm/tcs.

- Goal 10. Involve dm/tcs researchers with Other Fields of Research
  - Goal 10a. Encourage some leading researchers in dm/tcs to become involved in applications of these fields.
  - Goal 10b. Expose many outstanding young people in the fields of dm/tcs to the applications of these fields.
  - Goal 10c. Involve outstanding young people in the fields of dm/tcs with the applications of dm/tcs in a serious way, and train them so they are prepared to make an impact in other fields of science, the mathematical sciences, or industry.

- Goal 11. Educate and Train Outstanding Young People in dm/tcs
  - Goal 11a. Involve graduate students and undergraduates in programs that expose them to the latest ideas, problems, and methods of dm/tcs and expose them to the excitement of the field and give them unique research experiences.
  - Goal 11b. Involve graduate students in relationships with those in other communities (see goals 9 and 10).
  - Goal 11c. Provide opportunities for high school students to learn dm/tcs as a means of encouraging them to continue working toward careers in the sciences and mathematical sciences.

- Goal 12. Contribute to the Professional Development of Teachers and to the Inclusion of dm/tcs Topics in Curricula
  - Goal 12a. Expose K-12 teachers to the content and applications of dm/tcs, as well as to ideas from contemporary research in the field, so that they may in turn design or modify curricula and convey to their students the excitement of these fields and their applications.
  - Goal 12b. Expose undergraduate college faculty to new ideas, results, and applications of dm/tcs at the undergraduate level so that they may in turn design or modify curricula and convey to their students the excitement of these fields and their applications.
  - Goal 12c. Create lines of communication and interaction between classroom teachers and researchers for their mutual benefit as educators, and to improve mathematics and computer science education at all levels.

- Goal 13. Involve Broader Mathematical Community
  - Goal 13a. Inform the broader community – that is, practicing research mathematicians and computer scientists, educators in the mathematical sciences at all levels from college to junior college to precollege, and students in the mathematical sciences at all levels – of the important results and applications of dm/tcs, and encourage their interest in these fields.
  - Goal 13b. Promote the widespread inclusion of dm/tcs in the curricula of students at the precollege, college, graduate, and postgraduate levels.

- Goal 14. Involve the dm/tcs Research Community in Education

- Goal 14a. Involve senior and junior researchers in educational activities addressed to classroom teachers and high school students, with the goal of creating career-long commitments to participation in education at all levels.
- Goal 14b. Work to improve mathematics and computer science education both locally and nationally, in collaboration with the education community (teachers, schools of education, college and university faculty, and relevant professional organizations).
- Goal 14c. Serve as a national resource and center for discrete mathematics in K-12 education.

- Goal 15. External Support/Linkages
  - Goal 15a. Increase our effectiveness and impact by obtaining additional support for DIMACS' programs from government agencies (including parts of NSF) and the private sector.
  - Goal 15b. Increase our effectiveness and impact on our field and on others by creating linkages between our center and other organizations.
  - Goal 15c. Increase our effectiveness and impact on our field and on others by leveraging our resources.
  - Goal 15d. Demonstrate the value of the center beyond year 11 of our STC award by exploring and developing alternative means of external support and linkages.

- Goal 16. Institutional Support
  - Goal 16a. Increase our effectiveness and impact by obtaining support for DIMACS' programs from our member organizations.
  - Goal 16b. Demonstrate the value of the center beyond year 11 of our STC award by exploring and developing additional means of support from our present institutional members and exploring the possibility of adding additional institutional members.

- Goal 17. Develop Managerial Structures that Enable the Organization to Reach its Other Goals Efficiently
  - Goal 17a. Make use of the expertise/knowledge of the many leaders of the dm/tcs community who are part of DIMACS, in planning and running programs, setting priorities, and identifying major opportunities.
  - Goal 17b. Make use of the large number of DIMACS member scientists to plan, manage and organize its programs.
  - Goal 17c. Centralize decisionmaking so as to connect all decisions to center goals.
  - Goal 17d. Streamline center decisionmaking so as to be able to take advantage of opportunities in a timely manner.
  - Goal 17e. Take advantage of expertise in the broader dm/tcs and, more generally, scientific community.
  - Goal 17f. Develop careful budget management methods and oversight so as to maximize our effectiveness and accountability.
  - Goal 17g. Regularly evaluate programs and use the evaluation to improve them, so as to maximize the impact and success of the center's activities.
  - Goal 17h. Regularly evaluate management and make changes in policies or personnel as necessary, so as to maximize the impact and success of the center's activities.

– Goal 17i. Inform the general community about DIMACS' activities and accomplishments so as to develop constituencies that will help us to maximize our effectiveness in reaching our other goals.

2.3 DIMACS Programs

Based on its purposes and goals, and taking account of its assets, DIMACS runs a variety of programs.

2.3.1 DIMACS Assets

The main asset of DIMACS is its dedicated and talented group of permanent members, some 175 researchers from six member organizations – Rutgers, Princeton, AT&T Labs - Research, Bell Labs/Lucent Technologies, Bellcore, and NEC Research. The members include some of the most important leaders in the fields of dm/tcs in the world, and provide a critical mass of scientific talent in dm/tcs unmatched anywhere. DIMACS' six cooperating institutions bring institutional resources and commitment to it, with strong emphasis on the encouragement of member scientists to participate both in the scientific activities of the center and its organizational/management activities. The physical facilities at both Rutgers and Princeton provide us with office space, seminar, conference, and lecture rooms, and equipment that can be used in support of our many programs. These facilities have been provided by our member organizations. We also have an experienced staff, expert in running conferences, hosting visitors, and dealing with postdocs and students. DIMACS also counts as its assets linkages to important organizations both inside and outside of dm/tcs and all over the world, with particular emphasis on the Center for Math, Science, and Computer Education and on the Institute for Advanced Study. Less tangible but no less important is the reputation that DIMACS has developed over the years, as a world-class center in the fields of dm/tcs.

2.3.2 The Primary Activities of DIMACS

The primary activities that DIMACS has developed to accomplish many of its goals and best take advantage of its assets are "special years" and "educational outreach activities." These activities receive the large majority of our center's funds.

2.3.3 Special Years

A "special year" refers to a focused and coordinated collection of activities that span a period of time, usually at least a year.[1] More than one special year may occur at one time. The focus of each special year is an area of dm/tcs which is selected for special attention because it satisfies most of the following criteria:

- (Cr1) the area is attractive to and contains problems appropriate for high quality research mathematicans and computer scientists;
- (Cr 2) the area has strong potential for applications in other areas of the sciences, the mathematical sciences, and/or industry;
- (Cr 3) the area has strong potential for introducing researchers to problems that arise in other subject areas;

[1] We use the terms "special focus" and "special year" somewhat interchangeably. A special focus tends to be smaller in scope than a special year, but sometimes we rename something a special focus after it has been a special year and sometimes we continue a special year more than one year.

- (Cr 4) the area is ready for rapid development and would benefit from increased attention by researchers; and
- (Cr 5) the area attracts substantial interest from DIMACS members.

Thus, a special year is chosen primarily to address the first three clauses in the DIMACS Purpose. By its very definition, a special year is designed to meet the goals of choosing appropriate research as spelled out in Goal 2.

It is anticipated that designating a DIMACS special year will result in sharply increased attention to the designated areas by the research community both at the time of the special year and long afterwards. Our special years almost invariably involve an outreach component, catering to a different scientific community, a different part of computer science or mathematics, or to industry.

Most special years at DIMACS involve workshops, mini-workshops (one day conferences often planned on short notice in order to respond to an especially promising opportunity or to pursue an especially promising topic), lecture and seminar series, tutorials, visitors, postdocs, and graduate students. The workshops, mini-workshops, lecture and seminar series, and visitor programs are designed to maximize scientific interactions and usually include participants from other areas of science, the mathematical sciences, and industry. We involve postdocs and graduate students in our special years in order to involve outstanding young people. The postdocs stay for periods ranging from six months to two years; visitors and graduate students participate in workshops or spend time at the center for periods of a few days to a year. The emphasis in choice of visitors, postdocs, and students is national scope, but we seek ways to involve local students in our programs and we hire them to participate in research programs funded by our non-STC sponsored research and other non-STC support. Tutorial programs are aimed at introducing researchers in dm/tcs to areas outside their special interests. Special years include careful attention to dissemination of results and ideas through technical reports in the DIMACS series, papers published in the literature, and volumes in the DIMACS/AMS book series and special issues of journals. In addition, we utilize less traditional avenues of dissemination, in particular web pages associated with workshops or special years, both present and past.

In order to choose areas of emphasis appropriately, we have a strong program to identify and develop topics for future special years through exploratory workshops, lecture and seminar series, and visits from leaders in various areas of research.

2.3.4 Other Research Activities

In addition to the special year activities, in order to fulfill its role as a leader in dm/tcs and to reach out to other communities, DIMACS introduces activities devoted to research in topics satisfying the above list of criteria (Cr1-Cr5), but which may not lead to a special year. In particular, we do this through our workshop, implementation challenge, visitor, and postdoc programs. For example, because we think that the best young scientists in dm/tcs can have special impact on our search to identify important new areas, and because we think it is especially important to influence these young scientists who will become the leaders of our field in the future, we involve some postdocs simply because of their excellence, independent of specialization. Similarly, we sometimes hold a workshop on a topic that may or may not lead to a special year, because the topic can be chosen to achieve one of the first three clauses (P1-P3) in the DIMACS Purpose. In carrying out these non-special year research activities such as non-special-year workshops, as with special years, we take into account our goals of involving outstanding young researchers and talented graduate students, our goals of reaching out to other communities, and our goals with regard to dissemination. We use the various methods of dissemination discussed relative to special years.

2.3.5 Educational Outreach Activities

An educational outreach activity is designed to inform the broader community about research in dm/tcs and and/or to facilitate the use of contemporary topics of dm/tcs in classrooms. Such activities are usually chosen to further the final two clauses in the DIMACS Purpose (P4 and P5) and to some extent the second clause (P2), specifically relative to the goal of exposing leading researchers outside of dm/tcs to these areas and their potential for usefulness. We run education, training, and "public awareness" programs directed to college and high school teachers, university and industry scientists, and undergraduate and graduate students. These programs focus on the advances being made in dm/tcs and the applications and impact of this research. More intensive programs enable DIMACS to train outstanding high school and undergraduate students in areas of dm/tcs and to familiarize experienced teachers with dm/tcs and provide them with materials so that they can incorporate them into their classrooms and curricula. To maximize our long-term impact, we seek ways to disseminate the results of our educational outreach activities by disseminating materials, through newsletters, edited volumes, modules, and the internet. The specific programs we have developed along these lines include our leadership programs for K-12 teachers; our young scholars program for high school students; our REU program for undergraduates; and our outreach activities aimed at two- and four-year college teachers through designated visitor slots for such teachers, hosting and organizing meetings of the local chapter of the Mathematical Association of America and the Two Year College Math Association, and providing speakers at meetings and conferences for college and high school teachers and students.

DIMACS serves as a national resource and center for discrete mathematics in K-12 education by publishing a newsletter "In Discrete Mathematics: Discrete Mathematics in the Classroom" for teachers; providing workshops in discrete mathematics in individual districts, with trained leaders provided through the leadership program; and sponsoring conferences and meetings for leaders in the educational community to share information and develop strategies for bringing discrete mathematics into classrooms and curricula. In addition, we publish curriculum materials for K-12 teachers on topics in discrete mathematics, including articles written by participants in our teacher programs (for example, our "Greatest Hits in Discrete Mathematics" publication); and we host and facilitate an email network of K-12 teachers who are using discrete mathematics in their classrooms. We also facilitate and support the development and use of technology which can enhance the learning of discrete mathematics, in particular LINK, a project for develping software for discrete mathematics research and education.

We affirm the importance of improving mathematics education locally and nationally by joining in partnerships with the education community – with teachers, schools of education, college and university faculty, and relevant organizations. Our long-standing partnership with the Center for Math, Science, and Computer Education at Rutgers forms the foundation for these activities. We pursue this goal by participating actively in the New Jersey Statewide Systemic Initiative, in the development of the New Jersey Mathematics Standards and the New Jersey Mathematics Curriculum Framework, and providing the mathematical sciences research community with models for collaborations and partnerships between the research and education communities, models for successful and portable projects that can be replicated nationally. One example of such a partnership is the DIMACS Research and Education Institute described in the next subsection.

2.3.6 Integrating Research and Educational Activities

DIMACS feels that its goals are best achieved by coordinating its many activities. In particular, the educational outreach programs and special year programs are linked through the incorporation of educational outreach activities in the planning and implementation of special years, and the direct participation of special year researchers in educational outreach activities. The DIMACS

Research and Education Institute, DREI, is aimed at integrating researchers, postdocs, graduate students, and precollege teachers, with the long-term goal of making education more relevant and alive, and creating long-term partnerships between researchers and educators. Our "Reconnect" Program is aimed at bringing current research topics to the attention of two- and four-year college teachers who might use the ideas in their teaching and who might get "reconnected" to research by participation in a world-class research center. Our "Modules" series integrates research and education by bringing out in readable written or electronic form current research material that is relevant to the college or precollege classroom. We involve our permanent members, postdocs and graduate students in the educational programs of DIMACS, in order to help instill in them a career-long commitment to helping in the education of young people.

2.3.7 Continuing or Returning to Successful Activities

The success of special years, educational outreach activities, and other projects at DIMACS is monitored both while they are in progress and afterwards. Because of our desire to have a long-term impact on a developing field, we take one of several actions if an activity has been successful: 1) help to make sure that it leads to an activity that is self-standing (and perhaps informally connected to DIMACS); 2) return to it with occasional future programs; 3) develop it into a regular, continuing activity of DIMACS; 4) seek renewed funding for continuing it over a period of time.

An example of the first is the annual CAV conference that was born of an early workshop at DIMACS. Examples of the second are the second workshop on "Combinatorics of Statistical Mechanics Models" to be organized by DIMACS in the Spring of 1999 and the research program on graph theory in the Summer of 1998 at the DIMACS Research Education Institute, following up on an early special year. The third is illustrated by the continuing series of DIMACS algorithm implementation challenges and the continuing series of workshops and postdoc programs in mathematical support of molecular biology. The latter also present examples of the fourth, as do our Leadership and Young Scholars Programs in K-12 education; the DIMACS Research and Education Institute; and our project on simulations of integrated communications systems.

2.3.8 Linkages

DIMACS wants to have a maximum impact both within dm/tcs and on fields outside. In order to accomplish this, we need to obtain additional support from government agencies and the private sector in order to leverage our STC funds, and to establish linkages, both formal and informal, with a variety of organizations both inside the field of dm/tcs and outside it. The STC funds do not begin to pay for the wide ranging, ambitious agenda we have set for ourselves. We seek supplementary funds for our special years, workshops, educational outreach activities, postdoc, and graduate student programs, and for our regular, continuing activities. We also seek additional funds from our member institutions to support these kinds of programs and to use as matching funds in obtaining further external funding. We seek linkages to other organizations so as to maximize exchange of ideas, methods, concepts, and people. Sometimes these linkages involve formal arrangements such as affiliated membership, subcontracting, or partnerships in supporting special focus programs or postdocs. Our relationship with the Institute for Advanced Study illustrates both our subcontracting and partnership links, the former through our DIMACS Research and Education Institute program and the latter through a cooperative postdoc program and a collaboration on a special focus in discrete probability. Our relationship with the Center for Math, Science, and Computer Education is another example of a formal partner, as is DIMATIA, the Center for Discrete Mathematics and Theoretical Informatics in Prague, Czech Republic. Other examples of formal linkages are service on external advisory boards of other organizations. Less formal linkages are those in which members of other organizations, to some extent representing their organizations, serve on various DIMACS committees. Examples of these kinds of linkages are the service of representatives of the Waksman Institute for Molecular Biology, the Center for Molecular Biophysics and Biophysical Chemistry, and the Center for Theoretical and Applied Genetics at Rutgers, Center for Molecular Biotechnology at the University of Washington, and the Genome Center at the University of Pennsylvania on various steering committees for the Special Year in Mathematical Support for Molecular Biology. Less formal linkages can also involve collaborative research between members of different organizations, hosting of meetings of other organizations such as the Mathematical Association of America New Jersey Section, the International Conference on Combinatorial Pattern Matching and the International Seminar on Average Case Analysis of Algorithms.

2.3.9 Management

In order to run our large array of programs, pick and choose programs carefully, allocate resources, and determine priorities, we seek to develop managerial structures that enable the organization to operate efficiently. One of our primary assets is the large number of leaders in the field of dm/tcs who are permanent members of DIMACS. We make use of their ideas, experience, connections, and insight in planning programs and setting priorities. To do so, we involve a large number of them in an advisory group called the DIMACS Council that meets several times each academic year for exchange of ideas and discussion. In addition, we involve a large number of permanent members of DIMACS from all of our member organizations in the planning and organization of different programs. We do this by establishing a strong system of committees. A typical DIMACS committee has representatives from each DIMACS constituent unit (meaning AT&T Labs, Bellcore, Bell Labs, NECI, Princeton CS, Rutgers CS, Rutgers Math, and Rutgers OR). There are committees devoted to postdoc selection, graduate student selection, visitor selection, projects (identification of workshop, special year, and research project topics), educational programs (chaired by our Associate Director for Education), equipment, and strategic planning. Decisionmaking is centralized in order to connect decisions to center goals. Thus, each committee reports to the director and/or to the managing Executive Committee.

Center decisionmaking is primarily the function of an Executive Committee, which is designed to be as small as possible but still have one representative from each constituent unit. The decisionmaking process between a strong director and the EC is clarified through a set of bylaws. In order to make our center national in scope and take advantage of the expertise in the dm/tcs community nationally and internationally and more generally of the national scientific community, we make heavy use of our External Advisory Board, in particular making use of them in evaluating our programs.

In order to run our center well, we centralize all budget management, and appoint a budget committee consisting of senior representatives from our member organizations who have budget oversight and serve to advise the director about problems and opportunities and about budget reallocations. We regularly evaluate all programs and center management, making sure that our programs and our leadership are maximizing our chances of achieving our organizational goals.

In order to maximimize our effectiveness and impact, we seek to develop programs that publicize DIMACS and its successes. We publicize DIMACS on electronic nets and newsgroups, scientific news magazines and newsletters of both the mathematical sciences and other areas, through sessions at meetings of professional societies, displays and brochures, and through the world wide web.

ABBREVIATIONS USED THROUGHOUT

ATTL = AT&T Labs - Research
BC = Bellcore
BL = Bell Labs
dm = discrete mathematics
GS = graduate student
IAS = Institute for Advanced Study
NEC = NEC Research Institute
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.

3 Research Accomplishments and Plans

3.1 Special Year on Large Scale Discrete Optimization, 1998-99

Overview:

The Special Year on Large Scale Discrete Optimization began in September 1998. Some pre-special year activities were held during 1997-98. The year builds on the 1992-93 Special Year on Combinatorial Optimization.

Organizers:

- Endre Boros, Rutgers University (local chair)
- William Pulleyblank, IBM
- Michael Trick, Carnegie Mellon University (general chair)
- Bob Vanderbei, Princeton University
- Vijay Vazirani, Georgia Tech

Sponsors and Collaborating Institutions:

The National Science Foundation
The New Jersey Commission on Science and Technology
AT&T Labs - Research
Bell Labs
Bellcore
Carnegie Mellon University
Georgia Institute of Technology
IBM Research
Institute for Advanced Study, Princeton
NEC Research Institute
Rutgers University
Princeton University

Theme of Program:

This special year is motivated by developments in the past ten years. There have been three simultaneous advances: at the algorithmic level, with new and interesting algorithms for solving large scale discrete problems both exactly and approximately; at the implementation level, providing codes and systems for solving such problems; and at the practical level, with interest in the fast solution of real world problems. Examples of such applications include airline crew scheduling, chemical process design, and telecommunication network design. Our goal in this year is to unite groups working at these three levels. The confluence of exciting recent work on approximation and other algorithms and greatly improved software/hardware for optimization makes large scale discrete optimization a practical, useful approach to solving problems of interest in applications.

The special year organizers have provided forums to increase research contacts between groups that do not frequently come in contact - for example practitioners and academics, operations researchers and computer science researchers, theoretically oriented researchers (in areas such as interior point IP methods, combinatorial geometry approaches, approximation methods, and group formulations) and computationally oriented researchers, including those interested in AI techniques such as constraint programming.

Workshops and Special Events, Feb. 1998 - Nov. 1998

- **Joint DIMACS-CMU-Georgia Tech. Workshop on Large Scale Discrete Optimization**
  - Dates: May 27-29, 1998
  - Location: Rutgers University
  - Attendance: 113
  - Organizers: R. Ravi, Carnegie Mellon University, and Vijay Vazirani, Georgia Tech

Very large problems in a variety of areas ranging from telecommunications to airline scheduling map to questions in discrete optimization. With a growing number of applications seeking to solve very large discrete optimization problems, impressive advances in the design of algorithms for solving such problems exactly and approximately, and the development of large computer systems that can support such activity, there is tremendous interest in this topic in industry and academia. The aim of this workshop then was to bring practitioners and theoreticians with insight into specific areas where large scale discrete optimization methods are used, to talk about central issues as well as outstanding open questions in their respective areas. Examples of such areas include telecommunication network design and routing, airline crew and flight scheduling, scheduling and production. The focus was to identify problems for which fundamental insight into their structure and solution will translate into practical gains in their application areas. Keeping such high impact problems in mind, a mix of leaders from different areas of application and development were invited to lecture in the workshop.

The talks were roughly organized under two broad themes, "Applications" and "Tools." While the Applications talks introduced several traditional areas where discrete optimization is applied in practice (e.g., airline industry, production planning, network survivability) as well as advances in several methodological areas of development in optimization (e.g., flows, Markov chain methods), the Tools talks focused on design and functionalities in popularly available tools (e.g., CPLEX, LEDA) to solve generic optimization problems.

In the Applications area, George Nemhauser and Ellis Johnson surveyed their long-standing work on applying optimization in airline integrated planning and operations and Jitendra Malik introduced graph partitioning and matching problems arising in computer vision, and current methods used to tackle them. Methodological advances in optimization Algorithms were featured when Prabhakar Raghavan described an exciting microeconomic view of customer segmentation and collaborative filtering and the problems arising therein; Rajeev Motwani described recent work on high-dimensional similarity search in large databases; and Eva Tardos surveyed methods for regulating flows in communication networks. Laurence Wolsey described large production planning problems focusing on modeling and software issues; Gary Miller discussed the use of graph embedding techniques in scientific computation; and Dana Randall surveyed Markov chain methods in physics. Finally, Craig Tovey described his intriguing work on self-organizing social structures in fish groups and his attempts at modeling and predicting their behavior; Michael Trick described his adventures in sports scheduling; and Milena Mihail surveyed optimization issues in network survivability and network evolution.

In talks devoted to Tools, Kurt Mehlhorn, a co-developer of the popular LEDA software package, talked about Secondary Memory Algorithms for Networks and Geometry, and capabilities to handle such methods in LEDA; Andrew Goldberg reported on insights on implementing maximum flow methods and describing new techniques and algorithms that arose from such implementations; John Hooker talked about the connections between logical and linear programming and proposed a hybrid stronger than either of the two. Also, Bob Bixby described recent developments in the new version 6 of the popular CPLEX package; Martin Savelsbergh described the central design issues in PARINO, a PARallel INteger Optimizer developed at Georgia Tech; and Sebastian Ceria described the idea of lift-and-project cuts, an efficient solution method for mixed integer programs.

- **DIMACS Workshop on Constraint Programming and Large Scale Discrete Optimization**
  - Dates: September 14-17, 1998
  - Location: Rutgers University
  - Attendance: 75
  - Organizers: Richard J. Wallace and Eugene C. Freuder, University of New Hampshire

  This workshop was concerned with all aspects of constraint programming related to large scale discrete optimization, including contributions from artificial intelligence, discrete mathematics, operations research and programming languages. The intention of the organizers was to bring together people interested in constraint programming approaches to large scale discrete optimization (LSO) problems, coming from the areas just mentioned and from both academia and industry, so that new results could be presented, problems posed, and the current state of the field clarified. To further these ends, the workshop included tutorial talks from well-known figures in key areas related to constraint programming and LSO, who reviewed work within their fields of expertise.

  There were tutorial talks on "Constraint programming languages for combinatorial optimization" by Pascal Van Hentenryck and "Integration of artificial intelligence and operations research techniques" by Ken McAloon, Carol Tretkoff and Carla Gomes.

  There was a session on local search methods for LSO and invited talks on optimization methods in NASA's autonomous Deep Space probe by Brian Williams and on search methods for quantum computing by Tad Hogg. Some sessions were devoted to an important European project, CHIC2, that has pioneered the development of "hybrid algorithms" that combine procedures derived from constraint satisfaction algorithms and from operations research (integer and linear programming). Papers were given by Tibor Kokeny and Vassilis Liatsos on applications for two different large-scale industrial problems, involving scheduling and, in one case, inventory management, and by Carmen Gervet on guidelines, or "methodologies", that were spawned by the CHIC2 project, for developing new hybrid algorithms for industrial problems. There was an invited talk by David Johnson on "Average case behavior of bin packing algorithms," which covered an important case of approximation methods for NP-hard problems. A session presented AI, or constraint satisfaction, techniques for complex applications ("course of action analysis" in a military context), as well as new techniques for solving constraint satisfaction problems. A session organized by Thomas Schiex was a review of the "valued constraint satisfaction" framework that can represent various types of optimization problems as overconstrained constraint satisfaction problems. There was an "industrial session" with presentations by James Crawford of I2 Technologies, Mark Boddy of Honeywell, and Helmut Simonis of Cosytec on LSO problems that these companies have been involved with and some of the applications that have been developed for these problems. Finally, a half-day was devoted to presentations on applications based on constraint programming for scheduling problems, including sports scheduling, crew scheduling for train crews, bus drivers, and nurses, satellite scheduling and scheduling of broadcast commercials.

- **DIMACS/IAS Workshop on Randomized and Derandomized Algorithms for Discrete Structures**
  - Dates: November 2-4, 1998
  - Location: Institute for Advanced Study
  - Attendance: 41
  - Organizers: Bruce Reed, University Pierre et Marie Curie; Noga Alon, Tel Aviv University and IAS; Alan Frieze, CMU

  The workshop was devoted to two techniques which have wide application in theoretical discrete optimization: the semi-random method and Szemeredi's Regularity Lemma. The semi-random method is a technique for generating combinatorial objects via repeated applications of the probabilistic method; a typical example of applying the method is the so-called Rődl Nibble. The Regularity Lemma states roughly that the vertices of any large enough graph can be split up into pieces so that the bipartite graph spanned by the edges between each pair of pieces behaves essentially like a random bipartite graph. Both of these techniques apply mainly to large structures and have numerous recent algorithmic applications.

  The workshop began with a historical survey of the uses of the Regularity Lemma by Miki Simonovits. Gabor Sarkozy showed how together with the "blow-up" lemma it was being used to solve several graph packing problems. Raphy Yuster discussed decomposing hypergraphs with Hypertree. Andrzej Rucinski discussed some packing problems and Yoshiharu Kohayakawa discussed some Ramsey problems in random graphs. Jeong Han Kim showed how the semi-random method could be used to find small complete arcs in projective planes and Vijay Vazirani talked about the approximation of #P-complete problems. There was a talk by Andrew Thomason on generalized graph colouring. Penny Haxell discussed the use of the algorithmic version of the Regularity Lemma in finding integer and fractional packings in dense graphs and Lubos Thoma discussed dependent arcs in acyclic orientations. Joel Spencer analyzed an on-line algorithm for packing random rectangles. Sarmad Abbasi discussed some questions related to spanning subgraphs of dense graphs. Ravi Kannan described randomized algorithms for finding low rank approximations to matrices and applications thereof. Van Vu showed how the semi-random method could be used to bound the list-chromatic number of locally sparse graphs. Michael Krivelevich described an efficient algorithm for testing bipartiteness in dense graphs. Also, Janos Komlos discussed tiling Turan theorems, Hanno Lefmann talked about algorithms for finding independent sets in graphs and hypergraphs and Andrzej Czygrinow discussed the algorithmic aspects of the hypergraph regularity lemma.

Workshops and Miniworkshops Planned for the Rest of 1998-99

- **DIMACS Workshop on Semidefinite Programming and Large Scale Discrete Optimization**
  - Dates: January 7-9, 1999
  - Location: Princeton University
  - Organizers: Robert Vanderbei, Princeton University and Yinyu Ye, University of Iowa
- **DIMACS Workshop on Large Scale Discrete Optimization in Logistics**
  - Dates: February 8-10, 1999
  - Location: Rutgers University

- Organizers: George Nemhauser and Martin Savelsbergh, Georgia Tech

- **DIMACS Workshop on Large Scale Discrete Optimization in Robotics and Vision**
  - Dates: March 21-23, 1999
  - Location: Rutgers University
  - Organizers: Jitendra Malik, University of California - Berkeley; Jean Ponce, University of Illinois - Urbana Champaign

- **DIMACS Workshop on Large Scale Discrete Optimization, the State of the Art**
  - Dates: Tentative July 1999
  - Location: Rutgers University (tentative)
  - Organizers: Peter Hammer, Rutgers University

- **DIMACS - IBM Conference in Honor of Ralph Gomory's 70th Birthday**
  - Dates: August 2-4, 1999
  - Location: IBM Watson Labs
  - Organizer: Bill Pulleyblank

- **Computational Challenge Preworkshop**
  - Dates: Tentative TBA
  - Location: TBA
  - Organizers: David Johnson, AT&T Labs - Research

Special Year Tutorial Series
Tutorials Feb. 1998 - Nov. 1998

Models and Algorithms for Railway Crew Management
Paolo Toth, DEIS, University of Bologna, Italy
October 30, 1998

Crew Management is concerned with building the work schedules (rosters) of crews needed to cover a given set of timetabled trips. Toth gave a general definition of the problem and showed possible operational constraints and objective functions arising in real-world applications. In practice, the overall problem is decomposed into two subproblems, called Crew Scheduling (CSP) and Crew Rostering (CRP). In CSP the short-term schedule of the crews is considered, and a convenient set of duties covering all the trips is constructed. In CRP the selected duties are sequenced to obtain the final rosters. Several mathematical programming and graph-theory formulations, and different exact and heuristic algorithms proposed in the literature for the solution of the two subproblems, were presented. Advantages and disadvantages of the proposed formulations and algorithms were discussed. Two main solution approaches were illustrated in more detail. CSP is solved by generating a very large number of potential duties, and by solving an associated Set Covering Problem to select a minimum-cost set of duties covering all the trips. CRP is solved by using a Lagrangian relaxation of a graph-theoretical formulation of the problem. Extensive computational results on real-world instances arising in the Italian railway company, Ferrovie dello Stato SpA, were presented.

Tutorials Planned for the Rest of 1998-99

Progress in Linear Programming Based Branch-and-Bound: Modeling and Methodology
George L. Nemhauser, Georgia Institute of Technology
December 9 and 10, 1998

In the last decade, the use of mixed-integer programming (MIP) models has increased dramatically. Fifteen years ago, main frame computers were required to solve problems with a hundred integer variables. Now it is possible to solve problems with thousands of integer variables on a pc and to obtain provably good approximate solutions to problems such as set partitioning with millions of binary variables. These advances have been made possible by developments in modeling, algorithms, software and hardware. In the first of these two expository talks, Professor Nemhauser will focus on applications, effective modeling and preprocessing. The second talk will deal with the methodologies of branch-and-cut and branch-and-price, which are the techniques that make it possible to treat problems with either a very large number of constraints or a very large number of variables.

Metaheuristics
Fred Glover, Jim Kelly, Manuel Laguna,
University of Colorado, Boulder
April 29 and 30, 1999

An Introduction to Metaheuristics will place a primary focus on Tabu Search, from the standpoint of adaptive memory methods, and on Scatter Search and Path Relinking, from the standpoint of evolutionary methods. Comparisons and contrasts with Genetic Algorithms and other approaches will be included. An Introduction to OptQuest will highlight demonstrations of real world applications, especially those applications that involve significant elements of uncertainty and nonlinearity, including discreteness.

Optimal Connected Partitions of Graphs
Bruno Simeone, La Sapienza University, Italy
Spring 1999

This will be a series of five lectures on the topics: Path equipartition: trimming the shortest path network; greedy algorithms, semimodular lattices, and supermodular functions; shifting algorithms for tree partitioning; how to make a polynomial movie out of pseudo-polynomially many frames: continuous shifting algorithms; and connected k-partitions in k-connected graphs: a topological approach.

Discrete Optimization in Telecommunications and VLSI Design
Michel Minoux, Universitè Paris 6, France
May 1999

Eigenvalue and Semidefinite Techniques in Discrete Optimization
Henry Wolkowicz, University of Waterloo, Canada
May 1999

Local Search Algorithms in Discrete Optimization
Toshihide Ibaraki, Kyoto University, Japan
May 1999

Branch and Bound Algorithms on Distributed Processors
Ewald Speckenmeyer, University of Köln, Germany
June 1999

Special Year Graduate Course

Professor Vijay Vazirani, visiting DIMACS from Georgia Tech, has been giving a special year graduate course on Approximation Algorithms. The course has been attended by about 50 people from Rutgers, Princeton, AT&T Labs, Bell Labs, Bellcore, and NEC Research. Besides presenting key results in the area of approximation algorithms, the course has been giving a feeling for the important open problems. Topics include combinatorial algorithms, primal-dual schema based algorithms, multicommodity flows, semidefinite programming, estimating network reliability, PTAS for scheduling and Euclidean TSP, FPAS for knapsack, and hardness of approximation. The course is based on a preliminary draft of a book on Approximation Algorithms written by the instructor.

Selected Research Results

Although the year has just begun, there are already a number of exciting research results.

A Self Organizing Bin Packing Heuristic Suppose there is a random process that generates items of various integer sizes which we must pack into fixed size bins, and we wish to minimize the total wasted space in the non-empty bins of the packing. If N is the number of items generated, it is known that the total waste in the optimal packing must grow either as $\Omega(N)$, $\Omega(\sqrt{N})$, or $O(1)$, depending on the probability distribution. Standard bin packing algorithms such as Best Fit do not typically perform optimally. This is true even for offline algorithms such as Best Fit Decreasing. Although in general it is NP-hard to determine the optimal growth rate given a distribution, Janos Csirik, David Johnson (PM, ATTL), Claire Kenyon (V), Peter Shor (PM, ATTL), and Richard Weber showed that this problem can be solved in pseudo-polynomial time (polynomial in the bin size) using linear programming. They provided a remarkably simple online algorithm that is conjectured to perform "optimally" for ALL distributions, based on extensive empirical tests, which demonstrates its apparently "self-organizing" method of operation.

A Data Structure for Maintaining a Partition in VLSI Design Gruia Calinescu (PD) worked on data structures for maintaining a partition. In VLSI design, it would be useful to solve efficiently the following problem: update a partition of the ground set $\{1, 2, ..., n\}$ after adding or deleting a bipartition (a bipartition is a partition of the ground set into two parts). Using the credit technique, Calinescu proved that a straightforward data structure obtains the best possible amortized time bound.

The Effect of Slack on Competitiveness for Admission Control Michael Goldwasser (PD) worked on the effect of slack on competitiveness for admission control. Imagine that a company rents use of a single resource to customers, charging a fixed amount per minute of use. Job requests arrive in an online manner, each request specifying the length of time for which the customer will use the resource as well as a deadline by which the job should be completed. The company is not required to complete each job. However, it will only be paid for jobs which are completed on or before their deadline and it may not interrupt a job once started. Therefore, a scheduling algorithm provides admission control as it decides which jobs to run, as well as when to run them, and the goal is to maximize the successful use of the resource and thus the profit for the company. Goldwasser analyzed the performance of an algorithm using competitive analysis, comparing the schedule produced by an algorithm to the optimal schedule produced by an algorithm with full knowledge of future jobs. This model arises in many applications of real-time systems, with the most typical motivations involving communication-based applications. In this setting, the resource represents a channel of communication and a job (or call) represents a request for delivery of text, sound, or video through the channel. The *slack* of a job is equal to the amount of time between the job's arrival and the last possible time at which it could be started while still meeting its deadline. Goldwasser focused on the effect of slack on the competitiveness of scheduling algorithms in this model. His main result was that if each job is willing to wait an amount of time at least proportional to the job's length, the competitiveness for online scheduling algorithms can be significantly improved. When arbitrary slacks are allowed, previous work has shown that no bounded competitiveness is possible if arbitrary job lengths are allowed. However, Goldwasser showed for this model of slack that a simple deterministic algorithm is $(2 + \frac{1}{\omega})$-competitive, thereby providing a constant competitive ratio which does not depend on the range of job lengths. He also provided a matching lower bound showing that this is the best possible deterministic result, even when all jobs have one of three distinct lengths. In the special case where all jobs are required to have the same length (e.g., packets in an ATM network), the natural greedy algorithm is $(1 + \frac{1}{\omega+1})$-competitive, and again that is the best possible deterministic result. This result generalizes a previous bound for the case of arbitrary slack. In the case where jobs have one of two distinct lengths, Goldwasser gave tight upper and lower bounds which improve slightly on the $(2+\frac{1}{\omega})$ bound for arbitrary lengths.

A Polyhedral Analysis of the p-cycle Polytope Ozgur Ozluk (GS, U. North Carolina) worked on a polyhedral analysis of the p–cycle polytope which is the convex hull of the incidence vectors of all the p–cycles (simple cycles consisting of p arcs) of the complete directed graph. His main motivation was to use the results of this study as a first step in a polyhedral approach to the Traveling Circus Problem (TCP) which combines the location aspects of the p–median problem with the routing aspects of the Traveling Salesman Problem (TSP). The TCP has many potential applications. One such application is the design of the distributed computer networks in which case cities on the tour would represent concentrators connected by a tour (a ring network). A very interesting feature of the TCP is that it can be used to solve many extensions and generalizations of the TSP such as the Time-Dependent TSP and Generalized TSP. The p–cycle polytope should not be expected to have a simple characterization, since it contains the Asymmetric TSP polytope on p nodes as a face. It is also clear that the problem of finding a minimum length p–cycle is NP-hard. Ozluk developed an integer programming formulation of the p–cycle polytope, and derived some useful properties of its set of equations. He then determined the dimension of the p–cycle polytope, and proved two lifting results.

3.2 Special Year on Networks, 1996-00

Overview:

The Special Year on Networks began in July 1996 and will continue through August 2000 with the help of a three year grant received from the National Security Agency. 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 - Research

- Mihalis Yannakakis, Bell Labs

Scientific Advisory Board: Baruch Awerbuch, Johns Hopkins; Badri 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 - Research; 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

Sponsors and Collaborating Institutions:
The National Science Foundation
The New Jersey Commission on Science and Technology
The National Security Agency
AT&T Labs - Research
Bell Labs
Bellcore
Institute for Mathematics and its Applications, University of Minnesota
Rutgers University
Princeton University
NEC Research Institute

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.

The special year's topics can be grouped into four overlapping areas:

- **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, Miniworkshops and Special Events Feb. 1998 - Nov. 1998:

- **Workshop on Ethical and Social Dimensions of Information Technology**
  - Dates: February 28-March 1, 1998

- Location: Princeton University
- Attendance: 51
- Organizers: Helen Nissenbaum, Princeton University; Bernard Chazelle, Princeton University; Joan Feigenbaum, AT&T Labs - Research

The Workshop was dedicated to a discussion of how computer and information systems are shaped by societal and ethical values, including broadly encompassing values such as fair distribution of goods and power, freedom, autonomy, sovereignty, and privacy as well as more specific human ends such as wealth, efficacy, and rights to free expression, association, privacy and property. Panel presenters, representing the fields of computer science, the social sciences, philosophy, and policy studies, discussed values embedded in specific systems, including but not limited to the net, encryption, security, autonomous agents, educational software, user-interfaces, and the structure of information systems. The panelists were guided by questions such as: How do values influence or determine the shape of computer and information systems? Whose ends, interests or values are best, and most frequently, represented in contemporary systems? By what means are values embedded in systems – public policy, markets, or the discretion of individual scientists and engineers? Are some of these sources more "legitimate" than others? What values ought to shape computer and information systems? Is there some shared sense of public, community or individual welfare that ought to drive the design of systems? Is it enough to "let the market decide"?

There were panels on philosophical perspectives, technical perspectives, social science perspectives, and policy perspectives. There were discussions on democratic implications of Internet privacy by L. Jean Camp; distance education (pedagogical versus economic values) by Marvin J. Croy; metadata platforms by Joan Feigenbaum; oppression, resistance and information technology by Lucas D. Introna; and anonymity, surveillance, and crypotography by Deborah G. Johnson. There were also presentations on how the marriage of management and computing intensifies the struggle for personal privacy by Rob Kling and Jonathan P. Allen; policy in public key infrastructures by B. A. LaMacchia; information technology and freedom by Abbe Mowshowitz; the digital dilemma by W. R. Neuman; and ethics and information technology from a social science perspective by Susan Leigh Star. Finally, Lodis Rhodes talked about choice and control as essential values in a community computer network; M.J. van den Hoven talked about moral responsibility, public office and information technology; and Lorrie Faith Cranor talked about bias and responsibility in neutral social protocols.

- **Workshop on Multichannel Optical Networks: Theory and Practice**
  - Dates: March 16-19, 1998
  - Location: Rutgers University
  - Attendance: 52
  - Organizers: Peng-Jun Wan, Ding-Zhu Du, Panos M. Pardalos

Recently, widespread interest in information infrastructures has heightened interest in the high performance achievable in multichannel optical networks. WDM networks offer potential advantages, including higher aggregate bandwidth per fiber, new flexibility for automated network management and control, noise immunity, transparency to different data formats and protocols, low bit-error rates, and better network configurability and survivability–all leading to more cost-effective networks. In the last several years, truly astounding progress has been made in optical device technology. The research challenges in multichannel optical networks involve novel approaches to build cost-effective networks from these available device components. The main focus of this workshop was to bring together researchers working on the theoretical and practical aspects of network design problems such as network architecture, interconnection design, medium access control, channel assignment, wavelength routing, scalability analysis, reconfigurability analysis, cost-performance analysis, and network management.

Tutorial survey talks were presented by John Strand on network architectures and transport network design and Biswanath Mukherjee on optical network building blocks, mainly for wavelength-division multiplexing (WDM) and on various architectures, protocols, and algorithms for two classes of optical WDM networks, broadcast-and-select networks for local applications and wavelength-routed networks for wide-area networks. Among the other speakers, G. Persiano and Gordon Wilfong discussed the benefit of wavelength converters and their impact on the wavelength routings in trees and rings; Peng-Jun Wan, Suresh Subramaniam and Thomas Erlebach gave algorithms for maximizing the number of connections with given set of communication requests and the number of available wavelengths; Ornan Gerstel raised the issue of minimizing the SONET ADMs in SONET/WDM ring networks; Rami Melhem, K. Sivalingam and Klaus Jansen discussed the dynamic wavelength/lightpath assignments; and Bruno Beauquier, Eric Harder, Jean-Claude Bermond and Chunling Zhou presented the routing and wavelength assignment for various multicasting/broadcasting communications patterns. Also, Biswanath Mukherjee and Olivier Crochat described how to support fault tolerance and protection; Jadranka Skorin-Kapov, Liwu Liu, Rocky KC Chang and Radim Bartos discussed various aspects of the virtual topology embedding into optical broadcast-and-select optical networks; Chunming Qiao presented the polymorphic control for cost-effective design of optical networks; and Cheng-Chi Yu described the support of real-time services.

- **Workshop on Robust Communication Networks: Interconnection and Survivability**
  - Dates: November 18-20, 1998
  - Location: Rutgers University
  - Attendance: 81
  - Organizers: Nathaniel Dean, Rice University; Frank Hsu, Fordham University; R, Ravi, Carnegie Mellon University

Architectural interconnection and survivability play important roles in the design, construction, operation and application of robust communications networks. Due to the advent of VLSI and fiber optics technologies, it has become possible and feasible to design and construct large scale high performance and high speed wireline and wireless communication networks that are also robust. This opens many challenging issues and problems for both the theory community and practitioners. Topics and problems discussed at the RCN:IS workshop included network structure and architecture (topology/architecture, reliability/latency, routing/switching, and bandwith); WAN backbone and access network design and construction (capacity efficiency, cost, restoration methods, congestion control, trade-offs between different architectures, and interconnection between LANs and the Internet); container construction and fault tolerant properties (multipath, all pair shortest paths, multicast, wide-diameter and fault-diameter); and primal dual approximation methods (survivable network design). Other topics included highly connected networks (including k-connected graphs, algorithms for special classes of graphs, etc.) and vulnerability (minimum degree problems, network upgrading, network attack and inhibition); survivability in optical networks (double coverings, SONET, T/W/CDM) and wireless networks (switch assignments, interconnections between base stations); and network design support systems (design heuristics, add/delete algorithms) and novel networks.

Events Planned for the Rest of 1998-99

- **Mobile Networks and Computing**
  - Dates: March 25-27, 1999
  - Location: Rutgers University
  - Organizers: Badri Badrinath, Rutgers University; Frank Hsu, Fordham University; Panos Pardalos, University of Florida; Sanguthevar Rajasekaran, University of Florida
- **Scaling Phenomena in Communication Networks**
  - Dates: October 22-24, 1999
  - Location: Institute for Mathematics and its Applications (IMA), University of Minnesota
  - Organizers: Ashok Erramilli, Netmetrix Inc.; Iraj Saniee, Bellcore Research; Walter Willinger, AT&T Labs - Research

Selected Research Results

Synchronization and Packet Routing in Networks Boaz Patt-Shamir (V) worked jointly with Rafi Ostrovsky (PM, BC) and solved an important open problem: Can optimal synchronization be attained efficiently when clocks are drifting (except for a "source" clock which provides standard time), if communication is constrained in some reasonable way? It was well known that if clocks are drift-free, then there is an optimal efficient solution; it is also known that if clocks can drift and communication is arbitrary, then no algorithm is both efficient and optimal. In this work Ostrovsky and Patt-Shamir gave an algorithm whose complexity is polynomial in the number of "pending messages" (messages the processor knows to be sent, but the processor does not know when received). This result has immediate applications, since virtually all synchronization algorithms in use today have bounded (typically linear in the number of processors) pending messages in any given time.

Efficient Scheduling Algorithms under the "Adversarial Queueing Theory" Model Bill Aiello (PM, BC, now ATTL), Rafi Ostrovsky (PM, BC) and Boaz Patt-Shamir (V) looked at the question of efficient scheduling algorithms under the "adversarial queuing theory" model. They partially settled the primary open problem in this field by proving that the "longest in system" policy builds up queues of only linear size when the paths used are shortest paths, and they gave an algorithm guaranteeing polynomial size queues for any paths.

Strictly Nonblocking Switching Networks Frank Hwang (PM) published a major new volume on switching networks for telephone traffic. Benes published *Mathematical Theory of Connecting Networks and Telephone Traffic* in 1965. Hwang's book is the first attempt to update the classical theory on nonblocking and rearrangeable networks with the thirty-some years of intensive research since 1965. But it also brings in new subject matter like multimedia and broadcast networks. In particular, the book includes a new sufficient condition on strictly nonblocking networks which was applied to the EGS photonic switching patented by AT&T. The paper describing this result was delivered in preliminary form by Richards and Hwang at the DIMACS workshop on Nonblocking Switching Networks.

Placement Algorithms for Hierarchical Cooperative Caching A basic problem in distributed memory environments is to provide efficient access to remote objects (e.g., words of memory, files). This problem arises in several applications including distributed file systems and the World Wide Web. For example, consider a hierarchical network of machines in which each machine periodically issues a request for an object drawn from a fixed set of unit-size objects. Assume that: (i) the frequency with which each machine accesses each object is known, (ii) each machine has a cache of known capacity, (iii) any cache can be accessed by any machine, and (iv) any request is satisfied by the closest machine with a copy of the desired object. In such an environment, it is desirable to fill the available cache space with copies of objects in such a way that the average access cost is minimized. Madhukar Korupolu, Greg Plaxton and Rajmohan Rajaraman (PD) obtained both exact and approximate polynomial-time algorithms for this hierarchical placement problem. Their exact algorithm is based on a reduction to min-cost flow, and does not appear to be practical for large problem sizes. Thus they are motivated to search for a faster approximation algorithm. Their main result was a simple constant factor approximation algorithm for the hierarchical placement problem that admits an efficient distributed implementation.

Dynamic Replication on the Internet Michael Rabinovich (ATTL) and Rajmohan Rajaraman (PD) analyzed a dynamic replication protocol designed by Michael and Irene Rabinovich. This protocol consists of an algorithm for deciding on the number and location of object replicas and an algorithm for distributing requests among currently available object replicas. Their analysis showed that if the frequency with which each machine accesses each object is fixed, then the system reaches a state in which the total number of replications, deletions, and migrations is bounded with high probability.

Contention Resolution in Multiple Access Networks Micah Adler (V), Phil Mackenzie, and Rajmohan Rajarman (PD) studied the problem of scheduling the transmission of packets from multiple senders over a number of Ethernet-like multiple-access networks. In particular, they were interested in exploiting the additional power allowed by the ability of a transmission to choose from multiple channels. They demonstrated a protocol that, when p senders transmit over $m \geq \log p$ channels, is stable and has an expected delay for any message of $O(1)$, if the arrival rate at each sender is at most $\lambda m/p$ for a constant λ sufficiently small. Their protocol has several advantages over existing protocols: it is considerably simpler and has much better constants than any previous $O(1)$ expected delay protocol. In addition, this is the first $O(1)$ expected delay protocol for asynchronous environments. All previous $O(1)$ delay protocols required a tightly synchronized clock and all previous protocols for asynchronous systems had an expected delay of $\Omega(\log p)$. This protocol works for pessimistic models of asynchrony. They also showed a lower bound on the delay of any protocol if the arrival rate is sufficiently high. In particular, if the arrival rate at each sender is at least $\lambda m/p$, where λ is a sufficiently high constant less than 1, then the average expected delay of any protocol is $\Omega(p)$. This bound is tight because a natural generalization of the round-robin protocol to multiple channels incurs $O(p)$ delay for any $\lambda < 1$.

Scheduling to Minimize Average Stretch Traditional metrics for analyzing on-line scheduling algorithms such as average completion time, average flow time, and maximum flow time do not take into account the differing job sizes in the objective function. That is, these metrics tend to treat the delay incurred by a job in a uniform manner regardless of the size of the job. Thus, the penalty for a large delay incurred by a small job is identical to the penalty for the same delay incurred by a large job. For several applications, including scheduling in web servers, this is undesirable. In order to address this problem, Bender, Chakrabarti, and Muthukrishnan introduced a new metric called stretch. The *stretch* of a job in a schedule is the ratio of its flow time (the time the job spends in the system) to its job size. Rajmohan Rajaraman (PD) and Anthony Shaheen (DIMACS REU student) studied the on-line complexity of minimizing average stretch. They showed that a well-known scheduling algorithm, the shortest remaining processing time algorithm (SRPT), is 4-competitive with respect to average stretch. S. Muthukrishnan (PM, ATTL) and Rajaraman are studying the complexity of average stretch in multi-processor and non-clairvoyant (where the size of a job is not known a priori) environments.

Allocation of Resources in a Multimedia Network Tina Heikkinen (GS, RU) worked on the modeling of a telecommunication network applying linear programming. Given a telecommunication system described in terms of a matrix G describing the link fading (gain) coefficients of the users of the network and a second matrix F describing the interference fading coefficients, the general problem is to define a resource allocation rule, i.e., an allocation of transmit powers $x = (x_1, .., x_n)$ of the n users, such that $(Gx)_i/(Fx)_i > a$, for all i. Heikkenen extended this framework to the case of multimedia. Two aspects of multimedia were modelled; first, by allowing for non-square matrices G and F each user was allowed to have access to a multiplicity of datatypes, modelled by varying the user-by-user data rates; second, temporal dynamics were introduced, to capture the dynamics of the resource allocation problem. The latter is quite new; the modelling of dynamics was done by defining a dynamic game against nature.

The Capacity of a Multimedia Network Continuing with the study of resource allocation in multimedia networks described in the previous example, Tina Heikkenen (GS, RU) took the novel approach of applying the theory of non-cooperative games to define a globally optimal capacity denoted by $a(*)$, i.e., formally, the von Neumann growth model was inverted to max a with respect to x, s.t. for all $i, (Gx)_i/(Fx)_i > a$. Various characterizations for $a(*)$ were obtained. In continuing work on this topic, allowing for incomplete knowledge about the coefficients of the matrices G and F, Heikkenen used game theory/linear programming, with the recognition that due to the interference matrix, the game is an externality game and the resource allocation problem is an allocation problem under externalities. Knowledge of economic resource allocation under externalites was applied. Two solutions for resource allocation under externalities were obtained. First, Heikkenen showed that there exist general classes of matrix games with the same integral optimal capacity. These games are called exchangeable. This result followed from the Leontief-property of the matrix G. Second, a Boolean formulation for the Leontief-von Neumann telecommunication system was obtained, allowing for nonexchangeable matrix games; this formulation established a connection between zero-sum noncooperative games and cooperative games described by Boolean (characteristic) functions.

Enforcing E-Commerce Policies Using Structured Messages Naftaly Minsky (RU) and Victoria Ungureanu (GS, RU) have been working on the formulation of policies for electronic commerce and their enforcement. They proposed a general mechanism that can be used to formulate and enforce a wide range of policies for electronic commerce in a unified manner. This mechanism makes a strict separation between the formal statement of a policy and its enforcement. The policy is stated explicitly and can be examined by anybody who wishes to engage in it, and it is enforced by a distributed set of trusted controllers that regulate the activities of each agent under this policy.

Fast, Fair and Frugal Bandwidth Allocation in ATM Networks In connection with ATM networks, convergence issues arise for various algorithms that have been proposed in the context of the rate allocation problem. In this problem the aim is to allocate the bandwidth requests of various sessions such that the allocations are fair and capacity constraints are met. ATM networks are used to carry a variety of types of traffic. For some types of traffic, in particular Available Bit

Rate (ABR) traffic, the bandwidth of a network is typically insufficient to satisfy the requests of all the sessions, and so some fair allocation scheme must be devised. The ATM Forum, the standards setting body for ATM networks, has specified that the fairness criterion for ABR traffic should be Max-Min Fairness, which intuitively means that raising the bandwidth of any session comes at the expense of some other session of no greater bandwidth. Protocols to allocate bandwidth to sessions in a max-min fair manner are an important part of a network design. For a protocol to be realistic, it must conform to the Resource Management (RM) Cell mechanism specified by the ATM Forum. Such RM Cells get sent as a constant fraction of all cells sent by the source. However, they have only a few fields. RM Cells are the only means of communication allowed between switches so any reasonable protocol is totally distributed and asynchronous, since the RM Cell mechanism does not easily lend itself to synchronization. Finally, RM Cells must be handled very quickly at the switches. Call a protocol frugal if the switch spends $O(1)$ computation on each RM Cell it receives, and it has $O(1)$ local space for each session through it. Recently, several frugal RM Cell protocols have been proposed for ABR traffic, but none have been shown to converge to the max-min fair state. Protocols which are known to converge in a linear number of maximum roundtrip times require RM Cell processing which is linear in the number of sessions at a switch. Y. Bartal (BL), Martin Farach (PM, RU), Shibu Yooseph (PD), and Lisa Zhang (PM, BL) gave a frugal RM Cell protocol for ABR that matches the convergence time of the fastest known non-frugal protocol. A second type of ABR traffic is the Minimum Cell Rate (MCR) type, where every session can specify a minimum amount of bandwidth. The max-min fair allocation should then respect these MCR requests. Bartal, et al. extended their results to give the first frugal RM Cell protocol for MCR and achieve a quadratic convergence rate.

Broadcast Disk Paging Broadcast disks are an emerging paradigm for massive information dissemination. In a broadcast disk, data is divided into n equal-sized pages and pages are broadcast in a round-robin manner by a server. Broadcast disks are effective because many clients can simultaneously retrieve any transmitted information. Paging is used by the clients to improve performance, much as in virtual memory systems. However, paging on broadcast disks differs from virtual memory paging in at least two fundamental aspects: A page fault in the broadcast disk model has a variable cost that depends on the requested page as well as the current state of the broadcast; prefetching is both natural and a provably essential mechanism for achieving significantly better competitive ratios in broadcast disk paging. Vincenzo Liberatore (GS, RU) designed a deterministic algorithm that uses prefetching to achieve an $O(n \log k)$ competitive ratio for the broadcast disk paging problem, where k denotes the size of the client's cache. He also showed a matching lower bound of $\Omega(n \log k)$ that applies even when the adversary is not allowed to use prefetching. In contrast, when prefetching is not allowed, no deterministic online algorithm can achieve a competitive ratio better than $\Omega(nk)$. Moreover, he obtained a lower bound of $\Omega(nlogk)$ on the competitive ratio achievable by any non-prefetching randomized algorithm against an oblivious adversary. These lower bounds are trivially matched from above by known results about deterministic and randomized marking algorithms for paging. An interpretation of these results is that in broadcast disk paging, prefetching is a perfect substitute for randomization.

3.3 Special Focus on Discrete Probability, 1996-99

Overview:

This program began during the 1996 - 1997 academic year and continued into 1997 - 1999 at a much lower level. It is jointly organized and sponsored with the Institute for Advanced Study, which has cohosted visitors and workshops with DIMACS.

Organizers:

- Enrico Bombieri, Institute for Advanced Study
- Jeff Kahn, Rutgers University
- Peter Winkler, Bell Labs

Sponsors:
The National Science Foundation
The National Security Agency
AT&T Labs - Research
Bell Labs
Bellcore
Georgia Tech
The Institute for Advanced Study
NEC Research Institute
National Institute for Standards and Technology (NIST)
New Jersey Commission on Science and Technology
Rutgers University
Princeton University

Theme of program:
Probabilistic methods and ideas have by now become central to large parts of discrete mathematics and theoretical computer science. 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 the dm/tcs community and the probability and statistics community. This was already happening to some extent prior to the special focus, 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, were still far more the exception than the rule when the program started. The program is intended to give a substantial push to a broader and more rapid development of these connections. The program is also reaching out from the mathematical sciences to the physical sciences, especially statistical mechanics.

Workshops Feb. 1998 - Nov. 1998:

- **Miniworkshop on Mathematical Aspects of Ordering and Phase Segregation Dynamics in Lattice and Continuum Systems**
  - Date: May 16, 1998
  - Location: Rutgers University
  - Attendance: 75
  - Organizers: Joel Leibowitz, Rutgers

  This miniworkshop took place just prior to the semi-annual Rutgers statistical mechanics conference which took place on May 17-18. It was devoted to mathematical aspects of phase segregation dynamics in both lattice and continuum systems, as well as to quasi crystals. These are both topics in which John Cahn, currently at NIST, has been a central figure. The statistical mechanics conference, as well as this meeting, honored his 70th birthday. This meeting focused more on the mathematical issues, while the statistical mechanics meeting focused more on the physical ones.

Workshops Planned for the Rest of 1998-99:

- **Workshop: Joint DIMACS-Georgia Tech Workshop on Combinatorics of Statistical Mechanics Models**
  - Dates: April 29-May 1, 1999
  - Location: Georgia Institute of Technology
  - Organizers: Dana Randall and Prasad Tetali, Georgia Tech

Selected Research Results

New Lower Bounds on the Decoding Error Probability Simon Litsyn (V) derived new upper bounds on the error exponents for maximum likelihood decoding in binary symmetric channels by using an analysis of possible distance distributions of codes. This is an improvement on the best earlier known bounds by Shannon-Gallager-Berlekamp (1967) and McEliece-Omura (1977) for the classical problem of information theory of estimating probabilities of decoding error when a block code is used for information transmission over a binary symmetric channel (BSC), in particular giving exponential bounds on the performance of the best codes. These bounds are of the form $e^{nE(R,p)+o(n)}$ where n is the length of the code, $o(n)/n$ tends to 0 when n grows and $E(R,p)$ is the error exponent depending only on the code rate R and the transition probability p. The function $E(R,p)$ has been thoroughly studied since 1948, when Shannon proved in his seminal paper that the error exponent for maximum likelihood decoding is positive in the interval $0 \leq R \leq C_p$ where C_p is the capacity of the channel. Thus for all rates less than C_p, the error probability can be made as small as one wishes by choosing an appropriate n. The problem of deriving lower and upper bounds on the error exponent has attracted a great deal of attention since then. Thus Litsyn's improvements are of great interest.

Distance Distribution of Codes The problem of estimating possible distributions and deriving upper and lower bounds on individual and integral components is of great importance in coding and information theory. Simon Litsyn (V) derived new lower bounds on interval distributions. The estimates guarantee existence of a distance component in a certain range having essentially (exponentially) large value. As a result of this, in joint work with A. Ashikhmin and A. Barg (PM, BL), he was able to attack two classical problems in coding theory: bounding generalized weights of binary linear codes and estimating the list radius of codes. In particular, a binary list of m radius r code is a subset of binary n-dimensional Hamming space with the property that every point of the space is covered at most m times by spheres of radius r centered at code vectors. If $m = 1$ the list code is just the conventional r-error correcting code. A new asymptotic upper bound on the size of the list codes with fixed m and r growing linearly in n was obtained.

An entropy approach to the hard-core model on bipartite graphs The *hard-core measure* with activity λ on a finite graph G is the measure $hc(\gamma)$ on the set of independent sets of G under which the probability of I is proportional to $\gamma^{|I|}$. One may take limits of such measures to obtain hard-core measures on infinite graphs, where uniqueness (for given γ) is a central issue. Jeff Kahn (PM, RU) introduced entropy as a tool for studying these measures on regular bipartite graphs of high degree (which includes the lattices $\mathbf{Z}^d$ for large d, the cases of greatest interest for the hard-core model). The entropy approach was used to establish, *inter alia*, asymptotically exact estimates for the expected size of an independent set chosen from $hc(\gamma)$ with $\gamma > 0$ fixed and G as above; for fixed $\gamma > 0$ and G the Hamming n-cube (n large), $hc(\lambda)$ exhibits a "phase transition," meaning that with probability close to 1, almost all vertices of the random independent set have the same parity. An analogue of the phase transition result for large, high-dimensional tori would prove a central conjecture in this area, that the *critical activity* for the hard-core model on $\mathbf{Z}^d$ (the value of γ below which one has a unique hard-core measure) tends to 0 as $d \to \infty$. Indeed, the phase transition result was suggested by J. van den Berg (V) as a test case for this conjecture.

A Fluctuation Theorem for Markov Processes Cohen and Gallavotti noticed that for time-reversible thermostatted dynamics the large deviations for the phase space contraction satisfy a certain symmetry. Their proof is based on the method of Markov partitions and thus assumes the thermostatted dynamics to be sufficiently chaotic. This necessary technical assumption can be proved only in a few cases, such as the case of Chernov, Eyink, Lebowitz, and Sinai. On the Markov partition level the dynamics are essentially described by a Markov chain. Thus it is natural to inquire whether and in what sense the Gallavotti-Cohen fluctuation theorem holds for stochastic Markov processes. Herbert Spohn (V) found that indeed the fluctuation theorem holds with great generality, provided a condition of local detailed balance is satisfied. He used techniques from the theory of large deviations and described systems in thermal equilibrium by reversible Markov processes. For them one has the Einstein relations and the Onsager symmetry. These notions found their natural extension to driven nonequilibrium systems.

A Central Limit Theorem for Convex Chains in the Square Points $P_1, \ldots, P_n$ in the unit square define a convex n-chain if they are below $y = x$ and, together with $P_0 = (0,0)$ and $P_{n+1} = (1,1)$, they are in convex position. Under uniform probability, I. Bárány (V), Gunter Rote, Bill Steiger (PM, RU), and Cunhui Zhang (RU) proved an almost sure limit theorem for these chains that used only probabilistic arguments and that strengthened similar limit shape statements established by other authors. An interesting feature is that the limit shape is a direct consequence of the method. The main result was an accompanying central limit theorem for these chains. A weak convergence result implied several other statements concerning the deviations between random convex chains and their limit. This was the first central limit theorem proved for any of the variations of limit shape that have been studied recently by Bárány, Sinai, and Vershik.

The Asymmetric Simple Exclusion Process Pablo A. Ferrari (V) worked on the asymmetric simple exclusion process in one dimension. It is known that for nearest neighbor jumps there are so-called blocking measures. These measures are products, and they approach densities one and zero as the space coordinate goes to plus or minus infinity, respectively. The proof of the existence of these measures uses the fact that they are reversible for the system. It is a striking open problem that the existence of these measures cannot be proven for general jumps. However, for very special values of the jump rates, it is possible to show that the above measures are still reversible for jumps larger than nearest neighbors. Ferrari developed a coupling between two exclusion processes with different rates. The method allowed him to extend the parameter region for which there are invariant blocking measures. The coupling showed that if certain inequalities hold for the jump rates (that essentially say that the first process has more drift than the second), then the existence of blocking measures for the first process implies the existence of an invariant measure for the second one.

DIMACS special foci cross-fertilize each other. The following is one example of a result that arose from the cross-fertilization of two such foci, the one on Discrete Probability (with its emphasis on statistical physics) and the one on Mathematical Support for Molecular Biology.

Statistical Properties of Contact Maps A contact map is a simple representation of the structure of proteins and other chain-like macro-molecules. This representation is quite amenable to numerical studies of folding. Michele Vendruscolo, Balakrishna Subramanian (GS, RU), Ido Kanter, Eytan

Domany and Joel Lebowitz (PM, RU) showed that the number of contact maps corresponding to the possible configurations of a polypeptide chain of N amino acids, represented by $(N-1)$ step half avoiding walks on a lattice, grows exponentially with N for all dimensions $D > 1$. They carried out exact enumerations in $D = 2$ on the square and triangular lattices for walks of up to 20 steps and investigated various statistical properties of contact maps corresponding to such walks.

3.4 Special Year on Mathematical Support for Molecular Biology, 1994-99

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

The program is now extending into a fifth year of activity. In the past year, we were aided in part by an additional grant from the National Center for Human Genome Research. We have had several postdoctoral fellows in a joint postdoc program with the University of Pennsylvania, several workshops, and an active seminar series. The first four years built up several collaborative relationships with other centers and departments in the biological sciences, and this year is building on those relationships.

DIMACS commissioned a report on the impact of our Special Year on Mathematical Support for Molecular Biology and the report, by Daniel Gusfield, Sampath Kannan, and Pavel Pevzner, was received in June. It says that "The special year has been enormously successful in meeting its goals." It emphasizes the role of DIMACS in building bridges between disciplines: "The special year helped launch a number of cross-disciplinary collaborations. Prior to the special year research in computational biology had been rather compartmentalized. The work of computer scientists did not usually find its way into the biologist's toolkit. In addition, computer scientists were working on a small set of clearly formulated problems without a full knowledge of their biological relevance. It was therefore important to greatly encourage collaborative research." The report also talks about the importance of the postdoctoral training program, calling this "the foremost accomplishment" of the special year and saying that "perhaps the most important function of DIMACS has been in training young computer scientists in computational biology so that they go on to establish collaborations later in their career." Discussing the many research results emanating from the special year, the report says that "the research program was an unqualified success." It concludes by saying that "DIMACS has played an important role in exciting interest in" computational biology and bioinformatics. The full report may be found at the following URL: http://dimacs.rutgers.edu/About/Reports/SYMB.html

Organizers:

- Fred Roberts (Chair), Rutgers
- Joachim Messing (Chair), Waksman Institute - Rutgers
- Lawrence Shepp (Co-Chair), Rutgers
- Michael Waterman (Co-Chair), University of Southern California
- Martin Farach-Colton (Assistant Chair), Rutgers

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, Bell Labs; and Andrew Yao, Princeton.

Collaborating Centers and Institutes:

- Ecology and Evolutionary Biology Department, Princeton University
- Computer Science Department and 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
- Green Plant Phylogeny Coordination Group
- Mt. Sinai School of Medicine
- Waksman Institute for Molecular Genetics, Rutgers University

Sponsors:

AT&T Labs - Research
Bell Labs
Bellcore
BIOSYM Tech
Centers for Disease Control
Ciba-Geigy Corp.
Department of Energy
Green Plant Phylogeny Coordination Group
International Union of Crystallography
Merck Research
National Center for Human Genome Research
National Institute of Allergy and Infectious Diseases
National Science Foundation
National Security Agency
NEC Research Institute
New Jersey Commission on Science and Technology
University of Pennsylvania
Princeton University
Program in Mathematics and Molecular Biology, UC-Berkeley
Promega Corporation
Roche Molecular Systems
Rutgers University
Sandia National Laboratories

The Sloan Foundation
Smith-Kline Beecham

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/tcs 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, Miniworkshops and Special Events 1998-1999:

- **Minisymposium on Advances in Phylogenetics Inference Methods**
  - Dates: March 20 1998
  - Location: Rutgers University
  - Attendance: 25
  - Organizers: Daniel Huson, Princeton University; Shibu Yooseph, Rutgers University

  Phylogeny inference plays an important role in several scientific disciplines such as evolutionary biology, molecular systematics, and historical linguistics. This minisymposium covered recent advances in phylogeny inference methods and focused on various issues including appropriate models for bio-molecular sequence data, sequence alignment techniques, and new techniques for handling highly divergent data sets. This minisymposium brought together a number of researchers who are studying advanced results and methods for phylogenetic analysis. The first topic was a new model of sequence evolution under which some of the sites are invariable in some parts of the tree and variable in others. A new divide-and-conquer approach to multiple sequence alignment was presented that heuristically divides the given sequences into shorter and shorter regions until the regions become short enough for the application of an exact algorithm to become feasible. Further, a heuristic for resolving phylogenetic trees was presented. Another talk presented a recently proven conjecture stating necessary and sufficient conditions for a set of quartet topologies to uniquely define a phylogentic tree. One presentation addressed the question of how to calculate the maximum likelihood time of divergence under a reversible continuous-time Markov chain. Another talk was on the application of trees as a datastructure for modelling the genetics of cancer tumor progression. Finally, a new divide-and-conquer method for tree-constructing trees on large datasets was presented.

- **DIMACS Workshop on Combinatorial Clustering and Multi-Domain Protein Structure Analysis**
  - Dates: June 26-27 1998
  - Location: Rutgers University
  - Attendance: 52
  - Organizers: Sylvia Spengler, Manfred Zorn, Inna Dubchak, Fred Roberts, Ilya Muchnik, Casimir Kulikowski

  This workshop was joint with the Special Year on Massive Data Sets and is described under that Special Year.

- **DIMACS Workshop on Estimating Large Scale Phylogenies: Biological, Statistical, and Computational Problems**
  - Dates: June 26-28 1998
  - Location: Princeton University
  - Attendance: 63
  - Organizers: Tandy Warnow, Junhyong Kim, Ken Rice

  Biological organization is fundamentally based on an evolutionary history of bifurcating descent-with-modification. Phylogenetic estimation is the inference of this genealogical history from present day data. Phylogenetic trees, the graph representation of the genealogical history, play a central role in evolutionary biology, and phylogenetic estimation techniques are being applied to a wide variety of computational biology problems. Accurate inference of phylogenetic trees poses surprisingly hard computational problems (not only NP-hard, but hard in practice, so that years of computation are used in the search for the best trees), as well as statistical and biological challenges. Progress towards accurate inference thus requires significant interactions among biologists, statisticians, mathematicians, and computer scientists. This workshop brought together an interdisciplinary mixture of researchers, to discuss phylogenetic tree estimation, but with special attention to the problem of recovering "large" phylogenetic trees. Over half of the participants were from the biological sciences.

  The size of a phylogenetic estimation problem is measured by the number of taxa and the number of characters. Until recently, computational and data limitations kept most phylogenetic estimation problems to small numbers of taxa. Now, however, the availability of computational resources and the influx of large molecular data sets are enabling researchers to tackle increasingly larger problems, and the analysis of large-scale data sets is rapidly becoming a central problem in phylogenetic biology.

  Recent experimental evidence has established the existence of large trees that can be estimated accurately as well as those that are difficult to accurately estimate with reasonable numbers of characters. Some of these examples have suggested that taxon sampling (increasing the size of the estimation problem through the addition of taxa rather than characters) might lead to more easily estimated trees. Conversely, it has been argued that big trees are hard to infer for a variety of reasons: NP-hardness of the optimization problems, properties of the search space, inadequacy of the heuristics, and even possible inadequacy of the optimization criteria. Unfortunately, very little actual evidence is available to support any conjectures about how the performance of estimators scale with respect to the size of the phylogenetic problem. In addition, the question of scaling is itself confused by poorly delineated notions. For example, the size of the tree also involves the maximum amount of divergence (not only the number of taxa and characters) and standards for measures of estimator performance have also not been agreed upon. The goal of this symposium was to precisely identify the key problems with respect to how the performance of phylogenetic estimators scale as with the size of the problem, and gather experimental and theoretical results addressing this problem.

Sessions at the workshop were devoted to biological problems, empirical results, algorithmic issues, and statistical problems.

Some of the highlight talks were on identification of new sources of biological data from protein structural modeling by M. Gerstein and G. Naylor; availability of a new molecular data set for the evolutionary tree of all green plants and the current difficulties stemming from this large data set, by P. Soltis, D. Soltis, and M. Chase and by B. Mishler; new algorithmic approaches to difficult optimizations including Monte-Carlo Markov chain, genetic algorithms, divide-and-conquer, and branch-and-bound, by L. Salter and D. Pearl, by P. Lewis, by T. Warnow and by P. Purdom; a new class of model family for the stochastic cladogenesis process by D. Aldous; empirical results on taxon sampling by J. Huelsenbeck and by K. Smith and T. Warnow; and new analysis of probabilistic foundations of evolutionary tree estimation by M. Steel and by J. Chang.

The meeting was especially enhanced by the presence of the Green Plant Phylogeny Coordination Group - an NSF funded organization for the coordination of research activities concerning the diversity of plants. One common agreement from the symposium was the recognition that both theoretical and empirical studies are hampered by the lack of a common framework for evaluating performance. There was a call for a standardization of the manner in which we can evaluate the performance of various methods and strategies for analyzing large-scale phylogenies. This call was followed up with a decision to create a newsgroup for discussion of such methodological concerns, and to establish a suite of model trees which could be used by this group for experimental purposes. Overall, the consensus of the attendees was that this symposium was extremely timely and unique in its synergistic inclusion of members of the biological, computational, and statistical communities.

- **Ninth Annual Conference on Combinatorial Pattern Matching**
  - Dates: July 20-22, 1998
  - Location: Rutgers University
  - Attendance: 60
  - Organizers: Martin Farach-Colton, Rutgers University; Udi Manber, University of Arizona

This symposium was joint with the Special Year on Massive Data Sets and is described under that Special Year.

Workshops Planned for the Rest of 1998-99

- **Workshop: Advances and Opportunities at the Biology/Math/Computation/Physical Sciences Interface**
  - Dates: March 6-7, 1999
  - Location: Rutgers University
  - Organizers: Wilma Olson, Fred Roberts, Rutgers University

Events Under Development

- **Workshop on Learning Stochastic Processes**
  - Location: Rutgers University
  - Organizers: Martin Farach-Colton, Rutgers University; Sampath Kannan, University of Pennsylvania

Selected Research Results

Better Methods for Solving Parsimony and Compatibility Evolutionary tree reconstruction is a challenging problem with important applications in biology and linguistics. In biology, one of the most promising approaches to tree reconstruction is to find the "maximum parsimony" tree, while in linguistics the use of the "maximum compatibility" method has been very useful. However, these problems are NP-hard and current approaches to solving them amount to heuristic searches through the space of possible tree topologies (a search which can, on large trees, take months to complete). Maria Bonet (former PD), Mike Steel (V), Tandy Warnow (V), and Shibu Yooseph (PD) developed a new technique, called optimal tree refinement, for reconstructing very large trees. Their technique was motivated by recent experimental studies which have shown that certain polynomial time methods reliably return contractions of the true tree. They studied the use of this technique in solving parsimony and compatibility, and obtained both hardness results and polynomial time algorithms.

Hybrid Tree Reconstruction A major computational problem in biology is the reconstruction of evolutionary trees for species sets, and accuracy is measured by comparing the topologies of the reconstructed tree and the model tree. One of the major debates in the field is whether large evolutionary trees can be even approximately reconstructed from biomolecular sequences of realistically bounded lengths (up to about 2000 nucleotides) using standard techniques (polynomial time distance methods, and heuristics for NP-hard optimization problems). Using both analytical and experimental techniques, Daniel Huson (PD), S. Nettles, K. Rice, Tandy Warnow (V), and Shibu Yooseph (PD) showed that on large trees, the two most popular methods in systematic biology, neighbor-joining and maximum parsimony heuristics, as well as two promising methods introduced by theoretical computer scientists, are all likely to have significant errors in the topology reconstruction of the model tree. They also found a new general technique for combining outputs of different methods (thus producing hybrid methods), and showed experimentally how one such hybrid method has better performance than its constituent parts.

The Disk-Covering Method for Tree Reconstruction The reconstruction of evolutionary trees containing significant amounts of divergence is especially challenging. Daniel Huson (PD), S. Nettles, L. Parida, Tandy Warnow (V), and Shibu Yooseph (PD) develped a new tree reconstruction method, the Disk-Covering Method, which can be used to recover accurate estimations of the evolutionary tree for otherwise intractable datasets. DCM obtains a decomposition of the input dataset into small overlapping sets of closely related taxa, reconstructs trees on these subsets (using a "base" phylogenetic method of choice), and then combines the subtrees into one tree on the entire set of taxa. Because the subproblems analyzed by DCM are smaller, computationally expensive methods such as maximum likelihood estimation can be used without incurring too much cost. At the same time, because the taxa within each subset are closely related, even very simple methods (such as neighbor-joining) are much more likely to be highly accurate. The result is that DCM-boosted methods are typically faster and more accurate as compared to "naive" use of the same method.

Cavender-Farris Trees Evolution is often modeled as a stochastic process which modifies DNA. One of the most popular and successful such models is the Cavender-Farris (CF) tree. The Phylogeny Reconstruction Problem can be described as follows, given k samples drawn from a CF tree, output a CF tree close to the original, by some metric on the space of probability distributions. Rick Desper (GS, RU) showed that a very simple, polynomial time algorithm converges to the true tree at a rate which differs from the optimum by a constant. He did this by analyzing upper and lower bounds for the convergence rate of learning very simple CF trees, and then showed that the learnability of each CF tree is sandwiched between two such simpler trees.

Models of the Progression of Cancerous Tumors Cancer progression is marked by large-scale over-duplication or disappearance of certain genes, a phenomenon known as a copy number aberration (CNA). Using hybridization methods, laboratory researchers have been able to detect and isolate CNAs, which has led to the discovery of cancer-related genes. It is a long-established fact that CNAs do not appear randomly. The Vogelstein model explained how colorectal cancer is marked by the appearance of a sequence of events appearing in a typical order. Rick Desper (GS, RU) developed a new model for the appearance of CNAs, based upon classical percolation models. Desper showed how a tree-based percolation model, the signal propagation tree (SP tree), generates a probability distribution similar to the Cavender Farris (CF) tree. He modified the CF-reconstruction algorithms to fit SP trees, and showed that the new algorithms are consistent, as well as calculating the rate of convergence to the true tree.

DIMACS special foci cross-fertilize each other. The following is one example of a result that arose from the cross-fertilization of two such foci, the one on Discrete Probability (with its emphasis on statistical physics) and the one on Mathematical Support for Molecular Biology.

Statistical Properties of Contact Maps A contact map is a simple representation of the structure of proteins and other chain-like macro-molecules. This representation is quite amenable to numerical studies of folding. Michele Vendruscolo, Balakrishna Subramanian (GS, RU), Ido Kanter, Eytan Domany and Joel Lebowitz (PM, RU) showed that the number of contact maps corresponding to the possible configurations of a polypeptide chain of N amino acids, represented by $(N-1)$ step half avoiding walks on a lattice, grows exponentially with N for all dimensions $D > 1$. They carried out exact enumerations in $D = 2$ on the square and triangular lattices for walks of up to 20 steps and investigated various statistical properties of contact maps corresponding to such walks.

3.5 Special Focus on DNA Computing, 1997-99

Overview:

The Special Focus on DNA Computing began in February 1997 and is continuing through 1997-99. It builds on two earlier DIMACS workshops in the area and on collaborations between computer scientists and biologists stimulated by the Special Year on Mathematical Support for

Molecular Biology.

Organizers:

- Laura Landweber, Princeton University
- Richard Lipton, Princeton University

Sponsors:
The National Science Foundation
The New Jersey Commission on Science and Technology
AT&T Labs - Research
Bell Labs
Bellcore
NEC Research Institute
Rutgers University
Princeton University
The Sloan Foundation

Theme of Program:
DNA computing presents great challenges both to computer sciences and molecular biology in such areas as using parallelism, designing new algorithms, and dealing with errors in molecular processes. The Focus on DNA Computing attempts to introduce researchers in discrete mathematics and theoretical computer science to parts of molecular biology relevant to computation. Molecular Biology is an area that is increasingly dependent on dm and tcs competencies and a field that has elements of being fundamentally an "information science," and so it is natural to explore connections between it and computing. DNA Computing has attracted interest from researchers from areas such as molecular biology, genetics, supercomputing, and parallel computing, to collaborate with scientists working in DIMACS core areas, on new models of computation.

Workshops, Miniworkshops and Special Events, Feb. 1998 - Nov. 1998

- **Workshop on Nucleic Acid Selection and Computing**
  - Dates: March 15-17, 1998
  - Location: Princeton University
  - Attendance: 118
  - Organizers: Laura Landweber and Richard Lipton, Princeton University; Robert Dorit, Yale University; Andrew Ellington, Indiana University

  Nucleic acid selection, or directed molecular evolution, allows the identification of molecules with unique functional properties from large pools of random sequences. This technology has pushed molecular biology towards DNA computing, which also uses combinatorial searches to find molecules that satisfy a unique set of filters. The goal of this workshop was to explore recent advances in test-tube evolution experiments that can be transported to DNA computing. This workshop assembled over 125 researchers from biology and computer science as well as the biotech industry to forge the link between directed evolution and DNA computing. Both experiments search through large combinatorial pools to discover molecules that satisfy a unique set of filters. The workshop successfully explored recent advances in test-tube evolution experiments that can be exported as components of both software and hardware for DNA computing, including molecular ribozyme switches, reversible DNA "chips", and continuous evolution.

Among the highlight talks were those by Andrew Ellington on natural and artificial RNA evolution; Robert Dorit on in vitro evolution of complexity; Michael Yarus on the hypothetical RNA world; and Pim Stemmer on DNA and protein shuffling. Also highlighted were talks by Barry Polisky on SELEX; Donald Burke on chimeric SELEX; Dinshaw Patel on comparative aptamer complexes; Niles Lehman on continuous ribozyme evolution; Dipankar Sen on DNA enzymes; and Peter Unrau on nucleotide forming ribozymes. And talks by Michael Hecht on protein evolution and design; Steven Benner on evolution of natural biopolymers; and Michael Heller on hybridization to electronic arrays. Speakers represented institutions such as Maxygen, Inc., NeXstar Pharmaceuticals, Memorial Sloan-Kettering, Nanogen, Inc., as well as a variety of academic institutions.

- **Fourth Annual DIMACS Meeting for DNA Computers**
  - Dates: June 16-19, 1998
  - Location: University of Pennsylvania, Philadelphia, PA
  - Attendance: 83
  - Organizers: Harvey Rubin (co-chair), University of Pennsylvania; David Wood (co-chair), University of Delaware; Anne Condon, University of Wisconsin; David Gifford, MIT; Lila Kari, University of Western Ontario; John Reif, Duke University; Grzegorz Rozenberg, Leiden University; Nadrian Seeman, New York University

The three previous meetings exclusively devoted to DNA based computers reflect the dynamic growth of the field as laboratory results have reinforced theoretical designs.

For this 4th international meeting on computing with DNA, topics included (but were not limited to): algorithms, applications, laboratory techniques, architectures, practical difficulties with DNA based computers, proposed solutions to such difficulties, computational processes in vivo, and relevant ideas regarding biological evolution.

Events Planned for the Rest of 1998-99

- **DIMACS Workshop on Evolution and Computation**
  - Dates: January 11-12, 1999
  - Location: Princeton University
  - Organizers: Laura Landweber, Princeton University; Erik Winfree, Princeton; Richard Lipton, Princeton University; Stephen Freeland, Cambridge University

Selected Research Results

Cellular Computables How do cells and nature "compute"? They read and "rewrite" DNA all the time, by processes that modify sequences at the DNA or RNA level. In 1994, Leonard Adleman reported an elegant solution to a seven-city Directed Hamiltonian Path problem using DNA. This launched the new field of DNA computing, which in four years has grown to international focus. However, unknown to Adleman and this field, ciliated protozoans of genus Oxytricha and Stylonychia had solved a potentially harder problem using DNA several million years earlier. The solution to this problem, which occurs during the process of gene unscrambling, represents one of nature's ingenious solutions to the problem of the creation of genes. Laura Landweber (PU) has studied the mechanism and evolution of these complex rearrangements in the DNA polymerase alpha gene, which she discovered is scrambled in 52 pieces dispersed over the germline genome of Stylonychia lemnae. In a formal model she also proved that this process has the computational power of a Turing machine.

DNA Computing and Chess Laura Landweber (PU) has expanded the field of "DNA Computers" to RNA to tackle a class of mathematical problems in chess. The versatility of RNA makes it even more suitable than DNA for molecular computation. For example, 'bit'-operations can be performed on RNA using specific ribonuclease digestions. Here, for example, is an algorithm to solve the knight problem. Simply put, this problem asks how many knights and what configuration can one stably place on an $N \times N$ chess board such that no knight is attacking any other knight on the board. This problem can in principle be solved using only biological molecules and enzymes as tools. For a simple case like a 3×3 chess board, the 9 spaces on the chess board correspond to 9 'bits' or place-holders in an RNA sequence. A bit flipped to 'on' or 1 represents a knight at that position, and 'off' or 0 means that position is empty. Each round of selection, performing single bit-operations, destroys those RNA molecules that do not satisfy a single criterion for this problem. The set of solutions can be read by PCR screening of colonies on a plate.

3.6 Special Year on Massive Data Sets, 1997-99

Overview:

The DIMACS special year on massive data sets began with prespecial year activities in 1996-97 and is continuing through 1999.

Organizers:

- Robert Sedgewick, Princeton University (chair)
- Jon Bentley, Bell Labs
- Martin Farach-Colton, Rutgers University
- Joan Feigenbaum, AT&T Labs - Research
- Helene Kulsrud, IDA-CCR, Princeton

Sponsors and Collaborating Institutions:

The National Science Foundation
The New Jersey Commission on Science and Technology
AT&T Labs - Research
Bell Labs
Bellcore
Institute for Defense Analyses/Center for Communications Research (IDA/CCR)
NEC Research Institute
Rutgers University
Princeton University

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 proliferation of massive amounts of data: Data from astrophysical models, genetic sequencing, geographical 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 is seeking to stimulate new research to address these problems. We are seeking to learn common properties among diverse applications, identify fundamental computational problems that transcend particular applications, develop and analyze new practical algorithms, and establish partnerships with researchers in other areas, with the goal of developing solutions for specific scientific and commercial problems.

Workshops, Miniworkshops and Special Events, Feb. 1998 - Nov. 1998

- **MiniWorkshop on Mathematical Methods for High Performance Data Mining**
  - Dates: April 27-28, 1998
  - Location: Princeton University
  - Attendance: 52
  - Organizers: Helene Kulsrud, Institute for Defense Analyses;
    Robert Grossman, University of Illinois

  The goal of the workshop was to survey some of the successful methods for Data Mining in Massive Data Sets. Unlike most of the other workshops in this series, applications were not restricted to a single field. Instead, in depth descriptions of five methods and the application areas to which they have been successfully applied were presented. This seemed particularly useful at this time when many of the well known methods are being extended to apply to applications different from those for which they were originally developed. There were several talks on Latent Semantic Indexing using Singular Value Decomposition (SVD), a method originally developed for problems in document retrieval. Michael Littman gave a general introduction to the method and described some exciting new applications such as scoring standard essay tests and modeling human lexical acquisition. Randall Dahlberg presented a new application to cross-language identification and emphasized the need for clustering techniques to identify the results. A significant difficulty in the use of SVD is the determination of the exact number of singular values to use in a given problem. Theoretical research on this problem is being carried out by Santosh Vempala and his colleagues and was presented by Vempala. As a result of Vempala's presentation, Littman is attempting to include these new methods into his suite of SVD programs.

  Robert Grossman led off the talks on the use of tree based classifiers. Due to the sensitivity of these methods to the exact training data, Grossman emphasized the new ideas for combining regression trees - bagging, boosting and attribute averaging. He touched on the problems of distributed data and the use of parallelism. Work on the application of regression trees to problems in the medical diagnosis field were cited as good examples of the need for these new methods. Shannon gave a more complete description of his experiences in using these combining methods to improve the applicability of regression trees. Volinsky showed the use of Bayesian Classifiers to solve similar problems. On the theory side, Jorma Rissanen presented some new ideas for the design of classification trees using minimum description length and discussed some interesting problems when the minimum code length is one.

  The status of the use of Association rules was covered by two papers directed at the use of computer parallelism. These studies are extremely significant since we will need all the parallelism we can muster to apply these methods to the petabyte data sets. Mohammed Zaki described an expansion of the "A Priori" method to work in a parallel computer. Beginning with "A Priori" assumptions, Vipin Kumar and his group have developed a hypergraph method which creates clusters of objects and allows for parallel computation, and produces better results than some standard clustering techniques such as Auto-class or K-means.

The current status of the statistical method called Acquaintance was presented by Claudia Pearce.

The talks of both Kumar and Volinsky indicated another area of progress. These presentations illustrated the use of combined methods and may indicate a trend to deal with the difficult aspects of individual methods. Volinsky begins with trees generated with uncertainty and applies Bayesian model averaging to produce better probabilities in the models.

The conference concluded with a startling and important paper presented by Michael Jordan of MIT. This work attempts to unify the entire field of graphical methods and set it on firm and consistent mathematical foundation. The graphical model formalism is an elegant combination of graph and probability theory that allows realistic and flexible models to be formed. In particular the problem of dense graphs, usually impractical to solve, is given a new framework. This paper encouraged the listeners to think about new methods.

- **Astrophysics and Algorithms: A DIMACS Workshop on Massive Astronomical Data Sets**
  - Dates: May 6-8, 1998
  - Location: Princeton University
  - Attendance: 72
  - Organizers: Michael Vogeley, Princeton University; David Spergel, Princeton University; Don Sun, Bell Labs; Tony Tyson, Bell Labs

This workshop was motivated by the "data avalanche" that astronomers now face, as several observational projects have produced or promise to yield Terabyte-class data sets. Examples include sky surveys in optical and infrared wavelengths, microwave background anisotropy satellite experiments, helioseismology data, gravitational radiation detection experiments, and results from N-body/hydrodynamical simulations. Processing of these data and subsequent statistical analyses clearly require advanced computational methods. Issues of scalability in database design and analysis algorithms that are well-known in, e.g., the telecommunications industry and defense establishment are now being grappled with by astronomers. A principal goal of this workshop was to foster interaction between the Astronomy community and Discrete Mathematics and Theoretical Computer Science communities. The participants were an unusual mix of mathematicians, astronomers, and computer scientists. The majority of the speakers were astronomers who are involved in collecting, archiving, and analyzing large astronomical data sets along with notable contributions from researchers in discrete mathematics and computer science.

The workshop began with keynote addresses by leaders in the respective fields of Astrophysics and Algorithms. Princeton University provost and astrophysicist Jeremiah Ostriker traced the evolution of computing in astrophysics from his own experiences as a graduate student student at Yerkes Observatory to his state-of-the-art N-body/hydrodynamical simulations on a 64-processor SGI Origin 2000. Robert Sedgewick noted the differing nature of computing in Astronomy and Theoretical Computer Science.

One session addressed analysis of large optical surveys. David Bennett described analysis of an imaging survey (500GByte) to look for evidence of MACHOS (Massive Compact Halo

Objects) that might comprise some of the so-called "dark matter" in the universe. David Monet detailed the effort to catalog objects detected on photographic plates (object detections now number in excess of 6 billion, stored on 581 CD-ROMs) and the challenge of archiving and analyzing data products that range from "G-rated" (object catalogs) to "X-rated" (the raw pixel database). Barry Lasker discussed all-sky image and catalog databases that have also been constructed from photographic plates, which are now on-line for use by the astronomical community. Lasker's talk highlighted database issues such as cross-matching, classification, and sky partitioning.

A second session moved to astronomical surveys at longer wavelength. Carol Lonsdale described the data processing and archiving of results from the 2 Micron All Sky Survey (2MASS) project which collects 19 Gbyte of data per night. Rick White summarized status, results, and analysis algorithms used for the FIRST radio survey, the results of which are also available to the astronomical community from a WWW page. The NASA/IPAC Extragalactic Database, which archives all published extragalactic object detections and serves this catalog over the Internet, was reviewed by Barry Madore. Returning to analysis of optical data, Steve Odewahn discussed his use of neural network methods for galaxy classification.

A third session focused on database methods and computing technology. Alex Szalay presented an overview of the Sloan Digital Sky Survey (SDSS) that is now beginning to image most of the Northern sky in five optical bandpasses and will also obtain spectroscopy of one million galaxies. The raw data for this sky survey will comprise more than 40TByte, reduced to 1TByte of processed data. Laney Kulsrud reviewed mathematical methods for data mining in massive data sets, particularly algorithms that scale to Petabyte-class data sets. Tom Prince proclaimed that "10 Terabytes is not a large data set" and charted the growth of computing technology, both in the past and predicted for the near future. Much discussion ensued about the need for people (and/or compilers) who know how to take advantage of the benefits of parallel computing.

A fourth session was devoted to issues of data analysis. Sarbani Basu spoke about methods for analyzing helioseismology data, in particular mathematical methods applied to invert the spectrum of modes in the Sun. William Press's talk "How to Connect the Dots, FAST, in a Noisy World" focused on the problem of curve fitting with noisy, sparse, irregularly sampled data with inhomogeneous errors, and promised code to do so from his WWW site. George Djorgovski reviewed scientific results from analysis of catalogs of objects found from digitization of optical plates from the second Palomar Sky Survey, emphasizing AI methods to classify the objects. Andreas Brieden discussed the mathematical problem of deriving approximate relations for properties of convex hulls in n-dimensional spaces. Sergei Shandarin presented a set of statistics to describe the shape of compact surfaces that are useful for description of the large-scale distribution of matter in the universe.

A fifth session was devoted to problems of analysis of the anisotropy of the Cosmic Microwave Background radiation (CMB). Francois Bouchet reviewed plans and status for future satellite missions and our current knowledge of anisotropy of the CMB, and described the computational challenge of reducing the one hundred billion data points to ten or so fundamental physical quantities. Regarding the latter, Bouchet pointed out the need for matrix inversion methods that are significantly faster than order n-cubed. Max Tegmark elaborated on mathematical methods for making maps of the CMB sky from such satellite data and described lossless methods for power spectrum estimation from these data. Dan Rockmore detailed methods for fast spherical transforms that compute in order n-$\log(n)$ time. His experience with different compilers/platforms led to discussion of analysis of algorithms viz. the dependence of performance on memory access and other factors. Julian Borrill discussed fast methods to bound the likelihood function of cosmological models that predict the observed power spectrum of the CMB anisotropy.

A final session dealt with classification problems and multiscale image analysis. Jon Hakkila described attempts to classify gamma-ray bursts, the fundamental properties of which are poorly understood, using data mining techniques and articial intelligence classification methods. Fionn Murtagh and Albert Bijaoui both spoke about multiresolution methods for image analysis in Astronomy. Renyue Cen discussed the computational challenges of computing and analyzing cosmological simulations that yield a Petabyte of stored data. Cen also described a novel method for out-of-core Fast Fourier Transforms.

Astronomers at this workshop were pleased to find that, when compared with some other fields, Astronomy does not yet have a massive data set problem and that many of the challenges faced by astronomers have been addressed in other contexts. The key to successful astronomical data analysis is clearly to collaborate with researchers in computer science and in fields where scalable algorithms have been developed. Although the computer scientists faced the daunting task of decoding much astronomical jargon, once this hurdle was mounted many productive discussions followed regarding how existing technology could be applied to Astronomy problems. Several speakers from the Astronomy side suggested problems that might be of interest for, e.g., a CS graduate student.

A common theme of this workshop was that Astronomy faces a paradigm shift, following which many researchers will spend their time data mining a “Digital Sky” compiled from the vast array of multi-wavelength sky surveys now underway, in contrast to individual efforts at a single telescope (though certainly that mode will continue to provide results at the cutting edge of research). Clearly, collaboration with the DIMACS community will be of increasing importance.

- **Workshop on External Memory Algorithms and Visualization**
  - Dates: May 20-22, 1998
  - Location: Rutgers University
  - Attendance: 54
  - Organizers: James Abello, AT&T Labs; A.Birilis, AT&T Labs; Steve Mahaney, NSF; Yossi Matias, Bell Labs; T.Munzner, Stanford; W. Sweldens, Bell Labs; Jeff Vitter, Duke University; Kiem-Phong Vo, AT&T Labs.

We are witnessing a proliferation of applications that need to process large amounts of information. I/O between main and external memory is becoming an increasingly significant bottleneck. How, where and in what circumstances to involve the user are central questions and visualization may provide some of the answers. This workshop aimed at fostering the cooperation among theoreticians and practitioners in the fields of External Memory Algorithms and/or Visualization. Papers were presented on the following topics as they related to the workshop theme: Computational Models and Complexity; Memory Allocation and System Integration; Indexing Methods, Data Compression and Succinct Data Structures; External Memory Algorithms in Graph Theory, Computational Geometry, Combinatorics, Data Bases, Linear Algebra and Statistics; Visualization Metaphors and Multiresolution Frameworks; User Interfaces for External Memory Algorithms; Integrated Systems: Examples and Applications.

- **DIMACS Workshop on Combinatorial Clustering and Multi-Domain Protein Structure Analysis**
  - Dates: June 26-27, 1998

- Location: Rutgers University
- Attendance: 52
- Organizers: Sylvia Spengler, Manfred Zorn, Inna Dubchak, Fred Roberts, Ilya Muchnik, Casimir Kulikowski

The combinatorics, graph theory and algorithms branches of discrete mathematics are powerful techniques for use in analysis of diverse sets of data. Combinatorics introduces clustering models, forms a basis for the interpretation of results, and brings new ideas to clustering methods. The area of combinatorial clustering provides new opportunities for systematic (routine) and comprehensive study of very large databases containing highly complex non-regular elements. However, researchers developing methods for combinatorial clustering often do not collaborate with specialists in the creation, management and analysis of large, complex databases. The goal of this workshop was to foster and facilitate such a collaboration by focusing on a particular application. The specific problem addressed was in the area of protein structure and analysis of multi-domain proteins (as an example for other type of interactions among large bio-molecules). By hosting this workshop, DIMACS' aim was not only to stimulate research at the interface between these disciplines, but also to stimulate development of ideas and methods in each field.

Papers presented at the workshop can be loosely grouped into three areas: academic research in clustering methods (combining discrete math, computer science, operations research, and statistics); academic research in applications (protein multidomain dissection and function identification, secondary structure prediction, optimal threading, etc.); and research related with real application problems (targeted expression of gene families, receptor-ligand interactions, biochemical maps of protein surfaces, etc.). The papers on methods covered the latest dm/tcs approaches to clustering (visual clustering, clustering with model-based neural networks, quadratic Euclidean clustering) and the applied papers related to challenging issues in bioinformatics (threading by combinatorial optimization methods, protein databases through hidden Markov models, distant evolutionary relationships from protein sequence and structure).

The general perspective of the field became more clear as a result of the workshop. It appears that there is not much experience in systematic C2P-supported handling of massive data sets in bioinformatics. The large data sets currently maintained are maintained primarily through clustering tools developed 20 to 30 years ago. It is necessary to bring the power of modern combinatorial clustering to all classification, search, and database design problems in bioinformatics. Bioinformatics presents a variety of challenges to computer science, in particular on geometrical classification of 3D-protein structures, pre/post-processing stages of data sets, and analyses such as recognizing which segment of an amino acid sequence is a separate domain.

- **9th Annual Conference on Combinatorial Pattern Matching**
  - Dates: July 20-22, 1998
  - Location: Rutgers University
  - Attendance: 60
  - Organizers: Martin Farach-Colton, Rutgers University; Udi Manber, University of Arizona

This conference emphasized combinatorial pattern matching and its applications. Specific topics covered included string algorithms, pattern recognition, applications in molecular biology, text searching, information retrieval, symbolic computing, natural language processing and data compression.

Proceedings were published in the Springer-Verlag series Lecture Notes in Computer Science and distributed at the meeting.

CPM is a small, specialty conference which alternates between Western Europe and North America. The North American conference had had declining attendance over the last few years, and this year reversed the trend. It is reasonable to believe that this was because of the strong synergy between the topic of CPM and various Massive Data and Computational Biology activities at DIMACS.

- **The Fourth International Seminar on Average-Case Analysis of Algorithms**
  - Dates: July 20-24, 1998
  - Location: Princeton University
  - Attendance: 47
  - Organizers: Philippe Flajolet, INRIA; Rainer Kemp, Johann Wolfgang, Goethe-Universitat; Hosam Mahmoud, George Washington University; Helmut Prodinger, Vienna University of Technology; Robert Sedgewick, Princeton University; Wojciech Szpankowski, Purdue University

The International Seminar on Average-Case Analysis of Algorithms is a biannual event. In 1998 DIMACS hosted the fourth of these seminars. The theme of the workshop was average-case analysis of algorithms, whereby one assesses the quality of an algorithm according to its underlying stochastic behavior. It is sometimes rather difficult to characterize the behavior of an algorithm under the various instances of its input. In such cases, experience shows that while the deterministic task may be formidable, by assuming a reasonable probability measure on the input, one can say something meaningful about the average behavior of the algorithm. More recently, research has gone into the more ambitious avenue of finding the complete probability distribution of characteristics of algorithms. Two major schools of thought approach the subject from two different angles. Given a problem, the "analytic school" sets out to represent the problem in terms of generating functions obtained from the structure of the problem. This approach usually leads to non-standard functional equations and recent developments indicate that a miscellany of integral transforms are best suited for the task. The "probabilistic school" seeks formulations in terms of sums of stochastic components to which laws of large numbers may be sought. This approach is sometimes simpler, but often gives less information than the analytic approach.

The algorithmic problems considered at the workshop can be broadly categorized as follows: Basic search and data structures (hashing tables, search trees, neighbor search, balanced trees); sorting and complex search problems (partial match queries, adaptive search); strings and textual data (exact and approximate string matching algorithms); graph algorithms and combinatorial optimization (spanning trees, graph partitioning, packing); systems algorithms (caching) and models of computation (circuits); seminumerical algorithms (greatest common divisors, polynomial factorization); information theory (entropy, faulty components detection, games).

The highlights of the meeting can be summarized in the following points. Efficient algorithms to cope with massive data sets should provide pragmatic guarantees of fast running time in practice. Traditional worst-case analysis should be a tool in every computer scientist's and practitioner's repertoire, but our area is providing and developing the most useful tools. It is better to have a practical algorithm with good expected performance and "low" variance than an impractical algorithm with worst-case guarantees. Constant factors and lower order terms *do* matter. A precise mathematical analysis may prove essential to improve an algorithm. The analysis points out the weak points of an algorithm and suggests ways to improve them, because the analysis requires deep understanding of the internal mechanics of the algorithm. The analysis can also predict what is to be quantitatively expected from the potential improvements. The community is developing more refined and realistic random data models. This is a key step towards the design of efficient algorithms (in pragmatic terms): one designs an algorithm to work well on typical instances; one does not spend one's efforts to cope with unlikely ill-behaved instances.

- **DIMACS Workshop on Codes and Trees: Algorithmic and Information Theoretic Approaches**
  - Dates: October 5-7, 1998
  - Location: Rutgers University
  - Attendance: 75
  - Organizers: Mordecai Golin, Hong Kong University of Science and Technology; Julia Abrahams, DIMACS, Rutgers University

This workshop had the aim of making connections between distinct research communities interested in code tree problems of which Huffman coding is the prototype. We were successful in bringing together groups representing the information theory, computer science algorithms, and formal languages communities, thereby promoting research linkages among people and perspectives which are typically disjoint.

Although Huffman coding is a very classical problem, interest in the area continues, prompted by its central role in data compression. In recent years, as the need to store and retrieve massive amounts of data increases, compression has become a key technology. In addition, many other problems in diagnostic testing, database search, and circuit design are isomorphic to the Huffman optimization problem and its variants. Furthermore, new developments in optimization theory provide techniques for the resolution of long-standing open questions in the design of optimal coding algorithms.

Three tutorials presented at the workshop were designed to acquaint representatives of the various distinct communities with issues and approaches they may not have been familiar with, from the discussion of Huffman problem variants typical of the information theoretic community as discussed by Julia Abrahams, to the discrete optimization techniques centered around Monge properties typical of contemporary approaches to optimization algorithms among the computer science community by Wolf Bein, to a discussion of automata theoretic problems and solutions by Veronique Bruyere.

Among the other speakers, Wojtek Szpankowski discussed the asymptotics of a tree structured approach to the problem of electing a leader. Kingo Kobayashi presented a method for representing the set of all k-ary trees through a hyper-tree encoding technique. Rudolf Fleischer discussed self organizing data structures, an area of on line algorithms, and its application in data compression. Mike Fredman described a problem in multiple parallel processors searching through a tree together with a reduction to a Huffman-like problem from which he was able to obtain a good but not optimal solution. Mordecai Golin presented a computationally efficient algorithm for the optimal solution of a constrained coding problem of significance in synchronization. Bob Sedgewick discussed quicksort, a different type of algorithmic problem. The program continued with Ruy Milidiu on maximum length constraints, Hiro Morita on using code trees for random number generation, Ian Munro and David Benoit on the succinct representation of trees, Karthik Visweswariah on parse trees, Erik Demaine on fast intersection for web search engine queries, Roberto De Prisco on the connection between redundancy and data expansion for in place Huffman encoding, Akiko Kato on infinite source alphabets, and Ion Mandoiu on the problem of optimally encoding source symbols in the presence of some prespecified codewords. Additional speakers included Andrej Brodnik on fast decoding techniques, Teresa Przytycka on heuristics for good trees within a constrained tree structure, Eduardo Laber on dynamic Huffman coding, Paul Tucker on applications in searching, Mihai Sipitca on applications in image compression, Amiya Bhattacharya on applications in wireless network location management, Helmut Jurgensen on language theoretic approaches to error correction, Yuriy Reznik on random search trees, Mamoru Hoshi on homophonic coding, Gil Shamir on context tree methods, and Artur Pessoa on efficient encoding techniques.

Events Planned for the Rest of 1998-99

- **Workshop on Algorithm Engineering and Experimentation (ALENEX99)**
  - Dates: January 15-16, 1999
  - Location: Omni Hotel, Baltimore, MD
  - Organizers: Catherine C. McGeoch, Amherst College; Mike Goodrich, Johns Hopkins University
- **Sixth Implementation Challenge Workshop on Near Neighbor Searches (a special session of ALENEX99)**
  - Dates: January 15-16, 1999
  - Location: Omni Hotel, Baltimore, MD
  - Organizers: Michael Goldwasser, Princeton University

Selected Research Results

Logical Analysis of Data and Clustering Peter Hammer (PM,RU) and Christian Lenart (V) investigated the relationship between logical analysis of data (LAD) and clustering. The objective of LAD is to produce a list of logical patterns (based on a given training set) which provide structural information about datasets; in particular, patterns are used to distinguish between so-called positive and negative observations (e.g., presence or absence of a certain disease). Based on experiments with various sets of data (taken from the Repository of Machine Learning Databases and Domain Theories maintained by the University of California at Irvine), they discovered that this relationship has two aspects. First of all, clustering can improve the performance of the LAD procedures developed by Hammer and his collaborators. The main idea is to apply a clustering procedure to the raw data and then to use the centers of the clusters to generate the patterns as logical descriptors of the clusters; the second step is based on a Boolean optimization procedure. Hammer and Lenart observed that this two-step procedure improves the results of the best algorithms for the diabetes dataset which appeared to be the most difficult for LAD. On the other hand, by a slight modification of the LAD procedures, one can obtain patterns corresponding to clusters whose overlap does not exceed a certain prescribed limit; hence LAD can be used for clustering purposes.

Pattern Generation is Hard Lijie Shi (GS,RU) examined the problem of creating a good pattern for a given set of data, where a pattern is defined as a collection of inputs that makes a given Boolean function take the value 1. This problem was proved to be NP-hard by Boros, Ibaraki, and Makino. Shi proved that this problem does not have any polynomial approximation algorithm with ratio $N^{1-\epsilon}$ unless P=NP, where N = the number of positive data points, and further proved that if the pattern must satisfy one of the positive data points, the maximum pattern is still hard to approximate unless P=NP. This result answered an open problem of Peter Hammer. From the approximation viewpoint, the greedy algorithm for this problem is optimal.

Finding Optimal Boolean Classifiers John Franco (V) worked on the problem of finding optimal Boolean Classifiers. From a given data set consisting of a number of 0-1 vectors, each with a value of 0 or 1, one can develop a Boolean expression F which evaluates to true iff an assignment of values to variables of F matches a vector of value 1. Well studied procedures for creating such Boolean classifiers are identified with a research area known as the logical analysis of data (LAD) – see first example of this section. These procedures can output many classifiers for a given data set. The question is, which is the best classifier to use from the perspective of trying to minimize the error in classification among future data points? Franco developed a statistical approach to solving this problem based on cross moments and covariances of existing data. The idea was to develop a reasonable parametrized model of data generation, use the data to set the parameters, then use one of a number of methods to generate an estimate, at each point, which minimizes the probability of error in some sense at that point. The need for a model seemed to be critical since it is easily possible to get poor results by estimating cross moments directly from the data if the underlying distribution of data points is not stationary. Hence he experimented with data sets in a quest for reasonable models.

Optimization Algorithms for Separable Functions with Tree-Like Conjugacy of Variables and their Application to the Analysis of Massive Sets of Seismic Data In work of Vadim Mottl (V) and Ilya Muchnik (V), a massive data set, in particular, that of seismic data, was considered as a set of experimentally acquired values of a number of variables each of which is associated with the respective node of an undirected conjugacy graph that presets the fixed structure of the data set. The class of data analysis problems under consideration can be described by the assumption that the ultimate aim of processing can be represented as a transformation of the original data array into a secondary array of the same structure but with node variables of, generally speaking, different nature, i.e. different ranges. Such a generalized problem was interpreted as the formal problem of optimization (minimization or maximization) of a real-valued objective function of all the node variables. The objective function wass assumed to consist of additive constituents of one or two arguments, respectively, node and edge functions. The former carry the data-dependent information on the sought-for values of the secondary variables, whereas the latter are meant to express the a priori model constraints. For the case when the graph of the pairwise conjugacy of the data set elements has the form of a tree, an effective global optimization procedure was proposed which is based on a recurrent decomposition of the initial optimization problem over all the node variables into a succession of partial problems each of which consists of the optimization of a function of only one variable like the Bellman function of classical dynamic programming. Two kinds of numerical realization of the basic optimization procedure were considered on the basis of parametric representations of Bellman functions, respectively, for discretely defined and quadratic node and edge functions. The proposed technique of optimization of separable functions supported by trees underlies the algorithmic solution of a group of seismic data analysis problems which includes the problems of segmentation of explorative seismic sections through layered underground rock mass, texture analysis of plane and three-dimensional explorative seismic data arrays, and detection of seismic signals.

A Functional Approach to External Graph Algorithms James Abello (PM, ATTL), Adam Buchsbaum (PM, ATTL) and Jeffrey R. Westbrook (PM, ATTL) developed a divide-and-conquer approach for designing external graph algorithms, i.e., algorithms on graphs that are too large to fit in main memory. Their approach builds a succession of graph transformations that reduce to sorting, selection and a recursive bucketing technique. They applied their techniques to devise external algorithms for computing connected components, minimum spanning trees (BMSTs), and maximal matchings in undirected graphs and multi-graphs. These algorithms are purely functional. That is, each algorithm is specified as a sequence of functions applied to input data and producing output data, with the property that information, once written, remains unchanged. The function is then said to have no "side effects." A functional approach has several benefits. External memory algorithms may run for hours or days in practice. The lack of side effects on the external data allows standard checkpointing techniques to be applied, increasing the reliability of any real application. A functional approach is also amenable to general purpose programming language tranformations that can reduce running time. Abello, Buchsbaum, and Westbrook also considered a semi-external model for graph problems, in which the vertices but not the edges fit in memory. This is not uncommon in practice, and when vertices can be kept in memory, significantly more efficient algorithms are possible. They designed new algorithms for external grouping and sorting with duplicates and applied them to produce better I/O bounds for the semi-external case of connected components.

External Memory Algorithms on Massive Data Sets Data sets in large applications are often too massive to fit completely inside the computer's internal memory. The resulting input/output communication (or I/O) between fast internal memory and slower external memory (such as disks) can be a major performance bottleneck and raises fascinating questions about the efficiency of algorithms on massive data sets. Jeffrey S. Vitter (V) conducted research on ways to solve massive data problems efficiently in external memory. He worked in particular on ways to do fast range searches (given a query rectangle, report all the points it contains) and on ways to do approximate queries in On-Line Analytic Processing (OLAP) in databases. He developed tree-based search methods that are efficient for various special cases of range queries, such as when two or three sides of the rectangle are specified and the other sides extend to infinity. The OLAP work deals with a new approach based on wavelet approximations of the data cube. The issue in practical use is to deal with the massive size of the data cube. Programming tools and environments are available for simplifying the programming task. Vitter did experiments on some newly developed algorithms for spatial databases incorporating these paradigms, implemented using TPIE (Transparent Parallel I/O programming Environment), and the results were very promising.

Amortized Analysis for Pairing Heaps The pairing heap was introduced in the mid 1980's as a self-adjusting and computationally more efficient alternative to Fibonacci heaps. Highly efficient in practice and well regarded in this respect, it is included in the GNU C++ library. Strikingly simple in design, the pairing heap data structure nonetheless seems difficult to analyze. It has been previously conjectured that the pairing heap provides an implementation of decrease key operations with constant amortized running time, the raison d'etre of Fibonacci heaps, and several published experimental studies have supported this conjecture. Efficient implementation of decrease key operations plays an important role in the design of many highly efficient network optimization algorithms, such as shortest path and minimum spanning tree algorithms. Michael Fredman (PM, RU) obtained a significant new result which demonstrates that the pairing heap requires more than constant amortized time to perform decrease key operations, contrary to conjecture. Nevertheless, although Fibonacci heaps are asymptotically superior, pairing heaps continue to be of great practical importance. Aside from the practical usefulness of the pairing heap, the question of its efficiency has theoretical interest deriving from the connection between pairing heaps and splaying. The mechanism underlying the pairing heap can be viewed as an application of the self-adjusting splay heuristic of Sleator and Tarjan. The splay heuristic has been demonstrated to have surprising power in the context of search trees. A recent and notable example concerns the work of Cole on the dynamic finger conjecture for splay trees. As an issue, the competitiveness of the pairing heap thus constitutes an interesting and significant challenge for splaying, testing the limits of this remarkable and powerful heuristic. More precisely, splaying is a general purpose technique having the surprising feature that in all of the domains in which it can be appropriately applied, it seems to be competitive with the alternative specialized techniques previously developed for that particular domain. For the domain concerning priority queues which efficiently implement decrease key operations (that for which Fibonacci heaps were developed) Fredman's new result shows that splaying is asymptotically dominated by the specialized method, and this is the first time a result of this kind has been established. The techniques developed to analyze pairing heaps may find application in the analysis of splay trees and other problems involving amortized analysis of algorithms. Fredman's approach was based on the following concept. An attractive aspect of splaying is that it selects the operands for the atomic operations that get performed (in this instance the operations are comparisons between quantities) in a highly uniform manner. The intuition behind the new result is to consider the information gain that these comparisons can potentially generate and then show that as a consequence of the uniformity intrinsic to the splaying technique, the resulting loss of flexibility, in choosing which comparisons to perform as the computation proceeds, inevitably creates inefficiencies, at least within an appropriately constructed probabilistic setting. However, in order to make this framework tractible, it was necessary to devise a combinatorial parody of efficiency, which highly distorts the usual notion, but nonetheless suffices for this purpose.

Compressing and Indexing Massive Text Archives Yossi Matias (PM, BL), S. Muthukrishnan (PM, BL), Suleyman Cenk Sahinalp (V), and J.Ziv (V, BL) worked on compressing and indexing massive text archives. This research resulted in a linear time algorithm for the HYZ compression method - which is in fact the first linear time algorithm for any context based compresion method. The algorithm was proven to outperform conventional compression schemes when the memory in use is limited, both in theory and practice. This work also provided insight into solving the document listing problem: given a set of documents to be preprocessed, index them to respond to list queries efficiently. The list query inquires the list of documents in which a given pattern occurs. The method is especially interesting for information retrieval applications in which many documents have to be preprocessed.

Bucket Sorting In the general area of distributions of sorting algorithms, one looks at the exact or limiting probability distributions of cost parameters of sorting algorithms (algorithms to order data either completely or partially in the case we are interested in finding order statistics). Exact distributions are sometimes too complex to deal with, in which case one seeks approximations by simpler functions that may be relevant only for very massive data sets. Knowing the distribution puts in good perspective the average-case analysis and tells us how well concentrated the algorithm's behavior is or how likely it is to deviate substantially from its expectation. Bucket sorting algorithms are sorting algorithms based on data classification according to "throwing data in buckets." Philippe Flajolet (V), Philippe Jacquet, Hosam Mahmoud (V), and Mireille Regnier used analytic probability as a general approach to bucket sorting. Functional equations for the cost parameters were set up and asymptotically solved by analytic techniques such as Poissonization and de-Poissonization and saddle point methods. This work was definitive because it dealt with all the possible flavors of bucket sorting as well as their adaptations for finding order statistics. For each case, Flajolet, et al. were successful in finding limit distributions of cost.

Stochastic Analysis of Shell Sort Hosam Mahmoud (V) and Robert Smythe conducted a stochastic analysis of shell sort. Shell sort is an algorithm that generalizes the method of sorting by insertion. It is essentially several stages of insertion sorting. The probabilistic analysis of shell sort stood as a formidable, tantalizing challenge. Most research on shell sort had previously been in the direction of making good choices for the parameters of the stages to obtain good worst-case behavior. Mahmoud and Smythe took the research in a different direction. They analyzed the stochastic structure of the algorithm and identified connections of the algorithm to order statistics and ultimately to standard stochastic processes such as the Brownian bridge process.

Visualization of Large Spatial Data Sets Octavian Procopiuc (V) worked on visualization of large spatial data sets with telephone data in mind. Telephone data can be thought of as a directed graph, with telephone numbers being vertices and calls being edges. A method of visualizing this graph is by constructing a height field based on the weights of the edges, as follows: the vertices are arranged in a certain order on the x-axis and on the y-axis of the 3-dimensional Euclidean plane, and each edge from a vertex on the x-axis to a vertex on the y-axis is represented by a spike of height given by the weight of the edge. This height field is obviously very large and sparse. The goal of the research was to design and implement an I/O efficient hierarchical data structure for such multidimensional data, which would increase the performance of a visualization program. The main features that can benefit immensely from such an indexing data structure are zooming in and out and panning. The main issues that Procopiuc had to deal with in designing such a data structure were: preserving the inherent hierarchy existent in the telephone numbers, preserving spatial locality, building the structure incrementally from the initial passes over the data in order to give the user immediate feedback. The data structure that resulted from extensive analysis of the problem and its peculiarities is based on the cross tree of Italiano and Grossi, modified to handle the hierarchy in vertices (telephone numbers) and simplified by observing that deletes don't need to be handled.

The Suffix Tree of a String The suffix tree of a string is a fundamental data structure for string processing. Recent focus on massive data sets has sparked interest in overcoming the memory bottleneck of known algorithms for building suffix trees. Paolo Ferragina (V), Martin Farach-Colton (PM, RU) and S. Muthukrishnan (PM, BL) showed how to choreograph almost all disk accesses to be via the sort-scan paradigm, thus achieving the first I/O-optimal algorithm for building a suffix tree on disk devices.

3-D Object Recognition One approach to 3-D object recognition is to represent an object by a set of 2-D views. To minimize the number of views required to encode an object, one seeks a 2-D representation that is highly viewpoint invariant. Sven Dickinson (PM, RU) recently proposed two new view-based object recognition approaches which offer high viewpoint invariance. The first approach was based on a hierarchical wavelet decomposition of an image into salient regions, giving rise to a directed acyclic graph, called the *saliency map graph.* Object recognition was formulated as matching an image graph to a database of model object graphs. He developed two coarse-to-fine matching algorithms based on a fast approximation to finding the maximum cardinality, minimum weight matching in a bipartite graph spanning an image graph and a model graph. One algorithm compares only the topological structure in a pair of graphs, while the other algorithm includes both topological and geometric structure. Initial results yielded a recognition accuracy of better than 95%. Dickinson's second approach to view-based object modeling was based on a skeleton-like decomposition of an object's silhouette according to a taxonomy of qualitative part classes. This description can also be represented as a graph which can, in turn, be mapped in linear time to a unique rooted tree. Object recognition, in this case, was formulated as matching an image shock tree to a database of model shock trees. In this work, he introduced a novel compact encoding of a tree's topology based on an eigenvalue characterization of the tree's structure. This topological signature not only provided a fast index into a database of model shock trees, but formed the basis of an efficient algorithm which computes, in the presence of occlusion and noise, the largest isomorphic subtree of two trees. Initial results showed that the matching and indexing algorithms support both generic recognition (focusing primarily on tree topology) and exemplar recognition (focusing on both topology as well as node geometry).

The List Update Problem Bela Csaba (GS, RU) and Sachin Lodha (GS, RU) investigated the list update problem. The problem can be briefly stated as follows: There are n different items in a doubly linked list. Every item has its own unique key. Access requests are coming for these items, one at a time, in the form of their key values. The request is satisfied by finding an item of the corresponding key value in the doubly linked list. If the position of the item requested is i in the list, then the cost of this operation is i units. While satisfying such a request, the requested item can be moved to any of the first i positions at no cost. The goal is to satisfy all requests at minimum possible cost. In the off line setting, the request sentence is known in advance. On the other hand, in the on line setting, a request must be satisfied without knowledge of any future request. Let ρ denote a finite length request coming from $\{1, 2, ..., n\}$. Let $C_{OPT}(\rho)$ denote the cost of the optimal satisfaction of ρ. Let A be a deterministic on line algorithm and let $C_A(\rho)$ denote the cost of A while satisfying ρ. If there exists a constant K such that $\forall \rho \;\; C_A(\rho) \leq cC_{OPT}(\rho) + K$ then A is said to be $c-$competitive. It is natural to ask whether knowing the structure of the request sequence could be of any help or not, i.e., beating the barrier of 2 on the competitive ratio of on line algorithms. This assumption may not be impractical, since in practice it may be reasonable to expect some kind of pattern in the incoming requests. Csaba and Lodha pursued this approach and considered two different kinds of restrictions on the requests. For these they gave two different optimal on line strategies.

3.7 Special Year on Computational Intractability, 1999-00

Overview:

We are currently planning the upcoming special year on computational intractability. The first half of this special year will be run under the auspices of our science and technology center grant from NSF. We have pledges of support for some of the second half activities from the DIMACS institutions and have applied for grants for additional support.

Organizers:

- Michael Saks (Chair), Rutgers University
- David Johnson, AT&T Labs - Research
- Christos Papadimitriou, University of California - Berkeley
- Eva Tardos, Cornell University
- Avi Wigderson, Hebrew University and Institute for Advanced Study
- Andrew Yao, Princeton University

Sponsors:

The National Science Foundation
The New Jersey Commission on Science and Technology
The National Security Agency
AT&T Labs - Research
Bell Labs
Bellcore
Rutgers University
Princeton University
NEC Research Institute

Theme of program:

The notion of computational intractability lies at the heart of computer science and is central to the interface between discrete mathematics and theoretical computer science. Research on intractability has two long range goals:

1. To identify which computational problems are tractable and which are intractable,
2. To develop algorithmic strategies for dealing with problems which are intractable.

These two aims have been the subject of intense research efforts over the past twenty five years, and these efforts have been extremely fruitful with a wealth of fundamental results, both positive and negative results in areas such as approximation algorithms, circuit lower bounds, and the relationship between intractability and randomized computation. The time is ripe for a refocusing of efforts, in order (i) to consolidate and build on the substantial body of knowledge and techniques that have been developed in the past, and (ii) to identify the next set of research directions and subgoals that will move us towards the ultimate goals of the field.

It is our goal to involve a large number of people in the special year, in the usual DIMACS mode of bringing together people with different backgrounds and points of view. Computational intractability is of direct interest to a large number of researchers working in the theory of algorithms and complexity theory. The field of complexity lower bounds has been an area of interest to discrete mathematicians and logicians. Algorithm design and analysis fosters interactions between theoreticians, discrete mathematicians, probabilists, and researchers in applied areas who need to use algorithms for hard problems. The search for new mathematical approaches in both complexity theory and algorithms is causing the field to branch out to other areas of mathematics, e.g., algebraic geometry and topology, and we will encourage interactions with these and other fields.

Workshops, Miniworkshops and Special Events Planned for 1999-00:

- **Pseudorandomness and Explicit Combinatorial Constructions**
  - Dates: October 4-6, 1999, tentative
  - Location: Rutgers University
  - Organizer: David Zuckerman, University of Texas
- **Probabilistic Analysis of Algorithms For Hard Problems**
  - Dates: November 8-10, 1999, tentative
  - Location: Rutgers University
  - Organizers: Alan Frieze, Carnegie-Mellon University and Martin Dyer, University of Leeds

- **Approximation Algorithms for Hard Problems**
  - Dates: February 7-9, 2000, tentative
  - Location: Princeton University
  - Organizers: Sanjeev Arora, Princeton University, and Eva Tardos, Cornell University
- **Faster Exact Algorithms for NP-Hard Problems (mini-workshop)**
  - Date: February 10, 2000, tentative
  - Location: Rutgers University
  - Organizer: Ramamohan Paturi, U.C. San Diego
- **Intrinsic Complexity of Computational Problems**
  - Dates: March 20-22, 2000, tentative
  - Location: Rutgers University
  - Organizers: Paul Beame, University of Washington, Steven Rudich, Carnegie-Mellon University, and Andrew Yao, Princeton University
- **Unexpected and Non-Standard Applications of Complexity (mini-workshop)**
  - Date: March 23, 2000, tentative
  - Location: Rutgers University
  - Organizer: Christos Papadimitriou, U.C. Berkeley

3.8 Exploratory and Specialty Workshops, Continuing Impacts of Earlier Special Years, Partnered Programs

Overview:

In addition to programs and projects 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 be inherently smaller in scale than a special year, that might develop into special focus topics, that "revisit" former special focus topics to keep up the scientific momentum established during the focus, or that involve collaborations with other units.

Workshops, Miniworkshops and Special Events, Feb. 1998 - Nov. 1998:

- **Workshop: Discrete Mathematical Chemistry**
  - Dates: March 23-25, 1998
  - Location: Rutgers University
  - Attendance: 59
  - Organizers: Pierre Hansen, Univ. of Montreal; Patrick W. Fowler, University of Exeter; and Maolin Zheng, Chemical Abstracts

  Discrete Mathematics has been used in chemistry for well over a century. Graph theory is a tool of choice to represent molecules (the name graph appears to have been coined by A. Cayley in the context of such studies), clusters or reaction paths. Group theory is much used for the study of molecular symmetries. Coding theory is basic in systematizing the enormous amount of chemical data available, in problems of enumeration of molecules, isomers, and families having various properties. Graph invariants are much used in Structure-Activity and Structure-Property Relationship studies. All these techniques are increasingly used by industry in the rapidly expanding fields of Computer-Assisted Molecular Design and Combinatorial Chemistry. These are the topics to which this workshop was devoted. While there have long been contacts between chemists and mathematicians or computer scientists (a few of whom do most of their research on problems from chemistry), much mathematical work is done independently by the chemists. This had led to new chemical theories but also to new concepts (e.g. the Randić number in graph theory) and new algorithmic ideas. However, there has been, and still is, some duplication of effort. The aim of the workshop was to bring together chemists working in the mathematical chemistry field; mathematicians and computer scientists working in that field; mathematicians and computer scientists who do not yet work on chemical problems but whose research topic and/or recent results hold much potential for solving problems in chemistry; and chemists, mathematicians and computer scientists working in the chemical industry. The workshop participants discussed theoretical problems of chemistry expressed by discrete mathematics; chemical graph algorithms; coding theory applied to chemistry; applications of discrete mathematics in the chemical industry; open problems and directions for future research in discrete mathematical chemistry; and software for discrete mathematical chemistry (with demonstrations).

Kurt Mislow presented a featured lecture entitled "Molecular Rubber Gloves" and Gunnar Brinkmann presented one on "Problems in Generating Chemical Structures."

Other interesting talks were by Andreas Dress on constructive molecular geometry; Dragos Cvetkovic on properties of some graph invariants related to electron charges in the Huckel molecular orbital theory; Petr Pancoska on spectra, graphs and proteins (discrete mathematical tools for structural studies of proteins); Philippe Vismara on abstract representation for molecular graphs; and Andreas Dietz on representation of stereochemistry using combinatorial maps. Also, Horst Sachs talked about hexagonal tilings of the torus, (3,6)-cages, and their spectra; Daniel Huson on the combinatorics of periodic tilings; Peter Rowlinson on star sets and star complements in finite graphs; and Troy Shinbrot on granular patterns and granular coarsening. Other talks were by Dan Rockmore on group transforms for chemistry; C.M. Quinn on group and graph theory perspectives on large icosahedral carbon cages; and Karen Collins on symmetry breaking in graphs. Other speakers spoke about chemical nomenclature, stable fullerenes, software to explore chemical graphs, carbon networks, molecular shape analysis, fast and simple embedding of planar graphs, path matrices, computer-catalyzed chemistry, and handling synthesis planning in combinatorial chemistry by group theory. Still other topics included ordering trees with a perfect matching by their energies, 2-d hyperbolic nets, combinatorial geometry problems motivated by Benzenoid Hydrocarbons, algorithms for determining fixed bonds and normal components in a bipartite molecular graph, integral polyhedra and maximum resonant sets, property correlations among small molecules, similarity expression, applications of category theory to theoretical biochemistry, large sphere packings and crystal growth, disordered cellular structures, and enumeration of fusenes.

- **DIMACS-DIMATIA Workshop: From Erdős to Algorithms: Applications of the "Probablistic Method"**
  - Dates: June 8-10, 1998
  - Location: Rutgers University
  - Attendance: 84
  - Organizers: Joel Spencer, NYU; Jarik Nesetril, DIMATIA

The probabilistic method is a legacy of Paul Erdős that continues to grow and flourish and have powerful applications in all parts of the mathematical sciences. This workshop explored current results in discrete mathematics that use probabilistic existence arguments and require the use of sophisticated probability concepts. It also explored the connection to modern issues in computer science by examining the possible implementations of probabilistic existence arguments by randomized or deterministic algorithms.

Speakers at the workshop reported on numerous areas of research that are part of Paul Erdős' legacy. Vera Sős talked about $n\alpha$ sequences, Joszef Beck about multidimensional tic-tac-toe, Lenore Cowen about compact routing with minimal stretch, Michael Molloy about algorithmic aspects of the Lovasz Local Lemma, Prasad Tetali about expansion and concentration, Ehud Friedgut about sharp thresholds for Ramsey properties of random graphs, and Joel Spencer about a new B-bound of Srinivasan and Radhakrishnanz. Also, Van Vu spoke about small complete arcs in finite projective planes, Josep Diaz about layout problems on random graphs, Laszlo Lovasz about Markov chains and random walks, David Karger about probabilistic methods for cuts in graphs, Alan Frieze about random minimum length spanning trees in regular graphs, Ludek Kucera about small random graphs, and Fan Chung about spanning trees and the zeta function of a graph. Pavel Valtr spoke on the Erdős-Szekeres theorem, Gregory Sorkin on the random assignment problem, Michael Krivilevich on the choice number of dense random graphs, Yoshiharu Kohayakawa on the width of random subsets of Boolean lattices, Jef Kahn on Entropy, Jeong-Han Kim on increasing subsequences of random permutations, Peter Winkler on random colorings of a tree, and Gyorgy Elekes on coincidence that cannot happen just by chance. The talks gave dramatic evidence for our view that Paul Erdős' work has had major influence on many areas of the mathematical sciences that Erdős himself was not involved in or even knowledgeable about.

- **Joint DIMACS-Erdős Institute Workshop on Probabilistic Combinatorics**
  - Dates: June 25-July 8, 1998
  - Location: Mathematical Institute of the Hungarian Academy of Sciences, Budapest, Hungary
  - Attendance: 75
  - Organizer: Bela Bollobas

This workshop was held in the framework of the Paul Erdős Summer Research Center of Mathematics and dealt widely with the field of probabilistic combinatorics. This was the first workshop held by the new Erdős Institute and inaugurated the partnership with DIMACS. Among the featured speakers were Noga Alon, Borris Pittel, Fan Chung, Michael Krivelevich, Gregory Sorkin, Vu Ha Van, John Wierman, Tamas Szabo, Joszef Balogh, Tibor Szabo, and David Weinreich.

- **DIMACS-DIMATIA Workshop on Combinatorial and Algorithmic Geometry**
  - Dates: August 18-21, 1998
  - Location: DIMATIA, Charles University, Czech Republic
  - Attendance: 30
  - Organizers: Bernard Chazelle, Princeton; Jiri Matousek and Pavel Valtr, DIMATIA

This workshop covered a variety of topics in combinatorial and algorithmic geometry, with emphasis on probabilistic techniques in computational geometry. The twin effect of randomization and derandomization in computational geometry has been spectacular. A whole new theory of deterministic sampling has been developed recently and, as a result, several outstanding open problems have been solved (e.g., optimal algorithms for convex hulls and fixed-dimensional convex programming). This was the background for this workshop.

Among the highlight talks were those by M. Sharir on arrangements, B. Gaertner on combinatorial linear programming, T. Dey on open curve reconstruction, G. Elekes on incidences, polynomials, and metric combinatorics, and I. Barany on spiralling shortest paths on convex surfaces. P. Valtr talked on generalizations of Davenport-Schinzel sequences, J. Solymosi on crossing patterns of segments, E. Szabo on how to find groups, P.K. Agarwal on the complexity of the union of cylinders in 3-space, and B. Chazelle on the square discrepancy for halfplanes. Finally, G. Karolyi talked about new results related to the Erdős-Szekeres theorem, H. Edelsbrunner on topology preserving edge contraction, and S. Har-Peled on planar cuttings.

- **DIMACS Tenth Birthday Conference**
  - Date: October 9, 1998
  - Location: Rutgers University
  - Organizer: Fred Roberts

DIMACS celebrated its tenth birthday with a major scientific conference. There were six invited speakers discussing current areas of research and tracing them to their roots in DIMACS programs. Building on the DIMACS Special Year in Mathematical Support for Molecular Biology, Richard Karp talked about sequencing the human genome and Pavel Pevzner about evolution in a talk entitled "Transforming Men into Mice." Referring back to the Special Year on Discrete and Computational Geometry, Ricky Pollack talked about algorithms in real algebraic geometry. Tracing her results to the Special Focus in Discrete Probability, Jennifer Chayes talked about finite-size scaling in percolation and satisfiability. Discussing work started during the Special Year on Logic and Algorithms, Moshe Vardi spoke about verification of open systems. Finally, Laszlo Lovasz, building on work started during the Special Year on Combinatorial Optimization, gave a talk about speeding up sampling by Markov chains.

A second feature of the conference was a panel discussion on academic/industry cooperation moderated by Fred Roberts and Bob Sedgewick and featuring remarks by Alfred Aho, Associate Research Vice President, Bell Labs/Lucent Technologies; Richard DeMillo, Vice President, Information and Computer Sciences Research, Bellcore; James Flanagan, Vice President for Research and Director of the Center for Computer Aids to Industrial Productivity, Rutgers University; C. William Gear, President, NEC Research; and David Johnson, Head, Algorithms and Optimization Department, AT&T Labs.

A third feature was a panel discussion on educational issues in discrete mathematics and theoretical computer science moderated by Joseph Rosenstein and featuring remarks by Margaret (Midge) Cozzens, Vice Chancellor for Academic Affairs, University of Colorado, Denver; Henry Pollak, Visiting Professor of Mathematics Education, Teachers College, Columbia, and Bellcore, retired; and Stephen Maurer, Professor, Department of Mathematics and Statistics, Swarthmore College.

Many distinguished visitors attended and there were remarks by Francis Lawrence, President of Rutgers University, Jeremiah Ostriker, Provost of Princeton University, John Cherniavsky, Director of the Division of Experimental and Integrative Activities of the National Science Foundation, and Robyn Klotsky, Associate Director of the New Jersey Commission on Science and Technology. Professor Felix Browder of Rutgers University made luncheon remarks in a tribute to the founding director of DIMACS, Danny Gorenstein.

Events Planned for the Rest of 1998-99

- **DIMACS-HKUST Far-East Workshop on Algorithms and Combinatorics**
  - Dates: December 18-20, 1998
  - Location: Hong Kong University of Science and Technology
  - Organizers: Mordecai Golin, HKUST; Fred Roberts, DIMACS; Ronald Graham, AT&T Labs-Research; Ottfried Cheong, HKUST; Siu-Wing Cheng, HKUST; Xiao Tie Deng, City University of Hong Kong
- **Workshop: On-line Decision Making**
  - Dates: March 8-11, 1999
  - Location: Rutgers University
  - Organizers: Yoav Freund, AT&T Labs - Research; Rakesh Vohra, Northwestern
- **DIMACS-Institute for Research in Cognitive Science Workshop on Logic and Cognitive Science**
  - Dates: April 16-18, 1999
  - Location: University of Pennsylvania
  - Organizers: Moshe Vardi, Rice University and Scott Weinstein, University of Pennsylvania
- **DIMACS-DIMATIA Workshop on Algebraic Methods and Arithmetic Circuits**
  - Dates: June 2-4, 1999
  - Location: Rutgers University
  - Organizers: Eric Allender, Rutgers; Jan Krajicek, Czech Academy of Sciences; Pavel Pudlak, Czech Academy of Sciences; Mike Saks, Rutgers University
- **DIMACS-Hungarian Academy of Sciences Workshop: Paul Erdős and his Mathematics**
  - Dates: July 4-11, 1999
  - Location: Mathematical Institute, Hungarian Academy of Science, Budapest, Hungary
  - Organizers: V. T. Sős, Hungarian Academy of Sciences
- **Workshop: Showcasing Women in Theoretical Computer Science**
  - Dates: TBA
  - Location: Rutgers University
  - Organizer: Julia Abrahams and Fred Roberts, Rutgers University
- **Workshop: Unusual Applications of Number Theory**
  - Dates: TBA
  - Location: Rutgers University
  - Organizers: Jeffrey Lagarias, Andrew Odlyzko and Ronald Graham, AT&T Labs - Research; Melvyn Nathanson, CUNY

- **Workshop: Stock Market Modelling and Mathematical Finance**
  - Dates: TBA
  - Location: Rutgers University
  - Organizer: Larry Shepp, Rutgers University

Selected Research Accomplishments

Our exploratory and partnered activities often lead to surprising and powerful results. Moreover, 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 founded or initated during a special year or workshop may lead to questions that influence the field for many years. A variety of examples continues to illustrate this point and we present some examples here along with examples of work that has come out of exploratory or partnered activities.

We continued our program of evaluating the impact of earlier special years by commissioning an evaluation of DIMACS' second special year, the 1990-91 Special Year on Complexity Theory of Interactive Computing as seen from the perspective of eight years. The report by Lance Fortnow, Michael Luby, and Noam Nisan, which can be found on the DIMACS website at the URL: http://dimacs.rutgers.edu/About/Reports/complexitytheory.html, provides an exceptionally strong piece of evidence of the long-term impact of DIMACS' progams. It talks about the special year work of Feige, Goldwasser, Lovasz, Safra, and Szegedy that resulted in "perhaps the most influential paper in theoretical computer science during the 90's," a paper that "with a sheer stroke of genius ... introduced the scientific framework for a general theory of hardness of approximation, and showed that it was as hard to approximate some natural problems as it is to find an exact solution." The report talks about how the paper "represents the ideal of the DIMACS model in terms of its combination of authors, bringing together far flung researchers under a common theme." The report talks about the "floodgate of research" that resulted from the special year and continues today, and says that "the special year was a huge success."

The first three results that we mention can be traced back to our Special Year on Logic and Algorithms, with the third of these also related to the DIMACS-DIMATIA partnership.

The Commercial Verification Tool FormalCheck Computer-aided verification was a central topic in the Special Year in Logic and Algorithms and work and collaborations tracing back to that year continue. The users of the commercial verification tool FormalCheck often inquire about the relationship between the queries that are expressible using the user interface of the tool, and CTL specifications commonly used in the context of other tools. It turns out that there are significant misunderstandings, based upon subtle distinctions which often are overlooked even by experts. Moreover, there are certain common properties expressible with FormalCheck which (often to the surprise of users) cannot be expressed in CTL, some which can be expressed in an easily augmented version of CTL (CTL + Booleans), and others which cannot be expressed in any reasonable augmentation of CTL. James Zijang (Rice University, summer student at BL), In-Ho Moon (summer student at BL), Christine Chung (summer student at BL) Neil Immerman (V), Thomas Wilke (V, former PD), and Bob Kurshan (PM, BL) have worked to clarify all these issues. The fact that strong fairness, a basic behavior type for FormalCheck users, is not expressible in CTL even if augmented by Booleans, turned out to be quite non-trivial to prove, and not to be found in available publications. A complicated proof based upon game-theoretic techniques produced by Immerman was reproved in a shorter fashion by Wilke by citing a (deep) result of Vardi. This represents both mathematically deep and very practical work: the document which explains this will be given to all FormalCheck customers.

Message Sequence Charts Anca Muscholl (V) worked with Doron A. Peled (PM, BL) on the complexity of some basic verification tasks concerning message sequence charts (MSCs). Using MSCs in the early design of communication protocols is advantageous due to the conceptual simplicity of both syntax and semantics. Thus, MSCs are widely used for describing the communication skeleton of systems. Mushcholl and Peled were basically interested in two questions in this area. The first question concerns race conditions in MSCs, which arise when the causal order associated with an MSC allows more executions than the specification (so-called visual order). Fortunately, (basic) MSCs race conditions can be checked in square time, as previously shown. For high-level MSCs, i.e. MSCs given by labelled graphs (or automata), Muscholl and Peled showed that detecting race conditions is undecidable. The second problem they looked at concerns branching (nondeterminism) in high-level MSCs. The question arises whether there is some branching state which generates two executions α, β which are synchronizable, but there is no execution γ which subsumes α, β (where synchronization is meant with respect to partial order). They showed that the branching problem for high-level MSCs is undecidable, too. Both the race problem and the branching problem become decidable when high-level MSCs are restricted to be loop-connected, which is a syntactical condition depending on the chosen semantics. Muscholl and Peled also looked at the complexity of both problems with this restriction and obtained an exponential-space lower bound.

An Improved Exponential-Time Algorithm for k-SAT Ramamohan Paturi, Pavel Pudlak, Mike Saks (PM, RU), and Francis Zane (PD, BL) proposed and analyzed a simple new algorithm for finding satisfying assignments of Boolean formulae in conjunctive normal form. The algorithm, ResolveSat, is a randomized variant of the DLL (Davis, Longeman and Loveland) or Davis-Putnam procedure. They showed that, for each k, the running time of ResolveSat on a k–CNF formula is significantly better than 2^n, even in the worst case. In particular, they showed that the algorithm finds a satisfying assignment of a general 3–CNF in time $O(2^{.446n})$ with high probability, where the best previous algorithm has running time $O(2^{.582n})$. For each k, the bounds for general k-CNF are the best known for the worst-case complexity of finding a satisfying solution for k-SAT. The ideas used to analyze the algorithm were also applied to obtain lower bounds on circuit size. Here, they exhibited a function f such that any depth–3 AND-OR circuit with bottom fan–in bounded by k requires $\Omega(2^{c_k n/k})$ gates (with $c_k > 1$). This is the first such lower bound with $c_k > 1$.

The next two results we mention arose from our continuing exploratory interest in quantum computing and quantum cryptography.

Security of Quantum Cryptographic Systems Dominic Mayers (V) and Andrew Yao (PM, PU) extended the security proof for the case of ideal quantum apparatus so that it works with practical error-correcting codes, not only with random linear codes. Quantum cryptography uses the quantum property of light to realize cryptographic applications such as a secret communication between two persons. The advantage of quantum cryptography is that it does not use any computational or technological assumption on the power of a cheater (that is, an eavesdropper in a secret communication). Quantum cryptographic systems were implemented in laboratories in the U.K., in Geneva and in the US. However, the proof that these systems are really secure turned out to be a difficult task, open for the last 10 years, even if it is assumed that the apparatus used by the honest participant works perfectly, which is not the case in practice. Mayers and Yao worked out the details of the proof including an important extension allowing the use of practical linear error-correction procedures in the protocol.

Security of Quantum Cryptographic Systems with an Imperfect Apparatus In work related to the previous example, Dominic Mayers (V) dealt with the question of how to deal with an imperfect apparatus. It is not sufficient that the ideal protocol be secure. It is necessary to show that the protocol that is actually executed by the participants using imperfect apparatus is also secure. Of course, some conditions on the apparatus must be assumed because clearly, with no assumptions, it is like having no protocol at all. Even in the non-quantum case, these assumptions are necessary. The objective is to use the same kinds of assumptions in quantum cryptography as in non-quantum cryptography, not more. Mayers proposed a new kind of protocol that almost respects this property, that is, only very few assumptions are required on the quantum apparatus used by the honest participants.

The next result that we describe is related to the long-standing interest at DIMACS in discrete and computational geometry, which has its roots in the first special year at the center.

The Erdős-Szekeres Theorem in Discrete Geometry Gyula Karolyi (V) and Geza Toth (PD, RU) worked on a problem in discrete geometry related to the Erdős-Szekeres Theorem. They improved upon former results obtained by Bialostocki, Dierker, and Voxman and by Caro, who had proved that for arbitrary positive integers q, n, $(n \geq q + 2)$, in any set of points, in general position in the plane, which is large enough, one can find n points which form the vertex set of a convex polygon whose interior contains 0 points modulo q. (A negative result, which motivated research in this direction, is due to Horton, who constructed arbitrarily large point sets containing no empty convex heptagon.) Both proofs were based heavily on the condition $n \geq q + 2$, although there were no counterexamples known for $n = q + 1$. Karolyi and Toth were able to relax the condition to $n > cq$ with some constant $c < 1$ and hope that it will eventually lead to the resolution of the general conjecture, according to which no restriction on n is necessary. As a byproduct, they found a non-trivial condition which guarantees the existence of a large empty polygon if the number of given points is sufficiently large.

The next result we mention arose from our exploratory forays into mathematical finance. This work has led to an upcoming DIMACS workshop on mathematical finance (see above).

Option Pricing without Completeness and Non-Arbitrage The celebrated Black-Scholes Theorem for pricing an option is based on two assumptions: The market is complete in the sense that everyone in the market has the same information about the future of each stock or asset (which is patently false since attempts are made to limit profits based on insider information); a player in the market can never find an *arbitrage* in the sense that one cannot both buy and sell some asset putting in zero money and extracting a guaranteed profit (which sounds more reasonable but amounts to the same as the first assumption). Dropping these assumptions allowed Larry Shepp (PM, RU) and Xin Guo (GS, RU) to incorporate the more realistic notion that "inside" information might sometimes arise. They presented a new model for the fluctuations of a stock price X_t and used the new model to re-price hedge options. In addition, they took a different approach to option pricing based on the ideas of von Neumann and Morgenstern. This was an approach advanced by Samuelson, but later retracted by him in favor of martingale methods. Shepp and Guo's new model was

$$dX_t = X_t(\mu_{\epsilon_t} dt + \sigma_{\epsilon_t} dW_t),$$

where ϵ_t is a stochastic process which represents the *state of information* in the investor community. This model allows the market to occasionally be incomplete and arbitrage to exist from time to time. They considered the case when ϵ_t is Markovian, and they were able to derive closed form solutions for the maximal expected values of the European option, the Russian option, and the

American option, under optional and optimal choice of time of execution of each option. They argued that these values ought to be considered to be the fair price of each option in the new and more accurate model.

The last result we mention is related to DIMACS' long-standing interest in graph theory, which was the topic of our third special year.

Crossing Numbers of Graphs Janos Pach, Joel Spencer (V), and Geza Toth (PD) considered a *drawing* of a graph G, a mapping which assigns to each vertex a point of the plane and to each edge a simple continuous arc connecting the corresponding two points. The *crossing number* of G is the minimum number of crossing points in any drawing of G. They defined two new parameters, as follows. The *pairwise crossing number* (resp. the *odd-crossing number*) of G is the minimum number of pairs of edges that cross (resp. cross an odd number of times) over all drawings of G. They proved that the determination of each of these parameters is an NP-complete problem. They also proved that the largest of these numbers (the crossing number) cannot exceed twice the square of the smallest (the odd-crossing number). Their proof is based on the following generalization of an old result of Hanani, which is of independent interest. Let G be a graph and let E_0 be a subset of its edges such that there is a drawing of G in which every edge belonging E_0 crosses any other edge an even number of times. Then G can be redrawn so that the elements of E_0 are not involved in any crossing.

3.9 Multi-Disciplinary, Multi-Institutional Projects

DIMACS has two multi-disciplinary, multi-institutional research projects in network modeling and parallel simulations, each funded separately from its main STC award, and another such project involving software for research and education. Our projects of this kind are scientifically exciting and have taken us into new and exciting research areas.

The participants in these projects, both DIMACS members and outsiders, have repeatedly emphasized that the projects would not have happened without the organizational infrastucture provided by DIMACS. 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. Importantly, DIMACS provides a neutral site for collaboration: Industrial researchers find it easier to work on a joint high risk precompetitive project under the umbrella of an academic center which enhances the opportunities for interactions.

3.9.1 Network Modeling and Parallel Simulations of Networks

This project, which consists of two funded research efforts 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, and are now closely related to our continuing Special Year on Networks. The first is funded by NSF and the second, followup project by DARPA. We describe these two projects together in this section. In 1998, WINLAB received an NSF Instrumentation Grant for a 14-processor SUN ultrasparc that is used in these simulation projects.

Simulations of Integrated Communications Systems

This project is funded by the National Science Foundation for a three year period over 1995 - 1998, for a total of $1.5M.

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, RU), Diane Souvaine (PM, RU), Don Towsley (University of Massachusetts), and Paul Wright (PM, ATTL).

Collaborations: This project brings together experts in network modeling, design and analysis, distributed algorithms, and high performance discrete-event simulation to focus on the development of high-performance simulation tools for broadband networks and wireless communication services. This project directly brings together the expertise of scientists from two of the nation's leading telecommunications companies, in partnership with leading academic sites; and includes graduate and postdoctoral training that benefits students through exposure to industrial scale problems and internships.

Scalable Self-Organizing Simulations (S3)

This project is funded by DARPA for a three year period over 1996-1999, for a total of $3.1M. It represents a highly successful transition of the NSF-supported networks simulations project to DARPA. The theoretical underpinnings developed under NSF support are leading, in this project, to significant practical code creation and applications.

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

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, networking, software engineering, and defense applications. We have put together a team of scientists and engineers from seven institutions to work on the project, including participants from two small companies. Academic collaborators involve in the project postdoctoral fellows and many graduate research assistants.

Objective of the Projects

The goal of these projects is to achieve a generational transition in speed, scalability and manageability of simulations of very large scale multiprotocol communication networks. The challenge of enabling discrete event simulations of hundreds of thousands to millions of communicating entities requires novel, scalable network modeling techniques, and new, lightweight architectures for ultra-fast parallel simulation kernels. Current practice of network modeling and simulation cannot even approach problems of such size and complexity.

Approach

The success of these projects requires a significant collaboration across three domains: parallel discrete event simulations, software engineering, and network modeling. Therefore, the high level view of the project structure is as follows:

First, there is core software research involving the design and implementation of innovative software capable of reaching the "million entities target", including: i) Scalable, simulator-independent and user-friendly network modeling framework; ii) Very high speed, process-oriented parallel simulation kernels capable of supporting millions of logical processes; and iii) Highly usable, extensible runtime software environment. These project components will deliver the tools capable of efficient modeling and fast interactive simulations of very large communication networks.

Second, there is network modeling research on efficient model representations for network elements and protocols that are natural for networking researchers and engineers and strongly promote scalability; and on automation of network topology and traffic aggregation methods that can reduce model complexity without loss of accuracy. Very large models used in networking research also require generation of network topologies that approximate the global Internet, and automation of address and bandwidth/delay assignment in such models.

Third, there is practical validation of the new techniques by demonstration of scalability of simulations for a wide range of network types and protocol classes. Within these projects, these include: i) the Internet and ATM/PNNI internetworks - protocols, topology, traffic sources; and ii) Mobile wireless networks - protocols, radio propagation. External users have been experimenting with a variety of other models. The testing process involving internal and external users feeds back corrections and new ideas to simulator and tool developers. This project component will deliver a database of important models together with canonical methods for efficient parallelization.

Recent Accomplishments

S3 parallel simulators have demonstrated event execution rates and entity counts on realistic models well in excess of those achievable using the best tools available elsewhere.

A new very high performance conservatively synchronized parallel simulator, Nops, was designed and implemented. Simulation speed exceeding million events/second, and scalability to several hundred thousand network elements were demonstrated.

Performance of the GTW optimistically synchronized parallel simulator improved over 100%, and scalability to over a million network elements was demonstrated.

An extensive suite of portable graphical control and visualization tools for TeD (Jane) was developed. TeD/GTW modeling capability was extended to arbitrarily nested process suspension. This capability is not available in any other optimistic simulator.

The next generation modeling framework SSF was designed together with the reference Java implementation atop the JTED simulator. Scalability in parallel and distributed simulations was demonstrated. High performance implementation of SSF atop the Nops simulator is nearing completion. Progress was made in implementation of MegEnta, the ultra-high performance successor to GTW that will implement the SSF interface.

Novel techniques for simulated packet traffic aggregation were implemented. A 10-fold increase of simulation speed and high accuracy were demonstrated in a model of ATM internetwork.

The model of long-range correlated (fractal) Web traffic generator Surge was implemented in SSF/JTED. Surge's original configuration is used worldwide, and has been adopted by the World Wide Web Consortium (W3C) as a core component of their HTTP-NG testbed.

Testing of complex models of mobile wireless networks was concluded. Parallel speedup and scalability with varying detail level were demonstrated.

Testing of ATM internetwork models with full PNNI protocols was concluded. Excellent parallel speedup and scalability with network size were demonstrated.

Porting of protocol models from other sequential simulators (ns) to TeD was demonstrated, and simulation scalability was shown.

The first prototype system for dynamic hierarchical visualization of very large network graphs was completed.

As a result of our work, this is the first time anybody can actually model and simulate the global Internet with a hundred thousand hosts and routers, running IP packets with real routing protocols at a rate of one million packet hops per second (more on newer machines).

Technology Transfer

The complete parallel simulation package Jane/TeD/GTW is being actively used on networking projects in over 20 corporations and universities, including Lucent Bell Labs, Bellcore, Newbridge Networks, Traveller Multimedia Networks, Mitre Corp., JPL, and also in projects in other areas, such as virtual factory simulations at Gintic Inst. of Manufacturing.

The GT-ITM network topology generation package is in use in several network research projects, including the ns v2 simulator (project VINT at UCB, USC, LBL), U. Oregon, and others.

Communications application requirements were coordinated with Air Force Research Laboratory (AFRL) focusing on the Air Force C2 initiatives for Global Awareness, Global Grid, and Dynamic Planning and Execution. Transitioning of the S3 technologies to Air Force simulation needs was defined.

Models of networks developed in project S3 for parallel simulation capability testing are being used in networking research. The WiPPET Wireless testbed is extensively used in research on wireless IPv6 protocols at WINLAB, Rutgers. Several wireless networking companies and ARMY CECOM have expressed interest in this package, and further distribution will be made.

3.9.2 The LINK Project

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 are 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 was funded by an NSF grant to DIMACS for a three year period over 1993 to 1996, and it has also received support from the Office of Naval Research and Los Alamos National Laboratory. Participating scientists in the first stage of LINK were Jonathan Berry (PD, Elon College), Nathaniel Dean (PM, BL), Mark Goldberg (Rensselaer Polytechnic Institute), Greg Shannon (Spanning Tree Technologies), and Steve Skiena (SUNY Stony Brook).

LINK played a central role in two summer programs at DIMACS in the summer of 1998. Berry and Dean were two of the principal lecturers in the summer conference aimed at "reconnecting" instructors at two and four year colleges to the mathematical sciences enterprise (see Section 4.6.2). Berry and Dean lectured on recent research in network visualization, and introduced LINK as a hands-on tool for participants to experiment with research problems and to help develop educational materials to be brought into the classroom. Berry was also instrumental in introducing LINK to both high school teachers and researchers at our DREI'98 program on graph theory and network optimization, where it was emphasized as a tool for mastering concepts in graph theory and a research tool related to graph theoretical conjectures. (See Section 4.6.1 for a description of the DREI program.)

DIMACS sponsored a one-semester course release for Jon Berry, currently the LINK project leader, to prepare the system for use in the Reconnect and DREI programs. Berry completed the port of LINK to Windows 95 and NT so that it could be used in our PC lab during our programs and developed a 50 page tutorial designed to teach the system to non-technical users.

DIMACS also supported three undergraduate students to learn how to teach the use of LINK to the Reconnect and DREI participants. These students became integral parts of both programs and contributed their youth, undergraduate experience, and new LINK expertise, to the educators who attended. Two of the three students plan to submit their work on educational modules in LINK to the National Conference on Undergraduate Research.

The Reconnect and DREI experiences exercised the LINK software as it had never been exercised before and highlighted several areas in need of improvement, helping Berry to arrange the project's priorities for its continued development. These improvements are being implemented during the 1998-1999 academic year.

4 Education and Human Resources

4.1 K-12 Programs

DIMACS' programs for precollege teachers and students go back to the very first year of the center. Through them, we have had significant influence on the precollege curriculum, the philosophy of mathematics education, and on the teaching and careers of thousands of teachers and students. In this section, we discuss the two main precollege programs of the center, the Leadership Program for K-12 teachers and the Young Scholars Program for high school students, as well as a variety of other programs and products and the impact that they have had.

4.1.1 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 (1995-1999). In the present phase, under a grant from NSF and support from DIMACS and from its STC grant, teachers attend a two-week program during the summer, attend four follow-up days (six 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 and a half hours during which participants are engaged in mathematical activities which illustrate and highlight the key conceptual issues discussed in the workshop.

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. Once a topic has been fully explored, the teachers at the **LP** work 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 **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 on discrete mathematics among their peers and that they are encouraged and expected to try this material in their classrooms, to introduce it into their curricula, and 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. A typical year's cohort in the K-8 **LP** program gives workshops attended by at least 5000 teachers.

The 1998 **LP** involved two two-week residential programs (both at Rutgers) and one twelve-day commuter program, in Scottsdale, AZ. Those programs represent exactly 100 teachers participating in the **LP** for the first time. In addition, there were three one-week second-summer follow-up programs for the 1997 participants – two at Rutgers (accommodating both residential and commuters), and one (for commuters) at Scottsdale, AZ. Altogether, 90 of the 125 teachers in the 1997 summer programs attended one of these follow-up programs. We also conducted a two-day "crash course" in discrete mathematics, which was attended by 39 high school teachers. Thus, we have been able to continue our high school teacher programs even though we did not have a specific extra grant to fund these activities.

In order to ensure that our off-site programs are comparable in quality to our programs at

Rutgers, we have developed high-quality materials which were used at all different sites by a number of different workshop leaders. These materials include detailed instructor's notes, overhead transparencies, handouts, homework problems, and supplementary resource materials. All new instructors are provided with some training in the use of these materials and in the philosophy of the **LP**. Moreover, in order to compensate for the extra time that is available to participants in the resident programs, the 1998 commuter program was extended by 2 days, so that it involved 12 days of activities, as noted above.

Although this phase of the **LP** was expected to end with the 1998 cohort, we anticipate receiving NSF approval to conduct additional institutes for new teachers this summer so that we can reach the intended 660 teachers. The 1999 schedule will involve two two-week residential programs at Rutgers and three twelve-day commuter programs in New Jersey, Rhode Island, and Arizona. We anticipate that altogether these five programs will have an enrollment of 158 teachers, which will bring the total number of participants to 660.

In the summer of 1999, we will also have the returning teachers from the summer of 1998, and veterans from all programs before that. There will also be another two-day "crash course" for high school teachers.

The materials that have been developed for the **LP** workshops described above are sufficiently detailed and have been sufficiently field-tested and evaluated, that it will be possible for other mathematicians to use them as the basis for independently-conducted ("franchised") parallel high-quality institutes in 1999, after suitable training provided by the leadership program.

The **LP** is directed by Joseph Rosenstein (PM, RU).

4.1.2 Young Scholars Program in Discrete Mathematics

In its initial eight-year phase, The Young Scholars Program in Discrete Mathematics **YSP** received substantial funding from the National Science Foundation. Because of the success of the program, DIMACS decided that it would attempt to continue the **YSP** in the same format, even though funding from NSF was no longer available for such programs. The following model was developed and implemented during the period from 1990-1997: The **YSP** is a four-week residential program for 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 close to a 50/50 balance between females and males, and approximately 20% of the students are minorities.

An important feature of the **YSP** (since 1994) is the research component 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. Project topics have included secret communication, scheduling, polyomino tilings, and text compression. Both students and researchers have found the projects valuable. The researchers participating have included Las Goonetilleke ("Simulating the Truth"), Robert Hochberg (PM, RU, "Dominos Vobiscum" and "Ominous Ominos"), Paul O'Donnell (GS, RU, "Coloring the Plane"), Gretchen Osteimer (RU, "A Prime Forgery"), George Popel (Middlesex County College, "Nones and Ones"), Devan Putnam (RU, "Stable Marriages"), 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"), Stephen Penrice (PD, "Can you Design a Fair Draft Lottery?"), Nick Reingold (PM, ATTL, "Dotland"), and Nina Shteingold (GS, RU "Trinkets and More").

For the past three years there has been a significant technology component to the Young Scholars Program. Each year an entire week is 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.

The second phase of the **YSP** began in 1998, with partial funding by DIMACS and a grant of $50,000 from the AT&T Foundation. The cost to the participants was $1,900, a substantial increase from the charges of $700 to $900 in previous years; moreover, a limited number of scholarships were available. DIMACS is seeking additional funding from its other partner companies to reduce the participant fees and provide additional scholarships, and is developing strategies for publicizing the program more widely. It is anticipated that, over the next few years, the **YSP** will enroll 30 students each summer. The program is co-directed by Michael O'Nan (PM, RU) and Joseph G. Rosenstein (PM, RU); the late Robert B. Davis also served as co-director.

4.1.3 Additional K-12 Programs and Accomplishments

Over the past six years, selected teachers in the **LP** have participated in the "Workshops in Your District" outreach program, travelling to many schools to inform their teachers about discrete mathematics. These workshops were appropriate for middle school and high school teachers. However, we have now developed workshops for K-8 teachers as well. During the summer of 1997, 24 of the 1995 and 1996 participants attended a one-week "workshop workshop" program, at which they suggested appropriate topics and activities for workshops that would be designed for K-8 teachers. These suggestions were reviewed and developed further during the past year and are now ready for delivery to K-8 teachers. During October and December 1998, training sessions were conducted for 35 K-8 teachers from the 1995-1997 **LP**, and a brochure was developed for national distribution about the "K-8 Workshops in Your District Program".

We also have a newsletter for high school teachers, "In Discrete Math," that is circulated free twice a year to approximately 3500 teachers.

DIMACS has played a central role in the adoption of a new Curriculum Framework in Mathematics in New Jersey. A DIMACS volume, *Discrete Mathematics in the Schools*, was published in 1997, and it received very strong reviews. Although neither the "Framework" nor the volume were published during the present year, both reached significantly increased visibility nationwide and were widely cited and used.

DIMACS has launched a new Modules Series that will include educational modules relating to current research that is amenable for discussion in the high school classroom. The Modules Series is discussed in Section 4.6.3.

Through the various programs mentioned in this section, we have reached thousands of teachers.

4.2 Undergraduate Education Initiatives

DIMACS runs a number of initiatives in undergraduate education that are placing it in a leadership position nationally relative to inclusion of dm/tcs in the undergraduate curriculum and in introducing undergraduates and college teachers to the topics of dm/tcs and their applications. These initiatives include the "Reconnect" Program, described in Section 4.6.2, and the new Modules Series, described in Section 4.6.3.

Among the other undergraduate initiatives at DIMACS are the following. We continued our successful REU program and made some changes in it (as discussed below). We continued our thrust in logic and computer science in the undergraduate curriculum (discussed below). We continued to host visitors from such four year colleges as Drew and Haverford. We involved in a leadership role at DIMACS faculty from many two and four year colleges, including Drew, Swarthmore, Wheaton, Amherst, Middlesex County College, Union County College, and County College of Morris. We discussed programs with the NJ Chapter of the Mathematical Association of America and the Mathematical Association of Two Year Colleges of New Jersey, and partnered with the latter organization in developing our undergraduate initiatives for two year colleges.

4.2.1 Logic and Computer Science in the Undergraduate Curriculum

In 1997, as a followup to the Special Year in Logic and Algorithms, DIMACS helped to organize a “working group” on curriculum issues in logic and CS under the direction of Moshe Vardi of Rice University. This group organized a symposium on logic and CS at the February 1998 meeting of the ACM Special Interest Group in Computer Science Education (SIGCSE). As an offshoot of this activity, Vardi and his colleagues Kathi Fisler and Matthias Felleisen have developed a project called “Integrating Logic into the Computer Science Curriculum” and have applied for NSF funding for it. The description of this project follows.

Computer science programs prepare students to construct systems. The construction of systems, whether hardware or software, involves designing, selecting, and assembling components to perform a particular task. This process requires reasoning about components to understand their interactions, capabilities, and limitations. A good degree program should help students develop these skills by teaching them the tools that assist in these tasks. Logic is the most fundamental such tool. Software developers analyze program behavior during design, debugging, and testing; logic explains this behavior in the guise of programming language semantics. Hardware designers use gates with a direct correlation to Boolean logic; circuit minimization relies on the logical equivalence of designs. Operating system designers propose protocols for routing, scheduling, and synchronization; establishing protocol correctness requires a specification language and corresponding proof framework, both of which arise from formal logic. A solid preparation in logic exposes students to these issues directly and increases their abilities to reason about complex systems. Unfortunately, few undergraduate computer science curricula prepare students adequately in logic. The typical student encounters a couple of months of truth tables, propositional logic, and basic proof techniques in a sophomore-level discrete mathematics course. Such a cursory introduction does not include topics that illustrate how applying logic helps with practical work. Furthermore, topics on logic are typically relegated to a single course; senior-level courses on applied topics neither discuss the potential applications of logic nor reiterate the logical concepts covered in earlier courses. As a result, students perceive logic as irrelevant to their careers. Departments need curricular measures to rectify this situation. The planned project is aimed at developing a series of modules that seamlessly integrate logic and logic-based software tools into existing, widely taught computer science courses. We anticipate that some of these modules will find their way into the DIMACS modules series described in Section 4.6.3. Integration melds theory and practice, thus giving students concrete examples of how logic serves as a valuable tool. Spreading integration over several semesters reinforces logical concepts over the course of the entire undergraduate curriculum, thus increasing students' ability to absorb the material. Emphasizing logic-based software tools turns the concepts into useful methods and makes the whole topic more appealing to today's student generation. Finally, ready-made teaching modules will allow departments with scarce resources to add material on logical reasoning without introducing new courses. In short, by supplying modules, complete with lecture notes, presentations, problem sets, and tools, the investigators hope to facilitate curricular treatment of applied logic at all levels of college education. This project is aimed at restructuring the basic course on discrete mathematics and developing modules across a diverse collection of common undergraduate courses.

4.2.2 REU Program

DIMACS has run a Research Experiences for Undergraduates (REU) program since 1993. The program introduces mathematically talented students to research in DIMACS-related fields under the guidance of research mentors, both to discover new results and to 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.

Our 1998 program was the largest so far, with 12 students from 10 schools. We continued our DIMACS REU speaker series of talks by distinguished scientists, including Jozsef Beck (PM, RU), Joel Spencer (New York University), Janos Komlos (PM, RU), and Fred Roberts (PM, RU). Each of these talks was followed by an informal "meet the speaker" dinner during which time the students (and other attendees) could interact with the presenter. We housed our students on-campus again this summer, and will continue this practice since it seems to be of great benefit to the students.

Here is a list of participants in the 1998 DIMACS REU program:

- Scott Gaul
  - Mentor: Adi Ben-Israel (PM, RU)
  - Project area: Cost and Hypothesis Testing
  - Home insitution: U. Chicago
- Stephen Grimes
  - Mentor: Charles Sims (PM, RU)
  - Project area: Presentations for Groups of Exponent 6
  - Home institution: Bucknell
- Sara Grossman
  - Mentor: Shibu Yooseph (PD)
  - Project area: Phylogenetic Tree Reconstruction
  - Home institution: Smith College
- Carly Klivans
  - Mentor: Sven Dickinson (PM, RU)
  - Project area: Mobile Robotics Navigation
  - Home institution: Cornell
- Philip Kwok
  - Mentor: Andy Ogielski (PM, RU)
  - Project area: Network Simulation
  - Home institution: Hampshire College
- Tanzy Love
  - Mentor: Peter Hammer (PM, RU)
  - Project area: Logical Data Analysis

  - Home institution: CSU Hayward

- Jaya Paliwal
  - Mentor: Jon Berry (Elon College)
  - Project area: Animating Planarity in Basic Algorithms
  - Home institution: Rutgers

- Gordon Ritter
  - Mentor: Sandeep Bhatt (PM, BC)
  - Project area: Network Visualization
  - Home institution: U. Chicago

- Tony Shaheen
  - Mentor: Rajmohan Rajaraman (PD)
  - Project area: Caching Algorithms
  - Home institution: Loyola Marymount

- Kenneth Tignor
  - Mentor: Wilma Olson (RU)
  - Project area: DNA Supercoiling
  - Home institution: U. Chicago

- Ashley Trent
  - Mentor: Jon Berry (Elon College)
  - Project area: Visualizing Network Flow
  - Home institution: Elon College

- Winnie Yau
  - Mentor: Endre Boros (PM, RU)
  - Project area: Logical Analysis of Data
  - Home institution: Columbia

Of the twelve students we had this year, four are still actively working on their projects, and five say they hope to be able to do more work as it fits their schedule. Most of the students have put their projects on the web: These can be seen at `http://dimacs.rutgers.edu/REU/index.html`. Some of the students have begun papers describing their summer research, and most have given talks or informal presentations at their schools or at undergraduate research conferences. One of our students, Carly Klivans, won the Alice T. Schaeffer award for undergraduate research given out by the Association for Women in Mathematics for her project under the direction of Professor Sven Dickinson. One of our 1996 students, Lynette Hoche, together with her mentor Wilma Olson of the Chemistry Department at Rutgers, has had a paper accepted by the *Proceedings of the National Academy of Sciences.*

In the past, we have designed our REU program for students completing their junior year in college (and in exceptional cases students completing their sophomore year). This year for the first time, we experimented by including one student who had just completed her freshman year. This experiment resulted from a collaboration between DIMACS and the Douglass College Program for Women in Mathematics, Science, and Engineering, and it was a success.

This year, two of the students (Paliwal, Trent) worked closely with the DREI and Reconnect programs, by developing and implementing the LINK software (see Section 3.9) and assisting high school and college teachers and researchers who were learning how to use LINK in our pc labs.

The REU program has also worked out a collaboration with our sister center DIMATIA which, pending funding, will bring two Czech students to DIMACS for the duration of the program next summer and will send six of the American participants over to Prague for participation in a conference there.

The REU program is evaluated every year by questionnaires to the participating students. This year, 78% of the participants rated the experience as a whole as 4 or 5 on a 5-point scale with 5 being the highest. The remaining students rated 3. The ratings reflected the overall level of satisfaction the students had with their success at research. 75% of the respondents rated their overall satifaction with their project as 4 or 5, with the others rating it 3. All but one said they would recommend the REU program to their peers, or already had, and that one said "maybe." The following comment from one of the participants captures what is unique about the DIMACS REU program:

"I think that any REU is a good experience, but the DIMACS one is better. Most students who go to an REU get to work with a few mathematicians, and the other students. But between the probabilistic methods conference at the beginning of the summer, and DREI at the end of the summer, I got to meet and talk to a lot of mathematicians and graduate students. This helped me a lot in figuring out what kind of math I like and where I might want to go to graduate school. Plus there were a lot of lectures that I was able to go to, and a lot of math to learn; I wouldn't have had this at a different REU. Also, since my project wasn't straight graph theory, but computational biology, it showed me areas of math I did even know existed, but enjoy now."

4.3 Postdoctoral Training Program

One of DIMACS' goals is to involve outstanding young people in the life of the center. Our postdoc program is designed to accomplish this. We have leveraged our funds for support of postdocs by sharing postdocs with the Institute for Advanced Study and with the University of Pennsylvania, obtaining grant support for postdocs, and including in our postdoc program postdocs appointed by our industrial partners. AT&T Labs has given us $100,000 for support of joint postdocs, another way we are now beginning to leverage our postdoc funds, and we have agreed with NEC Research Institute to share a postdoc for the academic year 1999-00. We have instituted regular meetings with our postdocs to discuss the job market. Continuing procedures developed in previous years, each postdoc is assigned a faculty mentor. Moreover, because we want to train our postdocs to take an interest in educational activities and in outreach beyond the mathematical sciences, we ask each postdoc to take part in such activities. Our postdocs meet this requirement in a variety of ways, such as helping with our K-12 and "Reconnect" programs or mentoring in our REU program. Our postdoctoral fellows from past years have been remarkably successful in obtaining positions, both in academics and in industry. They have also rapidly risen to the top of their profession.

DIMACS has pursued a program of evaluation of the longer-term impacts of its special years by asking committees of experts to evaluate these impacts. A committee report evaluating the 1994-99 Special Year in Mathematical Support for Molecular Biology, authored by Daniel Gusfield, Sampath Kannan, and Pavel Pevzner, was received this year. The report talks about the importance of the postdoctoral training program, calling this "the foremost accomplishment" of the special year and saying that "perhaps the most important function of DIMACS has been in training young computer scientists in computational biology so that they go on to establish collaborations later in their career." See Section 3.4 for further discussion of this report.

During the period covered by this report, we have had two "classes" of postdoctoral fellows, the ones here during 1997-98 and the ones here during 1998-99. A list of these fellows and their present affiliations follows:

1997-98 Postdocs

- Daniel Bleichenbacher
  - Bell Labs DIMACS Postdoc
  - Research Interests: security of cryptosystems based on number theory, efficient algorithms for cryptosystems, applications of cryptography in general
  - Current Position: continuing Bell Labs DIMACS Postdoc
- Michael Goldwasser
  - Massive Data Sets Postdoc, Princeton
  - Research Interests: design and analysis of algorithms, with emphasis on approximations, online computations and randomized algorithms; additional emphasis on computational geometry, with applications in robotics and assembly planning, and on Web search and clustering
  - Current Position: continuing Massive Data Sets Postdoc, Princeton
- Mehdi Hoseyni
  - DIMACS-WINLAB Postdoc on Simulations Project, Rutgers
  - Research Interests: applications of operations research methods in telecommunications services; analysis, modeling and performance evaluations of telecommunications networks; analysis and optimization of stochastic systems
  - Current Position: DIMACS-WINLAB Postdoc on Simulations Project, Rutgers
- Daniel Huson
  - Molecular Biology Postdoc, Princeton
  - Research Interests: combinatorial tiling theory and sequence analysis
  - Current Position: Continuing Postdoc, Princeton
- Haim Kaplan
  - AT&T Labs DIMACS Postdoc
  - Research Interests: Data structures, algorithmic problems, design and analysis of algorithms, combinatorial optimization, linear and mathematical programming, complexity theory
  - Current Position: Continuing AT&T Labs DIMACS Postdoc
- Owen Kelly
  - DIMACS-WINLAB postdoc on Simulations Project, Rutgers

- Research Interests: statistical signal processing and applications of nonparametric models
  - Current Position: Continuing DIMACS-WINLAB Postdoc on Simulations Project

- Michael Krivelevich
  - DIMACS - Institute for Advanced Study Postdoc at IAS
  - Research Interests: probabilistic combinatorics, hypergraphs, Ramsey theory, extremal graph theory
  - Current Position: Continuing DIMACS-IAS Postdoc at Rutgers

- Anuj Puri
  - Bell Labs DIMACS Postdoc
  - Research Interests: computer aided verification, real-time systems, hybrid systems, computer networks, transportation systems, optimization
  - Current Position: Continuing Bell Labs DIMACS Postdoc

- Rajmohan Rajaraman
  - Networks Postdoc, Rutgers
  - Research Interests: combinatorics, distributed network algorithms, on-line algorithms, parallel models of computation, randomness
  - Current Position: Assistant Professor of Computer Science, Northeastern University

- Mark Smith
  - AT&T Labs DIMACS Postdoc
  - Research Interests: distributed algorithms, network protocols and formal verification techniques, the design of network protocols
  - Current Position: Continuing AT&T Labs DIMACS Postdoc

- Elizabeth Sweedyk
  - DIMACS - University of Pennsylvania Postdoc in Molecular Biology, Rutgers and U. Penn.
  - Research Interests: complexity theory and algorithm design in computational biology, phylogenetic tree construction and sequence assembly
  - Current Position: DIMACS Molecular Biology Postdoc, Rutgers

- Geza Toth
  - DIMACS - Institute for Advanced Study Postdoc, Rutgers
  - Research Interests: Euclidean Ramsey theory, geometric graphs, unit distance graphs, art gallery theorems
  - Current Position: Assistant Professor, MIT

- Shibu Yooseph
  - Molecular Biology Postdoc, Rutgers

- Research Interests: design and analysis of algorithms, applying techniques in algorithm design to solve problems in computational biology, problems arising in phylogeny construction and consensus methods
- Current Position: Postdoctoral Fellow, University of Southern California

- Shiyu Zhou
  - Bell Labs DIMACS Postdoc
  - Research Interests: design and analysis of data structures and algorithms, computational complexity
  - Current Position: Assistant Professor of Computer Science, University of Pennsylvania

1998-99 Postdocs

- Daniel Bleichenbacher
  - Bell Labs DIMACS Postdoc
  - Research Interests: security of cryptosystems based on number theory, efficient algorithms for cryptosystems, applications of cryptography in general
- Gruia Calinescu
  - Large Scale Discrete Optimization Postdoc, Rutgers
  - Research Interests: algorithms, approximation algorithms
- Michael Goldwasser
  - Massive Data Sets Postdoc, Princeton
  - Research Interests: design and analysis of algorithms, with emphasis on approximations, online computations and randomized algorithms; additional emphasis on computational geometry, with applications in robotics and assembly planning, and on Web search and clustering
- Mehdi Hoseyni
  - DIMACS-WINLAB Postdoc on Simulations Project, Rutgers
  - Research Interests: applications of operations research methods in telecommunications services; analysis, modeling and performance evaluations of telecommunications networks; analysis and optimization of stochastic systems
- Daniel Huson
  - Molecular Biology Postdoc, Princeton
  - Research Interests: combinatorial tiling theory and sequence analysis
- Haim Kaplan
  - AT&T Labs DIMACS Postdoc
  - Research Interests: Data structures, algorithmic problems, design and analysis of algorithms, combinatorial optimization, linear and mathematical programming, complexity theory

- Owen Kelly
  - DIMACS-WINLAB postdoc on Simulations Project, Rutgers
  - Research Interests: statistical signal processing and applications of nonparametric models
- Michael Krivelevich
  - DIMACS - Institute for Advanced Study Postdoc at Rutgers
  - Research Interests: probabilistic combinatorics, hypergraphs, Ramsey theory, extremal graph theory
- Anuj Puri
  - Bell Labs DIMACS Postdoc
  - Research Interests: computer aided verification, real-time systems, hybrid systems, computer networks, transportation systems, optimization
- Alex Samorodnitsky
  - DIMACS - Institute for Advanced Study Postdoc, Rutgers
  - Research Interests: combinatorics, especially extremal set theory and coding theory
- Mark Smith
  - AT&T Labs DIMACS Postdoc
  - Research Interests: distributed algorithms, network protocols and formal verification techniques, the design of network protocols
- Elizabeth Sweedyk
  - Molecular Biology Postdoc, Rutgers
  - Research Interests: complexity theory and algorithm design in computational biology, phylogenetic tree construction and sequence assembly
- Luca Trevisan
  - Large Scale Discrete Optimization Postdoc, Rutgers
  - Research Interests: algorithms and complexity theory, especially approximation algorithms for combinatorial optimization problems; hardness of approximation; computational randomness
- Kasturi Varadarajan
  - Large Scale Discrete Optimization Postdoc, Rutgers
  - Research Interests: algorithms, data structures, computational and combinatorial geometry
- Francis Zane
  - Bell Labs DIMACS Postdoc
  - Research Interests: complexity theory, circuit complexity, satisfiability algorithms, combinatorics

- Leonid Zosin
  - Networks Postdoc, Rutgers
  - Research Interests: approximation algorithms, combinatorial optimization, distributed computation, cryptography

The postdocs were asked to evaluate their experiences in DIMACS programs. Here are some selected comments from the postdocs:

"I was able to contact and collaborate with other visitors and permanent members of DIMACS and to participate in various activities in the area. This in my opinion has had a significant impact on my research and has resulted in several papers co-authored by other visitors to DIMACS/IAS." – DIMACS/IAS postdoc

"My focus has become far more practical." – Molecular Biology SY postdoc

"I think the DIMACS program has given me a very good opportunity to reshape my research directions. Specifically, I came here immediately after graduate school with the idea of doing research in new problems, entirely unrelated to my thesis. Eventually, I will certainly do more work related to my thesis, but I wanted to make sure that I also had other areas of computer science to pursue, and this has been a major benefit from my position at DIMACS." – Massive Data Sets SY postdoc

"The official DIMACS programs, such as the special year on massive data sets, and various workshops have been a great way to be introduced to new directions of research, as well as ways of coming into contact with other researchers." – Massive Data Sets SY postdoc

"My mentor was ... very helpful." – DIMACS/IAS postdoc

4.4 Graduate Student Program

Graduate students are a central part of a science and technology center. Our local graduate student program consists of small DIMACS fellowships to support the research of local students, small awards for summer research projects, and jobs for which graduate students are paid. The response to this program has been highly enthusiastic. We supported 27 students this way this year, and this support has already led to rather impressive scientific payback. We have also continued our programs of sponsoring DIMACS visits by non-local graduate students and continued our long-standing policy of allocating supplemental funding to each workshop for support of non-local graduate students. Scientific results of our graduate students are reported on in our discussion of research thrusts in Section 3.

Graduate Students Supported Feb. 1998 - Jan. 1999

- Bernardo Abrego
  - Affiliation: RU, Mathematics
  - Research Thrust: Discrete and Computational Geometry
  - Current Position: GS, RU
- Di Bai

- Affiliation: RU, OR
- Research Thrust: Networks, Large Scale Discrete Optimization
- Current Position: GS, RU

- Amit Chakrabarti
  - Affiliation: PU, CS
  - Research Thrust: Networks
  - Current Position: GS, PU
- Bela Csaba
  - Affiliation: RU, CS
  - Research Thrust: "General": tcs
  - Current Position: GS, RU
- Rick Desper
  - Affiliation: RU, CS
  - Research Thrust: Molecular Biology
  - Current Position: GS, RU
- Paul Dreyer
  - Affiliation: RU, Math
  - Research Thrust: Molecular Biology
  - Current Position: GS, RU
- Silvia Fernandez
  - Affiliation: RU, Math
  - Research Thrust: Discrete and Computational Geometry
  - Current Position: GS, RU
- Xin Guo
  - Affiliation: RU, Math
  - Research Thrust: "Exploratory": Financial Modeling
  - Current Position: GS, RU
- Tina Heikkenin
  - Affiliation: RU, OR
  - Research Thrust: Simulations
  - Current Position: GS, RU
- Jonathan Hope
  - Affiliation: RU, CS
  - Research Thrust: Simulations

- Current Position: GS, RU

- John Iacono
  - Affiliation: RU, CS
  - Research Thrust: Logic and Algorithms
  - Current Position: GS, RU
- Kamal Jain
  - Affiliation: Georgia Tech.
  - Research Thrust: Large Scale Discrete Optimization
  - Current Position: GS, Georgia Tech.
- Jie Lai
  - Affiliation: RU, WINLAB
  - Research Thrust: Simulations
  - Current Position: GS, RU
- Vincenzo Liberatore
  - Affiliation: RU, CS
  - Research Thrust: Networks, Massive Data Sets
  - Current Position: Assistant Professor, University of Maryland
- Daniel Lloyd
  - Affiliation: RU, CS
  - Research Thrust: "General": tcs
  - Current Position: GS, RU
- Sachin Lodha
  - Affiliation: RU, CS
  - Research Thrust: "General": tcs
  - Current Position: GS, RU
- Sofus Macskassy
  - Affiliation: RU, CS
  - Research Thrust: Simulations, Discrete and Computational Geometry
  - Current Position: GS, RU
- Unsal Ozdogru
  - Affiliation: RU, OR
  - Research Thrust: Large Scale Discrete Optimization
  - Current Position: GS, RU
- Ozgur Ozluk

- Affiliation: University of North Carolina, OR
- Research Thrust: Combinatorial Optimization
- Current Position: GS, University of North Carolina

- Jignesh Panchal
  - Affiliation: RU, WINLAB
  - Research Thrust: Simulations
  - Current Position: Lucent Technologies
- Christian Pedersen
  - Affiliation: BRICS (University of Aarhus)
  - Research Thrust: Molecular Biology
  - Current Postion: GS, Aarhus
- Gregory Perkins
  - Affiliation: RU, CS
  - Research Thrust: Simulations
  - Current Position: GS, RU
- Li Sheng
  - Affiliation: RU, OR
  - Research Thrust: Molecular Biology
  - Current Position: Assistant Professor, Drexel University
- Lijie Shi
  - Affiliation: RU, OR
  - Research Thrust: Massive Data Sets
  - Current Position: GS, RU
- Yueyue Song
  - Affiliation: RU, CS
  - Research Thrust: Networks
  - Current Position: GS, RU
- Balakrishna Subramanian
  - Affiliation: RU, Physics
  - Research Thrust: Discrete Probability (Statistical Mechanics)
  - Current Position: GS, RU
- Victoria Ungureanu
  - Affiliation: RU, CS
  - Research Thrust: Networks

  - Current Position: GS, RU

- Tonguc Unluyurt
  - Affiliation: RU, OR
  - Research Thrust: Massive Data Sets, Combinatorial Optimization
  - Current Position: GS, RU

- Song Wang
  - Affiliation: RU, WINLAB
  - Research Thrust: Simulations
  - Current Position: GS, RU

- Xuerong Yong
  - Affiliation: Hong Kong University of Science and Technology
  - Research Thrust: Massive Data Sets
  - Current Position: GS, Hong Kong University of Science and Technology

- Yang Yu
  - Affiliation: RU, Math
  - Research Thrust: "General": Combinatorics
  - Current Position: GS, RU

- Yuan Ping Zhang
  - Affiliation: Hong Kong University of Science and Technology
  - Research Thrust: Massive Data Sets
  - Current Position: GS, Hong Kong University of Science and Technology

The graduate students were asked to evaluate their experiences in DIMACS graduate student programs. All rated it 4 or 5 on a scale of 1 (low) to 5 (high), and 86% rated it 4 or 5 as far as level of influence/help with their research. Here are some selected comments from the graduate students:

"I met my Ph.D. advisor mostly through DIMACS." - Rutgers Math student.

"I have also gotten two jobs thus far with people who have spent time at DIMACS. ... I would say that DIMACS was invaluable in this regard." – Rutgers Math student.

"I accomplished a summer project sponsored by DIMACS by myself. I gained a lot of experience both theoretical and practical." – Rutgers WINLAB student.

"The Special Year in Biology definitely influenced my choice to work on math/bioinformatics problems." – Rutgers Math student.

"DIMACS experience has added a new research direction ... to my current research." – Rutgers OR student

4.5 Involvement of Underrepresented Groups

A large percentage of the participants in center programs are US citizens. Of the 1912 visitors, workshop, and program participants at DIMACS over the course of the year, 1240 were from the US, 406 were foreign, and 266 were of unidentifiable nationality. (Unidentifiable nationality (and gender) are primarily due to the lack of detailed data from off-site events.) US citizens were also involved in reasonable numbers in our training programs involving undergraduates, graduate students, and postdocs. Of the total of 64 individuals involved in these programs, 23 were US citizens.

The center has a long history of participation by women. Almost every leading woman researcher in discrete mathematics/theoretical computer science has been a visitor at DIMACS and many have taken leadership roles in the center's programs. Our External Advisory Board has four leading women researchers on it. Of the 1912 visitors, workshop, and program participants at DIMACS over the course of the year, 319 were women, 1327 were men, and 266 were of unidentifiable gender. In the undergraduate, graduate, and postdoc training programs mentioned above, a total of 14 of the 64 participants were women. We also have formalized a partnership with the Douglass College Program for Women in Mathematics, Science, and Engineering by mentoring Douglass College undergraduates and offering advice about undergraduate research. Fred Roberts was formally invited to join the Program's Advisory Board. Douglass College is the women's undergraduate college at Rutgers University. As the first formal arrangement in this partnership, we sponsored a Douglass College woman in our summer REU program. This was a great success, and we plan to continue doing this. Two of the women students supported by DIMACS have won major awards. Carly Klivans, a Cornell University student participant in our summer REU program, won the excellence in undergraduate research award from the Association of Women in Mathematics. Victoria Ungureanu, a Rutgers graduate student involved in the Networks Special Year, won a second-place award from IBM in their competition for best thesis proposal in E-Commerce.

DIMACS is currently in the planning process for a two day workshop, "Showcasing Women in Theoretical Computer Science," to be held during spring 1999. Theoretical computer science is now an area in which relatively many extremely distinguished senior women are active. Our goal is to "showcase" the accomplishments of this group in a strikingly visible fashion to the larger scientific community, both men and women, and while so doing provide an opportunity for women graduate students, postdocs, and junior researchers to become better acquainted with a segment of the theoretical computer science community likely to offer compelling role models and mentors. The workshop will consist of research talks by senior women in theoretical computer science ranging across a broad group of topics; the talks will include more background and tutorial material than is typical for research talks in order to make the workshop interesting and accessible to a scientifically diverse audience. In addition we plan two panel discussions, one on linkages between theory and other areas of computer science, and one on career issues in theory. Both should be of interest to both men and women audience members. However, since all panelists will be women in theoretical computer science, we expect the careers panel in particular to offer useful guidance to some of the more junior women in attendance.

The center continues to make major efforts in the direction of involvement by African Americans and Hispanics. We have had a long history of success in involving African Americans in our Young Scholars Program. We hosted a visitor from the University of Puerto Rico and opened discussions about joint programs. We continue to discuss joint programs with Spelman College and Howard University. We maintain contact with Clark Atlanta University through Fred Roberts' service on the External Advisory Board for the Minority Excellence Center for Theoretical Studies in the Physical Sciences at Clark Atlanta. We offered a postdoctoral fellowship to a Hispanic woman, but she was unable to accept it. Among the leaders of DIMACS programs are a number of leading African American and Hispanic researchers. This includes the head of our postdoctoral fellow committee, a chief organizer of our Molecular Biology special year, a key figure in our Reconnect program, and a number of workshop organizers.

4.6 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 ask special year organizers to plan on educational components for their years if possible) and the direct participation of special year researchers in educational outreach activities. Our Modules series is aimed at bringing current research into the undergraduate and precollege classroom. Our "Reconnect" project has the same goal at the two- and four-year college levels. Both of these programs are described in this section.

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 undergraduates program and in our Young Scholars Program for high school students, by giving lectures or teaching courses in our Leadership Program for precollege teachers, by supervising undergraduate theses, and by serving as mentors in our Reconnect program. Participants in these programs have also included some of the most senior DIMACS leaders from all six institutions. A small portion of this involvement is described in our discussion of the Leadership, Young Scholars, and REU programs in Section 4.1 and Section 4.2.2.

In the summer of 1996, we inaugurated a four year DIMACS Research and Education Institute (DREI), under a 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 also described in this section.

4.6.1 DIMACS Research and Education Institute (DREI)

The concept behind the DIMACS Research and Education Institute is that research and education should work hand-in-hand. Collaborations between researchers and educators are formed by understanding each other's work and needs - to this end the institute is designed around separate programs that are focused on each group's interests, along with plenary sessions and evening lectures aimed at meshing the two groups with the planned outcome of stimulating discussion and problem-solving.

The DIMACS Research and Education Institute (DREI '98) was held at the DIMACS Center on the Rutgers University campus (July 20 - August 7, 1998). This year's topics were Graph Theory and Combinatorial Optimization. There was an overall total of 194 people registered for this program. Program attendees included 25 high school teachers; 30 resident researchers for week one; 45 resident researchers for week two; 53 resident researchers for week three; and 55 local participants (which included undergraduate students, faculty and researchers) who attended one or more days of the program.

Linda Lesniak (PM, Drew University) was both overall program director and director of the research program. Glenn Hurlbert (Arizona State University) was the education program director. Henry Kepner (University of Wisconsin - Milwaukee) was education program consultant. L.

Charles Biehl (teacher, The Charter School of Wilmington), Ethel Breuche (consultant, Pennsylvania), and Patrick Carney (teacher, The Montclair Kimberley Academy) were the lead teachers for the education program who assisted Hurlbert in running homework sessions, professional plan development, teacher implementation sessions and overall program planning. Additional classroom presentations were given by Ruth Haas (Smith College), Garth Isaak (Lehigh University) and Nancy Eaton (University of Rhode Island).

The core of the education program was a series of eight weekly lectures given by Glenn Hurlbert. Every teacher who was accepted into the program was also included in the credit program in which we introduced the new concept of teacher professional plans of activities to be implemented during the subsequent academic year with status updates/presentations at the follow up sessions. Six of the teachers made presentations at our first follow up on November 7. Another innovation was the idea of teacher presentations on a topic based upon the classroom lecture materials/plenary sessions that were given to the entire group during the summer. Also, a library of books and materials was set up for the teachers and researchers to use as reference materials and for developing professional plans and presentations. As in previous years, plenary speakers spent time with the teachers after their presentations to answer questions, etc. Participants were introduced to the LINK software that had been developed at DIMACS and were given extensive time in the DIMACS PC labs to use LINK to experiment with graph-theoretical ideas and try it out on homework problems. They were assisted by two DIMACS REU students. Presentations were also given on Tom Sawyer graph drawing software and web-based graph drawing software which demonstrate how to visualize graph drawing and use available technology. This software was also helpful in doing some of the homework problems.

The lead teachers' responsibilities increased compared to last year. They participated in program and schedule planning from the very first meeting. Lead teachers were able to do one "implementation" lesson for the teachers each week during the three week period. The participants' reaction to these implementations was undeniably enthusiastic with a desire to have seen more. The lead teachers shared a plenary session for teachers and researchers in which each of them presented a mini implementation lesson. They also participated in one pedagogy plenary session. Each week they had an opportunity to work with teachers to help them devise their individual "Professional Plan."

The fact that the research program each week corresponded similarly to the topics of study in the education program helped overwhelmingly in terms of integration. The teachers were encouraged to ask the researchers for help with their homework problems or with their professional plans. The researchers were encouraged to discuss with the teachers matters of pedagogy. Both types of discussions occurred during the day at breakfast, lunch and breaks and continued in the evening at the hotel at dinner and after dinner. The "research scavenger hunts" were another vehicle for integration. These were intended, among other things, to introduce teachers to the activity of mathematical research, using books and journals in the library, MathSciNet on the computer, and research participants as resources, and to introduce participants to each other. The scavenger hunts were extremely successful, generating positive feedback from both teachers and researchers. Each week, a group of teachers was asked to do some "research" on an extended topic in graph theory. They had to do a presentation towards the end of the week for both the researchers and educators. Teachers used the researchers as resources. The researchers were impressed and delighted to see these presentations. Many of the afternoon sessions involved learning to use the LINK software.

The research program consisted of three week-long research conferences. The topic of week one was paths and cycles, chaired by Ron Gould (Emory University) and Ralph Faudree (University of Memphis); the topic of week two was topological graph theory, chaired by Robin Thomas (Georgia Institute of Technology) and Paul Seymour (PM, PU); the topic for week three was graph coloring, also chaired by Paul Seymour and Robin Thomas. High quality formal research presentations were given and throughout the conference we noted a high level of research activity (work sessions, small group sessions, presentations) taking place.

The research program led to numerous new research collaborations, the solution of several conjectures, and to a number of papers that have already been written. Here is just a small selection of these types of outcomes. Luis Goddyn solved a conjecture regarding circular chromatic numbers and forbidden minors. Thanks to discussions with Dan Archdeacon and Henry Glover, Petr Hlineny made great progress towards generalizations of Negami's planar covering conjecture and has already submitted a paper on this subject. Dominic Welsh collaborated with Paul Seymour, Dirk Vertigan, Robin Thomas, and Dan Archdeacon to solve a problem on knots that he had been working on for a long time. DREI also led to a collaboration on generalized vector space formulations of graph coloring between Tommy Jensen and Enrique Moreno at Cinvestav in Mexico City and to one among Jensen, Reinhard Diestel, Bjarne Toft and Alexander Kelmans from University of Puerto Rico.

Something that should be noted, as the organizers heard this repeated many times, was that many of the researchers came away with an increased respect for the individuals and the genuine professionalism/enthusiasm of the high school teachers, and they valued their own experience at DREI for its uniqueness.

The DREI program was evaluated by professional evaluator Dr. Norman Webb. DREI has been running for three years now and each year we have made changes and improvements. DREI'98 was clearly the best DREI yet, and the evaluation shows it. We cite the following excerpts from Dr. Webb's report. Participants indicated on both the Exit Questionnaire and in interviews a high level of satisfaction with the Institute. Over 90% of the responding participants (95% of the high school teachers, 100% of the researchers, and 92% of the graduate students) reported that their overall satisfaction with the Institute was high to extremely high. As Dr. Webb's report says, "The level of interaction among those from the different groups was significant." Over 86% of the high school teachers reported they had at least daily interactions with researchers. Sixty three percent of the researchers reported that they had at least daily interactions with high school teachers. Eighty-five percent of the researchers reported their interactions with high school teachers at the Institute were moderate to very valuable. Ninety-five percent of the high school teachers reported their interactions with researchers were moderate to very valuable. Interactions between teachers and researchers were much more significant than simply exchanging pleasantries or ideas about education. All but one of the Education Program participants who responded to the questionnaire reported they had collaborated with a researcher on a project, problem, or work more than once during the three weeks. Two-thirds of the respondents reported they had planned future joint work efforts with a researcher at least once during the Institute. Researchers confirmed what teachers reported. Sixty six percent of the researchers who responded to the questionnaire reported they had exchanged information on mathematics with someone from the Education Program two or more times a week. Researchers also reported their conversations with high school teachers went further than just the exchange of information. Over half of the Research Program respondents reported that they had collaborated on a project, problem, or work with a high school teacher at least once. One-third reported they had planned future joint work efforts with a high school teacher. We think this is dramatic evidence of true integration.

The following are some comments from the participants.

"I have attended several conferences that attempted to combine both a mathematics and education theme, none of which came close to the kind of interaction that took place between researchers and educators at DREI '98." I am taking back home with me "a notebook half filled with research projects and ideas at ALL levels. These include not just professional-level problems but at least half a dozen projects suitable for undergraduate student research ... and graduate student research." –

researcher from Department of Mathematics & Computer Science, Shippensburg University

"I've frequently told my colleagues how much I enjoyed the focus of the DREI program. Our department is currently discussing how to more closely interact with high schools in different ways. ... While the DREI program was not directly responsible for all of the above initiatives, it did stimulate several hours of discussion between me and my department chair about how to promote (what you call) vertical integration." – researcher from Department of Mathematics, University of Vermont.

"I benefitted pedagogically from the contact with the teachers. For example I now use the approach to Steiner trees that I learned at the meeting from L. Charles (Chuck) Biehl [one of the lead teachers] in my own teaching here." – researcher from University of Odense.

"The DREI '98 program was extremely valuable to my professional growth." – high school teacher

"I found the DREI '98 program extremely valuable to me as an educator. I've incorporated much of what I learned this summer into my classroom lessons." – high school teacher

"The researchers were terrific. I talked to several of them and they added tremendous insights into the happenings of current graph theory. They also seemed curious as to how I was using graph theory in my classes." – high school teacher.

"The researchers who attended the DREI'98 were among the most friendly and outgoing bunch of college professors I have ever met. I was never made to feel inferior. They were always willing to listen to my ideas. Many of them (who teach DM at the college level) were extremely interested in knowing what books and materials I used in my high school course. Their interest was genuine." – high school teacher

Plans are now underway for DREI'99. The theme is "Graph Theory and its Applications to Problems of Society." There will be three research workshops on both pure and applied graph theory, with the teacher program and the joint plenary sessions emphasizing the applications. The research workshops will be as follows: Matchings in Graphs and their Applications, organized by Brian Alspach; Concepts of Domination in Graphs and Directed Graphs, organized by K. Brooks Reid; and Graphs, Posets, and Algorithms, organized by Tom Trotter. Linda Lesniak continues as the overall program director and director of the research program and Barry Tesman will be the education program director.

4.6.2 The Reconnect Program

In 1997, DIMACS initiated a project that seeks to "reconnect" to the mathematical sciences research enterprise two and four year college faculty who haven't had the time to keep up with research developments. It aims to expose them to current research topics in discrete mathematics and theoretical computer science that are relevant to their teaching by involving them in a research center where much of the relevant research is being carried out. We aim to enhance the ability of teachers to transform their classrooms into places that connect up to modern uses of mathematics and computer science and to help them produce classroom materials that reflect current research. We are doing this through summer conferences and a sequence of two-day and one-day conferences. The summer conferences, national in scope and directed at two- and four-year college faculty with some prior exposure to discrete mathematics and theoretical computer science, present recent research results in topics such as computational molecular biology, network visualization, clustering, and visibility in geometry, and divide the participants into writing groups. The two-day conferences, regional in scope and directed at two year college faculty with little prior exposure to discrete mathematics and theoretical computer science, present an introduction to these fields, with connection to topics of current research interest, and also involve the participants in writing materials in a six month period between two two-day conferences. Participants in both the two-day and two-week conferences are expected to use the materials they have written as vehicles to bring discrete mathematics and theoretical computer science into their classrooms. Participants are encouraged to make their materials available to a broader audience through developing and publishing them in the DIMACS Module Series. Participants are also encouraged to become involved in other DIMACS activities, such as workshops and the DIMACS Research and Education Institute (DREI).

Reconnect Summer Program

The Reconnect '98 2-week conference in July 1998, organized by Fred Roberts (PM, RU) and Rochelle Leibowitz (Wheaton), with the advice of Reggie Luke (Middlesex County College) and Theresa Michnowitz (Jersey City State College), had 16 participants from across the United States. The two topics for this year's conference were computational biology and network visualization. Both topics have seen rapid growth in research results in the last few years, both topics have many real-world applications, and both topics can be brought into the undergraduate math and computer science classroom. The first week's principal speakers were Martin Farach-Colton (PM, RU) and Ramamoorthi Ravi (Carnegie-Mellon) in computational biology, and Nathaniel Dean (PM, Rice) and Jonathan Berry (Elon) in network visualization. These speakers were responsible for presenting the core material in their respective areas. In addition to the formal presentations by the speakers, we encouraged our participants to be active rather than passive learners by scheduling time for homework review, computer labs, library reading, question-and-answer periods with the speakers, and informal get-togethers with speakers and other researchers at DIMACS not directly connected to the conference. In particular, Berry introduced the participants to the LINK software developed at DIMACS, and they had an opportunity to experiment with it in the DIMACS PC labs, with the assistance and advice of several of our REU students. Another goal of the conference was for participants to take the material they learned and write undergraduate classroom modules so that not only would they be able to expose their college students to exciting new research in discrete math and theoretical computer science, but that the participants could expose other college faculty to these exciting areas, these college faculty in turn would bring it into their own classrooms, thus broadening the impact of the conference. Participants were placed in writing groups in the middle of the first week and were told to start thinking about a module topic. The speakers were helpful in suggesting topics and offering advice to these groups.

A different model was used during the second week of the conference. More time was allowed for writing and instead of having four principal speakers, we invited industrial and academic researchers to give talks on current hot research topics in computational biology and network visualization. Some of the speakers were Roberto Tamassia (Brown) and Stephen North (PM, ATTL) in network visualization, and Dannie Durand (V, PU) and Sridhar Hannenhalli (SmithKline Beecham Pharmaceuticals) in computational biology. In this way, the participants were exposed to a variety of current research problems. Again, time was scheduled for question-and-answer sessions and informal get-togethers with the visiting speakers. The visiting speakers also stayed around to help give advice and be the "experts" for the writing groups. Writing group presentations were given during the last two days of the conference. The conference ended with a charge to each group to polish its module and submit it to the DIMACS Educational Module Series for possible publication. Stephen Maurer (PM, Swarthmore), the editor of the Series, and the associate editors are working with the writing groups to help produce modules that will reach other undergraduate math and computer science faculty and ultimately, more undergraduate math and computer science classrooms.

The two-week conference was evaluated by consultants Regina C. Gray and Dr. Linda Tappin. The participants were generally very pleased with the organization of the conference, the topics covered and its relevance to the undergraduate classroom, and their exposure to the world of mathematical research. The participants suggested that we provide notes ahead of time and we will endeavor to do so in time for next year's conference. To be more specific, all of the participants responding rated the quality 4 or 5 on a 5-point scale, with 5 being excellent; 67% rated the quality excellent. As to the applicability of the materials presented to their classrooms, 57% of the respondents rated the applicability as 4 or 5 on a 5-point scale with 5 meaning very applicable, and 93% rated it at least 3. One hundred percent of those responding rated the "reconnecting" aspect of the program as good or excellent, i.e., 4 or 5 on a 5-point scale, with 85% rating this as excellent.

Here is a selection of comments from the participants:

"Reconnect 98 is the best conference of this type I've ever attended or been involved in. The most important factor is the high quality of the research results presented. The second factor is the good presentations and references. The surprising factor is the quality of the participants. Their competency as mathematicians, their ability to write and communicate. The organizers and staff are highly dedicated people who run the program efficiently and professionally."

"From start to finish, I felt welcomed, informed, and part of the DIMACS community."

"I benefited tremendously in knowledge, in contacts, in amount of resources that I can bring back to my students."

"I would like to thank [DIMACS] for this great idea."

Several immediate outcomes of this conference are: one participant and a co-organizer of the 2-week conference are planning a contributed papers session "Computational Biology in the Undergraduate Discrete Mathematics Curriculum" for MathFest '99; one participant working with a biologist is running a computational biology seminar at her institution; several participants have introduced material that they learned this summer into their classrooms. Other participants have indicated that further impact will be felt as they have more time to assimilate the material into their classroom.

Plans are well underway for Reconnect '99. The principal speakers for the two-week conference will be Pierre Hansen (Univ. of Montreal) in clustering and Joseph O'Rourke (Smith) in visibility in geometry.

Reconnect Two-Day and One-Day Workshops for Two-Year College Faculty

These workshops are geared towards introducing two-year college faculty to discrete mathematics and to some current research topics relevant to the classroom.

The 1998 workshops were organized by Jean Lane (Union County College), Fred S. Roberts (PM, RU) and Joseph Rosenstein (PM, RU). An advisory board with local and national members offered suggestions on content and publicity. The workshops were publicized through the American Mathematics Association of Two Year Colleges (AMATYC), the Mathematics Association of Two Year Colleges of New Jersey (MATYC-NJ), the Mathematical Association of America (MAA), as well as on the DIMACS home page on the internet and through mailings to Mathematics Department chairs in New Jersey. Applications were screened by Lane and local advisory board members Charles Miller (Camden County College) and George Popel (Middlesex County College).

Nineteen participants representing seven states met on May 18-19 at the DIMACS Center, Rutgers University. Presentations were made by Joseph Rosenstein (PM, RU), L. Charles Biehl (The Charter School of Wilmington, Delaware), Elizabeth Sweedyk (PD) and Stephen Maurer (PM, Swarthmore College). By the end of the sessions, participants had formed six groups to begin work on writing classroom materials. Many met again in June for dinner and in September for lunch; at both meetings, each group gave a progress report on its project.

All participants convened again at Rutgers University on November 14, 1998. Dr. Mordecai Golin (V) of the Hong Kong University of Science and Technology was the keynote speaker, speaking on Huffman codes, and five DIMACS members and postdocs attended to act as mentors for the participants. Each writing group made a presentation on its project, to be critiqued by the other participants. The DIMACS mentors will continue to advise the writing groups via e-mail and telephone consultation. A follow-up "wrap-up" meeting is planned for mid-January for presentations of the final versions of the classroom materials.

The program was evaluated by a series of questionnaires. 95% of the participants felt that the material was applicable or very applicable to their classrooms. 90% felt that the subject matter was about right. About half of the participants would have preferred more in-depth coverage and we will make this adjustment when we repeat the program. 95% of the respondents felt that the quality of the program was good or excellent. The participants also asked for more computational biology and more game theory.

Plans are already well underway to repeat this program starting in the Spring of 1999. In our first year, we emphasized New Jersey participants. In the coming year, we will seek to broaden the participation beyond New Jersey.

4.6.3 Modules Series

DIMACS has commenced a module series: the DIMACS Educational Modules in Discrete Mathematics and Theoretical Computer Science.

This series is part of the DIMACS effort to enrich undergraduate education, community college education, and precollege education and to integrate research and education. The modules will be self-contained material that can be used for one or more class meetings, or for independent study or outside reports.

DIMACS modules will emphasize topics that are not yet readily available in text material at the module's educational level. Almost every module will introduce some aspect of a topic undergoing active research development. For instance, a number of DIMACS modules will be developed in conjunction with the many workshops, tutorials, and other outreach activities at the Center. However, anyone can author a DIMACS Module.

DIMACS Modules will aim at students at various levels. Some modules will be aimed at the junior/senior undergraduate level, for students majoring in math, CS, or other technical fields. Some will be for technically oriented first year students. Some will be for liberal arts students and will provide a taste of what mathematics and computer science are like today. Some will be for high school students. Some topics will appear in both high-end and low-end modules.

Whatever the level, modules will present their topic carefully and develop some aspect of the topic in detail. At all levels, problem sets will reinforce and extend ideas and methods developed in the body of the module. Also, modules will provide pointers to full results and open questions.

DIMACS modules will take advantage of advances in technology. They will be available electronically, as well as on paper, and some electronic modules may be interactive.

Each module will be a refereed publication. But whereas journals are often forced to look for reasons to reject submissions, the goal of this series is to make submissions publishable, since our goal is to successfully inject as much current material into the "education stream" as possible. In addition to editorial help, help will be available to format submissions in a visually attractive way and to convert figures into an electronic form we can use. Modules may be bundled into DIMACS volumes. In preliminary discussions, the AMS has expressed interest in publishing them. AMS is interested in developing publications at the undergraduate level in its own effort to increase attention to educational issues. In addition, AMS wishes to experiment with custom publishing - mix and match bundling on request. The DIMACS Educational Module Series seems to present just the right context for these new ventures.

An Editorial Board of about 30 people has been chosen. It includes noted researchers and people with educational and writing experience as the various levels at which modules will be written. The Editor is Stephen Maurer (PM, Swarthmore), and the Associate Editors are:

- Susanna Epp, DePaul University
- Joseph Malkevitch, York College (CUNY)
- Catherine McGeoch, Amherst College
- Joseph Rosenstein, PM, RU
- William Steiger, PM, RU

So far two editors, Maurer and Malkevitch, have agreed to write sample modules, on mathematical phylogeny and fullerenes, respectively. Maurer has begun giving regular presentations to DIMACS conferences about the Series. He spoke to both branches of the Reconnect program (the weekend program in May, 1998, and the two-week program in July) and to each week of the DREI program (the research participants changed weekly). He stayed for several days of these conferences, and met with participants informally to discuss their ideas and look at early drafts of material. All of the participants in the 2-week Reconnect Conference are working in groups on module material to be completed during this academic year, with various members of the Editorial Board and speakers at the conference acting as mentors. Topics they have chosen include clustering, sequence alignment, DNA fragment reassembly, and graph crossing numbers. In addition, several participants at the weekend Reconnect conference and at DREI expressed interest in writing modules individually, and Professor Maurer is in contact with them.

A brochure about the modules series has been prepared and is being used as part of the editorial board's efforts to announce the series and entice authors at professional meetings and at DIMACS conferences.

The Editorial Board has identified many current research topics which seem ripe for modules. The Board has also identified researchers who may be receptive to writing modules, or serving as co-writers or advisors. The Module Editorial Board is especially eager to arrange for collaborations between researchers, expositors, and teachers. Once sample modules are in hand, a concerted effort will be made to attract writers.

5 Outreach and Knowledge Transfer

5.1 Industrial Interactions

In December 1997, the NEC Research Institute, Princeton, NECI, became a member institution in DIMACS. The first full year of NECI's formal involvement with DIMACS has proven fruitful for the center, with the active involvement of NECI scientists in an organizational and advisory role in many DIMACS projects. Particular thanks go to Dr. David Waltz, Vice President for Computer Science Research at NECI, for his extremely active involvement in the DIMACS Executive Committee.

DIMACS now has four industrial partners. Besides NEC Research, these are AT&T Labs-Research, Bell Labs (Lucent Technologies), and Bellcore. These institutions are truly partners in the center. Senior representatives from the institutions are heavily involved in planning DIMACS activities and play a major organizational role in the center's programs. Moreover, representatives of senior management are extremely active and visible supporters of DIMACS in numerous forums, such as in the recent visit of NSF Director Rita Colwell and at the DIMACS Tenth Birthday Celebration. DIMACS receives significant support from its institutional partners. They provide substantial support by appointing DIMACS postdocs, sponsoring workshops, and allowing release time for their researchers to actively participate in DIMACS administrative committees and to organize DIMACS workshops and Special Years. For the one year period beginning 7/1/97, the estimated value of their contributions are: Bell Labs $602,000; AT&T Labs - Research $515,000 and Bellcore $435,000.

All of our industrial partner 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 six 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. Our industrial partners have estimated the value of their in-kind contributions through the participation of their employees in DIMACS organizational activities as follows: Bell Labs $200,000; AT&T Labs $250,000; Bellcore $200,000. These amounts are included in the amounts in the previous paragraph. We have not attempted to put a value on the in-kind support provided by NECI as its involvement with DIMACS is still too new. However, this support is considerable.

IBM Research agreed to become an affiliated partner in connection with our application for another science and technology center award. We have discussed or planned a number of joint activities with IBM, including a joint workshop and a shared postdoc. The first joint workshop will take place in August 1999 at IBM Research and will culminate our Special Year on Large Scale Discrete Optimization with a major conference celebrating the 70th birthday of Ralph Gomory.

Biotech and pharmaceutical companies have been heavily involved in the Special Year in Mathematical Support for Molecular Biology. Participating companies include CuraGen, Seq Limited, Roche Molecular Systems, Hoffman LaRoche, Pharmagenics, Smith Kline Beecham, Johnson and Johnson, BIOSYM Tech., Ciba Geigy, Merck, Perkin-Elmer, Genome Therapeutics, and the Institute for Genomic Research. These companies have, over the years, provided financial support, speakers, committee members, organizers, attendees, and data and feedback for our algorithm implementation challenge. In particular, Merck Pharmaceuticals and Smith Kline Beecham Pharmaceuticals supported our workshop on Estimating Large Scale Phylogenies this past year. Also in the past year, we have worked hard to involve pharmaceutical companies in DIMACS molecular biology activities by working with Rutgers University to form an Industrial Advisory Board to provide advice on educational and research activities at the interface between the mathematical and physical sciences on the one hand and the biological sciences on the other hand. The companies that have provided senior level representations to this industrial advisory board include Parke-Davis, Novartis, R.W. Johnson Pharmaceutical, Wyeth-Ayerst Research, and Biogen.

Industrial participants have been very active in our workshops, as organizers, speakers, and attendees. This has been especially true in our ongoing special years on Molecular Biology, Networks, DNA Computing, and Large Scale Discrete Optimization.

Since DIMACS began, with the exception of one year, it has received support from the New Jersey Commission on Science Technology. The Commission supports programs that hold promise for the development of the economy of the State of New Jersey and strongly encourages collaborations with small companies in the State. DIMACS has been receiving an annual award of $100,000 from the Commission. In 1996, 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. All of the centers supported by the state have either been phased out or severely cut. 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 fiscal year beginning July 1998. While the amount of Commission support is not large, we consider it significant given the current situation.

5.2 Other Centers and Government Labs

A key goal of DIMACS is outreach to other communities, both within and without the mathematical sciences. In the past year, we have made major progress in enhancing and expanding our many partnerships.

A key partner of DIMACS is the Rutgers Center for Math, Science, and Computer Education. With CMSCE, we run the many K-12 educational programs described in Section 4.1, in particular the Leadership Program for precollege teachers and the Young Scholars Program for high school students. Our long-standing partnership with CMSCE has played a crucial role in our success in these programs.

The highly successful partnership with the Institute for Advanced Study has continued with joint workshops and shared postdocs. During 1997-98, we shared two postdocs, one at Rutgers and one at IAS, and we are sharing another joint postdoc in 1998-99. IAS has been a major partner in the continuing Special Focus on Discrete Probability. The Special Year on Large Scale Discrete Optimization has also been collaborating with IAS, and the special year workshop on "Randomized and Derandomized Algorithms for Discrete Structures" was held at IAS in November 1998. We are also discussing with IAS the extension of our 1999-00 special year on Computational Intractability to a second year to be centered there.

With the University of Pennsylvania, we have run a number of joint programs. In Molecular Biology, we shared a postdoc. We also ran a workshop on "Estimating Large Scale Phylogenies: Biological, Statistical, and Computational Problems". We are planning a joint workshop with our sister STC at U. Penn, the Institute for Research in Cognitive Science, on "Logic and Cognitive Science," to be held in April 1999.

At Rutgers, our simulations project has been carried out jointly with faculty at WINLAB, the Wireless Networking Laboratory, and we are discussing with WINLAB their involvement in future special years on wireless and mobile computing topics. Several postdocs, graduate students, and undergraduates at DIMACS have been mentored by WINLAB faculty.

In molecular biology and DNA computing, our special year activity has engaged us in serious collaborations with a number of centers both locally and around the country. The departments of Ecology and Evolutionary Biology and of Molecular Biology at Princeton have been involved since the beginnning of both of these activities. Professor Laura Landweber of Princeton is one of the organizers of the DNA Computing Special Focus and of a number of its workshops. Professor Lee Silver from Princeton hosted a DIMACS tutorial and this resulted in an interaction with DIMACS visitor Dannie Durand that produced very important scientific results on ultraselfish genes that we have reported on in earlier years. We have partnered with the Center for Molecular Biophysics and Biophysical Chemistry at Rutgers University in joint sponsorship of a series of workshops over the years, with another coming up in March 1999 on the topic "Advances and Opportunities at the Biology/Math/Computation/Physical Sciences Interface." The Director of the Center, Wilma Olson, has been mentoring in the DIMACS REU program for three summers. She and Fred Roberts have been appointed by Rutgers University to head up the new university BioMath - Biophysics Coordinating Council, which is to oversee interdisciplinary programs at the interface between the mathematical and biological sciences. This Council is the direct result of DIMACS-initiated programs in the molecular biology area. We have also partnered in workshops with the Program in Mathematics and Molecular Biology at UC-Berkeley. PMMB is also a sponsor of the March 1999 workshop. The Green Plant Phylogeny Coordination Group, an NSF funded organization for the coordination of research activities concerning the diversity of plants, collaborated with us on the June 1998 workshop on "Estimating Large Scale Phylogenies: Biological, Statistical, and Computational Problems." We continue our partnerships with the Center for Molecular Biotechnology, our sister STC at the University of Washington, whose director is on the steering committee for our molecular biology special year. Faculty members from the Waksman Institute for Molecular Genetics at Rutgers and the Rutgers Center for Theoretical and Applied Genetics are also involved in that steering committee and are participants in our special year seminar. Professor Joachim Messing, Director of Waksman, is a chair of the special year. With the Mt. Sinai School of Medicine, we have discussed shared postdocs and run some computational biology activities for REU students.

Several of our activities in the special year on Large Scale Discrete Optimization are jointly sponsored by Carnegie Mellon University and Georgia Tech. Faculty from those two universities are increasingly active in DIMACS activities. We will be holding a joint DIMACS - Georgia Tech workshop on the "Combinatorics of Statistical Mechanics Models" at Georgia Tech in April 1999. This is part of our special focus on Discrete Probability.

The Institute for Defense Analyses - Center for Communications Research (IDA-CCR) in Princeton has partnered with DIMACS in the past in shared postdoc programs. Helene Kulsrud from IDA-CCR is on the steering committee for our Massive Data Sets special year, and she organized one of the special year workshops. We are discussing with IDA-CCR the joint sponsorship of a future workshop.

Lawrence Berkeley National Lab has been a partner in our massive data sets and molecular biology activities. Sylvia Spengler and Manfred Zorn from Lawrence Labs were organizers of our workshop on "Combinatorial Clustering and Multi-Domain Protein Structure Analysis" that was organized in part to explore future joint research collaborations.

The Institute for Mathematics and its Applications, IMA, at the University of Minnesota, is one of NSF's mathematical sciences research institutes. We have planned, as part of our Networks Special Year, a joint workshop with IMA on the topic "Scaling Phenomena in Communication Networks."

At the January 1999 American Mathematical Society national meeting, a number of the research centers in the mathematical sciences will be holding an informational session to introduce the community to programs at these centers. DIMACS is sharing in the organization of that session with IMA, the Mathematical Sciences Research Institute (MSRI) at Berkeley, and the National Institute of Standards and Technology (NIST).

5.3 International Activities

DIMACS has been actively pursuing a variety of international partnerships. We now have ongoing shared activities started or planned with six international partners.

DIMATIA is the Center for Discrete Mathematics and Theoretical Informatics, based at Charles University in Prague, Czech Republic. With the help of a grant from International Programs at NSF, we have been running a series of joint programs with DIMATIA, starting in 1997 with the opening conference of DIMATIA on The Future of Discrete Mathematics. This conference has now led to a DIMACS volume, edited by Jan Kratochvil and Jarik Nesetril from DIMATIA and Ronald Graham and Fred Roberts from DIMACS. The volume will be in press shortly. During the present year, we ran two joint workshops, one on Combinatorial and Algorithmic Geometry in Prague in August 1998 and one on the probabilistic method, entitled "From Erdős to Algorithms," at Rutgers in June 1998. A fourth joint workshop, on Algebraic Methods and Arithmetic Circuits, is scheduled for Rutgers in June 1999. These specialized workshops return to continuing research thrusts at DIMACS, specifically the Special Year on Discrete and Computational Geometry, the Special Focus on Discrete Probability, and the Special Year on Logic and Algorithms, respectively. As part of the partnership with DIMATIA, we have had regular exchange of visitors, with an emphasis on young people. We are now planning a joint REU program, with several Czech students coming to DIMACS to participate in our summer REU program and a group of American students following up the summer DIMACS REU with a visit to DIMATIA.

The Erdős Institute is part of the Hungarian Academy of Sciences in Budapest. We have run one joint program with the Erdős Institute, a summer workshop on Probabilistic Combinatorics in July 1998, and are planning a summer workshop on Paul Erdős and his Mathematics in July 1999. We are planning a joint research effort centered around combinatorics and have been exchanging visits in planning that effort.

BRICS is the Center for Basic Research in Computer Science in Aarhus, Denmark. Erik Meineche Schmidt, the BRICS Co-Director, and Mogens Nielsen, BRICS Ph.D. School Director, visited DIMACS in March 1998 to discuss collaborative projects. We discussed a joint workshop, which is still being planned, and agreed to share a postdoc. We made an offer, but the offer was not accepted. Christian Pederson, a graduate student from BRICS, spent two months visiting DIMACS in the Fall of 1998.

HKUST is the Hong Kong University of Science and Technology. We have started a collaboration with the Computer Science Department at HKUST. This resulted in a planned December 1998 joint workshop on Algorithms and Combinatorics, in Hong Kong. Mordecai Golin from HKUST visited DIMACS in March 1998 to discuss mutual collaborations and Fred Roberts visited HKUST in May to prepare for the workshop and discuss future collaborations. DIMACS is currently hosting Professor Golin during his sabbatical, where he is interacting with us in the Massive Data Sets Special Year, and we are also hosting two students from HKUST this year.

We are beginning a collaboration with INRIA - Universite de Nice-Sophia Antipolis in France. Jean-Claude Bermond from Sophia-Antipolis visited DIMACS several times during 1998 and we are also currently hosting Herve Bronimann from Sophia-Antipolis. At a November 1998 meeting, we agreed to plan several joint workshops, starting with workshops related to our Networks special year, one to be held in France and one in the U.S. We also agreed to exchange graduate students.

DISCOM is the Center for Discrete Mathematics and Computer Science at Hebrew University in Jerusalem. We have been discussing joint programs and exchanging visitors and postdocs informally for a number of years. It is likely that the continued interaction will remain relatively informal for now.

The International Mathematical Sciences Institutes (IMSI) is a group of mathematical sciences research institutes which meet together at international congresses and have agreed to collaborate with each other. DIMACS has become a member of IMSI and Neil Sloane (PM, ATTL) represented

DIMACS at the IMSI meeting in Berlin in August 1998.

5.4 Visitor Program

A key to the success of the DIMACS programs is the interaction among a wide variety of people, permanent members of the center, visitors, postdoctoral fellows, and graduate students. DIMACS has almost two thousand visitors a year attending our many programs, and the visitors range from senior leaders in the area of a special year to outstanding young people. We put special emphasis on visitors from fields outside of dm/tcs, for visitors from companies and government labs, and for visitors from smaller institutions where research is not easily accomplished. Visitors come for periods of time ranging from a few days to a few months to a year. Visitors are central to the outreach and knowledge transfer programs of the center. They bring new ideas and new problems to participants in the center's programs. In turn, they bring back new methods, tools, algorithms, and questions to their home institution. A partial list of DIMACS visitors is included in Appendix B.2. The list does not include those who come only for a workshop or a seminar or a tutorial, which includes almost all of our industrial visitors. However, it gives a hint of the type of influence that outsiders have on DIMACS and that DIMACS has on outsiders.

DIMACS evaluates the visitor program by asking each visitor to fill out an evaluation form and also asks all non-workshop visitors to file reports on their work while at the center. Of the non-workshop visitors during the past year, 71% rated their experience at DIMACS 5 (the highest) out of a 5-point scale and 97% rated it 4 or 5. All said they would want to come back! Here are a few selected comments from the visitors:

"My stay at Princeton University changed my course on research."

"Being a theoretician working in academia I'm not so much exposed to applied problems. My visit at DIMACS (Bell Labs) was a perfect opportunity to catch up with current trends in industrial research and to get new inspiration for my future work, with regard to both research and (!)teaching."

"Even though I know most of the DIMACS members, every visit to DIMACS is valuable because of the number of new people you meet."

"DIMACS mission in the present form (support of discrete mathematics and theoretical computer science) is crucial to the existence of the foundation areas in many scientific fields."

DIMACS asks all attendees at workshops to fill out a questionnaire. Of the workshop attendees during the past year, 51% rated the workshop as 5 (highest) on a 5-point scale, while 88% rated it 4 or 5 and 98% rated it 3 or higher. 99% of the attendees said that they made at least one new contact at the workshop and 62% of those said they expected to collaborate with their new contact, with 94% saying they would or could collaborate. Here are a few noteworthy comments from workshop participants:

"I enjoyed immensely the workshop's interdisciplinary format, and I am eager to bring this flavor both more into my research and into the classroom." – astronomer

"I learned a great deal and, as a biologist, overcame some algorithm phobia. This conference had a tremendous impact on much interchange of ideas between disciplines."

"As a grad student I really appreciated the more informal/introductory/helpful/leading spirit of the meeting"

"I am learning a lot about many different topics by attending DIMACS workshops. This is opening up new areas of research for me."

6 Shared Experimental Facilities

The main offices of DIMACS are at Rutgers University. The university provides space for the Center in the CoRE Building on the Rutgers Busch campus in Piscataway, New Jersey. This space comprises the entire fourth floor of the building and amounts to 9,434 square feet, comprising 26 offices, a lounge, seminar room, kitchen, conference room, two PC labs, and a mailroom. The seminar room holds 75 people. The CoRE building has a 110-seat auditorium that is shared by the building's occupants and is available to the Center on a regular basis. Larger auditoriums are also available on campus.

The Center currently has a wide range of hardware which includes a central cluster of 6 Sun SPARC servers with over 40GB of storage capacity and 850MB of total RAM; 2 NT servers with 17GB of storage and 200MB of RAM; 7 Sun workstations, a Silicon Graphics Indigo workstation with 384MB of RAM; 2 PC Labs with a total of 14 workstations; and an additional 7 workstations, 7 PCs and 34 x-terminals used by staff, postdocs and visitors.

Access to Rutgers University Libraries is also available to Center visitors, postdocs, permanent members, and students. With holdings of more than three million volumes, the university's library system ranks among the top 25 research libraries nationally. The Mathematical Sciences Library, one of the system's eighteen units, has an extensive specialized collection, and is reachable from the CoRE building by enclosed ramp.

All DIMACS members are welcome at the Center's main offices at Rutgers. We have several offices set aside for visitors from our other partner institutions. Moreover, those visitors are provided accounts on the DIMACS machines as requested, access to the university's libraries, and staff support as needed. The same hospitality is offered to DIMACS postdocs, graduate students, and visitors, no matter where they are headquartered.

DIMACS also has a number of offices in the Princeton University computer science building. The university's computing facilities and library facilities are made available to DIMACS members, postdocs, students, and visitors who might wish to spend time in the Princeton CS Department either as short-term visitors while attending a workshop or as longer-term visitors.

Occasional DIMACS events such as mixers, seminars, and committee meetings take place at the facilities of our industrial partners. Since all of these partners require badged access to their facilities, this requires some advance planning. However, escorted access to events at these locations is easily arranged, and many of our permanent members, postdocs, students, and visitors are frequent guests for collaborative research at our industrial partners' locations.

7 Administration and Management

7.1 Organization of the Center

The center has the following management:

Management

- Director: Fred Roberts, RU
- Co-Director: Robert Sedgewick, PU
- Associate Director: Julia Abrahams, RU
- Associate Director for Education: Joseph Rosenstein, RU

Executive Committee

The major decisionmaking body of the center is the Executive Committee. This committee consists of the management plus one representative from each DIMACS unit, the units being the four industrial members plus Princeton CS, Rutgers CS, Rutgers Math, and Rutgers OR. The Executive Committee meets monthly during the academic year and otherwise as needed. Members of the Executive Committee are as follows:

- Julia Abrahams, RU
- Sandeep Bhatt, BC
- Mike Fredman, RU
- Peter Hammer, RU
- David Johnson, ATTL
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael Saks, RU
- Robert Sedgewick, PU
- Robert Tarjan, PU
- David Waltz, NEC
- Mihalis Yannakakis, BL

DIMACS Council

The center has a large advisory committee of local leaders, called the DIMACS Council. The Council consists of the Executive Committee, one additional representative from each DIMACS unit, and the chairs of all ongoing special years and DIMACS committees. The Council meets approximately twice a year to discuss general issues concerning the center and the fields of dm/tcs. The membership of the Council is as follows:

- Julia Abrahams, RU
- Eric Allender, RU
- Sandeep Bhatt, BC
- Endre Boros, RU
- Ken Clarkson, BL
- Martin Farach-Colton, RU
- Joan Feigenbaum, ATTL
- Steve Fortune, BL
- Michael Fredman, RU
- C. Lee Giles, NEC
- Peter Hammer, RU
- David Johnson, ATTL
- Jeff Kahn, RU
- Janos Komlos, RU
- Richard Lipton, PU
- Milena Mihail, BC
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Mike Saks, RU
- Robert Sedgewick, PU
- Peter Shor, ATTL
- Charles Sims, RU
- Robert Tarjan, PU
- David Waltz, NEC
- Peter Winkler, BL
- Mihalis Yannakakis, BL

DIMACS has a variety of other committees, each having representatives of each of the DIMACS units. The committees are as follows:

Postdoc Committee

- Julia Abrahams, RU

- Sanjeev Arora, PU
- Joszef Beck, RU
- Enrico Bombieri, Institute for Advanced Study
- Endre Boros, RU
- Martin Farach-Colton, RU
- David Johnson, ATTL
- Sanjeev Khanna, BL
- Richard Lipton, PU
- S. Muthukrishnan, ATTL
- Andy Ogielski, RU
- Fred Roberts, RU
- Mike Saks, RU
- Bob Sedgewick, PU

Projects Committee

- Julia Abrahams, RU
- Farid Alizadeh, RU
- Ken Clarkson, BL
- Mike Fredman, RU
- Janos Komlos, RU
- Joel Lebowitz, RU
- Dick Lipton, PU
- S. Muthukrishnan, ATTL
- Rafi Ostrovsky, BC
- Fred Roberts, RU
- Mike Saks, RU
- Bob Sedgewick, PU
- Rebecca Wright, ATTL

Education Committee

- Julia Abrahams, RU
- Bill Aiello, ATTL

- Adam Buchsbaum, ATTL
- Sanjeev Khanna, BL
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael Saks, RU
- Robert Sedgewick, PU
- Ann Yasuhara, RU
- Adi Ben-Israel, RU

Graduate Committee

- Julia Abrahams, RU
- Sanjeev Arora, PU
- Endre Boros, RU
- Leonid Khachiyan, RU
- Sivaram Rajagopalan, BC
- Fred Roberts, RU
- Martin Strauss, ATTL
- Eduardo Sontag, RU
- Sanjeev Khanna, BL
- Robert Sedgewick, PU
- Harold Stone, NEC

Visitor Committee

- Julia Abrahams, RU
- Farid Alizadeh, RU
- Steve Fortune, BL
- Andrew Odlyzko, ATTL
- Rafail Ostrovsky, BC
- Fred Roberts, RU
- Endre Szemeredi, RU
- Michael Saks, RU
- Robert Sedgewick, PU

Computing Equipment Committee

- Julia Abrahams, RU
- Vasek Chvatal, RU
- Nicholas Reingold, ATTL
- Fred Roberts, RU
- J.P. Singh, PU
- David Shanno, RU
- Charles Sims, RU
- Robert Sedgewick, PU

Strategic Planning Committee

- Julia Abrahams, RU
- Eric Allender, RU
- Sandeep Bhatt, BC
- David Johnson, ATTL
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael Saks, RU
- Robert Sedgewick, PU
- Dave Waltz, NEC

Budget Committee

- Ronald Graham, ATTL
- Clyde Monma, BC
- Fred Roberts, RU
- Robert Sedgewick, PU

DREI Steering Committee

- Julia Abrahams, RU
- Chuck Biehl, Charter School of Wilmington
- Ethel Brueche, consultant, Pennsylvania
- Pat Carney, Montclair Kimberly Academy

- Ronald Graham, ATTL
- Glenn Hurlbert, Arizona State University
- Henry Kepner, University of Wisconsin-Milwaukee
- Linda Lesniak, Drew University
- John Polking, Rice University
- Henry Pollak, Columbia University
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael Saks, RU
- Marjorie Senechal, Smith College
- Barry Tesman, Dickinson College
- William Velez, University of Arizona

7.2 Center Advisors

DIMACS External Advisory Board

- Peter Arzberger, University of California/San Diego
- Jennifer Chayes, Microsoft Corp.
- Cynthia Dwork, IBM Almaden Research Center
- Avner Friedman, University of Minnesota (Chair)
- Richard Karp, University of Washington
- Daniel Kleitman, M.I.T.
- Laszlo Lovasz, Yale University
- Jill Mesirov, Boston University
- Joseph O'Rourke, Smith College
- Christos Papadimitriou, University of California/Berkeley
- William Pulleyblank, IBM Watson Research
- Dana Scott, Carnegie Mellon University
- Eva Tardos, Cornell University
- Herbert Wilf, University of Pennsylvania

A Publications

A.1 American Mathematical Society Publications: DIMACS Volume Series:

Volume One: "Polyhedral Combinatorics" Editors: William Cook and Paul D. Seymour
Volume Two: "Distributed Computing and Cryptography" Editiors: Joan Feigenbaum and Michael Merritt
Volume Three: "Computer -Aided Verification '90" Editors: E.M. Clarke and R.P. Kurshan
Volume Four: "Applied Geometry and Discrete Mathematics" Editors: Peter Gritzmann and Bernd Sturmfels
Volume Five: "Reliability of Computer and Communication Networks" Editors: Fred Roberts, Frank Hwang and Clyde Monma
Volume Six: "Discrete and Computational Geometry: Papers from the DIMACS Special Year" Editors: Jacob E. Goodman, Richard Pollack and William Steiger
Volume Seven: "On-line Algorithms" Editors: Lyle A. McGeoch and Daniel D. Sleator
Volume Eight: "Mathematical Methods of Analysis of Biopolymer Sequences" Editor: Simon Gindikin
Volume Nine: "Planar Graphs" Editor: William T. Trotter
Volume Ten: "Expanding Graphs" Editor: Joel Friedman
Volume Eleven: "Groups and Computation" Editors: Larry Finkelstein, William M. Kantor
Volume Twelve: "Network Flows and Matching" Editors: David S. Johnson, Catherine C. McGeoch
Volume Thirteen: "Advances in Computational Complexity Theory" Editor: Jin-Yi Cai
Volume Fourteen: "Coding and Quantization" Editors: Robert Calderbank, G. David Forney, Jr., Nader Moayeri
Volume Fifteen: "Computational Support for Discrete Mathematics" Editors: Nathaniel Dean, Gregory E. Shannon
Volume Sixteen: "Quadratic Assignment and Related Problems" Editors: Panos M. Pardalos and Henry Wolkowicz
Volume Seventeen: "Langauage Computations" Editors: Eric Sven Ristad
Volume Eighteen: "Specification of Parallel Algorithms" Editors: Guy E. Blelloch, K. Mani Chandy, Suresh Jagannathan
Volume Nineteen: "Partitioning Data Sets" Editors: Ingemar Cox, Pierre Hansen, Bela Julesz
Volume Twenty: "Combinatorial Optimization" Editors: William Cook, Laszlo Lovasz, Paul Seymour
Volume Twenty One: "Workshop on Interconnection Networks and Mapping and Scheduling Parallel Computations" Editors: D. Frank Hsu, Arnold L. Rosenberg, Dominique Sotteau
Volume Twenty Two: "Workshop on Parallel Processing of Discrete Optimization Problems" Editors: Panos M. Pardalos, Mauricio G.C. Resende, K.G. Ramakrishnan
Volume Twenty Three: "Global Minimization of Nonconvex Energy Functions: Molecular Conformation and Protein Folding" Editors: P.M. Pardalos, D. Shalloway, and G. Xue
Volume Twenty Four: "Formal Power Series and Algebraic Combinatorics - Series Formelles et Combinatoire Algebrique" Editors: Louis Billera, Curtis Greene, Rodica Simion, and Richard Stanley
Volume Twenty Five: "Geometric and Computational Perspectives on Infinite Groups - Proceedings of a Joint DIMACS/Geometry Center Workshop January 3-14 and March 17-20, 1994" Editors: Gilbert Baumslag, David Epstein, Robert Gilman, Hamish Short, and Charles Sims.
Volume Twenty Six: "Cliques, Coloring and Satisfiability"
Editors: David S. Johnson and Michael A. Trick

Volume Twenty Seven: "DNA Based Computers - Proceedings of a DIMACS Workshop, April 4, 1995"
Editors: Richard J Lipton and Eric B. Baum
Volume Twenty Eight: "Groups and Computation II - Workshop in Groups and Computation, June 7-10, 1996," Editors: Larry Finkelstein and William Kantor
Volume Twenty Nine: "Partial Order Methods in Verification: DIMACS Workshop, July 24-26, 1996," Editors: Doron A. Peled, Vaughan R. Pratt, and Gerard J. Holzmann
Volume Thirty: "Parallel Algorithms: Third DIMACS Implementation Challenge,' October 17-19, 1994," Editor: Sandeep N. Bhatt.
Volume Thirty One: "Descriptive Complexity and Finite Models: Proceedings of a DIMACS Workshop, January 14-17, 1996," Editors: Neil Immerman and Phokion G. Kolaitis.
Volume Thirty Two: "The SPIN Verification System: The Second Workshop on the SPIN Verification System Proceedings of a DIMACS Workshop August 5, 1996," Editors: Jean-Charles Gregoire, Gerard J. Holzmann, Doron A, Peled
Volume Thirty Three: "Logic and Random Structures" Editors: Ravi Boppana and James Lynch.
Volume Thirty Four: "African Americans in Mathematics"
Editor: Nathaniel Dean.
Volume Thirty Five: "Satisfiability Problem: Theory and Applications" Editors: Dingzhu Du, Jun Gu, and Panos M. Pardalos.
Volume Thirty Six: "Discrete Mathematics in the Schools" Editors: Joseph G. Rosenstein, Deborah S. Franzblau, Fred S. Roberts. **Volume Thirty Seven:** "Mathematical Hierarchies and Biology" Editors: Boris Mirkin, F. R. McMorris, Fred S. Roberts and Andrey Rzhetsky.
Volume Thirty Eight: "Network Threats" Editors: Rebecca N. Wright and Peter G. Neumann.
Volume Thirty Nine: "Proof Complexity and Feasible Arithmetics" Editors: P. Beam and S. Buss.
Volume Forty: "Network Design: Connectivity and Facilities Location" Editors: Panos M. Pardalos and Dingzhu Du.
Volume Forty One: "Microsurveys in Discrete Probability"
Editors: David Aldous and James Propp.
Volume Forty Two: "Advances in Switching Networks" Editors: Ding-Zhu Du and Frank K. Hwang.
Volume Forty Three: "Randomization Methods in Algorithm Design" Editors: Panos Pardalos and Sanguthevar Rajasekaran.
Volume Forty Four: "DNA Based Computers II" Editors: Laura Landweber and Eric Baum.
Volume Forty Five: "Networks in Distributed Computing" Editors: Marios Mavronicolas, Michael Merritt, and Nir Shavitt.

A.2 1998 Technical Reports

- **98-01** Closed Sets and Generators in Ternary Hamming Spaces; Endre Boros, Peter L. Hammer, Frederica Ricca and Bruno Simeone.
- **98-02** On the Complexity of Generation in Ternary Hamming Spaces; Endre Boros, Peter L. Hammer, Frederica Ricca and Bruno Simeone.
- **98-03** Optimal Cell Flipping to Minimize Channel Density in VLSI Design and Pseudo-Boolean Optimization; Endre Boros, Peter L. Hammer, Michel Minoux and David Rader.
- **98-04** Measurements of Uncolorability; Eckhard Steffen.
- **98-05** The Set-Maxima Problem, an Overview; Richard Desper.
- **98-06** Key Management for Encrypted Broadcast; Avishai Wool.
- **98-07** A Characterization for Parity Graphs and a Coloring Problem with Costs; Klaus Jansen.
- **98-08** The Mutual Exclusion Scheduling Problem for Permutation and Comparability Graphs; Klaus Jansen.
- **98-09** Convexity and Logical Analysis of Data; Oya Ekin, Peter L. Hammer and Alexander Kogan.
- **98-10** Patience is a Virtue: The Effect of Delay on Competitiveness for Admission Control; Michael H. Goldwasser.
- **98-11** Nonexhaustive Generalized Fibonacci Trees in Unequal Costs Coding Problems; Julia Abrahams.
- **98-12** Tagged Probe Interval Graphs; Li Sheng, Chi Wang and Peisen Zhang.
- **98-13** On the Complexity of the Isomorphism Relation for Finitely Generated Groups; Simon Thomas, and Boban Velickovic.
- **98-14** Topology Aggregation for Directed Graph; Baruch Awerbuch and Yuval Shavitt.
- **98-15** Bellman Functions on Trees for Segmentation, Generalized Smoothing, Matching Multi-Alignment in Massive Data Sets; Ilya Muchnik and Vadim Mottl.
- **98-16** Routing Through Networks with Hierarchical Topology Aggregation; Baruch Awerbuch, Yi Du, Bilal Khan and Yuval Shavitt.
- **98-17** Online Channel Allocation in FDMA Networks with Reuse Constraints; Tomas Feder and Sunil M. Shende.
- **98-18** Functional Dependencies in Horn Theories; Toshihide Ibaraki, Alexander Kogan and Kazuhisa Makino.
- **98-19** On the Number of Equilateral Triangles in Euclidean Spaces; Bernardo M. Abrego and Silvia Fernandez-Merchant.
- **98-20** ATM Network Design: Traffic Models and Optimization-Based Heuristics; Iraj Saniee and Daniel Bienstock.

- **98-21** How tight is the Bollobas-Komlos Conjecture?; Sarmad Abbasi.
- **98-22** Geometric Graphs with Few Disjoint Edges; Geza Toth and Pavel Valtr.
- **98-23** Complexity Classes; Eric Allender, Michael C. Loui and Kenneth W. Regan.
- **98-24** Reducibility and Completeness; Eric Allender, Michael C. Loui and Kenneth W. Regan.
- **98-25** Other Complexity Classes and Measures
  Eric Allender, Michael C. Loui and Kenneth W. Regan.
- **98-26** Erdos-Szekeres-Type Theorems for Segments and Non-crossing
  Convex Sets; Janos Pach and Geza Toth.
- **98-27** Finding Convex Sets in Convex Position; Geza Toth.
- **98-28** A Formal Scenario and Metalanguage Support Means to
  Reason about It; Boris Galitsky.
- **98-29** Which Crossing Number Is It Anyway?; Janos Pach and Geza Toth.
- **98-30** Analysis of a Local Search Heuristic for Facility Location
  Problems; Madhukar R. Korupolu, C. Greg Plaxton and Rajmohan Rajaraman.
- **98-31** Approximate Coloring of Uniform Hypergraphs
  Michael Krivelevich and Benny Sudakov.
- **98-32** Coloring Graphs with Sparse Neighborhoods
  Noga Alon, Michael Krivelevich and Benny Sudakov.
- **98-33** Thin Complete Subsequence; Norbert Hegyvari.
- **98-34** On the Dimension of the Hilbert-Cubes; Norbert Hegyvari.
- **98-35** Do Answers Help in Posing Questions?; Sarmad Abbasi.
- **98-36** Broadcast Disk Paging with a Small Cache; Vincenzo Liberatore.
- **98-37** Scaling Dualities and Self-Concordant Homogeneous
  Programming in Finite Dimensional Spaces; Bahman Kalantari.
- **98-38** On the Arithmetic-Geometric Mean Inequality and Its
  Relationship to Linear Programming, Matrix Scaling, and Gordan's Theorem; Bahman Kalantari.
- **98-39** Crossing Properties of Graph Reliability Functions
  Alexander K. Kelmans.
- **98-40** An Adversarial Model for Distributed Dynamic Load Balancing
  S. Muthukrishnan and R. Rajaraman.
- **98-41** Strengthening the Varshamov-Gilbert Bound
  Alexander Barg, Sugi Guritman, Juriaan Simonis.
- **98-42** Note on the Permanent Rank of a Matrix; Yang Yu.

- **98-43** Limiting the Search for 2-Dimensional Optimal Alphabetic Trees
  Ahmed Belal, Magdy A.Ahmed, Shymaa M.Arafat.
- **98-44** List Update Problem; Bela Csaba and Sachin Lodha.
- **98-45** Partition-Based Goodness-of-Fit Tests on the Line
  and the Circle; Carol J. Feltz and Gerald A. Goldin.
- **98-46** On Crossing Sets, Disjoint Sets and the Pagenumber
  Farhad Shahrokhi and Weiping Shi.
- **98-47** Quasi-isometry Classification of Coarse Hadamard Manifolds
  Oliver Attie.
- **98-48** Poincare Duality for Lp Cohomology and Characteristic Classes
  Oliver Attie and Jonathan Block.
- **98-49** Phylogeny Numbers for Graphs with Two Triangles
  Fred S. Roberts and Li Sheng.
- **98-50** Extremal Phylogeny Numbers; Fred S. Roberts and Li Sheng.

A.3 Patents

In 1997-98, the center applied for one patent, jointly with the NEC Research Institute, for cryptographic software entitled "Identity Escrow" developed by postdoctoral fellow Erez Petrank with NECI researchers Joe Killian. The center did not apply for any patents during 1998-99.

B List of Center Participants

B.1 Permanent Members of the Center and their Research Interests

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

Julia Abrahams, Rutgers University
Research Areas: Lossless data compression; code and parse trees; combinatorial search algorithms; entropy bounds.

Alfred Aho, Bell Labs
Research Areas: Pattern matching algorithms; digital libraries; software engineering.

Bill Aiello, AT&T Labs - Research
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
Research Areas: Formal methods; verification; logic and concurrency.

Matthew Andrews, Bell Labs
Research Areas: Simulators for disk drive scheduling and ATM private network-network interface routing.

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

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

Alexander Barg, Bell Labs
Research Areas: Coding theory; combinatorics; finite algebra; computer science.

Eric Baum, NEC Research Institute
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
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
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
Research Areas: Algorithms; data structures; graph problems; combinatorics.

Richard Bumby, Rutgers University
Research Areas: Number theory.

Robert Calderbank, AT&T Labs - Research
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.

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

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

Edith Cohen, AT&T Labs - Research
Research Areas: Algorithm design and analysis.

William Cohen, AT&T Labs - Research
Research Areas: Machine learning and information retrieval.

John Conway, Princeton University
Research Areas: Combinatorics; games.

W.M. (Bill) Coughran, Jr., Bell Labs
Research Areas: Scientific computing.

Ingemar Cox, NEC Research Institute
Research Areas: Computer vision.

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

Nathaniel Dean, Rice University
Research Areas: Graph algorithms.

Richard DeMillo, Bellcore
Research Areas: Software reliability; security.

Sven Dickinson, Rutgers University
Research Areas: computational vision problems of object modeling, object recognition, attention, shape recovery and tracking; use of viewer-centered representations in landmark recognition and planning for mobile robot navigation.

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

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

Jonathan Eckstein, Rutgers University
Research Areas: Parallel branch and bound algorithms; proximal algorithms; monotone operators.

Kousha Etessami, Bell Labs
Research Areas: Descriptive complexity and finite model theory; applications of logic in computer science; temporal logic; model checking; computer-aided verification.

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

Joan Feigenbaum, AT&T Labs - Research
Research Areas: Security.

Anja Feldmann, AT&T Labs - Research
Research Areas: Analysis and design of algorithms; on-line and parallel algorithms; high-speed networks; traffic characterization.

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

Steven Fortune, Bell Labs
Research Areas: Computational geometry.

Matt Franklin, AT&T Labs - Research
Research Areas: Security.

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

Yoav Freund, AT&T Labs - Research
Research Areas: Machine learning theory and applications.

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

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

C. William Gear, NEC Research Institute
Research Areas: Numerical methods is computer vision; scientific computing; differential equations.

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

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

C. Lee Giles, NEC Research Institute
Research Areas: Novel applications of AI; neural networks; machine learning to the WWW; computer systems; telecommunications; finance.

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

Ram (Ramanathan) 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
Research Areas: Combinatorics; graph theory; number theory, theoretical computer science.

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

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

Adam Grove, NEC Research Institute
Research Areas: Artifical intelligence; machine learning; logic; probabilistic reasoning.

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, University of Delaware
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 Labs - 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.

Markus Jakobsson, Bell Labs
Research Areas: Cryptography/Security.

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

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

Bahman Kalantari, Rutgers University
Research Areas: Mathematical Programming; Global and Discrete Optimization.

Narendra Karmarkar, Bell Labs
Research Areas: Linear programming.

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

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

Sanjeev Khanna, Bell Labs
Research Areas: Design and analysis of algorithms; combinatorial optimization; application of combinatorial optimization techniques to computer systems; databases; network design, and robotics; structural complexity theory; combinatorics; graph theory.

Joe Kilian, NEC Research Institute
Research Areas: Computational complexity; cryptography.

Nils Klarlund, AT&T Labs - Research
Research Areas: Automatic verification; programming languages and programming logics.

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
Research Areas: Computer aided verification.

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

Kevin Lang, NEC Research Institute
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 Labs
Research Areas: Network protocols.

Linda Lesniak, Drew University
Research Areas: Graph Theory.

Richard Lipton, Princeton University
Research Areas: Architecture; complexity; VLSI.

Nicholas Littlestone, NEC Research Institute
Research Areas: Computational learning theory.

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

Richard Lyons, Rutgers University
Research Areas: Finite simple groups.

Dahlia Malkhi, AT&T Labs - Research
Research Areas: Security in distributed and networked systems; mobile code security; fault tolerant software solutions and analysis; distributed coding; advanced communication networks and protocols.

Yossi Matias, Bell Labs
Research Areas: Algorithms; parallel computation; query processing and optimization; data structures; internet applications; probabilistic methods.

Stephen B Maurer, Swarthmore College
Research Areas: Discrete math; algorithmics; modeling; mathematical exposition.

Alain Mayer, Bell Labs
Research Areas: Cryptography: electronic commerce; computer and network security; Internet technology; network and security management; network protocols; multimedia; distributed algorithms and systems.

Michael Merritt, AT&T Labs - Research
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.

S. (Muthu) Muthukrishnan, Bell Labs
Research Areas: Experimental algorithms; algorithms in compilers; programming languages etc; probabilistic methods; information retrieval; pattern matching; computational biology.

Arun Netravali, Bell Labs
Research Areas: Digital image processing.

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

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

Andrew Ogielski, Rutgers University
Research Areas: Simulation.

Rafail Ostrovsky, Bellcore
Research Areas: Cryptography; computer security; distributed protocols and algorithms.

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

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

Anuj Puri, Bell Labs
Research Areas: Computer-aided verification; real-time systems; hybrid systems; computer networks; transportation systems; optimization.

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

Sridhar Ramaswamy, Bell Labs
Research Areas: Databases performance; I/O algorithms; data mining.

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

Danny Raz, Bell Labs
Research Areas: Theory and applications of management related problems in IP networks, in particular: active networks, cache allocation, flow control, scheduling, theory of network management.

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

Marty Reiman, Bell Labs
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
Research Areas: Algorithms; network design.

Michael Reiter, AT&T Labs - Research
Research Areas: Research secure systems.

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

Fred S. 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, AT&T Labs - Research
Research Areas: Cryptography; network security; secure systems; web security.

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

Iraj Saniee, Bell Labs
Research Areas: Network and combinatorial optimization; SONET and ATM network design.

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

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

Rob Schapire, AT&T Labs - Research
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.

Yuval Shavitt, Bell Labs
Research Areas: networking, distributed algorithms, active networks, flow control, routing and reservation, network management, WWW caching.

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

Lawrence Shepp, Rutgers University
Research Areas: Probability; tomography; combinatorics.

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

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
Research Areas: Mathematics; statistics.

Warren Smith, NEC Research Institute
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: Algorithms; computational geometry; probability; parallel computation.

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

Harold Stone, NEC Research Institute
Research Areas: computer architecture; distributed processing image processing; parallel algorithms, microprocessor systems; operating systems; combinatorial algorithms; computer aided design of VSLI networks

Martin Strauss, AT&T Labs - Research
Research Areas: Massive data sets; security; complexity theory.

Torsten Suel, Bell Labs
Research Areas: Algorithm design and analysis; experimental algorithms; parallel computation; databases.

Don X. Sun, Bell Labs
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; graphical methods for large data analysis.

Ram Swaminathan, Bell Labs
Research Areas: Algorithms; combinatorics; probability; geometry; complexity.

Mario Szegedy, AT&T Labs - Research
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.

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

David Waltz, NEC Research Institute
Research Areas: Artificial intelligence; constraint propagation; memory-based reasoning; information retrieval.

Jeff Westbrook, AT&T Labs - Research
Research Areas: Data structures and algorithms; on line algorithms; network design.

Ward Whitt, AT&T Labs - Research
Research Areas: Operations research; telecommunications systems; probability and stochastic processes and queues.

Gordon Wilfong, Bell Labs
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
Research Areas: Discrete math; probability; theory of computing.

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

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

Rebecca Wright, AT&T Labs - Research
Research Areas: Cryptography; network security.

Mihalis Yannakakis, Bell Labs
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.

Peter Yianilos, NEC Research Institute
Research Areas: Machine learning; data compression; combinatorial optimization; electronic publishing.

Lisa Zhang, Bell Labs
Research Areas: Design, analysis and implementation of algorithms.

Shiyu Zhou, Bell Labs
Research Areas: Design and analysis of data structures and algorithms; computational complexity.

B.2 Center Visitors: February 1998 - January 1999

- **Micah Adler**; University of Toronto;
  Visit related to: Networks: Dates: 2/8/98 - 2/13/98
- **Ron Aharoni**; Technion;
  Visit related to: General: Dates: 9/26/98 - 10/10/98
- **Lars Arge**; Duke University;
  Visit related to: Massive Data Sets: Dates: 5/19/98 - 5/24/98
- **Adnan Aziz**; University of Texas;
  Visit related to: Logic & Algorithms: Dates: 8/3/98 - 8/7/98
- **Maury Bramson**; University of Wisconsin - Madison;
  Visit related to: Discrete Probaility: Dates: 3/8/98 - 3/14/98
- **John Byers**; Berkeley;
  Visit related to: Networks: Dates: 2/23/98 - 2/27/98
- **Ondrej Cepek**; Charles University;
  Visit related to: Large Scale Discrete Optimization: Dates: 9/20/98 - 9/26/98
- **Moses Charikar**; Stanford University;
  Visit related to: Massive Data Sets: Dates: 3/16/98 - 3/20/98
- **Yi-Jen Chiang**; SUNY - Stony Brook;
  Visit related to: Massive Data Sets: Dates: 5/19/98 - 5/23/98
- **Bryan Cooley**; University of Nebraska;
  Visit related to: Simulations: Dates: 5/9/98 - 8/7/98
- **Janos Csirik**; University of Szeged;
  Visit related to: General: Dates: 6/2/98 - 8/28/98
- **Gyorgy Elekes**; Eotvos University;
  Visit related to: General: Dates: 5/8/98 - 7/18/98
- **Funda Ergun**; University of Pennsylvania;
  Visit related to: Logic & Algorithms: Dates: 8/3/98 - 9/21/98
- **Raffaele Esposito**; Universita di L'Aquila;
  Visit related to: General: Dates: 12/1/98 - 12/23/98
- **Martin Evans**; University of Edinburgh;
  Visit related to: General: Dates: 3/18/98 - 4/8/98
- **Paolo Ferragina**; Max-Planck Institute fur Informatik;
  Visit related to: Molecular Biology: Dates: 7/20/98 - 7/31/98
- **Phillipe Flajolet**; INRIA;
  Visit related to: Massive Data Sets: Dates: 4/16/98 - 4/22/98
- **John Franco**; University of Cincinatti;
  Visit related to: Massive Data Sets: Dates: 1/1/98 - 4/30/98

- **Ari Freund**; Technion;
  Visit related to: Discrete Probability: Dates: 9/6/98 - 9/24/98
- **Leslie Goldberg**; University of Warwick;
  Visit related to: Massive Data Sets: Dates: 5/4/98 - 5/15/98
- **Mordecai Golin**; Hong Kong University of Science and Technology;
  Visit related to: Massive Data Sets: Dates: 9/1/98 - 2/1/99
- **Marty Golumbic**; Bar Ilan University;
  Visit related to: General: Dates: 7/2/98 - 7/17/98
- **Sudipto Guha**; Stanford University;
  Visit related to: Massive Data Sets: Dates: 2/16/98 - 2/20/98
- **Vladimir Gurvich**; Russian Academy of Science;
  Visit related to: General: Dates: 9/10/98 - 11/6/98
- **Donovan Hare**; Okanagan University College;
  Visit related to: Large Scale Discrete Optimization: Dates: 9/14/98 - 11/28/98
- **Norbert Hegyvari**; Eotvos University;
  Visit related to: General: Dates: 4/4/98 - 4/28/98
- **Petr Hlineny**; Charles University;
  Visit related to: DIMATIA: Dates: 6/8/98 - 6/10/98
- **Piotr Indyk**; Stanford University;
  Visit related to: Massive Data Sets: Dates: 3/16/98 - 3/20/98
- **Giuseppe Italiano**; Hong Kong University of Science and Technology;
  Visit related to: Massive Data Sets: Dates: 5/19/98 - 5/24/98; 9/26/98 - 10/10/98
- **Lila Kara**; University of Western Ontario;
  Visit related to: DNA Computing: Dates: 4/17/98 - 4/21/98
- **David Karger**; MIT;
  Visit related to: Massive Data Sets: Dates: 3/23/98 - 3/25/98
- **Howard Karloff**; Georgia Tech;
  Visit related to: Networks: Dates: 9/28/98 - 10/9/98
- **Gyula Karolyi**; Eotvos University and ETH Zurich;
  Visit related to: General: Dates: 3/13/98 - 3/22/98
- **Meir Katchalski**; Technion;
  Visit related to: Discrete Probability: Dates: 12/9/98 - 3/28/99
- **Gyula Katona**; Math Institute - Hungarian Academy of Sciences;
  Visit related to: Erdos Institute Partnership: Dates: 4/24/98 - 4/30/98
- **Sergei Kazakov**; Somagenics, Inc.;
  Visit related to: DNA Computing: Dates: 2/7/98 - 2/9/98
- **Myung Ho Kim**; University of Auckland;
  Visit related to: General: Dates: 12/1/98 - 12/31/98

- **Valerie King**; University of Victoria;
Visit related to: Massive Data Sets: Dates: 5/15/98 - 5/22/98

- **Yoshi Kohayakawa**; Instituto de Matemtica e Estatstica;
Visit related to: General: Dates: 1/5/99 - 2/2/99

- **Lucek Kucera**; Charles University;
Visit related to: DIMATIA: Dates: 6/8/98 - 6/12/98

- **Cristian Lenart**; MIT;
Visit related to: General: Dates: 7/22/98 - 8/5/98

- **Simon Litsyn**; Tel-Aviv University;
Visit related to: Discrete Probability: Dates: 9/1/98 - 2/28/99

- **George Lueker**; University of California - Irvine;
Visit related to: General: Dates: 10/1/98 - 12/20/98

- **Bruce Maggs**; Carnegie Mellon University;
Visit related to: Networks: Dates: 2/26/98 - 2/28/98

- **Hosam Mahmoud**; The George Washington University;
Visit related to: Massive Data Sets: Dates: 2/1/98 - 6/30/98

- **J.L. Marichal**; University of Liege;
Visit related to: General: Dates: 10/19/98 - 10/24/98

- **Rossana Marra**; Universita di Roma Tor Vergata;
Visit related to: General: Dates: 12/1/98 - 12/23/98

- **Yossi Matias**; Tel-Aviv University;
Visit related to: Massive Data Sets: Dates: 3/7/98 - 3/10/98

- **Jiri Matousek**; Charles University;
Visit related to: General: Dates: 1/25/98 - 1/31/98

- **Marios Mavronicolas**; University of Cyprus;
Visit related to: Networks: Dates: 6/8/98 - 8/31/98

- **Dominic Mayers**; University of Montreal;
Visit related to: General: Dates: 2/16/98 - 2/22/98

- **Angus McIntyre**; Oxford;
Visit related to: General: Dates: 9/6/98 - 10/1/98

- **Roy Meshulam**; Technion;
Visit related to: General: Dates: 8/30/98 - 9/27/98

- **Friedhelm Meyer auf der Heide**; University of Paderborn;
Visit related to: General: Dates: 3/6/98 - 3/11/98

- **Tamara Munzner**; Stanford University;
Visit related to: Massive Data Sets: Dates: 5/19/98 - 5/24/98

- **Anca Muscholl**; Universitat Stuttgart;
Visit related to: Logic & Algorithms: Dates: 9/12/98 - 9/19/98

- **Susana Nascimento**; University Nova de Lisboa;
Visit related to: Molecular Biology: Dates: 11/17/98 - 12/10/98
- **George Nemhauser**; Georgia Tech;
Visit related to: Large Scale Discrete Optimization: Dates: 12/9/98 - 12/10/98
- **Jarik Nesetril**; Charles University;
Visit related to: DIMATIA: Dates: 6/3/98 - 6/13/98
- **Boaz Patt-Shamir**; Tel-Aviv University;
Visit related to: General: Dates: 7/2/98 - 8/2/98
- **Christian Pedersen**; BRICS;
Visit related to: Molecular Biology: Dates: 9/2/98 - 10/23/98
- **Vladimir Pinsky**; Geographical Institute of Israel;
Visit related to: General: Dates: 9/23/98 - 9/26/98
- **Octavian Procopiuc**; Duke University;
Visit related to: Massive Data Sets: Dates: 5/11/98 - 6/26/98; 8/10/98 - 8/21/98
- **Jihad Qaddour**; Mesa State College;
Visit related to: General: Dates: 3/22/98 - 3/28/98; 6/18/98 - 8/15/98
- **Yuri Rabinovich**; Ben Gurion University;
Visit related to: General: Dates: 8/10/98 - 8/21/98
- **Nikolaus Rajewsky**; University of Cologne;
Visit related to: General: Dates: 11/1/98 - 12/31/98
- **Andrea Richa**; Carnegie Mellon University;
Visit related to: Networks: Dates: 2/24/98 - 2/28/98
- **Klaus Ritter**; Technische Universitdt Muenchen;
Visit related to: General: Dates: 4/13/98 - 4/20/98
- **Neil Robertson**; Ohio State University;
Visit related to: General: Dates: 4/15/98 - 6/15/98
- **Dan Roth**; University of Illinois;
Visit related to: General: Dates: 6/29/98 - 7/3/98
- **Michael Roytberg**; Institute of Math;
Visit related to: General: Dates: 10/1/98 - 12/15/98
- **Suleyman Cenk Sahinalp**; University of Warwick;
Visit related to: Massive Data Sets: Dates: 3/13/98 - 4/20/98; 9/1/98 - 9/30/98
- **Marti Sanchez**; Universitat Politecnica de Catalunya;
Visit related to: General: Dates: 2/3/98 - 6/1/98
- **Farhad Shahrokhi**; University of North Texas;
Visit related to: General: Dates: 5/18/98 - 6/9/98; 10/2/98 - 10/14/98
- **Narayan Shivakumar**; Stanford University;
Visit related to: Massive Data Sets: Dates: 8/31/98 - 9/4/98

- **Craig Silverstein**; Stanford University;
  Visit related to: Massive Data Sets: Dates: 3/24/98 - 3/27/98
- **Bruno Simeone**; La Sapienza University;
  Visit related to: General: Dates: 4/23/98 - 5/14/98
- **J. Simonis**; Delft University of Technology;
  Visit related to: General: Dates: 7/20/98 - 8/16/98
- **Miklos Simonovits**; Mathematical Institute of the Hungarian Academy of Sciences;
  Visit related to: General: Dates: 9/30/98 - 10/31/98
- **Meera Sitharam**; Kent State;
  Visit related to: General: Dates: 2/2/98 - 2/14/98
- **Jozsef Solymosi**; Hungarian Academy of Sciences;
  Visit related to: General: Dates: 5/30/98 - 6/10/98; 10/2/98 - 10/25/98
- **Herbert Spohn**; Technisce Universitat Munchen;
  Visit related to: General: Dates: 10/10/98 - 10/24/98
- **Benjamin Sudakov**; Tel Aviv University;
  Visit related to: Discrete Probability: Dates: 9/28/98 - 10/20/98
- **Sandra Sudarsky**; Texas A&M;
  Visit related to: Massive Data Sets: Dates: 6/15/98 - 6/17/98; 7/24/98 - 8/21/98; 9/10/98-9/30/98
- **Alain Tapp**; University of Montreal;
  Visit related to: General: Dates: 3/24/98 - 3/31/98
- **Robin Thomas**; Georgia Tech;
  Visit related to: General: Dates: 7/20/98 - 9/6/98
- **Paolo Toth**; University of Bologna;
  Visit related to: Large Scale Discrete Optimization: Dates: 10/28/98 - 10/30/98
- **Gyorgy Turan**; University of Illinois at Chicago;
  Visit related to: General: Dates: 6/15/98 - 6/18/98
- **Giuseppe Valla**; University of Genova;
  Visit related to: General: Dates: 1/12/98 - 4/23/98
- **Pavel Valtr**; Charles University;
  Visit related to: DIMATIA: Dates: 6/1/98 - 7/3/98
- **Remi Vanicat**; Ecole Normale Superieure de Lyon;
  Visit related to: Logic & Algorithms: Dates: 6/6/98 - 8/29/98
- **Vijay Vazirani**; Georgia Tech;
  Visit related to: Large Scale Discrete Optimization: Dates: 8/20/98 - 12/31/98
- **Suresh Venkatasubramanian**; Stanford University;
  Visit related to: Massive Data Sets: Dates: 2/2/98 - 2/6/98

- **V. Vinay**; Indian Institute of Science-Bangalore;
  Visit related to: Logic & Algorithms: Dates: 6/7/98 - 6/13/98
- **Jeff Vitter**; Duke University;
  Visit related to: Massive Data Sets: Dates: 5/19/98 - 5/24/98
- **Rafe Wenger**; Ohio State University;
  Visit related to: General: Dates: 1/20/98 - 4/3/98
- **Thomas Wilke**; Christian-Albrechts-Universitat zu Kiel (CAU Kiel);
  Visit related to: Logic & Algorithms: Dates: 8/24/98 - 9/25/98
- **David Wonnacott**; Haverford College;
  Visit related to: Large Scale Discrete Optimization: Dates: 9/10/98 - 12/18/98
- **Xuerong Yong**; Hong Kong University of Science and Technology;
  Visit related to: Massive Data Sets: Dates: 8/25/98 - 2/1/99
- **Yuan Ping Zhang**; Hong Kong University of Science and Technology;
  Visit related to: Massive Data Sets: Dates: 8/25/98 - 2/1/99

C Awards and Honors

- Sanjeev Arora, Princeton University: Packard Foundation Fellowship; Plenary lecture at ACM Symposium on Theory of Computing, 1998
- Robert Calderbank, AT&T Labs: 1998 IEEE Fellow
- Jonathan Eckestein, Rutgers University: Board of Trustees Research Fellowship for Scholarly Excellence
- Joan Feigenbaum, AT&T Labs: Invited Speaker, 1998 International Congress of Mathematicians; invited Plenary Speaker, 1999 American Mathematical Society Joint Winter Meetings; invited Plenary Speaker, 1999 SIAM Annual Meeting; also the subject of a chapter in *Notable Women in Mathematics: a Biographical Dictionary*, ed. by C. Morrow and T. Perl, Greenwood Press, Westport, CT, 1998
- Lee Giles, NEC: Paper listed for best paper award at 3rd ACM Conference on Digital Libraries.
- Simon Gindikin, Rutgers University: State Award of Russia 1998
- Jerry Goldin, Rutgers University: Humboldt Research Prize from the Alexander von Humboldt Foundation
- Peter Hammer, Rutgers University: Honorary Doctorate (Laurea Honoris Causa) from La Sapienza University 1998
- Frank Hwang: Joint winner (with Dr. Dingzhu Du) of the CSTS (computer science technical section) award at INFORMS, April 1998
- David Johnson, AT&T Labs: One of two co-winners of the first AT&T Mentoring Award.
- Jeff Kahn, Rutgers University: Rutgers Board of Trustees Award for excellence in research, May 1998
- Janos Komlos, Rutgers University: Elected to the National Academy of Hungary, May 1998
- Joel Lebowitz, Rutgers University: Received the AAAS Scientific Freedom and Responsibility Award, 1998
- Andrew Odlyzko, AT&T Labs: Keynote lecture at the 10th anniversary of Mathematica Conference; invited hour lecture at the 1998 summer meeting of the MAA
- Michael Rothkopf, Rutgers University: 1997 George E. Kimball Medal from INFORMS for service
- Saharon Shelah, Rutgers University: Received the Israel Prize in math 1998
- Lawrence Shepp, Rutgers University: Chosen as a Miller Research Fellow at Univ. of CA/Berkeley
- Jaswinder Pal Singh, Princeton University: Presidential Early Career Award for Scientist and Engineer (PECASE); Sloan Research Fellowship

- Neil Sloane, AT&T Labs: Elected to National Academy of Engineering, Feb. 1998; 1998 Claude E. Shannon Award, IEEE Information Theory Society; invited speaker, International Congress of Mathematicians, Berlin, 1998
- Peter Shor, AT&T Labs: Received the 1998 Rolf Nevanlinna Prize given by the Univ. of Helsinki which honors the work of young mathematicians (less than forty years of age) in the mathematical aspects of information science.
- David Waltz, NEC: President of AAAI (American Association for Artificial Intelligence) 1997-99; invited keynote address at AAAI-98 Madison, WI
- Andrew Yao, Princeton University: Elected member of the US National Academy of Sciences in April 1998
- Mihalis Yannakakis, Bell-Labs: ACM Fellow, 1998; Bell Labs Fellow, 1998; invited Plenary lecture at the joint MFCS/CSL conference