D I M A C S

Center for Discrete Mathematics and Theoretical Computer Science

ANNUAL REPORT

December 1999

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, Bell Labs, NEC Research, and Telcordia Technologies

Contents

1 Summary

In this report, we summarize the highlight accomplishments at DIMACS in the period February 1, 1999 to January 30, 2000 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 eleventh 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. 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. 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. As special years tend to last longer and longer, sometimes as long as six years, we are in the process of calling all of them special foci. During the academic year 1998-99, we ran six special focus activities. We introduced one major new special focus, on Large Scale Discrete Optimization, and continued our special foci on Networks, Massive Data Sets, DNA Computing, and Mathematical Support for Molecular Biology. We also continued our Discrete Probability Special Focus in a very modest way. In the Fall of 1999-2000, we launched our new Special Focus on Computational Intractability and continued Networks, Mathematical Support for Molecular Biology, and Large Scale Discrete Optimization. In Spring 2000, we will launch a new Special Focus on Next Generation Networks Technologies and Applications (NGN), a major new effort that is expected to run until at least Summer 2003. Our Networks special focus is being supported in part by the National Security Agency, allowing us to extend it until the year 2000. NSA is also supporting our new NGN activity, as is ONR. We received an NSF grant to support Computational Intractability.

We have worked very hard to plan future special foci. In Fall 2000, we will start a new Special Focus on Computational Molecular Biology, the successor to our six-year focus on Mathematical Support for Molecular Biology. We have received an award from the Alfred P. Sloan Foundation to fund this new special focus for three years starting Fall 2000. In addition, we have very detailed plans for three other future special foci. Computational Information Theory and Coding is expected to start in Fall 2001 and last for two years. Data Analysis and Data Mining will also start in Fall 2001 and last for three years. Computational Geometry and Applications, a return to DIMACS' first special year, but with a more applied orientation, will start in Fall 2001 or Fall 2002.

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. They can also be found at our website at http://dimacs.rutgers.edu/About/index.html#reports

Multi-disciplinary, Multi-institutional Research Projects

In addition to special years, DIMACS has continued major multi-disciplinary, multi-institutional research projects. One such project on "Scalable Self-Organizing Simulations" (S3) has been carried out under DARPA support. DARPA has extended its support for a year and we are hopeful of additional funding for this and its successor project. Its predecessor project, "Simulations of Integrated Communications Systems" (SICS), supported by an NSF grant, was concluded in 1999. As an outgrowth of these two projects, we received an award for our work on "C41SR Global

Network Simulator." The award is a subcontract from the Ultra Corporation, one of the small companies that we have partnered with on the DARPA project. See Section 3.10 for a description of the research in our S3, SICS, and Ultra Corporation projects.

Continuing and Exploratory Activities

It is the plan of DIMACS to return from time to time to successful past foci. The 1995-96 Special Year on Logic and Algorithms is a case in point. We returned to it in 1999 with a workshop on Algebraic Methods and Arithmetic Circuits as part of our international partnership with the DIMATIA Center in the Czech Republic and a workshop on Logic and Cognitive Science in partnership with the science and technology center, the Institute for Research in Cognitive Science at the University of Pennsylvania. The 1991-92 Graph Theory and Algorithms Special Year was followed up by an emphasis on graph theory in the summers of 1998 and 1999 at the DIMACS Research and Education Institute (DREI) and similar emphasis in our upcoming DIMACS Connect Institute. One way to develop topics for possible future special foci is to hold exploratory workshops. We explored a variety of such topics between February 1999 and January 2000 with workshops on financial modeling, codes and association schemes, graph theory and computer vision, and unusual applications of number theory, to give a few examples.

Evaluation Reports on Special Years

We continued our program of evaluating the impact of earlier special years by commisioning an evaluation of DIMACS' fourth special year, the 1992-93 Special Year on Combinatorial Optimization, as seen from the perspective of six years. We have now received six such reports. The report, which is discussed in Section 3.9, provides an exceptionally strong piece of evidence of the long-term impact of DIMACS' progams. This year, we have also commissioned a report on our 1996-99 Special Focus on Discrete Probability.

We often face the challenge of how to let people know how well we are doing. Reports like these really help make the case. We have been developing other materials that tell the story of the achievements of DIMACS. Reports of accomplishments of particular special years are on our website at the address http://dimacs.rutgers.edu/About/index.html#reports. This year, we completed a compilation of Scientific Achievements at DIMACS in its second four years. This and our compilation of Scientific Achievements at DIMACS in its first four years are at the same web address. We have also revised and expanded a brochure summarizing the activities and achievements of DIMACS, in layman's terms.

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

1998 brought to fruition the hard work of previous years to develop new and innovative programs in undergraduate education. We launched a new modules 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 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 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 Modules Series. Our traditional undergraduate program, the undergraduate research program, continued with 12 students. We won an award from NSF to combine our traditional REU program with an international experience, inviting some students from our sister center DIMATIA in Prague, Czech Republic, to join us at DIMACS for five weeks, and then taking five of our REU students to DIMATIA for three weeks. 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 (LP) 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. We have received from NSF a one-year supplement to keep our LP going for the summer of 2000 and launch our franchise model. NSF has phased out support of Young Scholars programs (YSP's) like ours. We turned instead to a combination of tuition payments and gifts to run the program, and were pleased to receive a series of major gifts from the AT&T Foundation to support it. Indeed, an additional $40,000 has just been received from that organization, guaranteeing that our summer 2000 YSP will run successfully. In earlier years, we played a central role in the adoption of new "Standards" for mathematics and a new Curriculum Framework in Mathematics in New Jersey. The "Framework" and the 1997 DIMACS volume, *Discrete Mathematics in the Schools*, viewed as exemplary models, have begun to have a real impact on mathematics education in the schools locally, nationally and internationally.

Graduate Student Programs

Graduate students are a central part of a science and technology center. Our *local* graduate student program consists of small DIMACS awards for travel and software to support the research of local students, small awards for summer and winter break salary to support research projects, and jobs and projects for which graduate students are paid. The response to this program has been highly enthusiastic. We are also very pleased at the number of DIMACS researchers, including individuals from our industrial partners, who are participating in this program as research mentors. We supported 42 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. The impact of our program on non-local students has been very high and the research of some of our non-local graduate student visitors is also included in Section 3.

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, in the past, with the University of Pennsylvania, obtaining grant support for postdocs, and including in our postdoc program postdocs appointed by our industrial partners. We have reached agreement with the Institute for Advanced Study to continue our joint postdoc program for three more years, and have identified resources for our share of the cost, with special help from Rutgers University. AT&T Labs has given us a gift for support of two joint postdocs, another way we are now beginning to leverage our postdoc funds, and NECI has done the same for one joint postdoc. We have plans for a similar arrangement with IBM Research. 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 modules series and the "Reconnect" project.

We also ran the fourth and final year of our DIMACS Research and Education Institute (DREI), under a $1.9M grant from NSF. This program has been 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. Because DREI was so successful, we wanted to find a way to continue it. We spent a great deal of time this year designing a successor program, the DIMACS Connect Institute, DCI. DCI has much the same goals as DREI, but it involves a two-summer commitment by teachers and puts more emphasis on a true research experience for the teachers. There will also be more emphasis on an academic year in-classroom and outside-of-classroom experience for teachers and their students, following up our summer program. We have succeeded in getting an award from NSF to run DCI for three years, starting in summer 2000, with two additional years to be run under funds we are in the process of raising.

Institutional Membership

We currently have six institutional members. The most important thing our institutional members bring to us is the encouragement they give individual scientists at their institutions to be involved in DIMACS activities. This is an especially vital contribution of our industrial partners. However, in this time of transition from our STC award, we have asked our partner institutions to support us with increased financial contributions and we are pleased to report that they are coming through.

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 have run a joint postdoc and workshop program with the University of Pennsylvania, which continued with a joint workshop in 1999. We have continued several joint activities with the Center for Molecular Biophysics and Biophysical Chemistry at Rutgers and that center is a crucial player in our new Special Focus on Computational Molecular Biology. We held workshops sponsored jointly with a variety of institutions, including Georgia Tech, MIT, IBM Research, the Institute for Mathematics and its Applications (IMA) at University of Minnesota, and the Electric Power Research Institute (EPRI). Future joint workshops are in the works, with IBM and the Society for Industrial and Applied Mathematics as two of our future partners.

On the international front, our contacts with other institutions have been formalized and have matured into joint programs. During this year, we have run a workshop with our partner in the DIMATIA Center in Prague (with NSF support) and one with the new Erdős Institute in the Hungarian Academy of Sciences, and we have exchanged visitors with these partner institutions and others such as INRIA and the Hong Kong University of Science and Technology. Further international collaborations are in the process of being developed, especially with INRIA in connection with our planned Special Focus on Next Generation Networks Technologies and Applications.

Programs for Underrepresented Groups

We continued our collaboration with the Douglass College (Rutgers University) Program for Women in Mathematics, Science, and Engineering. Fred Roberts served on the Advisory Board for this program. 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. Three prominent African American scholars have agreed to help us recruit African American participants to our programs and in particular to work with us on our Reconnect Program. We opened discussions about joint programs with the University of Puerto Rico and Texas Southern University and continued discussions with Spelman College and Howard University about joint activities. In particular, we have arranged for Texas Southern to be involved in our Leadership Program as a franchise site.

Changes in Personnel

This has been a rather chaotic year for us in terms of the staff. Elaine Foley, our DREI and Reconnect coordinator and staff member in charge of our evaluation and some of our publicity programs left in April on short notice and was not replaced. On short notice, Chris Spassione, our Visitor Coordinator, stepped in and did a remarkable job in supporting our DREI and Reconnect activities. Julia Abrahams, our Associate Director, left on short notice in August and has also not been replaced. Many of her duties have been taken on by Director Fred Roberts and by Center Administrator Sarah Donnelly. Two DIMACS members, Professors Michael Fredman (PM, RU) and Michael Saks (PM, RU), are also taking on a few of the jobs that Dr. Abrahams had been doing. We do not feel we can afford the luxury of a full-time associate director at this time. In September Christine Houck, our publicity coordinator, went out on maternity leave (earlier than expected) and was away for almost four months. Also in September, our workshop coordinator, Patricia Pravato, passed away. She had worked for DIMACS since its inception and was personally responsible for making so many of our visitors feel like DIMACS was family to them. She is sorely missed. For fourth months in Fall 1999, therefore, we operated with four fewer staff members than we had been used to, with all of the remaining staff members carrying the burden. It is remarkable that we were able to continue all planned activities of the center. We have found a temporary Workshop coordinator and are in the process of interviewing permanent candidates for that position. Nevertheless, we will be operating with a smaller staff in the future.

DIMACS Book Series

The DIMACS series, published by the American Mathematical Society with high quality hard cover volumes, reached a milestone with the publication of the 50th book in the series in 1999.

Transition to the Post-STC Era

DIMACS was founded in 1989 as an NSF "science and technology center" with an 11 year grant that ends (with a recent extension) on July 31, 2000. The STC grants were originally intended to be nonrenewable and none of them, including ours, has been renewed. Programs under the DIMACS

11-year science and technology center award end on January 31, 2000, with the remaining six months of the STC award usable for STC closeout activities. We are very proud of the many accomplishments of DIMACS as an STC.

When the STC program was designed, it was envisioned that the science and technology centers would become "self-sufficient" after 11 years, through a combination of state, federal, university, and corporate support. During the past year, indeed during the past four years, we have devoted a great deal of time and energy to the transition to a post-STC life for the center. We have sought to raise funds for a variety of the programs of the center while making sure we have the resources to maintain the infrastructure upon which those programs depend.

The signature program of DIMACS is the special focus and our primary goal has been to make sure that we can continue to run special foci. Our strategic plan is to run a number of special foci at the same time and to run them for longer periods. Special foci can be run at different levels and we have worked hard to make sure that we at least have the resources to run special foci at a modest level. With the help of a variety of agencies, we have succeeded in this regard. NSF and the National Security Agency have helped us with the continuation of the Special Year on Computational Intractability. The Alfred P. Sloan Foundation has helped us with the Special Focus on Computational Molecular Biology. The National Security Agency and the Office of Naval Research have helped helped us with the Special Focus on Next Generation Networks Technologies and Applications. With this support and the help of our member organizations, we are assured of funding to run these major programs at some level through at least 2003. More is needed and we also need to find funds for the planned special foci that will commence in 2001 and later. However, we have made a strong start in assuring the continuation of DIMACS' most vital programs.

DIMACS has attained a position of leadership through its numerous highly successful programs in education and integrating research and education. We are very happy to have obtained NSF support for the continuation of one of our major programs through the funding of the DIMACS Connect Institute, support from the AT&T Foundation for the continuation of our exciting Young Scholars Program, and a supplement from NSF to continue our Leadership program and launch our franchise model. We are also delighted to have obtained NSF support for an international version of our REU program.

DIMACS will continue to have a variety of visiting opportunities. Our special focus activities will have funds for visitors. Other visitors will be accommodated as funds allow. We intend to continue to host graduate students from outside the center's institutions. And we will continue to have graduate student and postdoctoral programs through our special foci and through partnered programs with the Institute for Advanced Study and a variety of corporate partners.

The infrastructure of the center, its staff and facilities, are crucial to the success of our programs. Our university partners continue to make available to us the facilities and equipment that are needed to run our many programs. Some of the funding we have obtained from outside agencies will provide us with support for staff. However, finding funding for the staff needed to run our programs and for some of the general operating expenses of the center is perhaps our biggest challenge. Our partner organizations have made a commitment to the future of the center with increased contributions that can be used to support these infrastructure costs. These contributions will also be used to support some programs that are very important to the center such as the program of visitors not specifically tied to a special focus, the program of exploratory and specialty workshops, and some of our international partnerships.

The New Jersey Commission on Science and Technology continues to support DIMACS as a State of New Jersey "match" for our federal funds. The Commission money goes to support the whole range of DIMACS programs and is very important to DIMACS.

DIMACS continues to exist to serve its broad constituency, which includes researchers in discrete mathematics and theoretical computer science from around the world; workers in other fields of science, the mathematical sciences, or industry who work in fields that could benefit from or influence progress in dm or tcs; and teachers and students at all levels.

DIMACS owes a vote of thanks to all of those who have helped it in its transition from its status as an NSF science and technology center. We are particularly grateful to the many volunteers who have helped or will be helping to organize programs, work on proposals to funding agencies, and spread the word about DIMACS. We are also particularly grateful to the many institutions which are helping either financially or through encouraging their employees to volunteer to help at the center. With continued help and good will, we will be able to continue to serve our constituencies well in the future.

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, NEC Research Institute (NECI), and Telcordia Technologies (formerly Bellcore) and their scientific personnel, as well as their colleagues nationwide, working together to play a national leadership role in the development, application and dissemination of discrete mathematics (dm) and theoretical computer science (tcs) by

- (P1). promoting and supporting high quality research within these fields, especially research areas that are poised for rapid advancement and have a potential for high impact;
- (P2). promoting and supporting the application of the methods and results of these fields in other areas of science, the mathematical sciences, and industry;
- (P3). encouraging and facilitating efforts of researchers in these fields to draw more of their research inspiration from other areas of science, the mathematical sciences, and industry;
- (P4). informing the broader mathematical community about the research currently taking place in the fields of dm/tcs and involving them in these fields; and
- (P5). promoting and facilitating the incorporation of contemporary mathematical ideas and techniques in these fields into classrooms at all levels.

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, Bell Labs/Lucent Technologies, NEC Research, and Telcordia Technologies (formerly Bellcore). 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"/"special foci" and "educational outreach activities." These activities receive the large majority of our center's funds.

2.3.3 Special Years/Special Foci

A "special year" or "special focus" refers to a focused and coordinated collection of activities that span a period of time, usually at least six months. We use the terms "special focus" and "special year" somewhat interchangeably. As our "special years" are tending to last for more than one year, we are moving to call all such activities "special foci." 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;

- (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. Two examples of such partnerships are the DIMACS Research and Education Institute and the DIMACS Connect 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, and its successor program, the DIMACS Connect Institute, are 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 workshop on "Combinatorics of Statistical Mechanics Models" organized by DIMACS in the Spring of 1999 and extending our Special Focus on Discrete Probability, and the research program on graph theory in the summers of 1998 and 1999 at the DIMACS Research Education Institute, following up on an early special year on Graph Theory and Algorithms. 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 for 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 DIMACS Connect 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 and Applications 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, Bell Labs, NECI, Telcordia Technologies, 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 maximize 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
TC = Telcordia Technologies (formerly Bellcore)
tcs = theoretical computer science
V = visitor.

3 Research Accomplishments and Plans

3.1 Special Year on Large Scale Discrete Optimization 1999-00

Overview:

The Special Year on Large Scale Discrete Optimization began in September 1998 and continues through much of 1999-00 at a considerably lower level than in 1998-99. Some pre-special year activities were held during 1997-98. The year builds on the 1992-93 Special Year on Combinatorial Optimization. (In 1999, we continued our project of evaluating the impact of earlier special years by obtaining an evaluation of that Special Year on Combinatorial Optimization. We discuss the results of that evaluation in Section 3.9.)

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
Carnegie Mellon University
The Electric Power Research Institute (EPRI)
Georgia Institute of Technology
IBM Research
Institute for Advanced Study, Princeton
NEC Research Institute
Princeton University
Rutgers University
Telcordia Technologies

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 February 1999 - January 2000

- **DIMACS Workshop on Large Scale Discrete Optimization in Logistics**
  - Dates: February 8-10, 1999
  - Location: Rutgers University
  - Attendance: 79
  - Organizers: George Nemhauser and Martin Savelsbergh, Georgia Tech
  - Comment: Joint with Special Year on Networks

  An efficient and effective logistics system is the key to being competitive in today's global marketplace. This realization has fueled an enormous interest in new technology for designing and operating logistics systems. Discrete optimization frequently is an integral part of such technology. In particular, in the last decade, the use of large-scale discrete optimization models in transportation and production planning has increased dramatically. This has been made possible by developments in modeling, algorithms, software, and hardware. This workshop brought together leading researchers to present state-of-the-art research activities and to discuss innovative future directions of research for this very active field with significant practical importance. The talks covered a wide range of topics including modeling, algorithms, and applications. Application domains that were covered include network design, production planning, vehicle routing, and crew scheduling. The workshop opened with a talk by William Cook titled "On the Solution of Traveling Salesman Problems". Research on the solution of the traveling salesman problem, the prototypical discrete optimization problem, has stimulated enormous interest in discrete optimization and has helped significantly in the growth of the field. Several other talks focused on methodology, such as "Approximation Algorithms via Linear Programming: Rounding Algorithms" by David Shmoys, "An Overview of Classical and Modern Heuristics for the Vehicle Routing Problem" by Gilbert Laporte, "GRASP: Greedy Randomized Adaptive Search Procedures" by Mauricio Resende, "AI Techniques for Large-Scale Discrete Optimization" by Jimmy Crawford, and "Very Large Scale Neighborhood Search" by Ravi Ahuja. Furthermore, there were several talks related to the main application domains of discrete optimization. In the topic of Airline Optimization, we had "Solving Aircraft Maintenance Scheduling Problems" by Daniel Bienstock, "Optimization Problems in Air Traffic Management" by Mike Ball, and "Airline Integrated Recovery: AIR and SimAIR" by Ellis Johnson. In the topic Manufacturing and Production Planning, we had "BC-PROD and LOTSIZELIB" by Laurence Wolsey, "Optimization in IBM Manufacturing" by Brenda Dietrich, and "Painting Cars and the Traveling Salesman Problem" by Tom Magnanti. In the topic Supply-Chain Management, there was "Optimal and Approximation Algorithms for Large Scale Supply Chain Management Problems" by David Simchi-Levi and "Dynamic Programming Approximations for the Discrete Multi-commodity Network Flow Problem over a Dynamic Graph" by Warren Powell. Finally, in the topic Telecommunication Network Design there was "The Prize Collecting Steiner Tree Problem" by David Johnson. There was also a full group discussion on future directions for the field. The group was very encouraged by the increasing use of discrete optimization methods in industry and the need for the solution of larger more complex models that, for example, include stochastic elements. In conclusion, the workshop provided an excellent environment in which leading researchers in the field could share their ideas not only with each other, but with junior researchers and students as well.

- **DIMACS Workshop on Large Scale Discrete Optimization in Robotics and Vision**
  - Dates: March 21-23, 1999
  - Location: Rutgers University
  - Attendance: 72
  - Organizers: Jitendra Malik, University of California - Berkeley, and Jean Ponce, University of Illinois - Urbana Champaign

The theme of the workshop was to address the fundamental mathematical and computational problems underlying computer vision and robotics. We were able to attract many of the major researchers in the field to help give a perspective on the leading issues today and the likely directions for future progress. Points of contact with computer science theory and algorithms, probabilistic inference, and other areas of applied mathematics in general were stressed. Biological vision is used to guide a number of different aspects of behavior–guiding locomotion, controlling manipulation and recognition of objects. Computer vision seeks to develop models and algorithms for these competences, drawing heavily on geometry, probability and statistics and computation. Likewise, robotics is concerned with controlling mechanisms such as manipulator arms, grippers and mobile platforms using sensory input as well as a priori knowledge about the geometry and physics of these mechanisms and their environment. This discipline is at the boundary between fields as diverse as mathematics, mechanics, control theory and computer science. The morning and early afternoon of the first day was devoted to the modeling and inference challenges in vision. David Mumford articulated the case for probabilistic modeling in a Bayesian framework as the natural way to incorporate both low level and high level knowledge for solving vision problems. He argued for particle filtering as a convenient formalism for representing and computing with probability distributions. David Forsyth demonstrated the power of sampling-based methods with algorithms for three different computer vision problems–color constancy, structure from motion, and finding people in images. Jitendra Malik and Ramin Zabih presented solutions to problems of image segmentation and stereoscopic correspondence in graph theoretic formulations. It turns out that spectral graph theory and various minimum cut formulations, which have been much studied by the theoretical computer science community, can provide practical and efficient algorithms for these problems. The morning session of the second day was again devoted to vision, primarily to object recognition. Sven Dickinson presented a theory for the generic representation of 2-D shape, in which structural descriptions are derived from the shocks (singularities) of a curve evolution process acting on a silhouette. David Kriegman described recognition algorithms that are based on characterizing the set of images of an object in fixed pose but under a range of lighting conditions, and demonstrated their performance in face recognition tasks. Martial Hebert presented a new approach to object recognition in range data that deals particularly well with the clutter present in complex industrial environments. The afternoons of both the first and second day had a series of talks on robotics. Jean-Paul Laumond gave an overview of classical configuration space concepts and modern approaches to motion planning, then focused on his recent work using probabilistic roadmaps methods to solve a variety of hard robotics problems such as obstacle avoidance for manipulators and navigation for non-holonomic mobile robots. Bud Mishra surveyed the field of grasp planning, presenting its geometric and algorithmic foundations before concentrating on his recent work aimed at using simple devices in efficient algorithms for grasping applications. Greg Hager investigated new theoretical aspects of visual servoing. In particular, he provided some answers to the question: "When is it possible to decide, on the basis of the information contained in the images of an imprecisely modeled two-camera vision system, whether or not a prescribed robot positioning task has been precisely accomplished?" There were four talks on the concluding day of the conference. Jean Ponce and Mark Overmars led off with the application of computational geometry to design of fixtures for part manipulation. Mark Overmars addressed the problem of constructing efficient algorithms for automatically designing conveyor part sorting systems using belts equipped with fences and vibratory bowl feeders. Jean Ponce concentrated on a new approach to manipulation planning based on the concept of capture regions, and illustrated it with experiments using a reconfigurable gripper and a part feeding mechanism. Andrew Zisserman gave a throrough survey of the state of the art in structure from motion–both the theory of the problem and some impressive reconstruction results from video sequences. The concluding talk of the workshop was given by Jan Koenderink, who has been the pioneer in the study of shape and the application of differential geometry and singularity theory to vision. In this talk he presented his experimental results on the perception of pictorial relief. In addition to the series of talks, we had two panel discussions, one on the first day "Can Probabilistic Inference in Vision be Made Computationally Tractable" and one on the second day "Vision and Robotics: Growing Apart or Coming Together?" Both were lively affairs with much participation from the audience. We also had a submitted poster session with a dozen participants. This gave an opportunity for various graduate student and postdoctoral fellows to present their research and get feedback.

- **Mini-Workshop on Mathematical and Computational Approaches to Elections**
  - Date: June 18, 1999
  - Location: Rutgers University
  - Attendance: 38
  - Organizer: Bruno Simeone, La Sapienza University, Rome, Italy

In recent years, the debate on electoral reform has been very lively both in newborn and old democracies. Although the choice of an appropriate electoral system is affected by political and social factors, and as such it certainly belongs to the domain of social sciences, it is also evident that the formal aspects and procedures of voting naturally call for the use of quantitative methods for the analysis and design of such systems. Several quantitative approaches have been developed, such as axiomatics (in the more general context of social choice theory), game-theory, geometry, statistics, optimization, and discrete algorithms. The workshop aimed to discuss and compare these approaches and to pinpoint or foresee connections between them. Since electoral systems obviously involve only finite (although possibly quite large) sets and, on the other hand, they essentially provide procedures to transform votes into seats and to assign seats to elected candidates, combinatorial and algorithmic aspects are of paramount importance in their study. All of the speakers were well recognized for their individual contributions to voting systems, and altogether they indeed offered a comprehensive and broad view of this area. Fred Roberts opened the Workshop with a thorough discussion on the axiomatics of the mean and the median, two widespread social choice functions. He presented many elegant combinatorial results concerning these two functions in the areas of location theory, graph theory, and molecular biology. Steven Brams's talk, "Voting Models: Explaining Practice and Recommending Reform", dealt with three topics: resource-allocation models in US presidential campaigns, the establishment of approval voting procedures, and a new "paradox of multiple elections" occurring in referendums with multiple questions. His focus was on the implementation and validation of electoral procedures. Peter Fishburn talked about "Electoral Axiomatics and Combinatorics", demonstrating quite convincingly that elections and their axiomatics provide a rich source of intriguing combinatorial problems. Michel Regenwetter, in his talk on "Probabilistic Models of Social Choice and Preference", presented some preference and choice models blending probabilistic tools with ordinal utility theory. Here too significant combinatorial results are obtained. Donald Saari's keen and stimulating presentation "From Arrow's and Sen's Theorems to Resolving all Vote Paradoxes" offered some fresh insights on classical Impossibility Theorems. A careful reading of these results suggests one possible way out of related paradoxes. Bruno Simeone, in his talk "Discrete Optimization Models for Electoral Systems", pointed out that discrete optimization, on the one hand, provides a broad conceptual framework for electoral procedures; on the other hand, it yields a powerful tool-kit for designing different components of an electoral system. He illustrated these two points showing that all proportional formulas are obtainable via the minimization of suitable Schur-convex functions, or generalizations thereof; and describing a multi-criteria graph partitioning model for political districting. The last talk, Michael Trick's "Computational Complexity and Algorithmic Aspects in Voting Theory", focused on computational complexity issues in determining winners, manipulating elections, and strategic voting. The participants found the workshop interesting and stimulating. There was general agreement that the piece de resistence of the workshop was the excellent mix of different approaches.

- **Discrete Optimization '99 (DO '99)**
  - Dates: July 25-30, 1999
  - Location: Rutgers University
  - Attendance: 102
  - Organizers: Peter Hammer and Endre Boros, Rutgers University
  - Co-Sponsor: RUTCOR (Rutgers Center for Operations Research)

Discrete Optimization underwent dramatic development in the last half century, with particularly spectacular growth in the last few decades. The main goal of this workshop was to survey the state of the art in D.O. There were a series of expository lectures presenting the major subareas of the field, including its theoretical foundations, its methodology and applications, with the surveys presented by some of the most prominent researchers in the field. There was also a forum for the presentation of new developments in D.O. through a series of contributed papers sessions. D.O.'99 was held 22 years after D.O.'77, which had very similar goals, and whose collection of surveys (*Discrete Optimization I and II*, Annals of Discrete Mathematics, vols. 4 and 5, ed. by P.L. Hammer, E.L. Johnson, and B.H. Korte, North-Holland, Amsterdam, 1979) is still frequently used. On the first day, Tom Magnanti presented a survey on network design, Egon Balas one on "Lift and Project," and Paolo Toth one on vehicle routing. Surveys on the second day were on complex scheduling problems by Peter Brucker, graph connectivity by Toshihide Ibaraki, greedy algorithms in linear programming problems by Alan Hoffman, bin packing problems in two and three dimensions by Silvano Martello, and ideal clutters by Gerard Corneujols. Day three featured surveys by David Williamson on approximation algorithms, Dorit Hochbaum on "A Unified Technique for Approximation Algorithms for Integer Programs with Three Variables per Inequality", Andras Recski on combinatorics of grid-like frameworks, Jim Orlin on neighborhood search, and Fred Glover on Tabu search and evolutionary developments. On Day 4, Henry Wolkowicz presented a survey on semidefinite programming for D.O., Yves Crama one on optimization models in production planning, and Rainer Burkard one on assignment problems. The fifth day saw surveys by Bill Pulleyblank on "Hilbert Bases, Caratheodory's Theorem and Integer Programming," by Laurence Wolsey surveying mixed integer programming, Vasek Chvatal on the solution of the traveling salesman problem, Bernhard Korte on the subject "Gigahertz-processors Need Discrete Optimization," and Jakob Krarup on locational decisions. The final two surveys were given on day 6 by Endre Boros on pseudo Boolean optimization and Peter Hammer on Boolean optimization. Jakob Krarup led a lively roundtable discussion on perspectives and trends in D.O. The organization of the workshop was made possible by combining funding through a DIMACS allocation and an NSF grant to the Rutgers Center for Operations Research, the co-sponsor of the program, for the support of young participants. Arrangements have been made for the publication of two special issues of *Discrete Applied Mathematics*, one dedicated to the survey lectures, and the other to the contributed papers presented in the workshop. Moreover, the volume of survey lectures will be reprinted in book format and made available to the public at large.

- **DIMACS - IBM Conference on Theory and Practice of Integer Programming in Honor of Ralph Gomory on the Occasion of his 70th Birthday**
  - Dates: August 2-4, 1999
  - Location: IBM Watson Labs, Yorktown Heights, NY
  - Attendance: 120
  - Organizers: Bill Pulleyblank, IBM Watson Research Center, and William Cook, Rice University
  - Co-sponsor: IBM Research

Integer Programming is a discipline which underlies much of modern Operations Research, including scheduling, logistics, resource allocation and routing problems. It is also the subject of much theoretical research, combining the study of linear convex polyhedra with integer lattices and group theory. Because of its importance to applications, it is also the focus of significant amounts of computational research, leading to software which is much more powerful than what was available a few years ago, but which still is unable to deal with real world problems of the required size and complexity. Recently we have seen significant advances in linear programming, an underlying problem, as well as the emergence of totally new methods, based on geometric principles. In addition, the broad availability of high performance scientific workstations is now making it possible for a much broader community to take advantage of integer programming software. The workshop brought together a who-is-who collection of international experts covering all aspects of this field. It provided a unique opportunity for a focused study of these advances and for exploration of the synergy possible by combining the various subfields. The attendance each day exceeded one hundred people, including a substantial number of graduate students and young researchers. The feedback we received from the participants was extremely positive. One of the senior participants wrote: "It was the best workshop I ever attended and the atmosphere was terrific." Another wrote: "It felt like a watershed event didn't it?" A lecturer in the final session stated that it was the first time in many years that he attended every lecture of a workshop. Vasek Chvatal opened the conference by talking about Gomory cuts and combinatorial reasoning. Herbert Scarf described homotopy methods for computing test sets. A featured talk by Egon Balas revisited disjunctive programming. George Nemhauser described large-scale integer programming in airline scheduling. Balanced matrices was the subject of the featured talk by Gerard Cornuejols. These and 26 other lectures covered a good mix of theory, computation, and applications. Theoretical topics included homotopy methods, the geometry of numbers, commutative algebra, Hilbert bases, and the structure of 0-1 polytopes. Computational lectures covered large-scale tests of cutting planes, tests of relaxation methods for linear programming, and branch-and-cut algorithms to solve combinatorial problems. The applications discussed included airline crew scheduling, the design of telecommunication networks, sports scheduling, and the design of computer processors. One of the important themes that emerged in the lectures is the remarkable performance of modern implementations of Gomory's mixed integer cutting plane method. This created a natural highlight to the meeting in Ralph Gomory's lecture on the corner polyhedron and its relationship to cutting planes. The meeting included a dinner in honor of Dr. Gomory. Speeches by Alan Hoffman, T.C. Hu, Ellis Johnson, William Baumol, Herb Scarf, and Ralph Gomory made this a memorable event.

- **Diagonal Matrix Scaling and its Generalizations and Their Applications in Convex Programming Over Cones**
  - Dates: August 25 - 27, 1999
  - Location: Rutgers University
  - Attendance: 14
  - Organizers: Bahman Kalantari, Rutgers University, Uriel Rothblum, Technion-Israel Institute of Technology, and Alex Samorodnitsky, DIMACS, Rutgers University

The diagonal matrix scaling problem and its generalizations form a fundamental set of problems in mathematical programming and computer science. On the one hand, for nonnegative matrices the doubly stochastic scaling is a fundamental problem in itself with diverse applications, e.g., in estimation of transition probability in Markov chains, network optimization, planning of traffic and transportation, matrix preconditioning, and image reconstruction. On the other hand, the doubly quasi-stochastic scaling of positive semidefinite matrices and its generalizations has tremendous theoretical and algorithmic applications in linear programming, semidefinite programming, and more generally self-concordant conic convex programming. Indeed diagonal matrix scaling has played a key role in Karmarkar's famous LP algorithm. The matrix scaling problem also finds application in combinatorial problems as well as in the approximation of NP-complete problems. The aim of this workshop was to bring together researchers working on theoretical and practical aspects of matrix scaling problems (including matrix balancing), and their generalizations. It was the hope of the organizers that the workshop would give convincing evidence of the broad significance of the matrix scaling problems to researchers within this area, as well as those interested in convex programming, self-concordance theory, duality theory, complexity of algorithms, combinatorial optimization, homogeneous inequalities, numerical analysis, etc. Moreover, the workshop was aimed at defining new research topics and describing open problems. The emphasis of the talks on the first day of the workshop was on topics related to the structural properties of problems pertaining to diagonal scaling. The emphasis of talks on the second day was on algorithms and applications in other fields such as convex programming. Lev Bregman spoke on generalized matrix scaling, combinatorial games, and other applications. Uriel Rothblum dealt with existence and computation in matrix scaling, with emphasis on complexity analysis of methods that compute approximate solutions. Michael Neumann talked about spectral radii of fixed Frobenius norm perturbations of nonnegative matrices. A talk about minimization of spectral radius and norm of a nonnegative matrix diagonal equivalence scaling to a multiple of a doubly stochastic matrix was given by Hans Schneider. Bit-Shun Tam spoke on matrices with cyclic structure. His talk was on structure of nonnegative matrices and generalizations. Alexander Barvinok talked about randomized polynomial time algorithms to approximate the permanent within a simply exponential factor. Leonid Gurvits presented the topic "Operator Matching, Scaling, and Sinkhorn's Iteration," dealing with various convergence properties of the operator Sinkhorn iteration and its applications, such as approximations of mixed determinants, non-convex optimization and generalized Alexandrov-Fenchel inequalities. He also introduced a certain operator/quantum analog of the usual permanents. Alex Samorodnitsky spoke on scaling in approximation algorithms and described two applications of scaling - in approximation of the permanent and in approximation of the mixed discriminant. Shmuel Friedland spoke on computing the Hausdorrf dimensions of subshifts using matrices. Bahman Kalantari concluded the meeting by speaking about matrix scaling over convex cones and dualities and their applications in mathematical programming.

- **Next Generation of Unit Commitment Models**
  - Dates: September 27 - 28, 1999
  - Location: Rutgers University
  - Attendance: 66
  - Organizers: Mike Rothkopf, Rutgers University, Benjamin F. Hobbs, Johns Hopkins University, and Hung-Po Chao, Electric Power Research Institute
  - Co-sponsor: Electric Power Research Institute (EPRI)

For many years, the electric power industry has been using a variety of technical methods to help them solve the unit commitment problem. The result has been savings of tens and perhaps hundreds of millions of dollars in fuel costs. However, things are changing. Optimization technology is improving and the industry is undergoing radical restructuring. Consequently, the role of unit commitment models is changing and the value of the improved solutions that better algorithms might yield is increasing. The purpose of this conference was to explore the technology and needs of the next generation of computer models for aiding unit commitment decisions. The unit commitment problem can be defined as the scheduling of production from power generation units over a daily to weekly time horizon in order to accomplish some objective. The problem solution must respect both generator constraints (such as ramp rate limits and minimum up or down times) and system constraints (reserve and energy requirements and, potentially, transmission constraints). The objective function should account for costs associated with energy production and start-up and shut-down decisions, along with possible revenue or customer impacts of those decisions. The resulting problem is a large scale nonlinear integer program. Because of the problem's size and complexity, and because of the large economic benefits that could result from its improved solution, considerable attention has been devoted to algorithm development. Heuristics such as "priority lists" have long been used by industry; but in the last three decades, more systematic procedures based on a variety of algorithms have been proposed and tested. These techniques have included dynamic programming, branch-and-bound mixed integer programming, linear and network programming approaches, and Benders decomposition methods, among others. Recently, metaheuristic methods have been tested, such as genetic programming and simulated annealing, along with expert systems and neural networks. However, the solution approach that has been most successful, and which is most widely used at present, is Lagrangian relaxation. This procedure decomposes the problem by multiplying constraints that couple different generators (such as energy demand and reserve constraints) by Lagrangian multipliers and placing them in the objective function. Given a set of multiplier values, the problem is then separable in the generating units, and a dynamic program of low dimension can be used to obtain a trial schedule for each unit. A process of multiplier adjustment is used to search for feasible near-optimal solutions. Lagrangian relaxation has proven useful for quick development of good, if not optimal, generator schedules. However, recent improvements in integer programming codes and other algorithms suggest that it may be possible to find better solutions more rapidly. Meanwhile, restructuring has sharpened the appetite of generation owners for more efficient operation. In the past, utilities have been regulated on a cost-plus basis, and so may not have put as much priority on squeezing out the last few percent improvements in the objective function. Furthermore, system operators realized that cost functions were approximate and so were perhaps more likely to be satisfied with good solutions that were technically suboptimal. Now, with restructuring, we have schedule coordinators making commitment decisions in a market environment, and ISOs dealing with bids. Bids are precise, and small improvements in solutions can result in significant changes in payments to bidders. Further, the fact that optimization models are, in some cases, being used to determine which generators will be operated and thus paid implies that there is a greater incentive to get exact answers to make the bidding process legitimate (and to disempower the bid taker). In other words, it's a whole new world, and a reconsideration of how unit commitment models are solved and the purposes for which they are used is needed. The goal of this workshop was to bring together those who know the problem and what is needed with those who know what improvements in algorithms are really possible. The workshop started with a series of in-depth presentations by top experts in algorithms and utility systems operations. Then working groups convened to define the new roles that unit commitment models may play and promising directions in algorithm development and application. Richard O'Neill, Director of the Office of Economic Policy of the Federal Energy Regulatory Commission, opened the workshop with a talk entitled "Why are We Here?" Andrew Ott from the Pennsylvania-NewJersey-Maryland ISO presented the point of view of an independent systems operator, Fred Lee, Art Breipohl, and Walter Hobbs spoke about unit commitment in a competitive environment, David Sun from ALSTOM ESCA spoke about current experience in solving unit commitment problems, and Joseph Graves from PHB Hagler Bailly spoke about simultaneous clearing and optimization of energy and ancillary service markets in the unit commitment context. There were presentations by Ross Baldick on alternative formulations, multiplicity of solutions, and unpriced products, by Robert Bixby from Rice and ILOG entitled "MIP: Closing the Gap between Theory and Practice," and Sebastian Ceria from Columbia and Dash Optimization, Inc. on new developments in integer programming. Panels discussed the definition of the problem to be solved and what we can expect from integer programming. Parallel sessions allowed many contributed talks. One highlight was the demonstration during the workshop by Bob Bixby of CPLEX of the capabilities of new integer programming codes to solve problems that were previously thought impossible to solve to optimality. A new theoretical result on the relation between solutions to the unit commitment model and the problem of deriving prices in daily power auctions was a direct result of the workshop and is described below. The workshop was partly supported by the Electric Power Research Institute (EPRI). Among the products of the workshop will be an EPRI report summarizing the deliberations and presenting the conclusions of the workshop and a book to be published by Kluwer or IEEE Press, with EPRI agreeing to purchase up to 1000 copies of the book for widespread distribution. The intent of the workshop sponsors is that this workshop will ultimately result in better algorithms yielding more value to their users.

Workshops and Miniworkshops Planned for the Rest of 1999-00

No workshops are planned for the rest of this academic year. However, the "Algorithm Implementation Challenge" associated with this special year, on semidefinite programming, will reach its culmination in a workshop in September:

- **Workshop: Computational Challenge on Semidefinite Programming**
  - Dates: September 13 - 15, 2000
  - Location: Rutgers University
  - Organizers: David Johnson, AT&T Labs - Research, Farid Alizadeh, Rutgers University, and Gabor Pataki, Columbia University

Special Year Tutorial Series
Tutorials February 1999 - January 2000

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

An Introduction to Metaheuristics placed 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 were included. An Introduction to OptQuest highlighted 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, University of Rome, La Sapienza, Italy
April 9, April 15, April 21, April 28 and May 5, 1999

This was 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 7 and May 13, 1999

This was a series of two lectures on the topics: generalized graph coloring and the frequency assignment problem; and some combinatorial problems in optimal telecommunication network design. The first lecture investigated generalized graph coloring problems arising in the context of operating large radio link telecommunication networks, or cellular radio communication networks. The second talk concentrated on minimum cost multicommodity flows as basic models frequently arising in the context of topological network optimization, optimal network design, and dimensioning.

Local Search Algorithms in Discrete Optimization
Toshihide Ibaraki, Kyoto University, Japan
June 1, 1999

Local search is a well-known basic tool for constructing heuristic algorithms for discrete optimization problems. It repeats the operation of replacing the current solution with a better solution in its neighborhood (if there is any) until no better solution can be found in the neighborhood (i.e., local optimality is attained). This lecture discussed generalizations, improvements, and ramifications of local search, including effect of neighborhood sizes, adaptive usage of neighborhoods of various sizes, and combination with metaheuristics.

Dynamic Load Balancing on Heterogenous Networks
Ewald Speckenmeyer, University of Kőln, Germany
June 10, 1999

This lecture presented a dynamic load balancing algorithm for a network of heterogeneous computers. The approach was based on a fast distributed precomputation step calculating the necessary load movements. When dealing with slow network connections or incomplete load information often a perfect balanced load is not necessary. Therefore the allowed degree of imbalance is adjustable. A detailed average case analysis was carried out concerning the total load transfer and the number of transfer rounds. Its performance was empirically demonstrated to solve the satisfiability problem on a heterogeneous network of workstations.

Selected Research Results

Connection between the Unit Commitment Problem and Daily Power Auctions The unit commitment problem was described above in the discussion of the workshop on Next Generation of Unit Commitment Models. As a direct result of that workshop, Michael Rothkopf (PM, RU), Ben Hobbs (V), Bill Stewart (V), and Richard O'Neill (V) demonstrated theoretically that the relaxed LP with cutting planes that solves certain forms of the unit commitment problem has a dual interpretation useful for deriving prices in daily power auctions. In particular, the mechanisms used by the New York and PJM system operators to "make whole" producers who don't recover their costs were shown to be closely related to the dual variables for cutting planes that force those producers into the optimal solution.

A Combinatorial Algorithm for Pfaffians The Pfaffian of an oriented graph was introduced by Kastelyn in 1967 to count the number of dimer coverings of a lattice graph. It is closely linked to Perfect Matching and one of the motivations for studying Pfaffians is to understand the complexity of Perfect Matching. Perfect matchings have many important applications in assignment problems, chemistry, etc. The Pfaffian is also naturally related to the determinant of an appropriately defined matrix. This relation between Pfaffian and determinant is usually exploited to give a fast algorithm for computing Pfaffians. Meena Mahajan (V), P. R. Subramanya and V. Vinay (V) found the first completely combinatorial algorithm for computing the Pfaffian in polynomial time. The algorithm works over arbitrary commutative rings. Over integers, it can be computed in the complexity class GapL, the class of all functions that are logspace reducible to computing the integer determinant of a matrix; this result was not known before. As a corollary, it follows that under reasonable encodings of a planar graph, Kasteleyn's algorithm for counting the number of perfect matchings in a planar graph is also in GapL. The combinatorial characterization of Pfaffian also makes it possible to directly establish several algorithmic and complexity theoretic results on Perfect Matching which otherwise use determinants in a roundabout way.

Unique Maximum Matching Algorithms Harold Gabow, Haim Kaplan (PD), and Robert Tarjan (PM, PU) considered the problem of testing the uniqueness of maximum matchings, both in the unweighted and in the weighted case. For the unweighted case, they obtained two results. First, given a graph with n vertices and m edges, they found a way to test whether the graph has a unique perfect matching, and find it if it exists, in $O(m \log^4 n)$ time. This algorithm uses a recent dynamic connectivity algorithm and an old result of Kotzig characterizing unique perfect matchings in terms of bridges. For the special case of planar graphs, they improved the algorithm to run in $O(n \log n)$ time. Second, given one perfect matching, they found a way to test for the existence of another in linear time. This algorithm is a modification of Edmonds' blossom-shrinking algorithm implemented using depth-first search. A generalization of Kotzig's theorem proved by Jackson and Whitty allowed them to give a modification of the first algorithm that tests whether a given graph has a unique f-factor, and find it if it exists. They also showed how to modify the second algorithm to check whether a given f-factor is unique. Both extensions have the same time bounds as their perfect matching counterparts. For the weighted case, they were able to test in linear time whether a maximum-weight matching is unique, given the output from Edmonds' algorithm for computing such a matching. The method is an extension of their algorithm for the unweighted case.

Improvements in Throughput Maximization for Real-Time Scheduling Piotr Berman and Bhaskar DasGupta (V) considered the problem of off-line throughput maximization for job scheduling on one or more machines, where each job has a release time, a deadline and a profit. They provided algorithms that do not use linear programming, are simple and much faster than the corresponding ones proposed earlier in the literature, while either having the same quality of approximation or improving it. More precisely, compared to the known results, their pseudo-polynomial algorithm for multiple unrelated machines and all of their strongly-polynomial algorithms have better performance ratios, all of their algorithms run much faster, are combinatorial in nature and avoid linear programming. Finally, they showed that algorithms with better performance ratios than 2 are possible if the stretch factors of the jobs are bounded.

An Artificial Intelligence Approach to Forecasting of the Event Rate for an Aircraft Operator The Event Rate is an aviation safety metric that records the number of occurrences, incidents, and accidents (or "mishaps") associated with the operations of an aircraft. The calculated value is the total number of these events for a one-month period divided by the aircraft's fleet size for that same period. The value is adjusted by the fleet size to consider that an operator's exposure to events depends on the number of aircraft in its fleet. Atul Deshmukh (GS, RU) and James Luxhoj (RU) investigated the use of an Artificial Intelligence technique, such as an artificial neural network, to develop a forecasting model for the Event Rate for a major aircraft operator. Inputs to the neural network included data on airworthiness (i.e., maintenance-related) and operations surveillance inspections as well as financial data, obtained through collaboration with a project sponsored by the Federal Aviation Administration. A variety of different learning architectures, such as the backpropagation, general regression neural network, etc. were evaluated across a variety of learning parameters. For the aircraft operator in this study, the preliminary conclusion was that the Multiple Regression (MR) models for Event Rate forecasting are sometimes more effective than the Neural Networks (NN) models.

DIMACS special foci cross-fertilize each other. There was considerable cross-fertilization between our Massive Data Sets Special Year and our Large Scale Discrete Optimization Special Year.

The following two results are examples of such cross-fertilization.

Generating Transversals Generating transversals to a hypergraph (or the dual of a Boolean function) is a well-known problem, which became very important recently due to its data-mining applications ("association rules" or more specifically "frequent sets" are transversals to a certain hypergraph defined by the input data set.) This problem is also interesting since its true complexity (in the incremental sense) is not known. Due to an incrementally quasi-polynomial algorithm (Fredman and Khachiyan, 1996) it is very unlikely to be an NP-hard problem, but as of today we do not know any polynomial algorithm, either – it could even be the core to a separate new class, laying somewhere between P and NP-hard (unless of course these two coincide). Endre Boros (PM, RU), Vladimir Gurvich (V), Leo Khachiyan (PM, RU), and Kazuhisa Makino (V) showed that several open versions of transversal generation are in fact equivalent to Boolean dualization, and hence treatable with an incrementally quasi-polynomial algorithm. Their results are based on several new inequalities relating the number of maximal false points to the number of minimal true points of a Boolean function. For instance, they proved that the number of maximal infeasible binary vectors to a monotone system of linear inequalities is polynomially (almost linearly) bounded in the number of minimal feasible vectors. Based on this result, they showed how to generate all minimal feasible solutions to a monotone system of inequalities in incremental quasi-polynomial time.

Attribute Selection in Learning and Data-Mining Problems In the logical analysis of data, one of the first and most important steps is the selection of a small subset of binary attributes to a given input data set which will be used in later steps of the learning process. Such a reduction step, if done carefully, is useful to eliminate noise and to reduce dimensions, making possible the application of more demanding computational methods. Recently similar considerations became popular in the learning community – so called "attribute efficient learning." Endre Boros (PM, RU) and Toshihide Ibaraki (V) considered such an attribute selection problem for the case of binary data, and based on various measures of the "separating power" of the individual attributes and their small subsets. They showed that a wide variety of such models are in fact linearly constrained maximization problems with submodular objective functions, and hence efficient approximation algorithms with a guaranteed performance rate can be applied. They considered continuous relaxations of these problems, which turn out to be smooth convex minimizations over a convex set, and hence can also be solved efficiently.

Inapproximability Results for Bounded Occurrence Satisfiability Problems Piotr Berman and Marek Karpinski (V) obtained a number of improved inapproximability results, including the best up to date explicit approximation thresholds for bounded occurrence satisfiability problems like MAX-2SAT and E2-LIN-2, and the bounded degree graph problems, like MIS, NodeCover, and MAX CUT. They also proved for the first time the inapproximability of the problem of sorting by reversals and displayed an explicit approximation threshold for this problem.

Sequential Testing of Series-Parallel Systems of Small Depth Endre Boros (PM, RU) and Tonguc Unluyurt (GS, RU) considered the problem of testing sequentially the components of a multi-component system when the testing of each component is costly. They considered a polynomial time testing policy for series-parallel systems proposed in the literature. Generalizing earlier results, they proved that the considered policy is cost-minimal in the average case sense for two sub-families of series-parallel systems. They also demonstrated via examples that neither this algorithm nor some of its improved versions are optimal for general series-parallel systems, disproving some published claims.

Randomized Complexity of Linear Arrangements and Polyhedra Marek Karpinski (V) found new applications of recent results on the complexity of recognizing n-dimensional linear arrangements and convex polyhedra by randomized algebraic decision trees. In particular, he derived the first nontrivial, in fact quadratic, randomized lower bounds on the problems like Knapsack and Bounded Integer Programming.

Partitionable Graphs. The famous Berge conjecture concerns perfect graphs, which have many important applications in biology, scheduling, traffic phasing, radio frequency assignment, etc. The nnclass of partitionable graphs includes all "counterexamples" to Berge's conjecture (if one exists), as was observed by Lovasz (1972) and by Padberg (1974). Still, there is very little known about this family. There are only a handful of infinite families known within this class, and even those are the ones with very high symmetries. Endre Boros (PM, RU), Vladimir Gurvich (V), and Stefan Hougardy found a recursion which generates several infinite families of partitionable graphs, based on the so called "critical edges" of some cliques of such a graph. For many of these generated families they were able to show that chordless odd cycles are inherited, hence proving that no counterexample to Berge's conjecture can exist within those.

The following six results connect up two special years, Large Scale Discrete Optimization and Networks.

The Element Connectivity Problem in Survivable Network Design The element connectivity problem falls in the category of survivable network design problems – it is intermediate to the versions that ask for edge-disjoint and vertex-disjoint paths. The edge version is by now well understood from the view-point of approximation algorithms, but very little is known about the vertex version. Vijay Vazirani (V) considered the problem where vertices are partitioned into two sets: terminals and non-terminals. Only edges and non-terminals can fail – we refer to them as *elements* – and only pairs of terminals have connectivity requirements, specifying the number of element-disjoint paths required. Vazirani found an algorithm that achieves an approximation guarantee of factor $2H_k$, where k is the largest requirement and $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Besides providing possible insights for solving the vertex-disjoint paths version, the element connectivity problem is of independent interest, since it models a realistic situation.

Primitive Ring Generation and Ring Grooming in Design of Optimal Networks A central problem in the design of optical networks is the placement of the ADM (Add/Drop Multiplexes) devices. Many restricted versions of this general design problem have been studied in the literature. The general version can be formulated as a discrete optimization problem. Due to the complexity of the problem, the current methodology recommends dividing the general version of the design problem into two phases, called Primitive Ring Generation and Ring Grooming. The General Design Problem, Primitive Ring Generation, and Ring Grooming are NP-Hard. We have obtained theoretical results on the quality of approximation algorithms. Primitive Ring Generation has two variants, depending on whether or not traffic demands can be split. Gruia Calinescu (PD) and Peng-Jun Wan designed new algorithms for Primitive Ring Generation guaranteed to produce an output of cost at most 25% bigger than optimum if traffic demands can be split, and at most 50% bigger than optimum if traffic demands cannot be split. All previously known algorithms either have bigger worst-case guarantees or are provably not better than the new algorithms. For Ring Grooming, the traffic granularity parameter seems to influence the quality of the approximation algorithms. Gruia Calinescu (PD), Li-Wu Liu and Peng-Jun Wan provided a nontrivial analysis of a natural algorithm, proving performance guarantees of 1.5 and 2.5 respectively, for the important case when traffic granularity is 4 and 8 respectively.

Directed Multicut vs. Multicommodity Flow Given a graph $G = (V, E)$ and k pairs of vertices $(s_1, t_1), ..., (s_k, t_k)$, a *multicut* is a subset of edges $F \subseteq E$ such that if all edges in F are deleted, none of the pairs (s_i, t_i) are in the same connected component in the remaining graph $G' = (V, E \setminus F)$. In the minimum multicut problem, we are also given a positive cost $c(e)$ for each edge $e \in E$ and we wish to find the multicut of minimal total cost. This problem is known to be NP-complete for $k \geq 3$, and therefore it is interesting to consider approximation algorithms for it. One way to do it is to state the multicut problem as an integer programming problem, relax it to a linear programming problem, and solve this problem or it's dual - the *maximum multicommodity flow problem.* The maximal multicommodity flow will always be at most as big as cost of the minimal multicut. On the other hand, it is easy to see that there are multicuts which cost at most k times the value of the maximal flow. Thus, the described approach gives a factor-k approximation algorithm for a minimal multicut. It turns out that when the underlying graph is *undirected* more can be said, and in fact the cut/flow ratio in this case can be at most as large as $O(\log(k))$. On the positive side, an algorithm of Garg, Vazirani and Yannakakis exists, which takes a maximal flow and constructs a multicut which costs at most as much as $O(\log(k))$ times the flow cost. Michael Saks (PM, RU), Alex Samorodnitsky (PD), and Leonid Zosin (PD) showed that in the more general *directed* case, the cut/flow ratio could be as large as the trivial upper bound of k. It seems, therefore, that to find a better than a factor-k approximation, one must look for other approaches than the simple linear relaxation technique.

A Unified Approach to Approximating Resource Allocation and Scheduling Amotz Bar-Noy (ATTL), Reuven Bar-Yehuda (V), Ari Freund, Joseph (Seffi) Naor, and Baruch Schieber developed a general framework for resource allocation and scheduling problems. Given a resource of fixed size, they developed algorithms that approximate the maximum throughput or the minimum loss by a constant factor. Their approximation factors apply to many prominent problems, among others: (i) Real-time scheduling of jobs on parallel machines. (ii) Bandwidth allocation for sessions between two endpoints. (iii) General caching. (iv) Dynamic storage allocation. (v) Bandwidth allocation on optical line and ring topologies. For many of the problems, this is the first constant factor approximation algorithm. The algorithms are extremely simple and efficient. They use the local-ratio technique but they can also be equivalently interpreted within the primal-dual schema.

Load Balancing in a Distributed Computing Environment Endre Boros (PM, RU) and Ewald Speckenmeyer (V) studied load balancing in a distributed computing environment. They showed that the minimum message length to re-balance the load of n processors can be computed in linear $O(n)$ time, even though the actual balancing may require $O(n^2)$ messages exchanged between the processors.

A Bottleneck Load Balancing Minimization Problem Consider a processor network with n overloaded and m underloaded processors. The overloaded and the underloaded nodes are connected by a completely bipartite network $K(n, m)$. Ewald Speckenmeyer (V) considered the problem of achieving a uniform load-distribution among all processors via sending appropriate amount of load from overloaded to underloaded nodes so that the maximum of load sent via a link is minimized. He showed that this problem can be solved in time $O(mn)$.

The Zero-Extension Problem. The Zero-Extension Problem, a natural generalization of multiway cut, is defined as follows: given a metric space (T, m) with at least three points called terminals, and a graph $G = (V, E)$ whose edges have nonnegative cost c, and such that T is a subset of V, find a mapping from V to T which keeps T invariant and minimizes the summation over all the edges of E of the product between the cost of an edge and the distance in between the terminals where the endpoints of the edge are mapped. Multiway Cut corresponds to the case when the metric space is uniform. We denote by k the number of terminals. Kleinberg and Tardos studied a generalization of Zero-Extension, called Metric Labeling, and provided an approximation algorithm with performance guarantee of the order of $\log k \log\log k$. The Zero-Extension problem has a natural linear programming relaxation. Karzanov gave a characterization of graph-induced metric spaces for which the linear programming relaxation equals the value of the Zero-Extension problem. Griua Calinescu (PD), Howard Karloff (PM, ATTL) and Yuval Rabani (V) found approximation algorithms for the Zero-Extension problem: a factor of $\log k$ rounding procedure for the linear programming relaxation, a $15/7 > 2$ lower bound for the integrality ratio of the linear program, and a dual-based constant-ratio rounding procedure for the case when G is a tree.

Formulating Semidefinite Programming Problems as Smooth Convex Nonlinear Optimization Problems Consider the diagonal entries d_j, $j = 1, 2, \ldots, n$, of the matrix D in an LDL^T factorization of an $n \times n$ matrix X. As a function of X, each d_j is well-defined on the closed domain of positive semidefinite matrices. Robert Vanderbei (PM, PU) and Hande Yurttan Benson (GS, PU) showed that these functions are twice continuously differentiable and concave throughout the interior of this domain. Using these facts, they showed how to formulate semidefinite programming problems as standard convex optimization problems that can be solved using an interior-point method for nonlinear programming. To summarize, these results allow one to marry implementations of SDP with implementations of general nonlinear programming solvers thereby creating the possibility to solve problems from both of these problem classes individually with one software system but more importantly to solve hybrid problems that contain a mix of traditional NLP stuff with some SDP components.

Approximation of the Permanent and of the Mixed Discriminant It is $\#P$ hard to compute the permanent of a nonnegative matrix, but it seems like a realistic goal to approximate the permanent efficiently. A strongly polynomial algorithm achieving approximation up to a multiplicative factor of e^n was given in 1998 by Linial, Samorodnitsky, and Wigderson. Alex Samorodnitsky (PD) has extended these results to some families of $0-1$ matrices, such as *dense matrices* (more than half the entries in each row and column are 1) and *random matrices* (each entry is assigned zero or one randomly and independently). As it turns out, the earlier results may also be generalized to approximation of a *mixed discriminant* of an n-tuple of positive semidefinite (PSD) matrices. This line of research has been pursued by Leonid Gurvits (NEC) and Alex Samorodnitsky (PD). They defined notions of *double stochasticity* and of a *doubly stochastic scaling* (generalizing the usual matrix scaling) for an n-tuple of PSD matrices, and constructed a polynomial-time algorithm for such a generalized scaling. This together with bounds on the mixed discriminant of doubly stochastic n-tuples leads to a deterministic polynomial-time approximation algorithm for a mixed discriminant. It also provides a deterministic polynomial-time approximation algorithm for the mixed volume of n convex bodies, answering a question of Dyer, Gritzmann and Hufnagel. Following this work, very recently Leonid Gurvits proved a generalization of the Van der Waerden conjecture for permanents to the mixed discriminant case. This result implies, in particular, that the Gurvits-Samorodnitsky algorithm gives a polynomial-time deterministic approximation algorithm for the mixed discriminant whose performance is of the same order as of the best randomized algorithms (and also a near-best possible approximation factor for the mixed volume).

Linearity Tests and Probabilistically Checkable Proofs Probabilistically checkable proofs (PCPs) are a model of computation that helps in understanding the approximability properties of combinatorial optimization problems. In 1998, Luca Trevisan (PD) developed a general analysis technique and a linearity test (a probabilistic algorithm to verify whether a given Boolean function is linear) achieving *amortized query complexity* of 1.5 bits. (A testing algorithm uses q amortized bits if, for some constant k it reads qk bits and has error probability at most 2^{-k}.) (One should mention that amortized query complexity couldn't go below 1, unless $P = NP$.) Linearity testing is interesting because of its applications in the PCPs. Sudan and Trevisan had succeeded in 1998 in adapting the 1998 approach of Trevisan to the PCP setting, obtaining a PCP characterization of NP using, essentially, 1.5 amortized query bits. Trevisan conjectured in 1998 that for any $\epsilon > 0$ there exists a linearity testing algorithm achieving amortized query complexity of $1+\epsilon$ bits, and that NP can be (PCP)-characterized using $1+\epsilon$ amortized query bits. He made the following conjecture, which implies the possibility of $1+\epsilon$ amortized query bits linearity testing: For any $f \ : \ \{-1,1\}^n \rightarrow \{-1,1\}$, for any nonempty graph $G = ([k], E)$,

$$\mathcal{E}_{x_1,\ldots,x_k} \prod_{(i,j)\in E} f(x_i)f(x_j)f(x_ix_j) \leq max_\alpha|\hat{f}(\alpha)|,$$

where $\hat{f}$ is the Fourier transform of f. Alex Samorodnitsky (PD) and Trevisan (PD) proved this conjecture and showed, moreover, that it can be adopted to the PCP setting, implying that NP can be (PCP)-characterized using $1+\epsilon$ amortized query bits. In particular this provides a significantly easier proof of a famous result of Hastad on hardness of approximation of the size of the maximal clique in a graph.

Generalized Matrix Scaling and Strategy Properties in Combinatorial Games Lev Bregman (V) considered a game with the set $\{X_1, X_2, .., X_p\}$ as the pure strategies set for the first player and the set $\{Y_1, Y_2, .., Y_q\}$ as the pure strategies set for the second player (X_i are non-negative vectors in R^k and Y_j are non-negative vectors in R^l). Let H be a $k * l$- matrix and define the payoff for the first player as (X_i, HY_j), if the players use their strategies X_i and Y_j respectively. We usually apply this framework to games with strategy sets of a combinatorial nature and thus term such games "combinatorial games". The numbers p and q are assumed to be much bigger than k and l. We are interested in finding optimal mixed strategies of the game which may be described by a small number of parameters. In particular, a product form is considered when the probability of a pure strategy is represented as a product of factors dependent on only one component of the strategy. To get such a representation, Bregman used a generalized matrix scaling and generalized permanents. The necessary and sufficient condition for the existence of such representations was established.

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 reach a broad segment of the dm and tcs research community as well as engage practitioners.

Organizers:

- Stuart Haber, 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, Poly Univ.;
- Jack Holtzman, Rutgers;
- Frank Hsu, Fordham;
- Srinivasan Keshav, Cornell;
- Arjen Lenstra, Citibank;
- Udi Manber, Arizona;
- Ray Miller, Maryland;
- Debasis Mitra, Bell Labs;
- Clyde Monma, Telcordia Technologies;
- 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
Institute for Mathematics and its Applications, University of Minnesota
NEC Research Institute
Princeton University
Rutgers University
Telcordia Technologies
WINLAB (Wireless Networking Lab), Rutgers University

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 rapidly 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 February 1999- January 2000

- **Large Scale Discrete Optimization in Logistics**
  - Dates: February 8-10, 1999
  - Location: DIMACS Center, CoRE Building, Rutgers University
  - Attendance: 79
  - Organizers: George Nemhauser and Martin Savelsbergh, Georgia Tech
  - Comment: Joint with Special Year on Large Scale Discrete Optimization

An efficient and effective logistics system is the key to being competitive in today's global marketplace. This realization has fueled an enormous interest in new technology for designing and operating logistics systems. Discrete optimization frequently is an integral part of such technology. In particular, in the last decade, the use of large-scale discrete optimization models in transportation and production planning has increased dramatically. This has been made possible by developments in modeling, algorithms, software, and hardware. This workshop brought together leading researchers to present state-of-the-art research activities and to discuss innovative future directions of research for this very active field with significant practical importance. The talks covered a wide range of topics including modeling, algorithms, and applications. Application domains that were covered include network design, production planning, vehicle routing, and crew scheduling. The workshop opened with a talk by William Cook titled "On the Solution of Traveling Salesman Problems". Research on the solution of the traveling salesman problem, the prototypical discrete optimization problem, has stimulated enormous interest in discrete optimization and has helped significantly in the growth of the field. Several other talks focused on methodology, such as "Approximation Algorithms via Linear Programming: Rounding Algorithms" by David Shmoys, "An Overview of Classical and Modern Heuristics for the Vehicle Routing Problem" by Gilbert Laporte, "GRASP: Greedy Randomized Adaptive Search Procedures" by Mauricio Resende, "AI Techniques for Large-Scale Discrete Optimization" by Jimmy Crawford, and "Very Large Scale Neighborhood Search" by Ravi Ahuja. Furthermore, there were several talks related to the main application domains of discrete optimization. In the topic of Airline Optimization, we had "Solving Aircraft Maintenance Scheduling Problems" by Daniel Bienstock, "Optimization Problems in Air Traffic Management" by Mike Ball, and "Airline Integrated Recovery: AIR and SimAIR" by Ellis Johnson. In the topic Manufacturing and Production Planning, we had "BC-PROD and LOTSIZELIB" by Laurence Wolsey, "Optimization in IBM Manufacturing" by Brenda Dietrich, and "Painting Cars and the Traveling Salesman Problem" by Tom Magnanti. In the topic Supply-Chain Management, there was "Optimal and Approximation Algorithms for Large Scale Supply Chain Management Problems" by David Simchi-Levi and "Dynamic Programming Approximations for the Discrete Multi-commodity Network Flow Problem over a Dynamic Graph" by Warren Powell. Finally, in the topic Telecommunication Network Design there was "The Prize Collecting Steiner Tree Problem" by David Johnson. There was also a full group discussion on future directions for the field. The group was very encouraged by the increasing use of discrete optimization methods in industry and the need for the solution of larger more complex models that, for example, include stochastic elements. In conclusion, the workshop provided an excellent environment in which leading researchers in the field could share their ideas not only with each other, but with junior researchers and students as well.

- **Mobile Networks and Computing**
  - Dates: March 25-27, 1999
  - Location: Rutgers University
  - Attendance: 48
  - Organizers: Badri Badrinath, Rutgers University, Frank Hsu, Fordham University, Panos Pardalos, University of Florida, Sanguthevar Rajasekaran, University of Florida

Advances in the technologies of networking, wireless communications, and miniaturization of computers lead to the rapid development in mobile communication infrastructure and have engendered a new paradigm of computing. Users carrying portable devices can freely move around while still being connected to the network. This provides flexibility in accessing information anywhere and at any time. On the other hand, the price of this flexibility has been the introduction of new levels of complexity that were not encountered in software and protocol design in wired networking. The new challenges in designing software systems for mobile networks include location and mobility management, channel allocation, power conservation, information dissemination, and sensor networks, among others. This workshop focused at bringing together researchers from academia as well as industry who are working on various aspects of mobile computing. One of the key successes of the workshop was that collaboration started among researchers from different institutes. In addition, several graduate students participated (and some presented papers) in the workshop. A refereed volume with papers presented at the Workshop will be published early next year by the American Mathematical Society. Among the presentations were discussions of DataSpace - querying and monitoring deeply networked collections of physical objects by Tomasz Imielinski; of improving the performance of TCP over wireless links, by Zygmunt J. Haas and A. Warkhedi; software implementation strategies for power-conscious systems by David Wei; and multicriteria optimization frequency assignment by Robert A. Murphy, Panos M. Pardalos, and Eduardo Pasiliao. In addition, Jean-Claude Bermond talked about routing in satellite constellations, Jeng-Ping Lin, Sy-Yen Kuo, and Yennun Huang spoke about a cluster-based checkpointing scheme for mobile computing on wide area networks, Stephen Olariu spoke about efficient initialization and leader election in mobile radio networks, and Ibrahim Korpeoglu, Pravin Bhagwat, Chatschik Bisdikian, and Mahmoud Naghshineh presented a talk about the challenges of mobile and wireless networking support for palmtop devices. Ravi Prakash and Mukesh Singhal made a presentation on the impact of unidirectional links in wireless ad-hoc networks, Javier Gomez-Castellanos and Andrew T. Campbell talked about supporting integrated-QoS in wireless networks, Sanguthevar Rajasekaran dealt with frequency assignment in cellular networks, and Sumi Helal, Jin Jing, and Ahmed Elmagarmid spoke about supporting application-level handoff in overlay networks. Brigitte Jaumard, Christophe Meyer, and Tsevi Vovor presented a paper entitled "How to Combine a Column and Row Generation Method with a Column or Row Elimination Procedure – Application to a Channel Assignment Problem", Anup K. Talukdar, B.R. Badrinath, and Arup Acharya spoke about "Rate Adaptation Schemes in Networks with Mobile Hosts", Sajal K. Das and Sanjoy K. Sen spoke about quality of service provisioning in next generation wireless networks, and Kshirasagar Naik presented a talk about energy-conserving software design for mobile computers. A talk on agent-based architectures for securing mobile IP was given by Xun Yi, one on adaptive detection of multi-level modulation formats for DC/CDMA systems in a slowly fading channel by Ali F. Almutairi, Scott L. Miller, and Haniph Latchman, and one entitled "Online Protocols for Channel Assignment in Wireless Networks" was presented by Sunil Shende. The meeting concluded with talks by Sandeep K.S. Gupta and Pradip K. Srimani on using self-stabilization to design adaptive multicast protocols for mobile ad hoc networks, Anupam Joshi on mobility and agents, Xin Liu, Panos M. Pardalos, and Sanguthevar Rajasekaran on frequency assignment problems in mobile networks, and Julio C. Navas and Tomasz Imielinski on the topic "Geographic Routing: Improving Routing through Approximation"

- **On-Line Decision Making**
  - Dates: July 12 - 15, 1999
  - Location: Rutgers University
  - Attendance: 89

– Organizers: Yoav Freund, AT&T Labs - Research, and Rakesh Vohra, Northwestern University
– Comment: This was both a Special Year Networks workshop and an exploratory workshop, investigating a potential new special focus.

In recent years there has been increasing interest in the analysis of algorithms for making repeated decisions within an unknown environment. It has been realized that interesting performance guarantees can be obtained without assuming that the environment's behavior is generated by a stochastic process. Work along these lines has been going on in several different fields under a variety of names, for example: information theory: Universal coding of individual sequences, universal prediction, universal portfolios ...; game theory and decision theory: Adaptive game playing, learning in games ...; machine learning: Aggregating strategies, using expert advice, prediction in the worst case ...; statistics: Prediction in the worst case, universal calibration, prophet inequalities ...; computer science: Competitive analysis of Online Algorithms. The goal of the workshop was to bring together researchers from these different areas to identify common questions as well as explore different perspectives on the same questions. The talks were (by design) mostly survey talks of various areas with a minority of talks on more focused or specialized issues. We have not created a printed proceedings for this workshop; instead, a web page which summarizes the meeting was created and is accessible at http://www.research.att.com/~yoav/DIMACS.workshop/index.html. Rather than offer a talk by talk description of what transpired it might be more useful to provide an overview of what seem to be the important issues. The basic problem that everyone was focused on is what we call "the online decision problem." This problem considers a scenario in which an agent performs some kind of repeated interaction with its environment. On each iteration of this interaction the agent chooses one out of many actions whose consequences/costs depend on the unknown environment. The action must be chosen *before* its cost is known but can depend on the outcomes of previous actions. Unlike other frameworks for studying this problem, we do *not* assume that the environment is an oblivious stochastic process. Rather, we allow the environment's behaviour to be arbitrarily complex; in particular, it can depend on the actions of the agent. There is general agreement that 'good' in this context means relative to some benchmark collection of procedures, but differences arise as to how the benchmarks should be chosen. These differences stem, in large part, from the context in which the online decision problem arises. To illustrate, we summarize a debate that took place at the workshop. In several fields it is common to use a benchmark based on "regret". A list of procedures (also called "experts", "models" or "strategies") is given a-priori. One does not know in advance which procedure will generate the lowest time average cost. However, one *can* describe adaptive methods that are guaranteed to do as well as the best procedure in hindsight. That is, one can, *after* using the adaptive method, go back and calculate the total loss (or gain) of each expert (by itself) against the sequence of outcomes one faced in the past. The regret is defined as the difference between the total gain of the adaptive algorithm and the total gain of the best procedure. (It is called regret because it quantifies the feeling of regret one suffer when one considers their total loss had they followed the recommendations of the best procedure). Methods developed in various fields have provable a-priori guarantees on the maximal value of the regret. However, such guarantees have no meaning in some contexts. The assumption implicit in the guarantee is that nature does not respond to the choice of procedure used. Specifically one assumes that had one used a procedure different from the one used initially, the sequence of outcomes stays the same. This does not make sense in a game theory context. Regret-type benchmarks have been popular in both computer science and game theory and a number of tools have been developed to analyze them. Two that were highlighted in the talks were Bregmann distances and Blackwell's approachability theorem. A tantalizing possibility is that Bregmann distances are related to the potentials associated with Blackwell's Approachability theorem. A number of the more focused talks discussed applications of these tools as well as generalizations. When one moves beyond regret-type measures of performance to measures that assume that nature chooses the outcomes as a function of the procedure one uses, guarantees appear impossible to come by. This was a point made in a series of talks on impossibility theorems. More importantly they highlight an important weakness of the Bayesian approach to decision making. Interestingly, this was a topic that that was discussed in a minor way among the participants and continued via email after the workshop.

- **Scaling Phenomena in Communication Networks**
  - Dates: October 22-24, 1999
  - Location: Institute for Mathematics and its Applications (IMA), University of Minnesota
  - Attendance: 90
  - Organizers: Ashok Erramilli, Netmetrix Inc., Vern Paxson, Lawrence Berkeley National Laboratory, Iraj Saniee, Lucent Technologies, Walter Willinger, AT&T Labs - Research
  - Co-sponsor: Institute for Mathematics and its Applications

The study of scaling phenomena in modern communication networks is another realization of Mandelbrot's vision of order in physical and social phenomena that are characterized by scaling laws. This exciting new multi-disciplinary field has attracted the attention of researchers from networking, mathematicians with interest in fractal geometry, physicists experienced in dealing with scaling laws, and computer scientists, economists and control theorists concerned with robustness and scaling issues associated with complex large-scale interacting systems. Moreover, developments in this field have been accompanied by the availability of extended, high quality data sets of network traffic measurements that are unprecedented in other disciplines. This 3-day workshop was intended to bring together the leading researchers in this emerging area, representing its various constituencies. Its main objective was to foster the exchange of ideas between leading networking experts and researchers in other fields, matching problem areas with solution methods. The workshop was structured around three fundamental aspects of the study of scaling phenomena in networks: description (e.g., empirical evidence, physical understanding, dynamical systems); analysis (e.g., network performance with fractal traffic flows, renormalization group techniques for large scale distributed systems, mean-field theory approaches for full-service networks); and control (e.g., self-organization; pattern formation, evolution and adaptation in spatially extended non-equilibrium systems). Participants were expected to contribute to this effort by giving a talk and/or actively engaging in the proceedings. In contrast to many other fields where scaling phenomena have had a long history but have not moved beyond the descriptive stage, this area shows great potential to apply the theory to analyze and control complex, large-scale networks such as the Internet. The workshop was intended to advance the study of scaling phenomena in networks from a descriptive theory to a prescriptive reality. A high point of the meeting was an IMA Public Lecture at the end of the first day by Benoit Mandelbrot: "Fractals/Multifractals in Finance, the Internet and Other 'Wild' Aspects of Man's Works." This talk (preceded by a reception for workshop participants) attracted an outside audience of about 300, including high school students and teachers and college students. The format of the workshop involved one-hour and half-hour talks interspersed with discussion, summary panel discussions that pointed out the importance of modeling and scaling issues for the future development of communications networks, and time off for informal interaction. In the first session on Networks as Large-Scale Complex Systems, Vern Paxson spoke about why analyzing the Internet is painfully hard, Mark Handley about architectural principles of the Internet, and John Doyle about HOT power laws in Networks and other complex systems. A second session on the same topic featured talks by Sally Floyd on the evolution of end-to-end congestion control in the Internet, Andrew Odlyzko on economics and its implications for the Internet, and Steven Strogatz on complex networks from the point of view of nonlinear dynamics. There was a session on scaling phenomena in the Internet, in which Scott Shenker talked about asymptotics in the Internet, Mark Crovella about self-similarity and powers laws in the web, and Don Towsley on network tomography. This was followed by a session on scaling phenomena in other branches of science. Here, Ed Waymire spoke about statistical estimation of multiscaling exponents in multiplicative cascades, Larry Liebovitch about scaling phenomena in biological and medical systems, Donald Turcotte on mechanisms for generating fractional noises and walks, and Didier Sornette on generic mechanism for hierarchies. In the session on Scaling Analysis and Wavelets, Patrice Abry, Patrick Flandrin, and Darryl Veitch presented a talk on wavelet tools for scale invariance, Anna Gilbert spoke about network inferences from scaling analysis, and Jim Roberts about flow aware networking for effective quality of service control. In the session on Network Performance Analysis and Scaling, Robert Savit gave a talk with the title "That Place is so Crowded, No one Goes there Anymore," which dealt with adaptive competition for scarce resources. Then Onuttom Narayan spoke about scaling and network performance. There was also a work in progress session for students and new researchers. The full program schedule, copies of the transparencies, PDF files and Power Point presentations for many of the talks, and the audio recording of the IMA Public Lecture are available online at http://www.ima.umn.edu/reactive/fall/networks.html The IMA provided office space, computer accounts and secretarial support to all registered participants. The event proved so popular that registration had to be closed at 90 persons in order to preserve the "workshop nature" of the meeting and maximize informal interaction.

Workshops and Miniworkshops Planned for the Rest of 1999-00

These are joint with Special Focus on Next Generation Networks Technologies and Applications as that special focus begins and this special focus winds down.

- **Data Processing on the Web: A Look into the Future**
  - Dates: March 6 - 7, 2000
  - Location: Rutgers University
  - Organizers: Mary Fernandez, S. Muthukrishnan, Dan Suciu, AT&T Labs - Research
- **Cryptography and Intractability**
  - Dates: March 20 - 22, 2000
  - Location: Rutgers University
  - Organizers: Moni Naor, Weizmann Institute of Science, Joe Kilian, NEC Research Institute, and Shafi Goldwasser, MIT and Weizmann Institute of Science
  - Comment: Joint with Special Year on Computational Intractability

- **Management of Digital Intellectual Property**
  - Dates: April 17 - 18, 2000
  - Location: Rutgers University
  - Organizers: Joan Feigenbaum, AT&T Labs - Research, Dan Boneh, Stanford University, and R. Venkatesan, Microsoft
  - Co-sponsor: Microsoft
- **Web Characterization, Metrics and Statistics** (miniworkshop)
  - Dates: Spring 2000
  - Location: Rutgers University
  - Organizers: Lee Giles, NEC, Steve Lawrence, NEC, Sam Weerahandi Telcordia Technologies
- **Data Mining in the Internet Age**
  - Dates: May 1-2, 2000
  - Location: IBM Almaden
  - Organizers: Joan Feigenbaum, AT& T Labs, Prabhakar Raghavan, IBM Research
  - Co-sponsor: IBM Research

Selected Research Results

Estimating Distances in Internet Applications Yuval Shavitt (PM, BL), Xiadong Sun (GS, RU), Avishai Wool (PM, BL) and Bulent Yener (PM, BL) have found a new algorithm for distance estimation in Internet applications. Distance estimation is important to a WWW client that needs to select a server among several potential candidates. To aid in estimating distances in the Internet, past research suggested placing tracer stations in key locations to measure distances among themselves and supply client applications with an approximated map of the Internet. The new algorithm extracts more data from the same measurements, namely, obtaining good estimates to locations where tracers are not placed, obtaining data from the same measurements as if the number of tracers were several-fold larger, and obtaining good estimates for distance between locations where we cannot place tracers. This is extremely important when unidirectional measurements are conducted using specialized equipment such as GPS that cannot be placed everywhere.

Special foci at DIMACS cross-fertilize each other. The following two results connect up our special years on Networks and Massive Data Sets.

Maximum Clique Problems in Very Large Graphs The proliferation of massive data sets brings with it a series of special computational challenges. Many of these data sets can be modeled as very large multidigraphs with a special set of edge attributes that represent special characteristics of the application at hand. Understanding the structure of the underlying digraph is essential for storage organization and information retrieval. Such problems arise for example from telephone billing records. One builds a multidigraph with vertices being telephone numbers and edges calls made from one number to another. James Abello (PM, ATTL) sudied the evolution of such a graph over one 20-day period and found that it grew to have 290 million vertices and 4 billion edges. In practice, this multidigraph is treated as if it were undirected and multiple edges are collapsed into a single edge. Even a computer with six gigabytes of main memory cannot hold the full call graphs arising in practice. Cliques in the call graph represent important structural information. Maximum cliques are important in such practical applications as project selection, classification theory, fault tolerance, coding theory, computer vision, economics, information retrieval, signal transmission theory, and molecular biology. Yet, identifying a maximum clique in a graph is computationally difficult even in a graph of modest size. Abello and collaborators Panos Pardalos (V), Mauricio Resende (PM, ATTL), and Sandra Sudarsky (GS, V) developed an approach to finding cliques in very large multidigraphs. They developed several graph decomposition schemes for breaking up the problem into manageable pieces and a semi-external greedy randomized adaptive search procedure for finding approximate solutions to the maximum clique problem and the maximum quasi-clique problem in very large sparse graphs. They experimented with their heuristic on real data, for instance a one-day call graph with 53,767,087 vertices and 170 million edges, finding large cliques of as many as 30 vertices.

Maintaining Hierarchical Graph Views In connection with a problem in visualizing large graphs, Adam Buchsbaum (PM, ATTL) and Jeff Westbrook (PM, ATTL) considered a hierarchy defined on the vertices, and how one would visualize aggregations of the graph by contracting vertices with respect to the hierarchy. For example, in the graph induced by telephone customers (vertices) and calls (edges), one can contract all the vertices in a single area code together. The problem then becomes one of maintaining this induced, contracted graph under expansion (area code into exchanges, for example) and contraction (the inverse) operations. They formalized the problem and provided data structures that allow navigation in optimal time (linear in the size of the change to the induced graph), although they use superlinear space.

Combinatorial Design of Multi-Ring Networks with Combined Routing and Flow Control Avishai Wool (PM, BL) and Bulent Yener (PM, BL) developed a novel design technique, and combined routing and flow control algorithms, for congestion-free packet switched networks. The design is based on the construction of multiple virtual rings, under the constraint that the path between any two nodes is either confined to a single ring, or traverses exactly two rings (passing through a single bridge node). Their best designs were constructed by using Finite Generalized Quadrangles of combinatorial design theory, together with a scaling algorithm for realizing networks of arbitrary size. The target topology was obtained by taking the edge union of the multiple virtual rings. Capitalizing on the underlying topological properties, they designed routing, flow and access control protocols. They proved that the proposed network architecture, coupled with their protocols, ensures that (i) no loss due to congestion occurs inside a network, under arbitrary traffic patterns; (ii) all the packets reach their destinations within bounded time; and (iii) the bandwidth is allocated fairly and no host is starved. Moreover, they achieved these desirable properties much more efficiently than earlier proposals.

Dynamic Behavior of Differential Pricing and Quality of Service Options for the Internet The simple model on which the Internet has operated, with all packets treated equally, and charges only for access links to the network, has contributed to its explosive growth. However, there is wide dissatisfaction with the delays and losses in current transmission. Further, new services such as packet telephony require assurance of considerably better service. These factors have stimulated the development of methods for providing Quality of Service (QoS), and this will make the Internet more complicated. Differential quality will also force differential pricing, and this will further increase the complexity of the system. The solution of simply putting in more capacity is widely regarded as impractical. However, it appears that we are about to enter a period of rapidly declining transmission costs. The implications of such an environment were explored by Peter Fishburn (PM, ATTL) and Andrew Odlyzko (PM, ATTL), considering models with two types of demands for data transport, differing in sensitivity to congestion. Three network configurations were considered: (1) with separate networks for the two types of traffic, (2) with a single network that provides uniformly high QoS, and (3) with a single physical network that provides differential QoS. The best solution depends on the assumptions made about demand and technological progress. However, they showed that the provision of uniformly high QoS to all traffic may well be best in the long run. Even when it is not the least expensive, the additional costs it imposes are usually not large. Their results concluded that in a dynamic environment of rapid growth in traffic and decreasing prices, these costs may well be worth paying to attain the simplicity of a single network that treats all packets equally and has a simple charging mechanism.

Orientations of Annular Networks and Communication in Interconnection Networks Orientations of undirected graphs arise in a variety of applications, including traffic flow in cities and communication in interconnection networks. If G is a graph, the minimal diameter among all orientations is useful in such applications. For example, gossiping (also called total exchange or all–to–all communication) in interconnection networks is the process in which initially each processor has an item of information that must be distributed to every other processor of the system. Gossiping arises in a large class of parallel computation problems, such as linear system solving, matrix manipulation, and sorting, where both input and output data are required to be distributed across the network. The process is accomplished by means of a sequence of *synchronous calls* between processors. Furthermore we suppose that during each call a processor can communicate with all its neighbors. If both adjacent processors can communicate, this communication model is usually referred to as *Full–Duplex* Δ*–Port* (F_*). In that case the gossiping time (minimum number of steps needed to complete the protocol) is equal to the diameter of the graph G representing the interconnection network. If we suppose that simultaneous exchange of a message on the same link is not authorized, we have another popular communication model called *Half–Duplex* Δ*–Port* (H_*). In that case, for a given orientation, we can consider the greedy protocol which at each step sends all the information known by a vertex to all of its out-neighbors according to the orientation; so we have the diameter of the orientation as an upper bound of the gossiping time. Since a lower bound is the diameter of G, for any network G that admits an orientation whose diameter is close to the diameter of G, the performance of the greedy protocol is provably close to an optimal performance. Motivated by this type of application, Jean-Claude Bermond (V), Johny Bond, Carole Martin, Aleksandar Pekec (V), and Fred Roberts (PM, RU) studied annular networks, graphs that can be represented as a 2-dimensional grid consisting of concentric circles around the center and straight lines crossing all the circles. They found bounds on the diameter of orientations of such networks and orientations that show that in most cases, these bounds are tight.

Deadlock-free Routing in Networks Jie Wu and Li Sheng (V) proposed a deadlock-free routing in irregular networks using prefix routing. Prefix routing is a special type of routing with a compact routing table associated with each node (processor). Basically, each outgoing channel of a node is assigned a special label and an outgoing channel is selected if its label is a prefix of the label of the destination node. Node and channel labeling in an irregular network is done through constructing a spanning tree. The routing process follows a two-phase process of going up and then down along the spanning tree, with a possible cross channel (shortcut) between two branches of the tree between two phases. Wu and Sheng showed that the proposed routing scheme is deadlock- and livelock- free. They also studied a broad class of deadlock-free wormhole routing algorithms for interconnection networks and found a simpler proof of a lower bound of Libeskind-Hadas on the number of channels required by such routing algorithms.

Characterizations of Consistent Marked Graphs A *marked graph* is a graph with a + sign or a

- sign on each vertex and it is called *consistent* if each cycle has an even number of - signs. This concept is motivated by problems of communication networks and social networks. For example, a communication network is consistent if whenever nodes x and y have the same sign, then if a binary message originating at x goes to y by two node-disjoint paths and if it is reversed at - nodes, the same message will arrive at y no matter which of the two paths is used. Fred Roberts (PM, RU) and Shaoji Xu (GS, RU) found characterizations of consistent marked graphs and quadratic algorithms to recognize such graphs.

An Active Network Approach to Efficient Network Management Active networks are a framework where network elements, primarily routers and switches, are programmable. Programs that are injected into the network are executed by the network elements to achieve higher flexibility and to present new capabilities. Danny Raz (PM, BL) and Yuval Shavitt (PM, BL) described a novel active network architecture which primarily addresses the management challenges of modern complex networks. Its primary component is an active engine that is attached to any IP router to form an active node. The active engine, designed and implemented, executes programs that arrive from the network and monitors and controls the router actions. The design is based on standards (Java, SNMP, ANEP over UDP) and can be easily deployed in today's IP networks. The contribution of Raz and Shavitt was the introduction of novel architectural features such as: isolation of the active mechanism, the session concept, the ability of active session to control non-active packets, and blind addressing. Implementing these ideas, Raz and Shavitt built a system that enables the safe execution and rapid deployment of new distributed management applications in the network layer. This system can be gradually integrated in today's IP networks, and allows smooth migration from IP to active networking.

The following result was obtained as part of our project on simulation of the global Internet described in Section 3.10.

Efficient Radio Wave Propagation for Wireless Radio Signal Prediction Due to the ever increasing complexity of wireless network design, wireless network simulation engines have become a necessary tool for wireless network engineers. A signal-strength prediction system lies at the foundation of every wireless simulator, making signal-strength prediction for microcells and picocells an important research topic in the radio field. Efficient and accurate methods are needed for use in current wireless network design and future wireless network research. Greg Perkins (GS, RU) and Diane Souvaine (PM, RU) developed a system, PREDICT, designed for use in wireless radio signal prediction research. PREDICT tracks the expansion of a transmitter's wavefront by modeling the wavefront as an ever-expanding sphere centered upon the transmitter. The sphere's expansion is tracked by discretizing the sphere's surface with a triangulation and then propagating each triangular wavefront through the environment. In the propagation prediction of these triangular wave fronts, PREDICT supports reflection, transmission and diffraction based upon the geometric optics (GO) model of electromagnetic waves. Signal prediction is performed through a special summation of the predicted multi-path at a receiver. PREDICT's algorithms were specifically designed for use in radio signal prediction. The current environments supported by PREDICT are urban with flat or varied terrain and a single floor of a building, where all obstacles in the environment have vertical or horizontal walls. The system design is flexible, allowing PREDICT to be expanded for use in many different environments. Modularity allows for easy modification to the GO model or implementation of alternate radio modeling methods (one can easily modify PREDICT, or use the results generated by PREDICT, with one's own radio wave modeling methods). PREDICT is publicly available for any non-commercial use.

Average-Case Analysis of the Sum-of-Squares Bin Packing Algorithm In a "discrete distribution" we are given a bin capacity B and a probability distribution over the item sizes from 1 to B-1. Such distributions model many real-world applications of bin packing, where our goal is typically to minimize the total number of bins used or, equivalently, to minimize the wasted space in the non-empty bins of the packing. If N is the number of items generated, it is known that for any discrete distribution, the total expected waste in the optimal packing must grow either as $O(N), O(\sqrt{N})$, or $O(1)$, depending on the distribution. Standard bin packing algorithms such as Best Fit can be far from optimum, for instance yielding linear waste under distributions where the optimal expected waste is $O(1)$. This is true even for offline algorithms such as Best Fit Decreasing. Janos Csirik (V), David Johnson (PM, ATTL), Claire Kenyon, Jim Orlin, Peter Shor (PM, ATTL), and Richard Weber described a remarkably simple $O(nB)$-time online algorithm, the "Sum-of-Squares" algorithm (SS), for which the expected wasted space is provably $O(\sqrt{N}$ whenever the optimal expected waste is sublinear. Based on this, they found an $O(nBlogB)$ algorithm that essentially matches the optimal expected waste growth rate for all discrete distributions. The asymptotic worst-case ratio for SS lies somewhere between 1.54 and 3, but SS can be incorporated into a randomized offline $O(nBlogB)$ algorithm whose worst-case expected ratio is asymptotically 1.

Low Discrepancy Sets Yield Approximate Min-Wise Independent Permutation Families Motivated by a problem of filtering near-duplicate Web documents, Broder, Charikar, Frieze and Mitzenmacher defined the following notion of *ϵ-approximate min-wise independent permutation families.* A multiset $\mathcal{F}$ of permutations of $\{0, 1, \ldots, n-1\}$ is such a family if for all $K \subseteq \{0, 1, \ldots, n-1\}$ and any $x \in K$, a permutation π chosen uniformly at random from $\mathcal{F}$ satisfies

$$\mid \Pr[\min\{\pi(K)\} = \pi(x)] - \frac{1}{|K|} \mid \leq \frac{\epsilon}{|K|}.$$

Michael Saks (PM, RU), Aravind Srinivasan, Shiyu Zhou (PM, BL) and David Zuckerman (V) showed connections of such families with *low discrepancy sets for geometric rectangles*, and gave explicit constructions of such families $\mathcal{F}$ of size $n^{O(\sqrt{\log n})}$ for $\epsilon = 1/n^{\Theta(1)}$, improving upon the previously best-known bound of Indyk. They also presented polynomial-size constructions when the min-wise condition is required only for $|K| \leq 2^{O(\log^{2/3} n)}$, with $\epsilon \geq 2^{-O(\log^{2/3} n)}$.

A Minimax Game of Power Control in a Wireless Network under Incomplete Information Tiina Heikkinen (GS, RU) studied optimal transmit power control and capacity of a wireless multimedia network, where there is incomplete information about the link gain coefficients. Previously, Perron-Frobenius theory had been applied to the study of optimal power control under complete information about the link gain matrix and the interference matrix. Heikkinen derived a characterization of the optimal power vector and capacity in terms of an incomplete-information-matrix-game against nature choosing the elements of the link gain and interference matrices. A game-theoretical rationale for not using power control was obtained.

High Performance Distributed Protocols on a Cluster of PCs Distributed Shared Memory (DSM) protocols involve a lot of communication and this communication often proves to be a bottleneck in the performance of such protocols due to decreased processor utilization. Liviu Iftode (RU) and Murali Rangarajan (GS, RU) developed a 2-pronged approach to improve the performance: Hide the communication latency by using multi-threading, and further improve the performance by using appropriate scheduling policies; evaluate the use of a co-processor on the network interface to perform protocol actions, thus freeing the host processor from being interrupted to perform these actions. They did this by using one of the processors in an SMP machine to emulate the coprocessor in the network interface, freeing the other processor from being interrupted for protocol actions. Iftode and Rangarajan developed a multi-threaded implementation of Home-based Lazy Release Consistency (HLRC) protocol, which is a DSM protocol. With an 8 node cluster of SMPs (two PII 300 MHz processors) connected by a high speed Myrinet (Gigabit) network, they obtained results to indicate performance gains from multi-threading by a significant reduction in the idle time, leading to increased processor utilization.

EmailValet: Learning User Preferences for Wireless Email Sofus Macskassy, Aynur Dayanik (GS RU), and Haym Hirsh (PM, RU) studied EmailValet, a system developed at RU that learns users' email-reading preferences on email-capable wireless platforms – specifically, on two-way pagers with small "qwerty" keyboards and an 8-line 30-character display. They gathered data on email-reading preferences over more than 8900 email messages received by their people for about three months. They compared the ability of different learning methods to form models that can predict whether a given message should be forwarded to the user's wireless device. Their results showed that the best performance of one method, over a range of established learning methods developed by the information retrieval and machine learning communities, was able to achieve a break-even point of over 53% for one user that had received over 5000 messages. They also found that, in general, all methods are able to achieve better performance than what would be achieved by a baseline of simply forwarding all messages to the wireless device, and that many methods are able to find procedures that, although they forward only a small fraction of the messages that a user would want, achieve 100% precision on those messages that they do actually choose to forward.

Exploiting Regularities in Web Traffic Patterns for Cache Replacement Caching web pages at proxies and in web servers' memories can greatly enhance performance. Proxy caching is known to reduce network load and both proxy and server caching can significantly decrease latency. Web caching problems have different properties than traditional operating systems paging, and cache replacement can benefit by recognizing and exploiting these differences. Edith Cohen (PM, ATTL) and Haim Kaplan (PD) addressed two aspects of the predictability of traffic patterns: the overall load experienced by large proxy and web servers, and the distinct access patterns of individual pages. They formalized the notion of "cache load" under various replacement policies and demonstrated that the trace of a large proxy server exhibits regular load. Predictable load allows for improved design, analysis, and experimental evaluation of replacement policies. They provided a simple and (near)-optimal replacement policy when each page request has an associated distribution function on the next request time for the page. Without the predictable load assumption, no such online policy is possible and it is known that even obtaining an offline optimum is hard. For experiments, predictable load enables comparing and evaluating cache replacement policies using *partial traces*, containing requests made to only a subset of the pages. Their results are based on considering a simpler *interval caching* model for cache replacement. They related traditional and interval-caching policies under predictable load, and derived (near)-optimal replacement policies from their optimal interval-caching counterparts.

Connection Caching Communication between clients and servers in the Web is performed using TCP (Transmission Control Protocol). TCP first opens a connection between the two nodes and then sends data packets via this open connection. The time needed to open a connection contributes to the latency of servicing a request. Hence, latency can be reduced if open connections are kept idle and re-used for future requests. Indeed, new versions of HTTP support that. Web servers and proxies typically maintain many open connections, many of them idle. The number of connections each node can keep open is limited. Hence, connection management is increasingly perceived and handled as a caching problem. A fundamental difference between *connection-caching* and standard caching is that connection caching involves multiple caches that affect each other. Each open connection is cached at two nodes. If one of these nodes decides to terminate the connection, the other node is forced to accept this decision. Edith Cohen (PM, ATTL), Haim Kaplan (PD), and Uri Zwick proposed a theoretical model for connection caching and studied the connection-caching problem in this model. While the ordinary off-line caching problem is easily solved in polynomial time, they showed that the off-line connection caching problem is NP-complete, and in fact also APX-complete, i.e., it is even hard to approximate. They complemented this result by showing that a distributed version of Belady's algorithm, which is optimal for a single cache, is a 2-approximation algorithm for off-line connection-caching. They provided examples that show that the analysis of this algorithm is tight.

Prefetching the Means for Document Transfer: A New Approach for Reducing Web Latency User-perceived latency is recognized as the central performance problem in the Web. Haim Kaplan (PD) and Edith Cohen (PM, ATTL) systematically measured factors contributing to this latency, across several locations. Their study revealed that DNS query times, TCP connection establishment, and start-of-session delays at HTTP servers, more so than transmission time, are major causes of long waits. Wait due to these factors also afflicts high-bandwidth users and has detrimental effect on perceived performance. They proposed simple techniques that address these factors: (i) *pre-resolving* host-names (pre-performing DNS lookup), (ii) *pre-connecting* (prefetching TCP connections prior to issuance of HTTP request), and (iii) *pre-warming* (sending a "dummy" HTTP `HEAD` request to Web servers). Trace-based simulations demonstrated a potential to reduce perceived latency dramatically. Their techniques surpassed document prefetching in performance improvement per bandwidth used and can be used with non-prefetchable URLs. Deployment of these techniques at Web browsers or proxies does not require protocol modifications or the cooperation of other entities. Applicable servers can be identified, for example, by analyzing hyperlinks. Bandwidth overhead is minimal, and so is processing overhead at the user's browser. Kaplan and Cohen proposed scalable deployment solutions to control the potential overhead to proxies and particularly to Web servers.

The following six results connect up two special years, Large Scale Discrete Optimization and Networks.

The Element Connectivity Problem in Survivable Network Design The element connectivity problem falls in the category of survivable network design problems – it is intermediate to the versions that ask for edge-disjoint and vertex-disjoint paths. The edge version is by now well understood from the view-point of approximation algorithms, but very little is known about the vertex version. Vijay Vazirani (V) considered the problem where vertices are partitioned into two sets: terminals and non-terminals. Only edges and non-terminals can fail – we refer to them as *elements* – and only pairs of terminals have connectivity requirements, specifying the number of element-disjoint paths required. Vazirani found an algorithm that achieves an approximation guarantee of factor $2H_k$, where k is the largest requirement and $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Besides providing possible insights for solving the vertex-disjoint paths version, the element connectivity problem is of independent interest, since it models a realistic situation.

Primitive Ring Generation and Ring Grooming in Design of Optimal Networks A central problem in the design of optical networks is the placement of the ADM (Add/Drop Multiplexes) devices. Many restricted versions of this general design problem have been studied in the literature. The general version can be formulated as a discrete optimization problem. Due to the complexity of the problem, the current methodology recommends dividing the general version of the design problem into two phases, called Primitive Ring Generation and Ring Grooming. The General Design Problem, Primitive Ring Generation, and Ring Grooming are NP-Hard. We have obtained theoretical results on the quality of approximation algorithms. Primitive Ring Generation has two variants, depending on whether or not traffic demands can be split. Gruia Calinescu (PD) and Peng-Jun Wan designed new algorithms for Primitive Ring Generation guaranteed to produce an output of cost at most 25% bigger than optimum if traffic demands can be split, and at most 50% bigger than optimum if traffic demands cannot be split. All previously known algorithms either have bigger worst-case guarantees or are provably not better than the new algorithms. For Ring Grooming, the traffic granularity parameter seems to influence the quality of the approximation algorithms. Gruia Calinescu (PD), Li-Wu Liu and Peng-Jun Wan provided a nontrivial analysis of a natural algorithm, proving performance guarantees of 1.5 and 2.5 respectively, for the important case when traffic granularity is 4 and 8 respectively.

Directed Multicut vs. Multicommodity Flow Given a graph $G = (V, E)$ and k pairs of vertices $(s_1, t_1), ..., (s_k, t_k)$, a *multicut* is a subset of edges $F \subseteq E$ such that if all edges in F are deleted, none of the pairs (s_i, t_i) are in the same connected component in the remaining graph $G' = (V, E \setminus F)$. In the minimum multicut problem, we are also given a positive cost $c(e)$ for each edge $e \in E$ and we wish to find the multicut of minimal total cost. This problem is known to be NP-complete for $k \geq 3$, and therefore it is interesting to consider approximation algorithms for it. One way to do it is to state the multicut problem as an integer programming problem, relax it to a linear programming problem, and solve this problem or it's dual - the *maximum multicommodity flow problem.* The maximal multicommodity flow will always be at most as big as cost of the minimal multicut. On the other hand, it is easy to see that there are multicuts which cost at most k times the value of the maximal flow. Thus, the described approach gives a factor-k approximation algorithm for a minimal multicut. It turns out that when the underlying graph is *undirected* more can be said, and in fact the cut/flow ratio in this case can be at most as large as $O(\log(k))$. On the positive side, an algorithm of Garg, Vazirani and Yannakakis exists, which takes a maximal flow and constructs a multicut which costs at most as much as $O(\log(k))$ times the flow cost. Michael Saks (PM, RU), Alex Samorodnitsky (PD), and Leonid Zosin (PD) showed that in the more general *directed* case, the cut/flow ratio could be as large as the trivial upper bound of k. It seems, therefore, that to find a better than a factor-k approximation, one must look for other approaches than the simple linear relaxation technique.

A Unified Approach to Approximating Resource Allocation and Scheduling Amotz Bar-Noy (ATTL), Reuven Bar-Yehuda (V), Ari Freund, Joseph (Seffi) Naor, and Baruch Schieber developed a general framework for resource allocation and scheduling problems. Given a resource of fixed size, they developed algorithms that approximate the maximum throughput or the minimum loss by a constant factor. Their approximation factors apply to many prominent problems, among others: (i) Real-time scheduling of jobs on parallel machines. (ii) Bandwidth allocation for sessions between two endpoints. (iii) General caching. (iv) Dynamic storage allocation. (v) Bandwidth allocation on optical line and ring topologies. For many of the problems, this is the first constant factor approximation algorithm. The algorithms are extremely simple and efficient. They use the local-ratio technique but they can also be equivalently interpreted within the primal-dual schema.

Load Balancing in a Distributed Computing Environment Endre Boros (PM, RU) and Ewald Speckenmeyer (V) studied load balancing in a distributed computing environment. They showed that the minimum message length to re-balance the load of n processors can be computed in linear $O(n)$ time, even though the actual balancing may require $O(n^2)$ messages exchanged between the processors.

A Bottleneck Load Balancing Minimization Problem Consider a processor network with n overloaded and m underloaded processors. The overloaded and the underloaded nodes are connected by a completely bipartite network $K(n, m)$. Ewald Speckenmeyer (V) considered the problem of achieving a uniform load-distribution among all processors via sending appropriate amount of load from overloaded to underloaded nodes so that the maximum of load sent via a link is minimized. He showed that this problem can be solved in time $O(mn)$.

The following three results are joint with the Special Focus on Computational Intractability.

Randomizing Low-Degree Polynomials and Round-Efficient Secure Computation Yuval Ishai (PD) and Eyal Kushilevitz introduced a notion of *randomizing low-degree polynomials* as a powerful and conceptually simple paradigm for obtaining round-efficient security. "Standard" low-degree multivariate polynomials can be securely evaluated in a small number of rounds by virtually any "off-the-shelf" secure multi-party computation protocol. However, most Boolean functions f on n input bits cannot be evaluated by a polynomial whose degree is smaller than n. Ishai and Kushilevitz got around this barrier by relaxing the requirement of evaluating f into a weaker requirement of *randomizing* f; that is, mapping the inputs for f along with independent random inputs into several outputs, whose *joint distribution* depends only on the value of f. They showed that degree-3 polynomials are sufficient to randomize *any* function f, where the number of outputs is at most quadratic in the branching program size of f, and where both input variables and random variables count towards the degree. As an immediate application, the secure evaluation of any function f reduces to a secure evaluation of degree-3 polynomials; this reduction is applicable in any reasonable setting for secure computation. One corollary is that two (respectively, three) communication rounds are sufficient to compute any n-argument function f with perfect information-theoretic $\lfloor \frac{n-1}{3} \rfloor$-privacy (resp., $\lfloor \frac{n-1}{2} \rfloor$-privacy), probabilistic correctness, and communication complexity which is at most quadratic in the branching program size of f. This improves over a previous constant-round protocol of Bar-Ilan and Beaver, which in a similar setting requires more communication rounds and a higher communication complexity.

Adaptive vs. Non-Adaptive Security In a *non-adaptive* attack against a cryptographic protocol, the adversary must choose in advance the set of parties it corrupts. A stronger type of attack, called an *adaptive* attack, allows the adversary to dynamically choose which parties to corrupt as the protocol proceeds. R. Canetti, I. Damgard, Dziembowski, Y. Ishai (PD), and T. Malkin clarified the relation between adaptive and non-adaptive security. They showed that under some previously studied settings for secure computation, such as perfect security against a passive adversary or a restricted type of an active adversary, any protocol which is non-adaptively secure must also be adaptively secure. On the other hand, they provided provable separations between non-adaptive and adaptive security in other settings.

Correlated Pseudorandomness and Compressing Cryptographic Resources A private-key cryptosystem may be viewed as a means by which a trusted dealer privately conveys a large, shared pseudorandom object to a pair of players, using little communication. Alternatively, the messages distributed by the dealer may be viewed as a secure *compression* of a pair of large identical random pads (or random functions) into a shorter shared "key" or "seed". Yuval Ishai (PD) extended this compression problem to more general correlation patterns among *several* players. Unlike the simple case of identical pads, where the main security concern is with respect to external eavesdroppers, in the case of general correlations participants also have to be protected from each other. That is, collusions of computationally-bounded players should gain no additional knowledge about the joint pads of the remaining players from the compressed messages they receive, other than what follows from the pads they generate and from knowing the joint distribution of all pads. Ishai developed a modular replication-based approach for securely compressing any *linear* correlation pattern, in which the random pads jointly form a random vector in a linear space. Specifically, the compressed messages received by the players consist of *independent* random seeds, where each seed may be replicated among several players. This approach allows one to use pseudo-random generators or pseudo-random functions in a black-box manner.

Mixing in Small Batches Markus Jakobsson (PM, BL) and Ari Juels developed Millimix, a mix network that is highly efficient on small input batches. The construction is conceptually simple, and both robust and private in the face of collusion by any minority set of malicious players. Additionally, Millimix has the feature of being publicly verifiable. In other words, the mixing operation yields a transcript that demonstrates correctness to a third party that the mix proceeded correctly.

Approximate Solution to the Directed Steiner Tree Problem The directed Steiner tree problem is the following: given a directed graph $G = (V, E)$ with weights on the edges, a set of terminals $S \subseteq V$, and a root vertex r, find a minimum weight branching T rooted at r, such that all vertices in S are included in T. This problem is known to be NP-hard. There are many papers that give approximation algorithms for this problem starting with the work by Zelikovsky, and more recently by Charikar, Chekuri, Cheung, Dai, Goel, Guha and Li. In these papers a polynomial time algorithm is developed with an approximation ratio of $O(k^\epsilon)$ for any fixed $\epsilon > 0$. This is extended to a poly-log approximation algorithm, but the running time is not polynomial. Samir Khuller and Leonid Zosin (PD) proved that the gap between a standard fractional formulation of this problem and the integral minimum Steiner tree is $O(\sqrt{k})$. Seffi Naor and Zosin found a new fractional formulation for this problem and proved that the gap for this formulation is at most $O(\log^3 k)$.

Testing Shared Memories Sequential consistency is the most widely used correctness condition for multiprocessor memory systems. Phil Gibbons (PM, BL) and Ephraim Korach (V) studied the problem of testing shared-memory multiprocessors to determine if they are indeed providing a sequentially consistent memory. Their work represented the first formal study of this problem, which has applications to testing new memory system designs and realizations, providing run-time fault tolerance, and detecting bugs in parallel programs. A series of results were obtained for testing an execution of a shared memory under various scenarios, comparing sequential consistency with linearizability, another well-known correctness condition. Linearizability imposes additional restrictions on the shared memory, beyond that of sequential consistency; these restrictions were shown to be useful in testing such memories. (Gibbons and Korach received the Yossi Levy Prize in Operations Research and Managerial Economics for the year 1999 from the Operations Research Society of Israel for this work.)

Placing Web Proxies on Tree-Broadcasting Networks Mordecai Golin (V) and Lixin Gao (V) designed an algorithm for placing a given number of proxies (copies) of a web page on nodes in a tree-topology network so as to minimize the average time any user would need to access a copy of the page. In graph-theoretic terms, this is a k-median problem.

New Upper and Lower Bounds on the Channel Capacity of Read/Write Isolated Memory In general, we may regard the set of all permissible binary memory configurations as a channel alphabet which obeys some restrictions, where, for each alphabet symbol, the rewriting restrictions (if they exist) dictate which symbols may succeed it. In this way, the rewritable memory can be viewed (over its lifetime) as a noiseless communication channel. Shannon defined the capacity of memory, in bits per rewrite, as $C_k = \lim_{r\to\infty} \frac{1}{r} \log_2 N(k, r)$, where k is the size of the memory in binary symbols, r is the liftime of the memory in rewrite cycles, and $N(k,r)$ is the number of distinct sequences of r characters. Shannon showed that $C_k = \log_2 \lambda_k$ where λ_k is the largest eigenvalue of the channel graph of order k, i.e., of its associated adjacency matrix. Xuerong Yong (GS, V), Yuanping Zhang (GS, V), Mordecai Golin (V), and Li Sheng (V) found some new recursive properties of λ_k and new lower and upper bounds on C_k.

Watermarking Maps: Hiding Information in Structured Data The storage of data in digital form makes reproduction of valuable data, and thus piracy, easier than ever. Watermarking techniques address this problem by embedding hidden information into the data, which either asserts a copyright or allows tracing of copies. Sanjeev Khanna (PM, BL) and Francis Zane (PD) examined the problem of protecting map data from piracy. Consider the following scenario: MapQuest commissions resources to produce an accurate map of the US, and sells modified, or watermarked, copies to web sites like Yahoo!. Each web site then uses this map to answer queries (e.g, what is the shortest route from A to B?). If a new map web site appears, MapQuest wants to know two things: Is this site using a pirated MapQuest map? If so, which copy is it using and who purchased that copy? To do this, they hide some information in each copy of the map (say, small changes in the distances), allowing them to distinguish the copies. However, these small changes can easily accumulate over long paths to yield unacceptably distorted distances. Khanna and Zane presented a method for making these changes which encodes the identities of many users while introducing a very small distortion. In addition, a web site with a pirated copy will naturally attempt to disguise its copy with misleading answers to queries. Under some natural assumptions (e.g., the pirate's answers cannot distort the graph too much), they presented a scheme which detects even such evasive pirates with high probability.

3.3 Special Focus on Next Generation Networks Technologies and Applications 2000-2003

Overview:

The Special Focus on Next Generation Networks Technologies and Applications will begin in January 2000 and go at least through August 2003. The year builds on the 1996-2000 Special Year on Networks. It will coincide with our Special Year in Networks as the latter activity winds down. Funds for this activity have been obtained from the National Security Agency and from the Office of Naval Research.

Organizers:

- B. R. Badrinath, Rutgers University
- Sandeep Bhatt, Akamai Technologies
- Joan Feigenbaum, AT&T Labs – Research
- S. Muthukrishnan, AT&T Labs – Research

Advisory Committee:

- Sajal Das, U. Texas
- Cynthia Dwork, IBM
- Lee Giles NEC Research
- David Goodman Poly Univ.
- Tomasz Imielinski, Rutgers CS & WINLAB
- David Johnson, AT&T Labs
- Tom Leighton, MIT and Akamai Technologies
- Cathy Meadows, Naval Research Laboratory
- Andrew Ogielski, Rutgers DIMACS & WINLAB
- Hilarie Orman, Novell
- Prabhakar Raghavan, IBM
- Nambi Seshadri, Broadcom
- Mihalis Yannakakis, Bell Labs
- Roy Yates, Rutgers WINLAB

Sponsors and Collaborating Institutions:

AT&T Labs - Research
Bell Labs
IBM Research
The National Science Foundation
The National Security Agency
NEC Research Institute

The New Jersey Commission on Science and Technology
The Office of Naval Research
Princeton University
Rutgers University
Telcordia Technologies
WINLAB (Wireless Networking Lab), Rutgers University

Theme of Program:

Next generation networks technologies offer opportunities to dramatically transform society. While the underlying technologies continue to advance at a rapid pace, the fundamental principles governing the design, deployment and use of next-generation networks of unprecedented scale, heterogeneity, and complexity are not entirely clear. How are systems involving the emerging technologies (such as cable or wireless) best designed? This question pervades all levels of networking: at the network level (providing a personal network space, designing QoS for heterogeneous networks such as wired-cable and wired-wireless); at the application level (providing a thin-client view of the web for mobile units, secure electronic commerce, etc.); at the system level (providing interoperability, a homogeneous view of the system in the presence of mobility, etc.); at the protocol level (how to modify TCP for wired-wireless networks, how to multicast in wireless networks, etc.); at the link level (scheduling coding and power level for packet transmissions). New network services raise new challenges: Privacy can be threatened; small faults can cause cascading network failures with severe economic consequences; the sheer number of requests for information can slow the networks and information servers to a crawl; protocols created for previous generation networks may not be appropriate for the scale and complexity of newer technologies; achieving effective interoperability among systems remains extremely difficult; managing distributed systems with mobile software is poorly understood; multiple networks and software systems are being interconnected without a full understanding of the ramifications. This "special focus" seeks to take advantage of new opportunities and to solve underlying problems by addressing the theoretical foundations on which the emerging environment of commingled communication, information, and computation is based.

The special focus will concentrate on the following:

- Algorithms to achieve cost- and service-effective use of the resources in the emerging environment in virtually all areas, including communications (mobile or stationary, network design, protocols, routing), large scale planning and scheduling, and economic decision making. This will involve expertise from fields such as optimization, approximation, probability and queuing theory, control theory and game theory.

- Models for massive data analysis of multimedia information (e.g., data warehousing and data mining) and traffic and information management in global intelligent networks. Because existing algorithms for data analysis are unlikely to scale, new methods are needed to organize, store, and retrieve data, discover patterns, and use them for planning and decision making. This involves expertise from diverse areas including data structures, databases, data compression and coding, statistics, learning, geometry and visualization.

- Fundamental theoretical work to address problems of reliability, security and privacy. The design of systems that maximize the potential for new technologies while protecting security and privacy requires expertise in fields such as verification, distributed computing, probability, coding theory, number theory, cryptography and security.

This special focus is intended to foster closer working relationships between the algorithms and networking technology communities. We will pursue the development of the fundamental theory and algorithms that cut through a variety of application domains. The goal of the special focus activities is to discover and understand the formal models and algorithmic principles that will enable us to manage the present technology and design the future technology, from the physical layer to the application layer. Foundational work is required to solve existing problems and to lay the groundwork for the next generation networks technologies and applications, and our special focus will be aimed at pinpointing the opportunities for theorists to make a contribution and at stimulating the research needed to take advantage of these opportunities.

Workshops and Special Events Rest of 1999-00

These are joint with Special Year on Networks as that Special Year winds up and this one begins.

- **Data Processing on the Web: A Look into the Future**
  - Dates: March 6 - 7, 2000
  - Location: Rutgers University
  - Organizers: Mary Fernandez, S. Muthukrishnan, Dan Suciu, AT&T Labs - Research
- **Cryptography and Intractability**
  - Dates: March 20 - 22, 2000
  - Location: Rutgers University
  - Organizers: Moni Naor, Weizmann Institute of Science, Joe Kilian, NEC Research Institute, and Shafi Goldwasser, MIT and Weizmann Institute of Science
  - Comment: Joint with Special Year on Computational Intractability
- **Management of Digital Intellectual Property**
  - Dates: April 17 - 18, 2000
  - Location: Rutgers University
  - Organizers: Joan Feigenbaum, AT&T Labs - Research, Dan Boneh, Stanford University, and R. Venkatesan, Microsoft
  - Co-sponsor: Microsoft
- **Web Characterization, Metrics and Statistics** (miniworkshop)
  - Dates: Spring 2000
  - Location: Rutgers University
  - Organizers: Lee Giles, NEC, Steve Lawrence, NEC, Sam Weerahandi Telcordia Technologies
- **Data Mining in the Internet Age**
  - Dates: May 1-2, 2000
  - Location: IBM Almaden
  - Organizers: Joan Feigenbaum, AT& T Labs, Prabhakar Raghavan, IBM Research
  - Co-sponsor: IBM Research

3.4 Special Year on Mathematical Support for Molecular Biology, 1994-00; Plans for Special Focus on Computational Molecular Biology, 2000-2003

Overview: The Special Year on Mathematical Support for Molecular Biology 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.

Because of continuing interest in it, the program is now extending into a sixth year of activity. While the level of activity in years 5 and 6 has not been as high as in earlier years, it has led us to initiate another high-visibility Special Focus on Computational Molecular Biology beginning in Fall 2000. (More on this below.) We have continued the program with several workshops, an active seminar series, a postdoctoral fellow, visitors, and the involvement of graduate students, both local and non-local. This year also saw the publication of a DIMACS book featuring surveys related to topics in our special year: *Mathematical Support for Molecular Biology*, edited by Martin Farach-Colton, Fred S. Roberts, Martin Vingron, and Michael Waterman. We have built up serious collaborative relationships with other centers and departments in the biological sciences and continue to take advantage of those relationships through this program.

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 (retired)
- Charles Cantor, Boston University
- Daniel Gusfield, UC - Davis
- Leroy Hood, Center for Molecular Biotechnology, Washington
- Richard Karp, UC Berkeley
- Eric Lander, MIT - Whitehead Institute
- Eugene Lawler, UC - Berkeley (deceased)
- Joshua Lederberg, Rockefeller
- Joseph Naus, Rutgers
- Michiel Noordewier, SmithKline Beecham
- Wilma Olson, Rutgers

- David B. Searls, SmithKline Beecham
- William Sofer, Rutgers
- Tandy Warnow, University of Texas
- Peter Winkler, Bell Labs
- 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, Florida State University
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
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, Florida State University
Promega Corporation
Roche Molecular Systems
Rutgers University
Sandia National Laboratories
The Alfred P. Sloan Foundation
Smith-Kline Beecham
Telcordia Technologies

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 February 1999 - January 2000:

- **Workshop: Advances and Opportunities at the Biology/Math/Computation/Physical Sciences Interface**
  - Dates: March 6-7, 1999
  - Location: Rutgers University
  - Attendance: 149
  - Organizers: Wilma Olson, Fred Roberts, Rutgers University
  - Co-sponsors: Center for Molecular Biophysics and Biophysical Chemistry, Rutgers; Program in Mathematics and Molecular Biology, Florida State University

  The workshop was co-sponsored by the Center for Molecular Biophysics and Biophysical Chemistry and the Program in Mathematics and Molecular Biology based at Florida State University. The goal was to introduce interested undergraduates, graduate students, postdoctorals and faculty members in mathematics, computational, and physical sciences to on-going activities and future opportunities in the biological sciences. The workshop drew approximately 150 participants from major institutions in the Northeast. Sixteen researchers presented invited talks on a broad range of topics, and 19 graduate and undergraduate students received travel awards to attend the workshop. A highlight of the event was a reception and poster session where 40 research and recruitment posters were presented. A session on protein folding featured talks by Lois Pollack, Mark Gerstein, and Dannie Durand. Speakers on the subject of DNA Topology were Matthew Young, David Swigon, and Alexander Vologodskii. Helen Berman chaired a session on Databases and Drug Design with talks by Tamar Schlick, Daniel Pilch, and Richa Agarwala. A session on Large-Scale Perspectives featured talks on microengineering cellular footprints on tissue analog polymers by Prabhas Moghe and one on topological assys for DNA structure and function by De Witt Sumners. Diane Zimmerman, Andrej Sali, and Craig Nevill-Manning spoke in a session on Protein Structure and Function. The conference concluded with talks by Robert Austin and Nadrian Seeman on Nanotechnology. There was a great deal of enthusiasm displayed by the all of the attendees.

- **Protein Structure and Function**
  - Dates: November 19 - 21, 1999
  - Location: Rutgers University
  - Attendance: 36
  - Organizer: Israel M. Gelfand, Rutgers University

The goal of the workshop was to discuss the problems of structural genomics: to classify the protein sequences and accurately predict their three-dimensional structures and functions. A broad spectrum of different methods of analysis of the relationship between sequences, secondary structures, three-dimensional structures and protein functions was discussed. The flood of new protein sequences demands new techniques to infer protein 3D structures from sequence alone. The workshop was devoted to the memory of Professor Oleg Ptytsyn, who greatly encouraged the idea of this workshop. We were fortunate to be able to obtain such authoritative speakers from different academic institutes and companies in the USA, Russia and England. C. Chothia spoke on protein families and gene duplications in 19 genomes. He showed how the major transitions in the evolution from prokaryotes to vertebrates affect protein repertoires. The analysis of 19 genomes revealed that the duplications of pre-existing genes, their divergence and recombination have an important role in the large increases in protein repertoires. The second part of the workshop was started with a talk by I. Gelfand entitled "The Problem of Adequate Language in Biology". Having an adequate language (an adequate set of notions and ideas) is a necessary prerequisite for development of any area of science. R. Russell described the problem of predicting function from protein structure comparison. Functional insights are based on the comparison of a protein with known structure, but of unknown function with other proteins of known structure. I. Gelfand and A. Kister discussed the problem of how unlike protein chains such as cadherins, immunoglobulin variable domains, fibronectins and other proteins of the Ig-like fold can share the same fold. They discovered a set of shared, structurally-similar amino acids which are common to all highly non-homologous proteins with immunoglobulin-like fold. A. Finkelstein presented the results of several new approaches to the problem of finding the folding nucleus in a given 3D protein structure. He showed that the participation of residues in the hydrophobic core and the secondary structure of native protein has a rather modest correlation with the experimentally found F values indicating participation of a residue in the transition state characterizing the participation of residues in the folding nuclei.

- **Discrete Mathematical Problems and Medical Applications**
  - Dates: December 8 - 10, 1999
  - Location: Rutgers University
  - Attendance: 35
  - Organizers: Ding-Zhu Du, University of Minnesota, Panos M. Pardalos, University of Florida, Jie Wang, University of North Carolina

Computer technologies have been widely used in medical research. For instance, virtual endoscopy is a new medical technology developed in recent years which allows physicians to examine computer simulations of patients' anatomy rendered from CT (computed tomography) scans, and provides the public with alternative medical procedures that are less painful, less costly, and less risky compared to conventional endoscopic procedures. While the methodology of continuous mathematics is used in most research on medical computing, discrete algorithmic techniques have also played an important role in developing various medical applications. For example, graph optimization algorithms have been used in developing virtual endoscopy and dynamic programming techniques have been used in finding an optimal radiosurgical treatment plan. Many medical applications call for solutions to various discrete algorithmic problems. For instance, in radiosurgical treatment planning, it is important to find the best positions to place radiation doses within a target image. In gamma unit treatment planning, the dose delivery is based on the unit shot, which is approximately a spherical dose distribution in 3D space. For irregularly shaped target volumes, it is a challenging task to pack a set of shots together to yield an acceptable cumulative dose distribution. This is a discrete optimization problem, in which the location, collimator size, and relative weight of each shot and the total shots being used are optimized at the same time. The study of computing in medical applications has opened many challenging issues and problems for both the medical computing community and the algorithm community. This workshop was intended to foster communication and collaboration between medical computing and algorithm researchers. The objectives of this workshop were to provide a forum where worldwide researchers and practitioners could meet and exchange research ideas and interests, as well as to discuss the new trends and identify open problems in the development and deployment of this area. The workshop brought together medical physicists, computer scientists, and mathematicians presenting research results and discussing discrete mathematical problems encountered in various fields of medical applications. Researchers from universities and research hospitals in Canada, Sweden, the UK, and the US attended. Jay Udupa opened the workshop by talking about discrete problems in practical image segmentation and visualization. Yuri Boykov discussed graph cut algorithms for fast approximate energy minimization in image analysis, Calvin Maurer talked about a linear time algorithm for distance transforms in arbitrary dimensions, Kaleem Siddiqi dealt with "Divergence-based Skeletons," Alan Sprague talked about frequent set generation for medical surveillance in the presence of clones, and Tao Jiang spoke about a Gibbs sampling approach to drug target search. Ingela Nystrom talked about analysis of magnetic resonance angiography images using skeletonization and distance transforms, work done cooperatively with Linkoping University Hospital in Sweden. A spectral encoding of the tree structure and its application to shape indexing and matching was the topic of Sven Dickinson's talk, while DNA screening was the topic of H. Ngo's talk and variants of interval graphs and their applications to DNA physical mapping were dealt with by Li Sheng. R. Chandrasekaran talked about isotonic separation, Michael Ferris about optimization of gamma knife radiosurgery, Danny Chen about optimal beam penetration problems in two and three dimensions, and Anthony Macula about two-stage group testing for complexes in the presence of errors. W. Jackie Wu described morphology-guided radiotherapy treatment planning and optimization, Jie Wang the complexity of min-max sphere packing with respect to treatment planning, Yaorong Ge discrete methods in virtual endoscopy, and Peter Hall a model for learning human vascular anatomy. Sanguthevar Rajasekaran talked about efficient algorithms for similarity search, Olvi Mangasarian described his work on breast cancer survival and chemotherapy analysis via support vector machines, Anand Rangarajan discussed an integrated pose and correspondence approach to image matching, and G.T. Herman spoke about medical applications in discrete tomography. The workshop concluded with presentations by Leon Iasemidis on an optimization approach to transition to epileptic seizures, by Guoliang Xue on fast algorithms for force-field computation in molecular clusters, and by Bhaskar DasGupta on computing distances between evolutionary trees. A proceedings in the DIMACS series is planned.

Plans for New Special Focus on Computational Molecular Biology 2000-2003

Starting in September 2000, the planned special focus is motivated by major developments at the interface between biology and information science. The year builds on the 1994-1995 Special Year on Mathematical Support for Molecular Biology.

Organizers:

- Martin Farach-Colton, Rutgers, Computer Science, Chair
- Craig Nevill-Manning, Rutgers, Computer Science, Assistant Chair
- Wilma Olson, Rutgers, Molecular Biophysics, Biochemistry, and Chemistry
- Fred Roberts, Rutgers, Mathematics

Steering Committee:

- Bonnie Berger, MIT
- Helen Berman, Rutgers University
- Douglas Brutlag, Stanford University
- Andrea Califano, IBM Research
- Nick Cozzarelli, UC Berkeley
- Jim Fickett, SmithKline Beecham
- Sorin Istrail, Sandia
- Thomas Lengauer, GMD
- Ron Levy, Rutgers University
- Michael Liebman, Roche Bioscience
- Mona Singh, Princeton University
- Martin Vingron, National Cancer Center of Germany
- Michael Waterman, University of Southern California

Sponsors and Collaborating Institutions:

The Alfred P. Sloan Foundation
The New Jersey Commission on Science and Technology
Rutgers University
Princeton University
AT&T Labs - Research

NEC Research Institute
Telcordia Technologies
IBM Research

Theme of Program:

Because of new technologies, we are receiving massive amounts of data and, moreover, the data are arriving in diverse forms and are often unstructured and non-homogeneous. This requires the intervention of experts at handling databases and processing massive amounts of information. Moreover, data are now available at much higher levels of organization than heretofore (for example at the whole genome level rather than at the gene level), which makes possible dramatic new biological understanding if we can find ways to handle the complex algorithmic questions that arise.

New biotechnologies such as gene expression arrays are powerful new experimental tools, but these tools and the development of new ones need to be integrated with fundamental research concerning algorithms. Thus, close collaborations between biological scientists and computer scientists need to be developed and/or strengthened.

New mathematical and computational tools have made it possible to understand biological processes at a much more complex level than before, for example at the level of cellular communities, the immune system, protein structure and design, and processes of learning and memory. This comes at a time when biologists are finding tools to explore how parts of a biological system (e.g., genes or molecules) interact, something that has long been understood to be just as important as understanding the parts themselves but which has been largely inaccessible to experimental or analytical techniques. These new tools make it possible to formulate powerful new mathematical models of biological systems, grounded in the communication and information transmission fundamental to those systems. This calls out for the expertise of mathematical modelers and information scientists who are best able to take advantage of the powerful new mathematical methods while at the same time keeping their models realistic through interactions with biological scientists.

The DIMACS Special Year on Mathematical Support for Molecular Biology commenced in 1994-95 with a very intense level of activity and has continued at a significantly lower level since then. Through that special year, we have focused the interests of a large number of computer scientists and mathematicans on the field of molecular biology, created real partnerships between mathematical and biological scientists, and trained outstanding young people to work in computational biology. We feel that the time is right for another highly visible special year thrust, like the one in 1994-95. At the time of our first special year, there was a need to address, review, and coordinate a variety of topics and methods, as well as to introduce many people to the field. Computational biology prior to the first special year was largely focused on a few problems such as alignments, physical mapping, and reconstructing phylogenies. Now, many of the topics emphasized in the first special year have become "standard" topics of research in computational biology and the field has become much broader in its focus. Other topics have begun to be investigated in a much more far-reaching way using methods of computer science, for example problems involving gene finding and motif recognition, protein and RNA folding, protein structure prediction, and linkage analysis. New biotechnologies such as SBH (sequencing by hybridization), optical mapping, and EST (expressed sequence tags) have been brought to the attention of a much larger number of computer scientists so that they can formulate and research the right questions concerning new technologies. But a whole host of new problems that lie at the interface between the computational/mathematical sciences and the biological sciences have arisen, stimulated, as we have noted, by the availability of massive amounts of new data, the integration of new experimental methods with algorithmic methods, and the development of powerful new mathematical tools for modeling ever-more-complex biological systems.

The title of the special focus, "Computational Molecular Biology," is certainly too broad to define its focus. We will concentrate on those areas where discrete mathematics and theoretical computer science seem likely to have a major impact. A fundamental premise of the special focus is that some of the most central problems in modern molecular biology are essentially problems involving the combinatorial and algorithmic questions of dm and tcs and it is a basic objective of the focus to create partnerships between biological and mathematical/computer scientists so that some of these central biological questions can be precisely formulated and analyzed. It is our expectation that by bringing together some of the world's leading mathematicians/computer scientists with leading biological scientists, major progress can be made. Moreover, it is our expectation that we can establish and nurture lines of communication and collaboration so that the new methods of information science that arise from the special focus can be quickly adapted to biological applications, and can be readily communicated to the biological community.

Workshops and Special Events Under Development

- Integration of Diverse Biological Data
- Sequence Motif Recognition
- Whole Genome Comparison
- Learning Stochastic Processes
- Protein Structure: Recent Computational Advances and Future Directions
- System-based Modeling in Bioinformatics
- Geometric Searching
- Analysis of Gene Expression Data
- Computational Issues in Signal Transduction
- Protein Structure Comparison and Prediction Assessment
- DNA Topology

Selected Research Results of Special Year on Mathematical Support for Molecular Biology

Cluster-preserving Embedding of Proteins Similarity searching in protein sequence databases is a standard technique for biologists dealing with a newly sequenced protein. Exhaustive search in such databases is prohibitive because of the large sizes of these databases and because pairwise comparisons are slow. Heuristic techniques, such as FASTA and BLAST, are useful because they are fast and accurate, though it has been shown that exhaustive search is more accurate. Therefore, there are times when one would like to perform an exhaustive search. Gabriela Hristescu (GS, RU) and Martin Farach-Colton (PM, RU) deveoped an efficient method, called SparseMap, for preprocessing a database of proteins to support efficient similarity searches using expensive but sensitive distance functions, such as those based on Smith-Waterman similarity. Their method is based on a low-dimensional Euclidean Eebedding approach. They compared their method with other embedding approaches, and showed that their method is faster and produces embeddings which preserve more biological information about the proteins, such as pairwise distance and biological clusters.

An Approach of Information Theory to Multiple Sequence Comparison Multiple sequence comparison is a basic problem for molecular biology and other sciences. Weiwu Fang (V), Fred S.

Roberts (PM, RU) and Zhengrong Ma (RU, Waksman Institute) introduced the concept of complete information set and some measurement principles for measuring multiple sequence discrepancy. Based on them, they developed a new measurement method satisfying some reasonable principles for comparing multiple sequences. They showed that this method can effectively distinguish different random sequences or DNA sequences, for example, distinguish DNA sequences of length 8000 by comparisons of 6-8 symbol strings or protein sequences of length 8000 by comparisons of 3-4 strings. It can also measure slight changes of a sequence, e.g., insertion or deletion of a symbol. Fang, et al. applied their method in the study of molecular evolution; the results show that there is a hierarchic relationship among the cytochrome C protein sequences of different species, much as that in taxonomy, and these results are consistent with previous studies.

The Shortest Superstring Problem Given a set S of strings over a finite alphabet, a superstring of S is a string that contains each string of S as a contiguous substring. The *Shortest Superstring* (SS) problem is to find a superstring of minimum length. This problem is motivated by one approach to the sequence assembly problem in shotgun sequencing. Each string in S models one of the sequenced DNA fragments created by a protocol for shotgun sequencing. The assembly problem is to deduce the original DNA string from the set S of sequenced fragments. Without sequencing errors, the originating string is likely to be a superstring of S and, under some assumptions, a shortest superstring. Problem SS also has applications in data compression. SS is MAX SNP-hard so it is unlikely that the length of a shortest superstring can be approximated to within an arbitrary constant. Blum, Jiang, Li, Tromp, and Yannakakis gave the first linear approximation bound for the shortest superstring problem in 1991; they described a 3-approximation algorithm. In 1996, Elizabeth Sweedyk (PD) described a 2.5 approximation algorithm; this remains the best known result. Researchers have worked for many years towards proving a conjecture that the length of the shortest superstring can be approximated to within a factor of 2. It had been widely believed that two greedy-like algorithms achieve an approximation factor of 2. However, Sweedyk constructed an example for which one of these algorithms does worse than 2.

Predicting Amino Acid Sequences of the Antibody Human VH Chains from its First Few Residues A new method for classification of immunoglobulin sequences was developed by Boris Galitsky (V), Israel Gelfand (PM, RU), and Alex Kister (V). The defining characteristic of a class is the presence of particular residues at several class-determining positions. Sequences within a class follow the same amino acid pattern, i.e., residues at identical positions are, in an overwhelming majority of sequences of that class, identical or chemically-related. Thus, once the class of a sequence is determined, one can predict the residue(s) at almost any position in the sequence. The method was applied to the analysis of 1172 human heavy chains. It was shown that a sequence can be assigned to one of six classes depending on which residues are found at its positions 1, 3, 5, 6, 7, 9, 10, 12 and 13. It is important to note that it is possible to achieve the same six-class classification of the human heavy chains on the basis of a different set of positions found not at the beginning but near the end of the sequence (around position 80). For every class, an amino acid pattern of the entire sequence (CDRs excepting) was determined. The approach allowed Galitsky, et al. to reconstruct the incomplete human heavy chains in which residues at certain positions at the beginning or end of the chain are known. They also developed a software tool for analysis, classification and prediction of residues in sequences of the immunoglobulin family.

Class-defining Characteristics in the Mouse Heavy Chains of Variable Domains Boris Galitsky (V), Israel Gelfand (PM, RU), and Alex Kister (V) carried out the analysis of residue correlation in over 2700 mouse heavy chains of the VH domains on three hierarchical levels. At the position level, statistical analysis revealed 45 positions that conserve similar residues in almost all chains. At the fragment level, the focus of investigation shifted to study of combinations of amino acids in strands and loops. It was found that no more than 10 patterns are sufficient for describing strands and loops in the chains. At the sequence level Galitdsky, et al. determined all possible combinations of these patterns and classified the mouse heavy chains. Comparison of the sequences in the eight classes revealed residues at the class-determining positions that are unique to each class. Because a strong correlation of residues was found, one only needs several residues to classify a sequence. It follows that not all residue alignment procedures are necessary to divide sequences into classes. An important corollary of the approach is the possibility of predicting residues in an incomplete sequence from a small sequence fragment. On the basis of the analysis of the mouse heavy chains, Galitsky, et al. hypothesized about the presently unknown mouse VH germline repertoire.

The Commonality of Cadherins Compared with Highly Non-homologous Proteins of Immunoglobulin-like Fold Alex Kister (V), Mikhail Roytberg (V), C. Chothia and Israel Gelfand (PM, RU) attempted to answer a question about how such 'unlike' protein chains as cadherins and proteins of other families of the Ig-like fold can share the same fold? They considered in this work 14 protein domains. The sequences in different domains showed a very low homology. For example, the sequences of the Cu, Zn superoxidase dismutase protein family revealed low 12% homology with proteins of the purple acid phosphatase. Despite large differences in sequences, the 3D structures of all domains share an immunoglobulin-like beta sandwich. Kister, et al. described a new approach to assigning a protein to a protein family. This method allowed them to classify proteins with very low homology. They were able to find a set of residues which would play similar structural roles within their respective Ig-like molecules' sequences. This set of shared, structurally-similar amino acids constitutes a necessary characteristic of immunoglobulin folds. It serves as the base of a structure of an Ig-folded molecule.

3.5 Special Year on Computational Intractability, 1999-00

Overview:

The Special Year on Computational Intractability commenced in Fall 1999 and will continue through late Fall 2000 and possibly beyond. The first five months of this special year are running under the auspices of our science and technology center grant from NSF. We have received a new grant from NSF to run the remainder of the activities, supplemented by funds from the DIMACS institutions and a small award from the National Security Agency. We are collaborating in this activity with a group at the Institute for Advanced Study which plans a 2000-2001 special focus on closely related topics.

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 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
Institute for Advanced Study
Microsoft
NEC Research Institute
Princeton University
Rutgers University
Telcordia Technologies

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 for February 1999 - January 2000:

- **Pseudorandomness and Explicit Combinatorial Constructions**
  - Dates: October 12-15, 1999
  - Location: Rutgers University
  - Attendance: 94
  - Organizers: David Zuckerman, University of Texas, Russell Impagliazzo, University of California at San Diego, and Amnon Ta-Shma, International Computer Sciences Institute

  The two topics appearing in the title of this workshop represent important and closely linked research directions within, respectively, computer science and combinatorics. Assessing the power of randomness in computation is a central theme within the theory of intractability. The central question is: "are there problems that are intractable using deterministic computation that are tractable using randomized computation?" Research on this question, which points to a negative answer, has focused around the study of the constructibility of pseudo-random generators, and the constuctibility of these generators has been shown to be closely connected to the existence of computational problems that are suitably intractable. The development of the probabilistic method in combinatorics has led to a variety of non-constructive proofs of existence of combinatorial objects satisfying certain specified properties: graphs whose clique and independence numbers are both very small (Ramsey graphs), graphs with strong connectivity properties, such as expanders, and large codes with optimal distances. It has long been of interest to find explicit constructions of objects with these properties. These two areas are brought together by the realization that the problem of designing pseudorandom generators, and related problems such as that of making randomized algorithms robust in the presence of imperfections in the random source, can be formulated as problems of explicit construction of combinatorial objects. Thus the study of such constructions has become an important part of the theory of probabilistic computation. The computational point of view has motivated new problems in combinatorics and also led to new techniques for constructing objects like expanders that are better than any known before. This workshop explored this common area of interest, focusing on problems and techniques for constructing combinatorial objects arising in computer science, graph theory, and coding theory. The workshop opened with the first invited talk by Luca Trevisan. He spoke about a new method for constructing extractors, based on pseudorandom generators. Omer Reingold next spoke about some improvements to the original method. Ran Raz talked about using extractors and further clever ideas to deterministically simulate space-bounded computation. Chi-Jen Lu discussed using generators for space-bounded machines and other ideas to construct pseudorandom generators for combinatorial rectangles. Finally, Aravind Srinivasan spoke about using pseudorandom generators for combinatorial rectangles to construct min-wise independent permutation families. Johan Hastad opened the second day with a talk on the security of all individual RSA bits. Moni Naor continued the cryptographic theme by discussing pseudo-random functions and factoring. Noga Alon gave the second invited talk on Ramsey type graphs. Tibor Szabo discussed the extremal properties of norm graphs. Alex Lubotzky and Igor Pak concluded by discussing a method for generating random group elements. Salil Vadhan opened the third day by discussing two methods for constructing pseudorandom generators from hard functions with using the XOR lemma. Madhu Sudan elaborated one of these methods by discussing list decoding. Amin Shokrollahi gave the third invited talk on algebraic-geometric codes. Juergen Bierbauer spoke about using such codes for epsilon-biased distributions. Finally, David Zuckerman led the open problems session. Russell Impagliazzo opened the last day by discussing derandomization under a uniform assumption. Ronen Shaltiel spoke about extending some of these ideas to build extractors. Avi Wigderson gave the last invited talk on some open problems. Peter Bro Miltersen next spoke about derandomizing Arthur-Merlin games. Adam Klivans concluded the workshop by discussing boosting and hard-core sets.

- **Probabilistic Analysis of Algorithms For Hard Problems**
  - Dates: November 1-3, 1999
  - Location: Rutgers University
  - Attendance: 78
  - Organizers: Alan Frieze, Carnegie-Mellon University and Martin Dyer, University of Leeds

The theme of the workshop was characterized by the following: The analysis of algorithms for NP-hard problems under distributional assumptions on the input is an important approach to dealing with hardness issues. The workshop focused on the average case analysis of algorithms in this framework. The main goal here is to obtain good estimates of various performance measures of algorithms such as solution quality and running time. A key objective is to explain why many simple algorithms, often used in practice, perform much better than predicted by worst-case analysis. Pathological cases are generally rare and simple algorithms can often be shown to perform well in the average case. This is especially true in the case of NP-complete problems where it seems that the probabilistic approach may be the best way and indeed may be the only way to understand the success of heuristic combinatorial algorithms. We were fortunate that many of the leading researchers in the above area were able to attend. Several speakers discussed algorithms on random graphs: Richard Karp spoke on algorithms for finding cores. Van Vu spoke on the problem of finding a large independent set. Tomasz Luczak spoke on random graphs and positional games. C.R. Subramanian spoke on graph coloring. Santosh Vempala discussed the use of random graphs in proving the existence of graphs which were then used in a hardness result for Travelling Salesman Problems. Several speakers spoke on the use of Markov chains for sampling and coloring: Mark Jerrum gave an overview of recent results on torpid mixing. Christian Borgs discussed the mixing rate of some algorithms on geometric lattices. Leen Stougie discussed an optimization algorithm which used a random walk. Greg Sorkin discussed Simulated Annealing. Leslie Goldberg spoke on an extension of path coupling. Several speakers spoke on the analysis of random satisfiability problems. Jeong Han Kim talked about a tight analysis of random 2-SAT. Dimitris Achlioptas described a greedy algorithm for random 3-SAT which strengthens the current lower bound on the satisfiability threshold. John Franco described several classes of satisfiability problem and attempted to estimate their likelihood. Packing and allocation problems were another area of emphasis. David Johnson spoke on the classical bin packing problem. Ed Coffman spoke on interval packing problems such as arise in reservation systems. Angelika Steger talked about load balancing. Andrzej Ruczinski discussed packing graphs into a larger graph. Peter Winkler discussed the optimality of a class of greedy algorithms. A variety of other problems were covered. Artur Czumaj discussed improvements in derandomizing the local lemma. David Matula discussed the practical issue of rounding error in floating point computation. Ramarathnam Venkatesan discussed random Cayley graphs and the discrete logarithm. Leonid Levin spoke about one-way functions and Santosh Venkatesh on information storage in finite memory. George Lueker gave new bounds on the length of the longest common subsequence of 2 random strings. Martin Furer spoke about approximating the permanent and Joel Spencer about the randomized greedy algorithm on hypergraphs. Joe Yukich gave a survey of random Euclidean optimization problems.

Workshops and Miniworkshops Planned for the Rest of 1999-00

- **Computing Approximate Solutions to NP-hard Problems**
  - Dates: February 20-22, 2000
  - Location: Princeton University
  - Organizers: Sanjeev Arora, Princeton University, Eva Tardos, Cornell University, and Luca Trevisan, Columbia University
- **Faster Exact Algorithms for NP-Hard Problems (mini-workshop)**
  - Date: February 23-24, 2000
  - Location: Rutgers University
  - Organizers: Ramamohan Paturi, U.C. San Diego and Richard Beigel, University of Illinois at Chicago
- **Cryptography and Intractability**
  - Dates March 20-22, 2000
  - Location: Rutgers University
  - Organizers: Moni Naor, Weizmann Institute of Science, Joe Kilian, NEC Research Institute, and Shafi Goldwasser, MIT and Weizmann Institute of Science
  - Comment: Joint with Special Year on Networks
- **Intrinsic Complexity of Computation**
  - Dates: April 10-13, 2000
  - Location: Rutgers University
  - Organizers: Paul Beame, University of Washington, Steven Rudich, Carnegie-Mellon University, and Andrew Yao, Princeton University

Selected Research Accomplishments

Optimal Proof Systems and Sparse Sets Harry Buhrman, Steve Fenner, Lance Fortnow (PM, NEC) and Dieter van Melkebeek (PD) investigated the existence of optimal proof systems for the set of propositional tautologies. They exhibited a relativized world where the class of sparse NP sets has no complete sets. This gives the first relativized world where no optimal proof systems exist. They also examined under what reductions the class of sparse NP sets can have complete sets and showed a close connection between these issues and reductions from sparse to tally sets.

Lower Bounds for the Satisfiability Problem One of the key problems in the study of computational intractability is the Satisfiability Problem (SAT). Dieter van Melkebeek (PD) studied lower bounds for the satisfiability problem both in the uniform and in the nonuniform setting. In the uniform setting, he found that no RAM can decide SAT in time n^a and space $n^{o(1)}$ for any constant a less than the golden ratio. This improves upon recent results by Fortnow and by Lipton and Viglas.

DIMACS special foci cross-fertilize each other. The following three results cross over between the Special Years on Computational Intractability and on Networks.

Randomizing Low-Degree Polynomials and Round-Efficient Secure Computation Yuval Ishai (PD) and Eyal Kushilevitz introduced a notion of *randomizing low-degree polynomials* as a powerful and conceptually simple paradigm for obtaining round-efficient security. "Standard" low-degree multivariate polynomials can be securely evaluated in a small number of rounds by virtually any "off-the-shelf" secure multi-party computation protocol. However, most Boolean functions f on n input bits cannot be evaluated by a polynomial whose degree is smaller than n. Ishai and Kushilevitz got around this barrier by relaxing the requirement of evaluating f into a weaker requirement of *randomizing* f; that is, mapping the inputs for f along with independent random inputs into several outputs, whose *joint distribution* depends only on the value of f. They showed that degree-3 polynomials are sufficient to randomize *any* function f, where the number of outputs is at most quadratic in the branching program size of f, and where both input variables and random variables count towards the degree. As an immediate application, the secure evaluation of any function f reduces to a secure evaluation of degree-3 polynomials; this reduction is applicable in any reasonable setting for secure computation. One corollary is that two (respectively, three) communication rounds are sufficient to compute any n-argument function f with perfect information-theoretic $\lfloor \frac{n-1}{3} \rfloor$-privacy (resp., $\lfloor \frac{n-1}{2} \rfloor$-privacy), probabilistic correctness, and communication complexity which is at most quadratic in the branching program size of f. This improves over a previous constant-round protocol of Bar-Ilan and Beaver, which in a similar setting requires more communication rounds and a higher communication complexity.

Adaptive vs. Non-Adaptive Security In a *non-adaptive* attack against a cryptographic protocol, the adversary must choose in advance the set of parties it corrupts. A stronger type of attack, called an *adaptive* attack, allows the adversary to dynamically choose which parties to corrupt as the protocol proceeds. R. Canetti, I. Damgard, Dziembowski, Y. Ishai (PD), and T. Malkin clarified the relation between adaptive and non-adaptive security. They showed that under some previously studied settings for secure computation, such as perfect security against a passive adversary or a restricted type of an active adversary, any protocol which is non-adaptively secure must also be adaptively secure. On the other hand, they provided provable separations between non-adaptive and adaptive security in other settings.

Correlated Pseudorandomness and Compressing Cryptographic Resources A private-key cryptosystem may be viewed as a means by which a trusted dealer privately conveys a large, shared pseudorandom object to a pair of players, using little communication. Alternatively, the messages distributed by the dealer may be viewed as a secure *compression* of a pair of large identical random pads (or random functions) into a shorter shared "key" or "seed". Yuval Ishai (PD) extended this compression problem to more general correlation patterns among *several* players. Unlike the simple case of identical pads, where the main security concern is with respect to external eavesdroppers, in the case of general correlations participants also have to be protected from each other. That is, collusions of computationally-bounded players should gain no additional knowledge about the joint pads of the remaining players from the compressed messages they receive, other than what follows from the pads they generate and from knowing the joint distribution of all pads. Ishai developed a modular replication-based approach for securely compressing any *linear* correlation pattern, in which the random pads jointly form a random vector in a linear space. Specifically, the compressed messages received by the players consist of *independent* random seeds, where each seed may be replicated among several players. This approach allows one to use pseudo-random generators or pseudo-random functions in a black-box manner.

SAT, MAX-k-SAT, and Formula Length For many problems, we do not expect to find polynomial time algorithms because the problem is NP- complete. However, this does not imply that exhaustive search is the most efficient algorithm, and indeed, for several such problems (k-SAT, Coloring, Independent Set), more efficient algorithms are known. However, many other natural NP-complete problems do not have such algorithms. Francis Zane (PD) proposed and analyzed algorithms for MAX-k-SAT (the optimization version of k-SAT) and SAT (satisfiability of general CNF formulae). Unlike the case of k-SAT, for these problems, no algorithms are known with running times of the form $2^c n$ with $c < 1$. For both SAT and MAX-k-SAT, several algorithms have been presented with running times of the form c^L for some constant c, where L is the length of the input formula. However, for $L = 10n$, all previous bounds exceed 2^n and become trivial. Zane presented simple algorithms which improve this dependence on formula length, providing exponential speed-up over exhaustive search for all linear sized formulae. The natural goal, posed in previous papers for both problems, would be a $2^c n, c < 1$ algorithm. Zane showed that if such an algorithm exists for the decision problem (SAT), then one exists for the optimization problem (MAX-k-SAT) as well.

3.6 Special Focus on DNA Computing, 1997-99

Overview:

The Special Focus on DNA Computing began in February 1997 and continued through Spring 1999. It built on three 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:

- Richard Lipton, Princeton University, Computer Science
- Laura Landweber, Princeton University, Molecular Biology

Sponsors and Collaborating Institutions:

The National Science Foundation
The New Jersey Commission on Science and Technology
AT&T Labs - Research
Bell Labs
Department of Molecular Biology, Princeton University
Massachusetts Institute of Technology
NEC Research Institute
University of Pennsylvania
Princeton University
Rutgers University
Telcordia Technologies

Theme of Program:

The field of DNA computing is concerned with the possibility of performing computations using biological molecules. It is also concerned with understanding how complex biological molecules process information in an attempt to gain insight into new models of computation. Cells and nature compute by reading and "rewriting" DNA by processes that modify sequence at the DNA or RNA level. DNA computing is interested in applying computer science methods and models to understand such biological phenomena and gain insight into early molecular evolution and the original of biological information procesing.

DNA computing presents great challenges both to computer science and molecular biology in such areas as using parallelism, designing new algorithms, and dealing with errors in molecular processes. The Special Focus is concerned with all areas that relate to computing with DNA including (but not limited to): algorithms and applications, analysis of laboratory techniques, error correction, sequence design and simulation software, useful DNA chemistry and motifs, computation by molecules other than DNA, computational processes in vivo, and fundamental limits to computation. The Focus also has attempted 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, and the special focus has sought to enable these collaborations.

Workshops, Miniworkshops and Special Events, February 1999 - January 2000

- **Fifth International Meeting on DNA Based Computers**
  - Dates: June 14 - 15, 1999
  - Location: Massachusetts Institute of Technology, Cambridge, Massachusetts
  - Attendance: 72
  - Organizers: David Gifford, MIT and Erik Winfree, Princeton University

The Fifth International Meeting on DNA Based Computers was held at MIT June 14-15, 1999, four and a half years after Leonard Adleman's landmark paper, "Molecular Computation of Solutions to Combinatorial Problems," appeared in *Science*. That paper, remarkable for both its theoretical insights and the accompanying experimental demonstration, spawned a series of meetings that brought together computer scientists, mathematicians, computational biologists, molecular biologists, chemists, and physicists to discuss the prospects of performing computations using biological molecules. The fact is that, despite all this discussion, no one is really sure just what the prospects are. Theoretical obstacles were quick to fall; Lipton, Boneh, and colleagues soon presented molecular algorithms for solving formula and circuit satisfiability problems, and several other groups showed that universal Turing machines can be simulated in parallel, with a single DNA molecule encoding the tape and head state of each machine. However, despite some very real experimental advances, laboratory demonstrations of DNA computing are still firmly in "Toyland." The question on everyone's mind during the meeting was, "Can we get out of Toyland?" The meeting was organized into four half-day sessions: Satisfying Experiments, Design and Simulation, Self-Assembly and Nanomachines, and Theoretical Perspectives. After each session, a panel discussion was held to bring into the open questions about where the research is going and what needs to be done next. The first session opened with three back-to-back talks each describing the successful solution of a small Formula Satisfiability problem using a different technique. The paper by Dirk Faulhammer, Anthony Cukras, Richard Lipton, and Laura Landweber described their work using RNA. The Princeton group's work is notable not only because it is the largest SAT problem yet attempted, with 9 input variables, but perhaps more importantly because they have made a considerable effort to characterize the error rates and error mechanisms in individual reactions. Hiroshi Yoshida and Akira Suyama described their work developing methods for solving 3-SAT by breadth-first search. The talk by Liman Wang described the Wisconsin group's recent success doing DNA computing on surfaces. The discussion panel brought up many tough questions about the feasibility of solving significant SAT problems with DNA. A critical question, if SAT is to remain a major target for DNA based computers, is whether there are any problems in the size range DNA based computers might be able to solve that would pose any real difficulty for electronic computers. It was pointed out that modern heuristic algorithms such as Walk-SAT can routinely solve "typical" thousand-variable 3-SAT problems as well as the hardest randomly-generated 250-variable problems, and there are other 3-SAT algorithms for which $O(1.61^n)$ bounds have been proved. But are there any hard 100 variable problems, of the scale that DNA computers can potentially handle? Going one step beyond Yoshida and Suyama on the road from rigorous algorithms to heuristics, Junghuei Chen, Eugene Antipov, Bertrand Lemieux, Walter Cedeño, and David Wood described preliminary results applying the logic of evolutionary computation to DNA-based computation. As hinted above, understanding and controlling error rates are central to solving large problems with DNA based computers. If at each step, a strand has even a 10% chance of being misprocessed or accidentally destroyed, after just 100 steps only 1 in

40,000 strands will have been processed correctly. Addressing this issue, Kevin Chen and Erik Winfree analyzed error-correcting algorithms for both separation errors and for strand loss errors, presuming that sequence independence of error rates can be achieved by biochemical means. However, errors are not the only potential bottleneck to scaling up DNA based computations to significant problems. The limiting factors could equally well be time if many steps are required, or strand length, or test tube complexity. It behooves us to figure out which are the critical bottlenecks. New implementations that allow more computation to occur in a single test tube, such as the use of liposomes to compartmentalize reactions as proposed by Brian Bloom and Carter Bancroft or the self-assembly approaches mentioned below, may make applications to larger problem instances feasible. Overall, things look pretty good. However, we are still very far from doing solid engineering, where we build systems with reliable performance by using well-characterized components; we do not yet have the intellectual framework for characterizing and predicting the behavior of molecular systems to the required level of accuracy. To make matters more complicated, the set of available molecular tools is ever increasing. In a captivating invited talk, Eric Kool described three molecular technologies he has been developing. Closest to the atomic scale is a new pair of artificial DNA "bases" that can be incorporated by DNA polymerases. The second session on Design and Simulation focused on how to design DNA computers within the limits set by physics. The paper by Allen Mills, Bernard Yurke, and Phil Platzman described how parallel DNA hybridization reactions could implement vector algebra operations, including neural networks. An unsolved problem is how to find a set of 10^7 oligonucleotides with no or very little cross-hybridization. Amit Marathe and Anne Condon addressed this problem by defining several criteria related to cross-hybridization, and finding bounds on the size of sets of strings that satisfy the criteria. Max Garzon, Russell Deaton, John Rose, and Donald Franceschetti discussed a different set of criteria for the same problem, which have been implemented in the software package `Edna`. Masahito Yamamoto et al. had their own criteria for designing non-interacting DNA sequences; their work is notable for having an experimental investigation of the sequences they designed. Alexander Hartemink, Tarjei Mikkelsen, and David Gifford described modular software for simulating biological reactions; submodules include prediction of hybridization and the action of enzymes such as polymerase and ligase. Panel discussion identified several challenges for simulation and design software. The first is how to assess validity and scope – that is, how to compare the simulation results to real experiments and understand the limits for accurate prediction of new experiments. There was discussion of what the fundamental data structures should be for such as simulation; should modelling be done at an abstract logical level, at a quantitative thermodynamic level, at the atomic or quantum level, or do all these levels of abstractions need to be combined to accurately model molecular systems? As a practical matter, there was general agreement that algorithms and code should be made available on the Web, to promote communication between labs; it is a bit frustrating that every lab has their own DNA design criteria, and no one can convincingly argue that one is unambiguously better than another. The third session on Self-Assembly and Nanomachines explored some alternative approaches to and alternative uses for molecular computation. Thomas LaBean, Erik Winfree, and John Reif reported experimental progress toward implementing the String Tile model, which holds promise using a single self-assembly reaction to construct a molecular lookup table for addition. Michail Lagoudakis and Thomas LaBean gave a theoretical example of performing more complex computations by 2D self-assembly: they showed how to construct a set of tiles that solve the satisfiability problem in a single self-assembly reaction. Takashi Yokomori proposed a new model of computation that involves a particularly simple experimental reaction: the self-assembly of linear chains of pre-formed components, followed by melting and re-hybridization into a low-energy structure.

Remarkably, this model is Turing universal. Andrew Turberfield, Bernard Yurke, and Allen Mills described a series of fiendishly clever, yet remarkably simple, experiments where DNA oligonucleotides are used as a catalysts to control the hybridization of other DNA molecules. A paper by Michael Robertson, Jay Hesselberth, Colin Cox, and Andrew Ellington described the group's work using *in vitro* evolution and rational modification to discover ribozymes that can ligate DNA under control of a complex switch that can depend on, so far, up to three inputs, such as a DNA oligonucleotide, adenosine, and theophyline. It is interesting to speculate on where this work is going. DNA and RNA can perform logical computations during self-assembly; it can catalyze and perform mechanical work on other DNA molecules; it can undergo conformational switching in response to effector molecules; it can catalyze chemical reactions. Are these not sufficient building blocks for constructing arbitrarily sophisticated biochemical systems and molecular devices? Might we be creating a "DNA world" as a stage in molecular engineering, regardless of whether such a stage existed in the evolution of life on Earth? These were all topics of the panel discussion. What theoretical constructs will be necessary for molecular computation and complex biochemical systems? What theoretical goals can help us guide experimental investigations into promising areas? This was the topic of the fourth and final session, Theoretical Perspectives. The paper by Andrzej Ehrenfeucht, Hendrik Jan Hoogeboom, Grzegorz Rozenberg, and Nikè van Vugt gave an elegant mathematical analysis of sets generated by simultaneous enforcing and forbidding conditions. Lila Kari and Laura Landweber asked whether concepts from computer science can help us understand biological phenomena, such as the mysterious gene unscrambling found in single-cell ciliates. Gheorghe Păun and Takashi Yokomori presented new results for a family of models inspired after the subcellular compartments found in Eukaryotic cells. Going ever smaller, Ehud Shapiro presented his model of computation in which a hypothetical holoenzyme – like the ribosome but tailored for computation – acts on a DNA-like polymer to simulate a Turing Machine, possibly producing molecules as output. The paper by Ashish Gehani, Thomas Labean, and John Reif analyzed a new and unexpected direction for DNA-based cryptography: rather than try to break electronic cryptosystems, use DNA to encrypt and transmit secret messages. During the panel discussion it was noted that in fields such as computer science and quantum physics, theory precedes experiment, which only goes so far as to find a way to obtain what the theory predicts; whereas in fields such as biology and solid-state physics, experiment precedes theory, which only manages to put the pieces together and explain them in a more understandable form. So far, DNA based computing has been driven by theory. But it is quite possible, as unexpected molecular computing components are discovered, as continually occurs in Ellington's lab, that the reverse trend will become dominant.

Selected Research Results

Mn-dependent Ribozymes Sergei Kazakov (V) and Laura Landweber (PU) studied very small ribozymes (catalytic RNA) which require the presence of D1 10 mM Mn2+ for optimal activity. These RNA's are capable of cleaving other RNA molecules, which is important for operating on RNA strands in molecular computing. In particular, they catalyze cleavage of the phosphodiester bond between G and A in the RNA sequence GAAA with formation of 5'-OH and 2',3'-cyclic phosphate termini. Because these and similar small catalytic RNA motifs can plausibly arise under fairly simple prebiotic conditions, they are interesting from a molecular evolution point of view. In this pilot project, Kazakov and Landweber sought to avoid Mn-induced aggregation of RNA by searching for conditions permitting the efficient cleavage of GAAA with low concentrations of Mn2+ ions. By manipulating variables such as ionic strength, type and concentration of monovalent salts and buffers, temperature, presence of organic and inorganic polycations, and RNA concentration, they achieved highly efficient GAAA cleavage at Mn2+ concentrations as low as 0.3 mM. This result is important because it immediately opens the way for structural studies of Mn-dependent ribozymes and their practical applications.

Counting DNA: Estimating the Complexity of a Test Tube of DNA Dirk Faulhammer (PU), Richard Lipton (PM, PU), and Laura Landweber (PU) considered the problem of estimation of the "complexity" of a test tube of DNA. The complexity is the number of different kinds of strands of DNA in the test tube. It is quite easy to estimate the total number of strands in the test tube, especially if the strands are all the same length. Estimation of the complexity is much less clear. Faulhammer, Lipton, and Landweber described a simple kind of DNA computation that can estimate the complexity.

Chess Games as a Model for RNA Based Computation Given an $n \times n$ chessboard, we are interested in determining the positions knights (or more generally any chess pieces) can occupy such that no knight can attack any other knight on the board. Though not NP-complete for a defined board, this problem is from the class of NP-complete SAT problems. Anthony Cukras, Dirk Faulhammer (PU), Richard Lipton (PM, PU), and Laura Landweber (PU) developed a theory of RNA computing and a method for solving the knight problem. Using only biological molecules and enzymes as tools, they developed and successfully executed an algorithm for solving this variant of the knight problem on a 3×3 chessboard using a 10-bit combinatorial pool and a protocol involving sequential R NASE H 'destroy' operations. The results of preliminary experiments revealed that the protocol recovers far more correct solutions than expected at random, and further work led to the successful completion of the largest molecular computation in any lab to date. (This work will soon be published in the Proceedings of the National Academy of Sciences.)

3.7 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. While there are currently no plans for workshops under this theme, there remains great interest in it among DIMACS members, visitors, postdocs, and students, as is reflected in the research described in this section. This special focus has been 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 and Collaborating Institutions:

The National Science Foundation
The National Security Agency
AT&T Labs - Research
Bell Labs
Georgia Tech
The Institute for Advanced Study
NEC Research Institute
National Institute for Standards and Technology (NIST)
Princeton University
Rutgers University
Telcordia Technologies

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 has been to foster interaction between the dm/tcs community and the probability and statistics community. This was already happening to some extent before 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 exception than rule. So the program has been intended to give a substantial push to a broader and more rapid development of these connections. The program also reaches out from the mathematical sciences to the physical sciences, especially statistical mechanics, and has been opening up and/or enhancing lines of communication and collaboration between mathematical and physical scientists.

Workshops Feb. 1999 - Jan. 2000:

- **Workshop: Combinatorial Methods for Statistical Physics Models**
  - Dates: April 29 - May 1, 1999

– Location: Georgia Institute of Technology
– Organizers: Dana Randall and Prasad Tetali, Georgia Tech
– Co-Sponsors: Georgia Institute of Technology, The Southern Applied Analysis Center (SAAC), The Algorithms, Combinatorics and Optimization Program (ACO)

This workshop focused on recent developments at the interface among combinatorics, statistical physics and theoretical computer science. Topics included Gibbs measures and phase transitions in various models (such as the Potts model, hardcore lattice gases and dimer systems), percolation theory, and mixing rates of finite Markov chains. It was a follow-up to the earlier activities of the Special Focus on Discrete Probability, during which we explored the interface among discrete mathematics, discrete probability, and statistical mechanics. The aim of the workshop was to address some of the fascinating questions in combinatorics and discrete probability theory that have arisen with somewhat independent motivations (from the above topics) and which seem to require techniques from several disciplines for solutions. There were introductory lectures intended for a wide audience as well as more focused talks highlighting recent research trends; beginners to the topic with a background in basic combinatorics and probability were encouraged to attend. Peter Winkler opened the workshop by speaking about percolation and collision. Jennifer Chayes followed with a discussion of the scaling window for 2SAT. Roman Kotecky presented an overview of the $q = 3$ Potts model, Sorin Istrail talked about the impossibility of exact solubility of the general 3-dimensional Ising model, David Wilson discussed scaling limits for minimal and random spanning trees in two dimensions, and Umesh Vazirani discussed group theory, Fourier transforms, and quantum computation. This was followed by a presentation on independence properties and combinatorial uss of monomer-dimer systems by Jeff Kahn, one on mixing properties of the monomer-dimer model, with applications to random sampling, by Rob van den Berg, and one on slow mixing for the independent set model and critical Swendsen-Wang by Christian Borgs. Mark Huber spoke about sampling independent sets, Gregory Sorkin about the random assignment problem, Rick Kenyon about long-range properties of spanning trees in Z^2, and Jeff Steif about amenability and nonamenability of graphs and connections with Ising. Finally, Claire Kenyon discussed mixing rates of block dynamics in one dimension, Eric Vigoda talked about improved bounds for sampling colorings, Stefan Boettcher described the virtues of large fluctuations in extremal optimization, and Kari Eloranta discussed bounded ice on planar lattices.

Selected Research Results

Testability of Graph Properties We have obtained several results in the field of *Combinatorial Property Testing*. Property testing deals with the question whether a given input x satisfies a prescribed property P or is "far" from any input satisfying it. A challenge here is to develop efficient randomized algorithms which will be able to answer the above question extremely quickly and yet reliably. Different choices of the property P lead to very diverse algorithmic problems, whose solution is based frequently on deep combinatorial results and techniques. Noga Alon (V), Eldar Fischer, Michael Krivelevich (PD) and Mario Szegedy (PM, ATTL) studied testability of graph properties. Let P be a property of graphs. An ϵ-test for P is a randomized algorithm which, given the ability to make queries whether a desired pair of vertices of an input graph G with n vertices are adjacent or not, distinguishes with high probability between the case of G satisfying P and the case that it has to be modified by adding and removing more than ϵn^2 edges to make it satisfy P. The property P is called *testable* if for every ϵ there exists an ϵ-test for P whose query complexity is independent of the size of the input graph. Alon, et al. studied testability of graph properties defined by first order statements, and showed that properties whose logical description does not contain a quantifier alternation of type $\forall\exists$ are always testable, while some properties containing this alternation are not. The proof is based on a quite involved application of the Szemerédi Regularity Lemma.

Testability of Formal Languages Noga Alon (V), Michael Krivelevich (PD), Ilan Newman and Mario Szegedy (PM, ATTL) studied testability of formal languages. Their main result is as follows. For a regular language $L \in \{0,1\}^*$ and an integer n there exists a randomized algorithm which always accepts a word w of length n if $w \in L$, and rejects it with high probability if w has to be modified in at least ϵn positions to create a word in L. The algorithm queries $\tilde{O}(1/\epsilon)$ bits of w. This result shows that all regular languages are quickly testable. They also studied testability of more complex languages and showed, in particular, that the query complexity of testing context-free languages cannot be bounded by any function of ϵ.

Monotonicity Testing Consider a problem of *testing whether a given Boolean function is monotone*. A function $f : \{0,1\}^n \to \{0,1\}$ is said to be *monotone* if $f(x) \leq f(y)$ for every $x \preceq y$, where $\preceq$ denotes the natural partial order among strings: $x \preceq y$ iff $x_i \leq y_i$ for all $1 \leq i \leq n$. The testing algorithm can request the value of the function on arguments of its choice, and is required to distinguish monotone functions from functions which are far from being monotone. More precisely, the testing algorithm is given a *distance* parameter $\epsilon > 0$, and oracle access to an unknown function f mapping $\{0,1\}^n$ to $\{0,1\}$. If f is monotone then the algorithm should accept it with probability at least $2/3$, and if f is at distance greater than ϵ from any monotone function then the algorithm should reject it with probability at least $2/3$. Distance between functions is measured in terms of the fraction of the domain on which the functions differ. The complexity measures that are usually focused on are the *query complexity* and the *running time* of the testing algorithm. Monotonicity testing is an example of combinatorial property testing. Improving on an earlier algorithm, O. Goldreich, S. Goldwasser, E. Lehman, D. Ron, and A. Samorodnitsky (PD) showed that there is an algorithm for Boolean monotonicity testing with query complexity and running time of order $O(n/\epsilon)$. A key step in the improved analysis is application of *shifting* of Boolean functions. This technique was extended further by Goldreich, et al. and then by Goldreich, E. Dodis, Lehman, D. Ron, S. Raskhodnikova, and Samorodnitsky (PD) to obtain monotonicity testers for wider classes of functions.

Combinatorial Aspects of Testing k-colorability of Graphs. Graph coloring is an old subject with many critical modern applications. Noga Alon (V) and Michael Krivelevich (PD) considered-combinatorial aspects of testing k-colorability of graphs and proved the following result. Let G be a graph on n vertices and suppose that at least ϵn^2 edges have to be deleted from it to make it k-colorable. Then most induced subgraphs of G on $O(k \ln k/\epsilon^2)$ vertices are not k-colorable. If G is as above for $k = 2$, then most induced graphs on $\tilde{O}(1/\epsilon)$ vertices are non-bipartite, and this is tight up to the poly-logarithmic factor. These combinatorial results lead immediately to algorithms for testing k-colorability whose query complexity is $O(1/\epsilon^4)$ for every fixed $k \geq 3$ and $\tilde{O}(1/\epsilon^2)$ for the case $k = 2$, thus improving results of Goldreich, Goldwasser and Ron.

Choosability Properties of Random Hypergraphs Michael Krivelevich (PD) and Van Vu studied choosability properties of random hypergraphs. Choosability problems arise when we specify in advance the set of allowable colors for an edge. A hypergraph $H = (V, E)$ is called k-choosable, if for every family of color lists $\mathcal{S} = \{S(v) \subset Z : v \in V\}$ satisfying $|S(v)| = k$ for all $v \in V$, there exists a choice function $f : V \to Z$ such that $f(v) \in S(v)$ for all $v \in V$, and no edge of H is monochromatic under f. The choice number of H, denoted by $ch(H)$, is the minimal k for which H is k-choosable. A random hypergraph $H(k, n, p)$ is a k-uniform hypergraph on the set

V of n labeled vertices, where each k-tuple of V is chosen to be an edge of H independently and with probability $p = p(n)$. The main result of Krivelevich and Vu states that for every $k \geq 2$ and for all values of p down to $p(n) = O(n^{-k+1})$ the ratio between the choice number and the chromatic number of $H(k, n, p)$ does not exceed $k^{1/(k-1)}$ asymptotically. Moreover, for large values of p, namely, when $p \geq n^{-(k-1)^2/(2k)+\epsilon}$ for an arbitrary positive constant ϵ, the choice number and the chromatic number of $H(k, n, p)$ have almost surely the same asymptotic value.

Packing Random Rectangles A random rectangle is the product of two independent random intervals, each being the interval between two random points drawn independently and uniformly from [0,1]. Ed Coffman (PM, BL), George S. Lueker (V), Joel Spencer (V), and Peter M. Winkler (PM, BL) proved that the number of items in a maximum cardinality disjoint subset of n random rectangles has the tight asymptotic bound $\Theta(n^{1/2})$. Although tight bounds for the problem generalized to $d > 2$ dimensions remain an open problem, Coffman, et al. showed that $\Omega(n^{1/2})$ and $O((nlog^{d-1}n)^{1/2})$ are asymptotic lower and upper bounds. In addition, they proved that $\Theta(n^{d/(d+1)})$ is a tight asymptotic bound for the case of random cubes.

Large Deviations for Ideal Quantum Systems Joel Lebowitz (PM, RU), Marco Lenci, and Herbert Spohn (V) considered a general d-dimensional quantum system of non-interacting particles in a very large (formally infinite) container. They proved that, in equilibrium, the fluctuations in the density of particles in a subdomain Ł of the container are described by a large deviation function related to the pressure of the system. That is, untypical densities occur with a probability exponentially small in the volume of Ł, with the coefficient in the exponent given by the appropriate thermodynamic potential. Furthermore, small fluctuations satisfy the central limit theorem.

An Entropy Approach to the Hard-core Model on Bipartite Graphs Jeff Kahn (PM, RU) introduced entropy as a tool for tackling the question of critical activity for the hard-core model on Z^d — a (still open) problem which was regularly under discussion during this special focus — and settled a related question of J. van den Berg (V). Specifically, Kahn used entropy ideas to study hard-core distributions on the independent sets of a finite, regular bipartite graph, specifically distributions according to which each independent set I is chosen with probability proportional to $\lambda^{|I|}$ for some fixed $\lambda > 0$. Among the results he obtained were rather precise bounds on occupation probabilities; a "phase transition" statement for Hamming cubes; and an exact upper bound on the number of independent sets in an n-regular bipartite graph on a given number of vertices.

Entropy, Independent sets and Antichains "Dedekind's Problem" of 1897 asks for the number $\psi(m)$ of elements in the free distributive lattice on m generators, or equivalently, of antichains in the Boolean algebra $\mathcal{B}_m$. In a 1969 breakthrough, Kleitman used an ingenious elaboration of the basic approach of Hansel to pinpoint the asymptotics of the logarithm of $\psi(m)$. Kleitman's argument was developed considerably further by Kleitman and Markowsky in 1975. For n-regular, N-vertex bipartite graphs with bipartition $A \cup B$, Jeff Kahn (PM, RU) gave a precise bound for the sum over independent sets I of the quantity $\mu^{|I \cap A|}\lambda^{|I \cap B|}$. (In other language, this is bounding the *partition function* for certain instances of the *hard-core model.*) He extended the result to graded partially ordered sets, which in particular provided a simple proof of the well-known bound for Dedekind's Problem given by Kleitman and Markowsky.

Proof of a Conjecture in Combinatorial Probability, with Consequences in Computational Complexity Jeff Kahn (PM, RU), Michael Saks (PM, RU) and Cliff Smyth (GS, RU) proved a well known conjecture of Steven Rudich in combinatorial probability. They formulated and proved a dual form of the well known Reimer - van den Berg - Kesten inequality, and used it to deduce

Rudich's conjecture. This conjecture is known to have some interesting consequences in computational complexity. Let X be a set of Boolean variables. A *term* is a product of literals (where a *literal* is a variable or its complement). Let T be a set of terms, and consider the graph with vertex set T where two terms are neighbors if they share at least one variable. For a term t, let $N(t)$ denote the neighborhood of t in T. Consider the probability space in which we select an assignment of the variables at random. For $S \subseteq T$, let p_S denote the probability that at least one term in S is satisfied and let u_S denote the probability, that *exactly* one term in S is satisfied. Rudich's conjecture says that there is an $\epsilon > 0$ such that if $u_T > 1 - \epsilon$ then there must be a term such that $p_{N(t)} > \epsilon$. The consequence (which is due to Rudich), concerns limitations on the construction of cryptographic primitives. The theory of cryptography involves a number of "primitives" which are used as building blocks to construct protocols. These include: one-way functions, one-way permutations, trap-door functions, pseudo-random generators. Considerable research has been done to determine which primitives (or sets of primitives) suffice to construct all of the others. Of particular interest are "black box" constructions of one primitive from another, which, roughly speaking, are constructions which build the target primitive from a given primitive making use only of the properties that define the given primitive, and not the actual realization of the given primitive. Rudich showed that his conjecture implies that if there exists a black box construction of one-way permutation from one-way function then $P \neq NP$. This means that finding such a construction is as hard as separating P and NP, and thus it is unlikely that such a construction is possible by any current techniques.

3.8 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 continued through 1999. The workshop activity has wound down, but there is still a great deal of interest among DIMACS members, visitors, postdocs, and graduate students in the topics of the special year. Massive data sets remain a major area of interest at DiMACS and many massive data sets problems fall at the interface of this special year and the Special Year on Next Generation Networks Technologies and Applications that we are launching early in 2000. Interest in this area has also led us to begin serious planning for a follow-up Special Focus on Data Analysis and Data Mining starting in Fall 2001.

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
Institute for Defense Analyses/Center for Communications Research (IDA/CCR)
NEC Research Institute
Princeton University
Rutgers University
Telcordia Technologies

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, credit card transactions, telephone calling data, 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, February 1999 - January 2000

- **Distributed Data and Structures**
  - Dates: May 10 - 11, 1999
  - Location: Princeton University
  - Attendance: 23
  - Organizers: Yuri Breitbart, AT&T Labs, Sajal Das, University of North Texas, Nicola Santoro, Carleton University, Peter Widmayer, ETH Zurich

The *1999 Workshop on Distributed Data and Structures, (WDAS'99)* continued the series successfully started with WDAS'98, which was held in Orlando as a satellite workshop of the *12th International Parallel Processing Symposium (IPPS'98).* The aims and the scope of WDAS'99 were characterized as follows: Currently, the research on distributed data structures is carried out in many fields, from parallel systems to distributed computing, from AI to GIS. Their systematic design and analysis has just started: in the database literature, dynamic file structures for distributed object management have attracted some attention, and in the algorithms literature, data structures have been studied from a complexity oriented point of view. However, this research is mostly "hidden", relegated to the side of each field, obscured by the weight of the application domain, especially in databases but also in the algorithms community. As databases are growing steadily, applications become more and more demanding, and distributed computer systems are becoming rather easily available, the problem of how to efficiently maintain large datasets gains importance. An important aspect of this problem is the design, implementation, and operation of a data structure in a distributed system. At the same time, in the constantly expanding net-centric universe, an increasing amount of data is available, distributed among sites. The structuring of the data for accessing, manipulation and processing is a crucial task which can ultimately affect the performance, integrity and usefulness of the entire system. These two developments bring the research on distributed data and structures to the forefront. This workshop was intended to bring together application oriented developers and theoretical researchers concerned with the maintenance of distributed data and the organization of the interaction between the computing nodes. Topics of interest included, but were not limited to: design, implementation and operation of distributed data structures in general; measures of efficiency for distributed data structures; complexity analysis (lower and upper bounds) for distributed data structures; particular data maintenance problems (distributed token, mutual exclusion, counter); connection between distributed data maintenance and mobile data maintenance; connection between the interaction structure of the nodes and the data maintenance; experience with distributed data structures on current systems. The program of the workshop included 18 presentations, either invited or contributed in response to the call for papers. As expected, a large majority of the talks addressed issues related to the design, analysis, implementation, application and testing of distributed data structures. Other topics covered in WDAS'99 were algorithms and data allocation problems for parallel disk systems, novel interconnection structures, and basic algorithmic primitives of distributed computing. The first day of the workshop was opened by Maurice Herlihy, who illustrated the design of a *directory*, a distributed data structure used to keep track of a mobile object. Some problems arising in practical scenarios were also at the core of Witold Litwin's talk. In the first of his two talks, Philip Gibbons introduced the concept of a *synopsis* as a small, cacheable "summary" of a large, remote data set. The first group of talks concerning distributed data structures was completed by Adriano di Pasquale, who discussed issues arising in the design and maintenance of logarithmic-cost, balanced, distributed search trees. External searching, sorting and data allocation were the subjects of the next three talks. First, Ekow J. Otoo discussed issues related to the implementation of *B+Tree indices* for storage systems made of multiple disks. Next, Sanguthevar Rajasekaran described the *Parallel Disk System (PDS)* model and gave an overview of a number of deterministic and randomized PDS algorithms for fundamental problems, such as sorting and selection. Finally, Maria Cristina Pinotti discussed an optimal disk allocation strategy for partial match queries on nonuniform cartesian-product files. In his talk, Nicola Santoro amply reported on the large body of past research on distributed sorting and selection, also highlighting several remaining open problems. Next, Paola Flocchini envisaged a distributed scenario where objects are dynamically added or removed, and there is the need for a consistent and efficient naming scheme so that the objects can cross-reference each other. The first day of WDAS'99 was closed by Nir Shavitt, who described two efficient and scalable implementations of priority queues for bounded-range keys, based on funnel-like structures. The second day was opened by Phil Gibbons' second talk. He presented "black-box" testing strategies to check the correctness of distributed data structures such as (priority) queues, stacks of counters. Linda Pagli introduced the *dense tree*, a recursive structure obtained as a natural extension of a tree. Data structures and their use in the realm of parallel computation was the focus of the next two talks. First, Geppino Pucci presented near optimal deterministic algorithms for the *branch-and-bound* problem on a distributed-memory, p-processor machine. Next, Andrea Pietracaprina explored the deterministic complexity of parallel *multisearch* on a distributed-memory, coarse-grained parallel machine. Three talks in WDAS'99 concentrated on the concept of *self-stabilization*, a feature that makes a distributed system tolerate transient faults by converging to the intended behaviour in finite time. In particular, Ajoy K. Datta concentrated on tree networks and derived optimal, self-stabilizing algorithms for the *Propagation of Information with Feedback (PIF)* problem. Next, Vincent Villain described a tool to achieve optimal, self-stabilizing algorithms in Dijkstra's *state model*. Finally, Franck Petit discussed the time and space optimality of self-stabilizing algorithms for distributed, depth-first token circulation. The proceedings of WDAS'99 will appear in the series "Proceedings in Informatics" of Carleton Scientific Press.

Selected Research Results

Engineering the Compression of Massive Tables: An Experimental Approach In a variety of problems, one has massive tables about which one can learn something by examining a small, fixed size of data. Adam Buchsbaum (PM, ATTL), Don Caldwell (ATTL), Ken Church (ATTL), Glenn Fowler (ATTL), and S. Muthukrishnan (PM, ATTL) showed how one can use such a training set to come up with a better way to organize the table, so that applying gzip (for example) to the newly reorganized table significantly outperforms just applying gzip to the table in the original order (by a factor of 2, in fact).

External Memory Graph Transversal Adam Buchsbaum (PM, ATTL), Michael Goldwasser (PD), Suresh Venkatasubramanian (PM, ATTL), and Jeff Westbrook (PM, ATTL) gave the first non-trivial external memory algorithm for *directed* breadth-first search and an improved external memory algorithm for directed depth-first search. They provided reductions between several common formulations of undirected bfs, showing them equivalent up to an external sort. Using those reductions, they gave the first optimal (constant sort) external memory algorithm for undirected bfs on trees. (This is trivial in core since a tree is a bfs tree, but externally, it's nontrivial to decide the bfs numbering if all one is given is a list of edges, even when told they form a tree.)

There was considerable cross-fertilization between our Massive Data Sets Special Year and our

Large Scale Discrete Optimization Special Year. The following two results are examples of such cross-fertilization.

Generating Transversals Generating transversals to a hypergraph (or the dual of a Boolean function) is a well-known problem, which became very important recently, due to its data-mining applications ("association rules" or more specifically "frequent sets" are transversals to a certain hypergraph defined by the input data set.) This problem is also interesting, since its true complexity (in the incremental sense) is not known. Due to an incrementally quasi-polynomial algorithm (Fredman and Khachiyan, 1996) it is very unlikely to be an NP-hard problem, but as of today we do not know any polynomial algorithm, either – it could even be the core to a separate new class, laying somewhere between P and NP-hard (unless of course these two coincide.) Endre Boros (PM, RU), Vladimir Gurvich (V), Leo Khachiyan (PM, RU), and Kazuhisa Makino (V) showed that several open versions of transversal generation are in fact equivalent with Boolean dualization, and hence treatable with an incrementally quasi-polynomial algorithm. Their results are based on several new inequalities, relating the number of maximal false points to the number of minimal true points of a Boolean function. For instance, we can prove that the number of maximal infeasible binary vectors to a monotone system of linear inequalities is polynomially (almost linearly) bounded in the number of minimal feasible vectors. Based on this result, they can generate all minimal feasible solutions to a monotone system of inequalities in incremental quasi-polynomial time.

Attribute Selection in Learning and Data-Mining Problems In the logical analysis of data, one of the first and most important steps is the selection of a small subset of binary attributes, to a given input data set, which will be used in later steps of the learning process. Such a reduction step, if done carefully, is useful to eliminate noise, and to reduce dimensions, making possible the application of more demanding computational methods. Recently similar considerations became popular in the learning community – so called "attribute efficient learning." Endre Boros (PM, RU) and Toshide Ibaraki (V) considered such an attribute selection problem for the case of binary data, and based on various measures of the "separating power" of the individual attributes and their small subsets. They showed that a wide variety of such models are in fact linearly constrained maximization problems with submodular objective functions, and hence efficient approximation algorithms with a guaranteed performance rate can be applied. They considered continuous relaxations of these problems, which turn out to be smooth convex minimizations over a convex set, and hence can also be solved efficiently.

The following two results come from cross-fertilization between our Special Years on Networks and Massive Data Sets.

Maximum Clique Problems in Very Large Graphs The proliferation of massive data sets brings with it a series of special computational challenges. Many of these data sets can be modeled as very large multidigraphs with a special set of edge attributes that represent special characteristics of the application at hand. Understanding the structure of the underlying digraph is essential for storage organization and information retrieval. Such problems arise for example from telephone billing records. One builds a multidigraph with vertices being telephone numbers and edges calls made from one number to another. James Abello (PM, ATTL) sudied the evolution of such a graph over one 20-day period and found that it grew to have 290 million vertices and 4 billion edges. In practice, this multidigraph is treated as if it were undirected and multiple edges are collapsed into a single edge. Even a computer with six gigabytes of main memory cannot hold the full call graphs arising in practice. Cliques in the call graph represent important structural information. Maximum cliques are important in such practical applications as project selection, classification theory, fault tolerance, coding theory, computer vision, economics, information retrieval, signal transmission theory, and molecular biology. Yet, identifying a maximum clique in a graph is computationally difficult even in a graph of modest size. Abello and collaborators Panos Pardalos (V), Mauricio Resende (PM, ATTL), and Sandra Sudarsky (GS, V) developed an approach to finding cliques in very large multidigraphs. They developed several graph decomposition schemes for breaking up the problem into manageable pieces and a semi-external greedy randomized adaptive search procedure for finding approximate solutions to the maximum clique problem and the maximum quasi-clique problem in very large sparse graphs. They experimented with their heuristic on real data, for instance a one-day call graph with 53,767,087 vertices and 170 million edges, finding large cliques of as many as 30 vertices.

Maintaining Hierarchical Graph Views In connection with a problem in visualizing large graphs, Adam Buchsbaum (PM, ATTL) and Jeff Westbrook (PM, ATTL) considered a hierarchy defined on the vertices, and how one would visualize aggregations of the graph by contracting vertices with respect to the hierarchy. For example, in the graph induced by telephone customers (vertices) and calls (edges), one can contract all the vertices in a single area code together. The problem then becomes one of maintaining this induced, contracted graph under expansion (area code into exchanges, for example) and contraction (the inverse) operations. They formalized the problem and provided data structures that allow navigation in optimal time (linear in the size of the change to the induced graph), although they use superlinear space.

A Multiple Prototype Model for Fuzzy Clustering In partitional fuzzy clustering, each cluster is characterized by two items, its centroid and its membership function, that are usually interconnected through distances between centroids and entities (as in fuzzy c-means). Susana Nascimento (GS, V), Boris Mirkin (V) and F. Moura-Pires proposed a different framework for partitional fuzzy clustering which suggests a model of how the data are generated from a cluster structure to be identified. In the model, they assume that the membership of each entity to a cluster expresses a part of the cluster prototype reflected in the entity. Due to many restrictions imposed, the model as is leads to removing of unneeded cluster prototypes and, thus, can serve as an index of the number of clusters present in data. A comparative experimental study of the method fitting the model, its relaxed version and the fuzzy c-means algorithm, suggested that their methods can be considered a model-based parallel to the fuzzy c-means approach. Moreover, their generic version can be viewed as a device for revealing "the natural cluster structure" hidden in data.

Massive Data Set Analysis in Seismic Explorations for Oil and Gas in the Crystalline Basement Interval Ilya Muchnik (V), Vadim Mottl (V) and Vladimir Levyant worked on an optimization-based approach to the analysis of massive ordered data sets. They introduced a new optimization method for computer-aided interpretation of seismic exploratory data from the so-called crystalline basement of the earth's mantle which underlies the relatively thin sedimentary cover which has been, up to now, essentially the only object of seismic explorations. The seismic exploratory data sets, seismic sections and cubes are a class of, respectively, two- and three-dimensional data arrays which are analyzed in the course of gas and oil reserves prospecting with the purpose of studying the structure of the underground rock mass. The seismic data sets consist of synchronous records of reflected seismic signals registered by a large number of geophones (seismic sensors) placed along a straight line or in the nodes of a rectangular lattice on the earth's surface. The vertical time axis forming the resulting two- or three-dimensional picture is identified with depth, so that the peculiarities of the reflected signal under the respective sensors carry information on the local properties of the rock mass at the respective points of the underground medium. The new method of seismic data analysis developed by Muchnik, Mottl, and Levyant is aimed at finding fractured zones of the basement rock mass capable of accumulating oil or gas. The essence of the method consists of numerically evaluating distinctions in the local spatial texture of the seismic picture that are caused by differences in physical properties of fractured and monolith rock.

Fully Consistent Extensions of Partially Defined Boolean Functions with Missing Bits The problem of finding patterns in data mining can be formalized as the problem of finding an extension of a partially defined Boolean function in which the input contains some missing bits. Endre Boros (PM, RU), Toshihide Ibaraki (V) and Kazuhisa Makino (V) considered four different definitions for an extension and showed that, for many general and reasonable families of function classes, three of these extensions are mathematically equivalent. They also demonstrated that such an equivalence does not hold for all classes.

Apparent Motion in Spatial Time-varying Data: A Variational Approach to Pointwise Tracking of Coherent Structures in Computational Fluid Dynamics A massive time-varying data set is an experimentally acquired or mathematically simulated function of spatial coordinates and time which is to be analyzed with the purpose of studying behavior of related distributed dynamical systems. Vadim Mottl (V), Alexander Blinov, Norman Zabusky and Ilya Muchnik (V) addressed problems of computational fluid dynamics that are concerned with the necessity to visualize and study overwhelmingly large time-varying data sets generated by numerical unsteady simulations. Of major interest is time tracking of persistent coherent structures of the fluid flow that arise, move in space for a time, fade, and are replaced by a new pattern building up from their remains. The spatial motion of persistent patterns in a time-varying data set is called apparent motion because it does not coincide, in the general case, with mechanical motion of any physical particles. With the purpose of tracking coherent structures, Mottl, et al. proposed to evaluate the apparent motion in terms of apparent velocity of space points. The apparent velocity is to be estimated in the form of a time-varying vector field as the result of processing the original time-varying data set or immediately in the course of numerically simulating the respective physical reality as the solution of an additional system of partial differential equations. In both versions, Mottl, et al. leaned upon the variational approach to the analysis of massive data sets and estimated the vector field of apparent velocity as a time-varying function that delivers the minimum value to a quadratic functional. In experiments with simulated two-dimensional time-varying spatial dynamics of the fluid's density and vorticity in shock wave interaction with an elliptical bubble, the estimated apparent velocity provided long-term pointwise tracking of persistent coherent structures.

A Geometric Model of Information Retrieval Systems This decade has seen a great deal of progress in the development of information retrieval systems such as Yahoo and Alta Vista. Unfortunately, we still lack a systematic understanding of the behavior of the systems and their relationship with documents. Myuung Ho Kim (V) developed a completely new approach towards the understanding of these information retrieval systems. Recently, it has been observed that retrieval systems in TREC 6 (TREC = Text Retrieval Conference) show some remarkable patterns in retrieving relevant documents. Based on the TREC 6 observations, Kim introduced a geometric linear model of information retrieval systems. He then applied the model to predict the number of relevant documents by the retrieval systems. The model is also scalable to a much larger data set. Although the model was developed based on the TREC 6 routing test data, Kim believes it can be readily applicable to other information retrieval systems.

Structural Properties and a Recognition Algorithm for Boolean Geodetic Functions A Boolean function is called *geodetic* if any pair of true points is connected by a shortest path consisting of true points (a path is a sequence of points of the Boolean hypercube B^n such that the Hamming distance between any two consecutive points of this sequence is 1). One of the main applications of the geodetic functions is in Logical Analysis of Data, since they provide models for the extensions of some data sets. Gabriela Alexe (GS, RU), Sorin Alexe (GS, RU), Peter Hammer (PM, RU), and Alexander Kogan (PM, RU) obtained structural characterizations of geodetic functions in terms of their set of true points and also in terms of their complete DNF. Based on these results, some recognition algorithms for the geodetic, co-geodetic and strongly geodetic functions were designed.

Data Access Hypergraphs and Applications in Database Modeling Gabriela Alexe (GS, RU) and Sorin Alexe (GS, RU) provided a hypergraph-based approach for the modeling of complex data structures and investigated some of its properties and operations. They defined the Data Access Hypergraphs, which are connected simple hypergraphs having the property that any two intersecting edges are comparable relative to set inclusion, and found several structural characterizations. They defined an abstract database model together with structural, transactional and mixed operations. By focusing on an object-oriented implementation model, and by defining data definition, data manipulation, and data query techniques, they found methods that can be efficiently used for the development of software applications.

Admission Control with Instant Notification When admission control is used, an online scheduler chooses whether or not to complete each individual job successfully by its deadline. An important consideration is at what point in time the scheduler determines if a job request will be satisfied, and thus at what point the scheduler is able to provide notification to the job owner as to the fate of the request. In the loosest model, often seen in real-time systems, such a decision can be deferred up until the job's deadline passes. In the strictest model, more suitable for customer-based applications, a scheduler would be required to give notification at the instant that a job request arrives. Michael Goldwasser (PD) and Boris Kerbikov (undergraduate, PU) studied the effect of the notification model on a specific problem, namely the non-preemptive scheduling of a single resource in order to maximize resource utilization. Although the immediate notification model is severely more restrictive, they were able to give alternate algorithms which provide such notification while matching most of the performance guarantees which are possible in the loosest model with no advance notification. In one case, they gave evidence that providing immediate notification may be more difficult, as they improved upon both the lower bound for the immediate notification model and the upper bound for the loosest notification model.

Concept Learning and Feature Selection Based on Square-error Clustering Based on a reinterpretation of the square-error criterion for classical clustering, Boris Mirkin (V) developed a "separate-and-conquer" version of K-Means clustering. The method determines a contribution weight for each variable of every cluster. The weight is used to produce conjunctive concepts that describe clusters and to reduce or transform the variable (feature) space.

Counting Special 0-1 Matrices, with Applications to Information Theory Mordecai Golin (V), Yuanping Zhang (GS, V) and Xuerong Yong (GS, V) examined 0-1 matrices which do not contain certain sub-matrices, for example do not contain any two consecutive horizontal or vertical ones. The problem they examined was how the number of such matrices grows with the size of the matrices. This problem has become of interest to the information theory community during the past few years because the matrices can model various physical constraints in 2-dimensional media and the growth of the number of such matrices can then describe the information carrying capacity of the media. Golin, et al. examined this quantity for various problems; in some they obtained better bounds than previously existed; in others, they derived new structural properties that will hopefully permit the derivation of better bounds in the future.

3.9 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 foci, 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 focus, 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 new collaborations with other units.

Workshops, Miniworkshops and Special Events, Feb. 1999 - Dec. 1999:

- **On-Line Decision Making**
  - Dates: July 12 - 15, 1999
  - Location: Rutgers University
  - Attendance: 89
  - Organizers: Yoav Freund, AT&T Labs - Research, and Rakesh Vohra, Northwestern University
  - Comment: This was both a Special Year Networks workshop and an exploratory workshop, investigating a potential new special focus.

  In recent years there has been increasing interest in the analysis of algorithms for making repeated decisions within an unknown environment. It has been realized that interesting performance guarantees can be obtained without assuming that the environment's behavior is generated by a stochastic process. Work along these lines has been going on in several different fields under a variety of names, for example: information theory: Universal coding of individual sequences, universal prediction, universal portfolios ...; game theory and decision theory: Adaptive game playing, learning in games ...; machine learning: Aggregating strategies, using expert advice, prediction in the worst case ...; statistics: Prediction in the worst case, universal calibration, prophet inequalities ...; computer science: Competitive analysis of Online Algorithms. The goal of the workshop was to bring together researchers from these different areas to identify common questions as well as explore different perspectives on the same questions. The talks were (by design) mostly survey talks of various areas with a minority of talks on more focused or specialized issues. We have not created a printed proceedings for this workshop; instead, a web page which summarizes the meeting was created and is accessible at http://www.research.att.com/~yoav/DIMACS.workshop/index.html. Rather than offer a talk by talk description of what transpired it might be more useful to provide an overview of what seem to be the important issues. The basic problem that everyone was focused on is what we call "the online decision problem." This problem considers a scenario in which an agent performs some kind of repeated interaction with its environment. On each iteration of this interaction the agent chooses one out of many actions whose consequences/costs depend on the unknown environment. The action must be chosen *before* its cost is known but can depend on the outcomes of previous actions. Unlike other frameworks for studying this problem, we do *not* assume that the environment is an oblivious stochastic process. Rather, we allow the environment's behaviour to be arbitrarily complex; in particular, it can depend on the actions of the agent. There is general agreement that 'good' in this context means relative to some benchmark collection of procedures, but differences arise as to how the benchmarks should be chosen. These differences stem, in large part, from the context in which the online decision problem arises. To illustrate, we summarize a debate that took place at the workshop. In several fields it is common to use a benchmark based on "regret". A list of procedures (also called "experts", "models" or "strategies") is given a-priori. One does not know in advance which procedure will generate the lowest time average cost. However, one *can* describe adaptive methods that are guaranteed to do as well as the best procedure in hindsight. That is, one can, *after* using the adaptive method, go back and calculate the total loss (or gain) of each expert (by itself) against the sequence of outcomes one faced in the past. The regret is defined as the difference between the total gain of the adaptive algorithm and the total gain of the best procedure. (It is called regret because it quantifies the feeling of regret one suffer when one considers their total loss had they followed the recommendations of the best procedure). Methods developed in various fields have provable a-priori guarantees on the maximal value of the regret. However, such guarantees have no meaning in some contexts. The assumption implicit in the guarantee is that nature does not respond to the choice of procedure used. Specifically one assumes that had one used a procedure different from the one used initially, the sequence of outcomes stays the same. This does not make sense in a game theory context. Regret-type benchmarks have been popular in both computer science and game theory and a number of tools have been developed to analyze them. Two that were highlighted in the talks were Bregmann distances and Blackwell's approachability theorem. A tantalizing possibility is that Bregmann distances are related to the potentials associated with Blackwell's Approachability theorem. A number of the more focused talks discussed applications of these tools as well as generalizations. When one moves beyond regret-type measures of performance to measures that assume that nature chooses the outcomes as a function of the procedure one uses, guarantees appear impossible to come by. This was a point made in a series of talks on impossibility theorems. More importantly they highlight an important weakness of the Bayesian approach to decision making. Interestingly, this was a topic that that was discussed in a minor way among the participants and continued via email after the workshop.

- **Workshop: DIMACS-Institute for Research in Cognitive Science Workshop on Logic and Cognitive Science**
  - Dates: April 16-18, 1999
  - Location: University of Pennsylvania
  - Attendance: 90
  - Organizers: Moshe Vardi, Rice University and Scott Weinstein, University of Pennsylvania
  - Co-sponsor: Institute for Research in Cognitive Science, University of Pennsylvania

This workshop was a follow-up to our 1995-96 Special Year in Logic and Algorithms. The goal of the workshop was to generate a lively discussion between researchers in finite model theory and descriptive complexity and investigators in various areas of cognitive science. Obvious areas of potential interaction include the study of constraint-satisfaction problems, the analysis and application of modal and temporal logics, and the construction of adequate languages for knowledge representation in areas from computational learning theory to robotics. It was our hope that the workshop would also bring to light further areas of potential interaction between finite-model theory and cognitive science. The program of the workshop consisted of hour-long survey talks as well as shorter 30-minute and 45-minute talks and the schedule allowed ample time for creative interaction. The program included invited presentations by Tom Dean on "Efficiently Identifying and Exploiting Structure in Dynamic Domains", Henry Kautz on "Compute Intensive Methods in AI", Bill Rounds on "Disjunctive Logic Programming and Default Reasoning with Clausal Partial Logic", and Gyorgy Turan on "Learning Theory and Logic." The workshop was preceded by a four-day program of tutorials on finite-model theory, descriptive-complexity theory, and their potential connections with areas of research in cognitive science. The tutorials were presented by Erich Graedel on Finite-Model Theory -Complexity, Phokion Kolaitis on Finite-Model Theory - Expressiveness, Leonid Libkin on Embedded Finite Model Theory, Maarten Marx and Yde Venema on Modal and Temporal Logics: A Finite Model Theory Perspective, Joel Spencer on Finite-Model Theory - Probabilities, and Moshe Vardi on Finite-Model Theory and Constraint-Satisfaction Problems. We are preparing a volume of papers based on the tutorials which were held in conjunction with the Workshop. The tutorials were a great success and we feel this volume will be a valuable addition to the literature on finite model theory and descriptive complexity.

- **Workshop: Kleitman and Combinatorics: A Celebration!**
  - Dates: April 16-18, 1999
  - Location: MIT, Cambridge, MA
  - Attendance: 100
  - Organizers: Bonnie Berger, MIT, Lenore Cowen, Johns Hopkins, Jerry Griggs, South Carolina, Tom Leighton, MIT, Mike Saks, Rutgers, Doug West, Illinois
  - Co-sponsors: Massachusetts Institute of Technology, National Security Agency

This conference in honor of the 65th birthday of Daniel J. Kleitman, a member of the DIMACS External Advisory Board since its inception, was hosted by the MIT Mathematics Department on the MIT campus. Over 100 people registered and attended the conference, coming from all over the United States, and as far away as Russia, India, and Israel. There were a total of twenty six invited talks. There was also an open problems session chaired by Professor Kleitman on Monday evening and an invited poster session on Tuesday evening. One problem posed at the open problem session has already led to a new result by Kleitman and Ron Holzman and others. Speakers at the conference represented an extremely broad variety of topics from combinatorics and included Noga Alon, Arnold Barnett, Bonnie Berger, Dimitri Bertsimas, Seth Chaiken, Lenore Cowen, Mike Fredman, Zoltan Furedi, Wayne Goddard, Michael Goemans, Curtis Greene, Jerry Griggs, Ron Holzman, Jeff Kahn, G.O.H. Katona, Nati Linial, Tom Leighton, Lior Pachter, Bruce Rothschild, Mike Saks, Leonard Schulman, James Shearer, Joel Spencer, Richard Stanley, Tom Trotter, Doug West, Herb Wilf, and Peter Winkler. The conference was supported by the MIT Department of Mathematics, a grant from the National Security Agency to be used exclusively for speaker and graduate student travel expenses, and funds from DIMACS to support student travel expenses. A subset of the organizing committee is now putting together a special issue of the journal *Discrete Mathematics*, which will be instead of a conference proceedings.

- **Workshop: New Market Models**
  - Dates: April 26-27, 1999
  - Location: Rutgers University
  - Attendance: 54

- Organizers: Larry Shepp, Rutgers, Chris Heyde, Columbia

The theme of the workshop was to debate the issue of whether a new model is needed for pricing hedge options in the aftermath of the failure of Long Term Capital Management, Inc. It was felt that the question of whether the Black-Scholes-Merton model for pricing of an option on a particular stock is accurate enough ought to be raised. This model makes the debatable assumption that arbitrage is impossible and that information is shared equally by everyone in the market. A tutorial presentation of the BSM model was presented by Stan Pliska. A panel consisting of Marco Avellaneda, Freddy Delbaen, Benoit Mandelbrot, Stan Pliska, Larry Shepp, with Chris Heyde (moderator), discussed the pros and cons of various models. The remainder of the meeting consisted of individual invited papers. Alexander Adamchuk spoke about "The Geometry of Financial Times", illustrating the multiscale nature of financial markets. Marco Avellaneda spoke about new techniques for calibrating multifactor asset-pricing models using weighted Monte Carlo. "Risk Constrained Dynamic Active Portfolio Management" was the title of the talk by Sid Browne of Firmwide Risk Management at Goldman Sachs & Co. There was a presentation by Jennifer N. Carpenter entitled "The Optimal Dynamic Investment Policy for a Fund Manager with an Incentive Fee". Freddy Delbaen developed some coherent risk measures, discussing underlying conceptual and mathematical problems, while Sergei Esipov of the Zurich Financial Services Group spoke about probabilities and pricing, pointing out that now financial math can focus on modeling the probabilities of different outcomes while pricing remains partially subjective. Benoit B. Mandelbrot spoke about the multifractal time model for financial prices, Thomas Noe and Larry Eisenberg of The Risk Engineering Company presented a paper entitled "Clearing Systems and the Transmission of Systemic Risk", and Larry Shepp presented a new model for stock price fluctuations based on delayed information. The elimination algorithm for the problem of optimal stopping was the topic of the presentation by Isaac M. Sonin. Robert Vanderbei talked about a Martingale system theorem for stock investments and Xiaolu Wang of Advanced Analytics, Inc. spoke about lessons from financial disasters and a dynamic asset pricing model, with reference to recent global financial storms in Asia, Latin America, and Russia. A final summing up session closed the meeting. We think the cause of getting past the standard model was advanced.

- **Workshop: Graph-theoretic Methods in Computer Vision**
  - Dates: May 20-21, 1999
  - Location: Rutgers University
  - Attendance: 39
  - Organizers: Sven Dickinson, Rutgers, Martin Farach-Colton, Rutgers, Lenore Cowen, Johns Hopkins

Graph algorithms are enjoying a renaissance in computer vision and are being applied to problems in image segmentation, motion tracking, stereo correspondence, mobile robot navigation, and object recognition, to name just a few. The goal of our workshop was to bring together two groups of researchers: those in computer vision working with graphs and those graph theoreticians working in graph algorithms. The format of the two-day workshop was a series of invited talks by computer vision researchers who are employing graph theoretic methods. Members of the graph theory community were invited to provide feedback and report on new results. It was our hope that during the course of the workshop, the vision researchers could offer to the graph theoreticians a novel set of problem formulations. In return, the graph theoreticians could provide important guidance in solving these problems. The opening session consisted of two talks. In the first talk, Dan Huttenlocher presented an approach to the problem of recognizing flexible objects (e.g., people or animals) that can be modeled as a collection of rigid subparts. Using a tree-structured object representation, Huttenlocher finds the best configuration of the parts given a particular model and image using maximum a posteriori (MAP) estimation. In the second talk, Ali Shokoufandeh presented the first of three talks that addressed the problem of 2-D silhouette matching based on a shock tree-based representation of a closed 2-D shape. The approach addressed both indexing and matching, unified by a novel eigencharacterization of a shock tree's topology. The Day-1 morning session concluded with two talks. In the first talk, Satish Rao discussed a number of important graph algorithms, e.g., shortest path, and their impact on computer vision. He then followed up with a discussion of his own work in applying graph algorithms to the problem of image segmentation. Kim Boyer concluded the morning session with a talk on inexact matching and the use of graph spectra in computer vision. The afternoon of Day-1 began with a talk by Kaleem Siddiqi on the use of association graphs to match hierarchical structures. This was the workshop's second talk on algorithms for matching shock trees, indicating a growing momentum in this area. Next, Hiroshi Ishikawa explored embedded graph algorithms for computer vision, in which graphs are embedded in Euclidean space and thereby endowed with important geometric interpretation. The final talk of the afternoon was given by Yali Amit, who presented a simple star type graphical model for detecting objects from a particular class in gray-level images. The first day ended with an outstanding panel discussion which addressed such topics as the role of graph algorithms in computer vision, what algorithms the vision community found particularly useful, and what problems were encountered in trying to use these algorithms. The second day began with a talk by Philip Klein, who presented a tree-edit based algorithm for matching shock trees, the third and final talk on the topic of shock graph matching. Next, Pedro Felzenszwalb presented a framework for region segmentation based on pairwise region comparisons. Olga Veksler presented her work on fast energy minimization for computer vision via graph cuts, followed by Edwin Hancock, who discussed a framework for embedding relational constraints in the EM algorithm. In the afternoon, Carly Klivans described an approach to automatic landmark acquisition for mobile robot visual navigation, which she formulated as a graph coloring problem. Anand Rangarajan discussed the use of weighted graphs and associated distance measures for non-rigid matching, and applied his approach to the problem of brain structure registration. The final talk of the workshop was given by Jianbo Shi, who presented his work on the use of normalized cuts in image segmentation, addressing such domains as region and contour segmentation, motion segmentation, and cue combination. No proceedings will be generated from the workshop, although the organizers are putting together a bibliographic website from the workshop which will include not only references to each speaker's work in the area, but other seminal references to graph theoretic methods applied to computer vision, and references to important graph algorithms, texts, etc.

- **Workshop: Joint DIMACS-DIMATIA Workshop on Algebraic Methods and Arithmetic Circuits**
  - Dates: June 2-4, 1999
  - Location: Rutgers University
  - Attendance: 45
  - Organizers: Eric Allender, Rutgers, Jan Krajicek, Czech Academy of Sciences, Pavel Pudlak, Czech Academy of Sciences, Mike Saks, Rutgers

- Co-sponsor: DIMATIA

This was the third workshop (and fourth event) organized jointly by DIMACS and DIMATIA (based at Charles University, Prague, Czech Republic). These workshops are organized by members of the two institutes and they alternate between being held at DIMACS and at DIMATIA. The aim of the workshop was to bring together mathematicians working in computational complexity who use algebraic methods. Algebraic methods are naturally used in the study of algebraic circuits, where the goal is to estimate the number of algebraic operations needed to compute a given algebraic function. In Boolean complexity, where combinatorial methods prevail, there have been a number of interesting results obtained using algebraic techniques. A special case of algebraic circuits are arithmetic circuits that compute over the integers. These are closely related to counting problems studied in circuit- and machine-based complexity. There are more connections than these; we shall mention some below. The talks presented at the workshop covered essentially all the main areas, with only one exception, the so-called BSS model of computational complexity. (Although there had been some initial interest among people working in the BSS model in attending the workshop, schedule conflicts unfortunately arose.) The workshop began with a survey talk by Amin Shokrolahi on the complexity of matrix multiplication. Shokrolahi is a leading expert in algebraic complexity and is a co-author of the most comprehensive monograph on this subject. This was followed by a talk on arithmetic branching programs by Amos Beimel and tensor formulae by Carsten Damm. The major part of the afternoon program was devoted to arithmetic circuits (Eric Allender, Samir Datta and Andris Ambainis), including a survey of the area, motivating interest in the model, and a more technical presentation of new upper and lower bounds in the arithmetic circuit model. The last talk of the first day was about circuits with modular gates given by Vince Grolmusz. The first talk of the second day, given by Meera Sitharam, was on the complexity of propositional proofs with modular connectives. She presented some partial results on this very difficult problem, based on application of the theory of representations of symmetric groups. Pierre McKenzie spoke about the complexity of a new class of counting problems (counting accepting subcircuits) and connections to polynomial replacement systems. In the afternoon there were two talks on another long standing open problem, the rigidity of matrices (Satyarayana Lokam and Pavel Pudlak). On the last day Amir Spilka presented his very interesting joint results with Avi Wigderson giving size bounds for depth 3 arithmetic circuits computing some explicit functions. Alan Woods spoke on a joint result with Petr Savicky on the number of Boolean functions computed by a formula of given size. The program closed with a lecture by Dave Barrington on the role of algebraic techniques clarifying the complexity of the planar circuit value problem. The workshop was well-attended, including many leading researchers in the field. A very positive feature of the workshop was a large proportion of young participants. This shows that the field has a promising future.

- **Workshop: Joint DIMACS-Hungarian Academy of Sciences Workshop: Paul Erdős and his Mathematics**
  - Dates: July 4-11, 1999
  - Location: Mathematical Institute, Hungarian Academy of Sciences, Budapest, Hungary
  - Attendance: 452
  - Organizers: Vera T. Sos, Hungarian Academy of Sciences
  - Co-Sponsor: Hungarian Academy of Sciences

The conference was held from July 4 to July 10, in Budapest, Hungary as a satellite conference to the UNESCO World Conference on Science. The site of the conference was the Main Building of the Hungarian Academy of Sciences. There were 452 participants, 145 Hungarian. From Monday to Friday plenary talks were held in the morning and were aimed at covering all aspects of the ouvre of Paul Erdős. New results and research directions motivated by the work of Paul Erdős were presented by invited speakers in the afternoons and Saturday morning, in parallel sections. In the plenary invited talks, Noga Alon (Israel) spoke about "Colorful Aspects of the Probabilistic Method." Bela Bollobás (USA, England) gave a talk entitled "From Prodigy to Grand Old Man, the Mathematical Journey of Paul Erdős". Paul Borwein (Canada) spoke about Erdős and his work on polyomials, Ron Graham (USA) about his work on Ramsey theory, and Petr Komjáth (Hungary) about his work on infinite graphs. P.D.T.A. Elliott (USA) gave a personal view of Erdős, D.S. Lubinsky (South Africa) recalled his work on interpolation and orthogonal polynomials, Carl Pomerance (USA) summarized his work on arithmetic functions, and Imre Ruzsa (Hungary) described his work on addition of sequences. M. Simonovits (Hungary) spoke about extremal graph theory and Joel Spencer (USA) talked about random graphs, sets, and tournaments. On each day, before the plenary talks, some old friends and colleagues of Paul Erdős spoke about their memories of him. George Szekeres, who worked with Erdős from the 1920's, gave a particularly interesting and moving talk. The more than 300 international participants were from more than 25 countries. Besides the 11 plenary invited speakers there were 130 invited section talks. Every participant had a possibility of showing her/his work either during the poster sessions or in the *Volume of Research Communications*, which was distributed at the beginning of the conference. There was a special session on Friday afternoon about the connection of Paul Erdős and high-school mathematics education. One of his lectures was recalled with the help of videotape. There will be a separate volume for the survey papers in the series *Bolyai Society Mathematical Studies*, and (in discrete mathematics) *Combinatorica* will have special issues for research papers (of invited speakers) dedicated to Paul Erdős. Both of these volumes are in preparation now. Furthermore, the CD-ROM version of collected publications of Paul Erdős, which will include **all** of his works, is in preparation.

- **Workshop: Codes and Association Schemes**
  - Dates: November 9-12, 1999
  - Location: Rutgers University
  - Attendance: 56
  - Organizers: Alexander Barg, Bell Labs, and Simon Litsyn, Tel Aviv University

Mathematical coding theory is related to the works of P. Delsarte in the field that became known as algebraic combinatorics. One might say that algebraic coding theory was defined in his works of 1972-73 and immediately became an established research area. Traditionally, applications of Delsarte's ideas in coding theory include bounds on the size of packings of polynomial metric spaces and structural properties of extremal configurations. Recently it was also observed that given the size of the code, it is possible to employ Delsarte's polynomial method to derive bounds on its distance distribution. In particular, this makes possible new bounds on reliability functions of memoryless channels that improve the results of Shannon and co-authors of 1959 and 1967. A specific goal of the workshop was to inform specialists in algebraic combinatorics of the latest applications in coding theory, and let coding theorists study more properties of Q-polynomial association schemes and spaces useful in the study of codes. The participants formed two groups, one with coding theory background and the other made of algebraic combinatorialists. The new results presented at the meeting fall into the following broad categories: applications of the polynomial (linear programming) method; classification and properties of association schemes; properties of orthogonal polynomials (zeros, asymptotics); properties of quantum codes and quantum information transmission; spherical codes and packings in Grassmannian spaces; properties of codes in discrete metric spaces. The workshop was introduced with a lecture by an expert with insights in both areas, Vladimir Levenshtein. He devoted his lecture to the multifold uses of association schemes in coding theory. E. Bannai drew interesting parallels in the theory of modular forms of half-integer weight inspired by properties of weight enumerators of self-dual codes. B. Martin introduced association schemes related to a special class of finite spaces equipped with a metric that depends on the ordering of coordinates. The second day opened with a talk "Polynomial Method in Coding and Information Theory" by Alexander Barg. He spoke of new uses of Delsarte's polynomial method, specifically on bounding different metric invariants of codes in Q-polynomial spaces. New asymptotic bounds on the distance distribution of codes obtained recently enable one to derive new upper bounds on the reliability functions of a number of classical information transmission channels. Vera Pless presented classification of self-dual codes over the field of 4 elements of length up to 16. Neil Sloane introduced the problem of packing the real Grassmannian manifold and presented a series of optimal packings and C. Bachoc suggested a new weight enumerator of linear codes that carries more combinatorial data about them than the standard weight function. The third day started with a lecture by M. Muzychuk on the theory of circulant association schemes. Subsequently P. Terwilliger gave a new characterization of schemes with the Q-polynomial property. Vladimir Levenshtein, in his second talk, defined a general notion of a t-design in measurable not necessarily homogeneous spaces and illustrated the definition by describing designs in the cube, on the torus and in the sphere in R^n. D. Jaffe discussed the McEliece-Rodemich-Ramsey-Welch upper bound on the size of binary codes with a given distance. He presented numerical evidence that this bound cannot be improved within the frame of the polynomial method. The first part of the last day was devoted to quantum information transmission. It was opened by a lecture of E. Knill offering a general view of classical and quantum codes. A. Ashikhmin introduced the problem of error detection with quantum codes and M. Grassl examined properties of classical codes suitable for constructing quantum codes. In the second half of the day Zoltan Furedi discussed a connection of Turan-type theorems with the covering property of codes and I. Honkala introduced codes that identify vertices of a graph and presented a series of examples. Several publication projects related to the workshop are being discussed. The one that has already taken shape is a collected volume of research articles in the AMS-DIMACS series.

Events Planned for the Rest of 1999-2000

- **Workshop: Unusual Applications of Number Theory**
  - Dates: January 10-14, 2000
  - Location: Rutgers University
  - Organizers: George Andrews, Penn State, David Chudnovsky, Poly Univ., Ron Graham, AT&T Labs, Jeff Lagarias, AT&T Labs, Victor Miller, IDA, Melvyn Nathanson, CUNY, Andrew Odlyzko, AT&T Labs, Carl Pomerance, Bell Labs

Selected Research Results

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.

Properties of k-critical Graphs A graph G is called *k-critical* if its chromatic number equals k, but the chromatic number of every proper subgraph of G is strictly less than k. The study of such graphs and the understanding of their properties is of fundamental importance in the theory and applications of graph coloring. Noga Alon (V), Michael Krivelevich (PD) and Paul Seymour (PM, PU) showed that any k-critical graph on n vertices contains a cycle of length at least $2\sqrt{\log(n-1)/\log(k-2)}$, improving a previous estimate of Kelly and Kelly obtained in 1954.

New Bounds on Crossing Numbers The *crossing number*, $\mathrm{cr}(G)$, of a graph G is the least number of crossing points in any drawing of G in the plane. Denote by $\kappa(n,e)$ the minimum of $\mathrm{cr}(G)$ taken over all graphs with n vertices and at least e edges. Jnos Pach, Joel Spencer (V) and Geza Toth (PD) proved a conjecture of Erdős and Guy by showing that $\kappa(n,e)n^2/e^3$ tends to a positive constant as $n \to \infty$ and $n \ll e \ll n^2$. Similar results were obtained for graph drawings on any other surface of fixed genus. They also proved better bounds for graphs satisfying some monotone properties. In particular, they showed that if G is a graph with n vertices and $e \geq 4n$ edges which does not contain a cycle of length four (resp. six), then its crossing number is at least ce^4/n^3 (resp. ce^5/n^4), where $c > 0$ is a suitable constant. These results cannot be improved, apart from the value of the constant. This settled a question of M. Simonovits.

Approximate Role Assignments and Social Networks Role assignments, introduced by Everett and Borgatti, who called them role colorings, formalize the idea, arising in the theory of social networks, that individuals of the same social role will relate in the same way to individuals playing counterpart roles. If G is a graph, a *k-role assignment* is a surjective function mapping each vertex into a positive integer $1, 2, \ldots, k$, so that if x and y have the same role, then the sets of roles assigned to their neighbors are the same. Aleksandar Pekec (V) and Fred Roberts (PM, RU) showed that all graphs G having no astronomical discrepancies between the minimum and the maximum degree have a k-role assignment. Furthermore, they introduced and studied a natural measure expressing how close an onto map $f : V(G) \to \{1, \ldots, k\}$ is to being a k-role assignment of a graph $G = (V, E)$, and showed that almost all graphs nearly have a k-role assignment.

Ramsey-type results for unions of comparability graphs It is well known that the comparability graph of any partially ordered set of n elements contains either a clique or an independent set of size at least $\sqrt{n}$. Adrian Dumitrescu (GS, RU) and Geza Toth (PD) showed that any graph of n vertices which is the union of two comparability graphs on the same vertex set contains either a clique or an independent set of size at least $n^{\frac{1}{3}}$. On the other hand, there exist such graphs for which the size of any clique or independent set is at most $n^{0.4118}$. Similar results were obtained for graphs which are unions of a fixed number k of comparability graphs. They also showed that the same bounds hold for unions of perfect graphs.

Results on the Bollobás - Eldridge Conjecture for Packing of Graphs into K_n Let G_1 and

G_2 be graphs on n vertices. If there are edge-disjoint copies of G_1 and G_2 in K_n, then we say there is a *packing* of G_1 and G_2. A conjecture of Bollobás and Eldridge asserts that, if $(\Delta(G_1)+1)(\Delta(G_2)+1) \leq n+1$ then there is a packing of G_1 and G_2. Bela Csaba (GS, RU), Ali Shokoufandeh (GS, RU), and Endre Szemeredi (PM, RU) proved this conjecture for $\Delta(G_1) = 3$, for sufficiently large n, and also for $\Delta(G_1) = 4$.

Efficient Dominating Sets in Cayley Graphs An independent set C of vertices in a graph is an efficient dominating set when each vertex not in C is adjacent to exactly one vertex in C. Italo Dejter (V) and Oriol Serra found an uncountable collection of countable families of nested graphs, each of which has efficient dominating sets. These families range from star graphs, for which the efficient domination property was proved by Arumugam and Kala, to pancake graphs. They extended the collection obtained to partitions of the graphs into efficient dominating sets. They showed that uniqueness holds for the given partitions whenever the countable family is not that of the star graphs. In this case, they established all the possible ways in which such a partition occur.

Tree Volume of a Graph The Prüfer code of a tree T with n vertices is a certain word of $n-2$ letters in the alphabet $V(T)$. The notion of the *Laplacian polynomial of a graph* (i.e. the characteristic polynomial of the Laplacian matrix of the graph) turns out to be very useful for various "spanning tree" problems. The *tree volume (or polynomial) of a graph* is, roughly speaking, the "sum" of the Prüfer codes of spanning trees of a graph. In other words, the tree volume is a generating function of spanning trees of the graph classified by their degree functions. The *forest volume of a graph G* is the tree volume of the graph obtained from G by adding a new universal vertex. Alexander Kelmans (V), I. Pak, and A. Postnikov showed that the forest volume of a graph is a natural generalization of its Laplacian polynomial, and many important properties of the Laplacian polynomials can be generalized to the forest volume. For example, the forest volume is also the characteristic polynomial of a matrix, it also has the inclusion-exclusion property, and the reciprocity theorem also holds for the forest volume of a graph.

The following four results reflect themes of an earlier special year, our 1995-96 Special Year on Logic and Algorithms. One concerns the interface between logic and complexity theory, a central theme of the special year. A second is concerned with computer-aided verification, another area of special year emphasis. The last two were not directly related to the special year, but reflect the interest of that special year in natural language processing, cognition, and human-machine interaction. This interest was also demonstrated by the workshop on logic and cognitive science described earlier in this subsection.

Parallel Complexity of Planarity Testing Eric Allender (PM, RU) and Meena Mahajan (V) studied the problem of identifying the parallel complexity of planarity testing. While there are several deterministic parallel algorithms around, the inherent complexity of this problem is not well-understood. The hitherto best known upper bound is the complexity class AC^1 (via a $\log n$ time CRCW PRAM algorithm due to Ramachandran and Reif in 1989). Allender and Mahajan showed that this algorithm can actually be implemented in symmetric logspace SL. They also showed that planarity testing is hard for deterministic logspace L. Given the widespread belief that $L = SL$, and the fact that in the non-uniform setting these classes do coincide ($L/poly = SL/poly$), this work thus essentially settles the question.

Formal Verification: A Model for Hierarchically Structured Designs with Heterogeneous Components Rajeev Alur (PM, BL) and Tom Henzinger (V) developed a formal model for hierarchically structured designs with heterogeneous components. The model, called Reactive Modules, is the foundation of a new model checker, Mocha, which attempts to exploit design modularity in formal verification. This is in contrast to traditional model checkers, which, before verification, transform designs into unstructured state spaces. Recently, Mocha has been used to verify a VGI parallel DSP chip with 96 processors and about 6 million transistors, which is far beyond the reach of direct state exploration. Mocha is still under development.

Multiagent Reasoning Boris Galitsky (V) and Maxim Grudin developed a question-answering system, functioning in the environment of multiple problem domains, aimed at building a natural language understanding system with high query complexity, involving up to 3-4 concepts from distinguishing domains. Each domain is represented by an agent, performing syntactic and semantic analysis and representing the knowledge, specific for a given domain. The agents have the domain-specific knowledge and the common knowledge of what the other agents represent. Receiving a question, the domain agent tries to understand it and redirects this question together with its formal representation to another domain agent if necessary. Each agent extracts the familiar concepts from the question and parameters associated with these concepts, searches for the rest of the parameters this agent is capable of identifying and sends the requests to the other agents. This request contains the original question with its formal representation by the current agent with the values of parameters this agent succeeded in finding. A problem domain manager controls the interaction between the problem domain agents, modeling their mental states based on the concepts of intention and knowledge. A system prototype was built to function in the following hierarchical set of domains: General geography, Middle Eastern geography, Iran-Taliban crisis. In addition, the prototype was used with such domains as chemical process control, air ticket reservation, autism training and the search of narratives. This work showed that multiagent architecture of natural language processing raises the flexibility of domain representation and allows for robust functioning in multi-domain question answering.

Agents with Adjustable Autonomy for Automatic Scheduling in a Competitive Environment Boris Galitsky (V) addressed the problem of constructing an adequate strategy for multiagent scheduling in situations with high uncertainty and inconsistent goals. In his approach, an automatic scheduling agent with adjustable autonomy builds the preliminary schedule and assists the human agents in modifying the schedule towards their specific goals. The degree of autonomy of an automatic agent is controlled via syntax limitations on formulas, which are the basis of reasoning about the mental states of the other (human) agents. Autonomy adjustment therefore could be implemented during the scheduling session. Galitsky used the reasoning tool of metalanguage support (MS) with the embedded metapredicates of mental state and action to provide the automatic agent with the ability to model the human agents with various degrees of reflection. His tool demonstrated the inference flexibility to represent the formal scenarios of multiagent interaction. Such MS capabilities as reasoning about action and change, belief, and temporal reasoning were illustrated by the scheduling system. The automatic, incremental, interactive and multiagent modes of the automatic agent demonstrated the increasing complexity of the scheduler operation.

The following result arose from our international partnership with DIMATIA in Prague, Czech Republic, and represents a unique collaboration among a senior researcher at DIMATIA, a Rutgers graduate student, and a summer undergraduate REU student that took place at DIMATIA.

Universal H-colorable Graphs Without A Given Configuration Paul A. Dreyer, Jr. (GS, RU), Christopher Malon (REU student), and Jaroslav Nesetril (V) solved an open problem regarding the chromatic number of a universal graph. For every pair of finite connected graphs F and H, and every integer k, they constructed a universal graph U with the following properties: (i) there is a homomorphism $\pi : U \to H$, but no homomorphism from F to U; (ii) for every graph G with maximal degree no more than k having a homomorphism $h : G \to H$, but no homomorphism from F to G, there is a homomorphism $\alpha : G \to U$, such that $h = \pi \circ \alpha$.

A New Upper Bound on the Reliability Function of the Gaussian Channel A. Ashikhmin (BL), Alexander Barg (PM, BL), and Simon Litsyn (V) derived a new upper bound on the exponent of error probability of decoding for the best possible codes in the Gaussian channel. This bound is tighter than the known upper bounds (the sphere-packing and minimum-distance bounds proved in Shannon's classical 1959 paper and their low-rate improvement by Kabatiansky and Levenshtein). The proof was accomplished by studying asymptotic properties of codes on the n-dimensional sphere $S^{n-1}(\mathcal{R})$. First, they proved that the distance distribution of codes of large size necessarily contains a large component. A general theorem establishing this estimate was proved simultaneously for codes on the Euclidean sphere and in real and complex projective spaces. Since on the average there are many code vectors in the vicinity of the transmitted vector x, they showed that the probability of confusing x and one of these vectors cannot be too small. This proves a lower bound on the error probability of decoding and a new upper bound.

Evaluation of Special Year on Combinatorial Optimization

We continued our program of evaluating the impact of earlier special years by commissioning an evaluation of DIMACS' fourth special year, the 1992-93 Special Year on Combinatorial Optimization, as seen from the perspective of six years. The report by Bill Cook, Bill Pulleyblank, and Mike Trick, which can be found on the DIMACS website at the URL:

http://dimacs.rutgers.edu/About/index.html#reports

provides an exceptionally strong piece of evidence of the long-term impact of DIMACS' progams. The report talks about the "prescience" of the organizers in choosing workshop topics, and says that "Overall, the Special Year hit two 'home runs' in working with computational combinatorial optimization and approximation algorithms before the large field realized the importance of these two areas." The report singles out the work of Neil Robertson (V), Daniel Sanders (V), Paul Seymour (PM, PU) and Robin Thomas (V), saying their "work on the four-color theorem is extremely important in that it greatly simplified the computer proof and decreased the computational requirements for verifying the proof. In doing so, this work has increased confidence in this fundamental result." The report also quotes Henry Wolkowicz as saying that the year started many people working on semidefinite programming and Panos Pardalos as saying that the special year played a key role in research on the quadratic assignment problem and nonlinear assignment problems. The subsequent work of the special year postdocs on graph entropy, integer programming and Groebner basis theory, and "colorful linear programming," is also discussed. The report concludes: "Overall, we believe that this year was very successful and occurred at an ideal point of time. Issues of approximation and computation in combinatorial optimization were ready to explode, and the DiMACS SY provided the organization for that to happen."

3.10 Multi-Disciplinary, Multi-Institutional Projects

During this period, DIMACS has had two multi-disciplinary, multi-institutional research projects in network modeling and parallel simulations, each funded separately from its main STC award. 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.

Network Modeling and Parallel Simulations of Networks

This program, 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 1993-94 Special Year on Massively Parallel Computing, and is now closely related to our continuing Special Year on Networks, forming a significant applied component of that activity. The first research effort was funded by NSF and the second, followup project by DARPA. We describe these two projects together in this section. At Rutgers, this project involves not only faculty, postdocs, and graduate students from DIMACS, but also a strong group from WINLAB (a leading NSF/industry/university research center on Wireless Networks). In 1998, WINLAB received an NSF Instrumentation Grant for a 14-processor SUN ultrasparc that is heavily used in these simulation projects by both local and remote project members. These projects have also led to a small subcontract from the Ultra Corporation for research on a global network simulator.

Simulations of Integrated Communications Systems

This project was funded by the National Science Foundation for a three year period over 1995 - 1998 with a one year extension through 1999, for a total of $1.5M.

Participating Scientists:
The original PI team was composed of Sandeep Bhatt (PM, BC/TC and RU), Richard Fujimoto (Georgia Tech), Albert Greenberg (PM, ATTL), Jim Kurose (University of Massachusetts), David Nicol (William and Mary / Dartmouth), Andrew Ogielski (PM, BC and RU), Diane Souvaine (PM, RU), Don Towsley (University of Massachusetts), and Paul Wright (PM, ATTL). During the lifetime of the project, Andrew Ogielski left Bellcore and joined Rutgers full-time, taking the leadership of the project. Nicol moved from William and Mary to Dartmouth.

The Centerpiece of the Project:
The centerpiece of the project has been the development of and experimentation with a network-centric simulation meta-language called TeD (Telecommunications Description Language). TeD provides language-independent constructs that may be embedded in a general computer language; in practice we have used C++. The TeD compiler is a source-to-source translator that transforms TeD code into pure C++ that is compiled and executed by the GTW (Georgia Time Warp) parallel simulation engine. An initial version of TeD was developed very early in the project, thus

allowing us to quickly begin the important cycle of model building and tool refinement. The official TeD home page http://www.cc.gatech.edu/computing/pads/teddoc.html contains information about TeD and instructions about how to obtain and install the core software.

Collaborations:

This project brought together experts in network modeling, design and analysis, distributed algorithms, and high performance discrete-event simulation kernels. Perhaps the most important contribution of this effort has been unfettered research and software design experimentation on most efficient, parsimonius and powerful methods to develop a user-friendly approach to harnessing parallel processing for large network simulation problems, a very difficult problem that has eluded the parallel simulation community for years. This required cross-disciplinary collaboration on implementation of a diverse suite of large and nontrivially detailed protocols and networks (traditionally not of interest in the community of parallel simulator software researchers), and very detailed parallel performance analyses of realistic network models (normally not of interest to network researchers).

Specific areas of network engineering of interest to us have included various aspects of multicast protocols, ATM networks, PNNI routing, and radio resource management in mobile wireless networks. Each of these areas developed substantial additional network modeling components and software. This project included a large number of graduate students whose educational benefits included exposure to industrial scale problems and internships.

Scalable Self-Organizing Simulations (S3)

This project was funded by DARPA for a three year period over 1996-1999, with an extension of the SSF sub-project (see below) through 2000, for a total of $3.1M. It represents a highly successful transition from the NSF-supported basic research on network modeling techniques for efficient parallel simulations to a very large network modeling software development effort.

Participating Scientists:

Andy Ogielski (PM, RU, leading PI), Narayan Mandayam (RU), Roy Yates (RU), Mark Crovella (Boston University), James Cowie (Cooperating Systems Corporation), Richard Fujimoto (Georgia Tech), Kalyan Perumalla (Georgia Tech), Ellen Zegura (Georgia Tech), Don Leskiw (Ultra Corporation), David Nicol (Dartmouth), with about thirty graduate students, postdocs, and part-time scientists at various times during the project lifetime.

Collaborations:

The complexity of modern protocols and networking technologies on the one hand, and of the large software systems on the other, demands that a project of this sort be carried out by an interdisciplinary team of experts in parallel simulations, networking, and industrial-strength software engineering. The latter in particular benefited greatly from the participation of Cooperating Systems, a small company specializing in high quality scientific software.

Objectives:

21st century global communications architectures will emerge from internetworking of millions of networks, as wireline, wireless, and satellite technologies interconnect hundreds of millions of diverse nodes. Verification of the design, operation, and management of these future architectures will rely on fast, accurate simulations. PITAC and NGI (Next Generation Internet) milestones called for a capability of simulations of 100,000 node, multiprotocol networks in year 2000, and simulations of 10,000,000 node networks in year 2001. At the beginning of this project the available network modeling tools and simulation techniques did not even approach systems of such complexity, falling short by many orders of magnitude.

In 1999 the network modeling software developed in this project, a direct descendant of the NSF-funded basic research described above, exceeded the year 2000 DARPA goal of modeling and simulations of 100,000-node multiprotocol Internet models. The project has developed and distributes a suite of detailed models of the major Internet protocols (IP, OSPF, BGP, TCP, plus sockets and assorted application level models), core modeling platforms, and other software subsystems, under the name of SSFNET (Scalable Simulation Framework). In June 1999 the SSF Research Network web site with scalable Internet modeling software and extensive documentation and tutorials went on-line at http://www.ssfnet.org.

Recent Accomplishments:

Our parallel network simulators have demonstrated execution rates and entity counts on realistic models well in excess of those achievable using the best tools available elsewhere. Using commodity multiprocessor machines (like SUN or SGI) the following results were established:

- November 1998: DaSSF simulations of the multi-domain Internet on a SUN Enterprise server deliver scalable performance for modeling networks of 100,000 multi-protocol hosts and routers with UDP Pareto-distributed traffic sources, or for multi-million-node networks of simpler elements, at the speeds of a million of IP packet hops per second.
- May 1999: DaSSF simulations on SGI Origin 2000 reached 100,000 concurrently active TCP connections (gigabit links, no losses); Java SSF simulations on SUN Enterprise servers (JDK 2) were closing in on this target. DaSSF multiprotocol Internet simulations ran faster than real time for several thousand hosts and routers.

These results exceeded the year 2000 goals of modeling and simulations of multiprotocol Internet models. By inherent, demonstrated scalability of SSF-based network models, and of the SSF parallel simulators, the SSF software can run unchanged on larger machines (as well as follow Moore's law), and reach many millions of hosts and routers in multi-domain, multiprotocol Internet models within the next couple of years or sooner.

The SSF event processing parallel kernels support high-speed interactive operation with real-time events and can be embedded in network monitoring and network management systems running on commodity multiprocessors (Linux, Solaris, SGI Irix), as well as in human-in-the-loop interactive applications.

Design and implementation of SSFNET, the first collection of SSF-based models for simulating Internet protocols and networks, was completed. Available in C++ and Java forms, the SSFNET models are open source software, distributed under the GNU General Public License. SSFNET models are self-configuring; that is, each SSFNET class instance can autonomously configure by querying a parameter database, which may be locally resident or available over the Web.

Development of a suite of core Internet protocol models for SSFNET was completed. They include IP, UDP, TCP, OSPFv2, BGP-4. Additional pseudo-protocols include network interface cards and queue management, bsd sockets, client and server applications. All protocols are new implementations, conforming to the SSF.OS "plug-and-play" self-configuring design. The protocols' conformance to IETF RFC requirements is rigorously tested.

Development of the first network design automation tools for SSFNET was completed. These include: automated IP address assignment following the CIDR guidelines; automatic verification of the network configuration syntax prior to model instantiation; verification of IP routing tables constructed by OSPF and BGP-4.

The SSF Research Network has been established, with the web site at http://www.ssfnet.org. It is an open consortium of SSF users, whose goal is to promote, develop, and validate SSF-based Internet models. This site contains substantial SSFNET documentation and tutorials, and will support automatic submission, search, and retrieval of a digital library of validated and experimental model components developed by the SSFNET community.

SSFNET 0.8, a major revision adding hierarchical Net specification, has been released.

DML 0.8, a major revision to the model configuration database code to support object-oriented attribute extension and import capabilities, has been released. Model database format was designed as a target for automated network design and validation tools.

Java SSF version 0.5 parallel simulator, a major revision supporting the full SSF 1.0 API, has been released.

Dartmouth SSF DaSSF 1.0.5 parallel simulator, a major revision supporting the full SSF 1.0 API was released.

The Surge workload generator was expanded to encompass a wider range of traffic types (HTTP 1.0 and HTTP 1.1), and both closed-loop and open-loop client-server interactions. We used an expanded workload generator to compare the performance of HTTP 1.0 vs. HTTP 1.1 over a wide range of system configurations. SURGE can be used for server benchmarking, and traffic generation both for networks and simulators.

Mobile wireless network modeling classes (WiPPET) were ported from TeD to SSF. Major new additions included: Wideband radio channel models (mobility; multipath; Rayleigh, shadow and distance fading; interference); wideband CDMA receiver signal processing (NTT DoCoMo / ARIB transceiver design and framing structure); multiple cells with antenna diversity and SIR based power control; soft hand-off protocol and the I/Q multiplexed framing structure. The granularity of these SSF models is at the baseband level (DS-CDMA chips).

We developed a new release of the TeD network description language compiler (predecessor of SSF) and the GTW parallel simulation engine, along with documentation, for public release. We developed a new release of the ATM Private Network to Network protocol (PNNI) models, along with documentation. We demonstrated near-linear speedup up to 16 processors clocking well over one million events per second, by using the novel "reverse computing" optimistic synchronization technique.

The next phase of SSF-based network modeling research and development will focus on a suite of scalable, multi-resolution tools for network design, validation, and analysis: Graphical editors for generation and validation of network topologies, scalable network data collection and multiresolution display facilities, statistical multiresolution tools for sensitivity analysis, etc.

We are planning the integration of SSF based Internet models (SSFNET) and physical/link wireless SSF models (WiPPET) for the first ever multi-scale baseband-to-IP wireless Internet modeling systems.

Technology Transition:

In August 1999, there was a public release of the SSFNET 0.8 scalable Internet modeling and simulation environment, comprising the parallel simulation kernel (JSSF), the model configuration database management system (SSF.DML), and the network hardware and Internet protocol models (SSF.Net, SSF.OS). Extensive documentation and tutorials are available on the SSFNET community web site http://www.ssfnet.org. Early adopters of SSFNET include:

- Internet traffic analysis (AT&T Labs - Research),
- Design of network security protocols and traffic engineering models (NIST)
- Design of Simulator of Air Force C4ISR aerospace operations within the Global Grid (Ultra Corporation)

- multicell GPRS (data for mobile GSM phones) models (University of Barcelona)

and a number of other projects in the USA and abroad are beginning.

In June 1999, there was a public release of Dartmouth SSF DaSSF 1.0.5 parallel simulator (all SSF simulators implement the same modeling API). Dartmouth SSF is ultra fast and demonstrably capable of simulating extremely large network models. See http://www.cs.dartmouth.edu/ jason-liu/projects/ssf/.

A portable version of Surge (web server workload generator) has been developed for public distribution. The Surge generator has been distributed to over forty sites, including:

- high-tech startups (Infolibria, Gigabit Wireless, Arrowpoint)
- telecom research labs (Lucent Bell Labs, AT&T Research, Telecom Italia, British Telecom, GTE Labs)
- industrial research labs (Xerox PARC, HP Labs, IBM T. J. Watson Labs, Tandem, Qualcomm, Motorola)
- over 20 universities, including MIT, CMU, Berkeley, UT Austin, and Cornell.

A professionally supported version of the C++ SSF parallel simulator was developed by Cooperating Systems Corp. The pre-release version was used for modeling and performance evaluation of the 3rd generation wireless networks (Wideband CDMA for IMT-2000) in a joint project of WINLAB and NTT DoCoMo. Modeling at the baseband signal level of detail ranges from multipath radio propagation with fading and interference, to radio receiver signal processing, to channel coding and link-layer protocols, and includes multi-cell mobility and signaling protocols such as power control and soft handoff. See

http://www.winlab.rutgers.edu/pub/projects/Spring99/p013.html

WINLAB adopted SSF as the platform for all wireless network projects. Several other wireless networking companies began negotiations to license this package for implementation of other receiver models (e.g., EDGE) and protocol suites (e.g., GPRS).

There was a new release of TeD-PNNI Models of ATM PNNI protocols for large-scale, high-fidelity simulation, comprising the protocol and network models, documentation, and the parallel simulation package Jane/TeD/GTW. The package is being actively used in networking projects in over 20 corporations and universities, including Lucent Bell Labs, Telcordia Technologies, Newbridge Networks, Traveller Multimedia Networks. See

http://www.cc.gatech.edu/computing/pads/pnnidoc.html

C41SR Global Network Simulator

In 1999, we received an award of $28,500 for our work on "C41SR Global Network Simulator." The award is a subcontract from the Ultra Corporation, one of the small companies that we have partnered with on the earlier DARPA project on telecommunications networks simulations, and comes from an Air Force contract aimed at design of a simulation of Air Force aerospace operations within the global grid.

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 also learn about how they can use LOGO to enhance their classroom discussions of discrete mathematics and where to find discrete mathematics resources on the Internet.

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 1999 **LP** involved two two-week residential programs (both at Rutgers) and three twelve-day commuter programs, in Scottsdale, AZ, in Warwick, RI, and at Rutgers. Those programs represent 138 teachers participating in the **LP** for the first time. In addition, there were two one-week second-summer follow-up programs for the 1998 participants – one at Rutgers (accommodating both residential and commuters), and the other (for commuters) at Scottsdale, AZ. Altogether, 77 of the 100 teachers in the 1998 summer programs attended one of these follow-up programs.

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, all of the commuter programs were extended by 2 days, so that they involved 12 days of activities rather than 10

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, after suitable training provided by the leadership program, and with the assistance of experienced lead teachers.

Phase III of the **LP** was scheduled to end with the 1999 cohort, but we have received a one-year supplementary NSF grant so that we can launch the franchise model. This will involve commuter programs held at the home sites of four mathematicians in the summer of 2000, and a residential program at Rutgers that will serve as a training site for the new program leaders. These five programs will involve 150 K-8 teachers in addition to the 640 teachers who have already participated in Phase III of the Leadership Program.

In the summer of 2000, we are also planning on four programs for the returning teachers from the summer of 1999, and a program for the veterans of previous summer institutes.

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** 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").

Since 1996 there has been a significant technology component to the Young Scholars Program. In addition to doing mathematics with computers and graphing calculators, students are involved in a week-long exciting "LEGO LOGO" project. Janice Kowalczyk (an educator from Rhode Island, now working with the Leadership Program full-time) runs four 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 grants of $115,000 (for 1998 and 1999 and 2000) 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; however, a limited number of scholarships were available. We are developing strategies for publicizing the program more widely. 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 seven 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 initially appropriate for middle school and high school teachers. However, we have now developed workshops for K-8 teachers as well and, as noted above, in a typical year, 5000 teachers attend workshops in our "Workshops in Your District" program.

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

Our LP for high school teachers has been replaced by a series of 2-day "crash courses", introducing those teachers to the excitement of discrete mathematics and its applications and to available resources and materials. Our summer 1999 crash course was attended by 36 teachers and the program will be repeated in the summer of 2000.

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 4.6.4.

In earlier years, DIMACS played a central role in the development of the New Jersey curriculum "standards" for mathematics and a curriculum "framework" that is based on those standards. The "standards" say in 30 pages what students at all grade levels are expected to achieve in all areas of mathematics. The "framework" is a massive piece of work, some 680 pages in all, that describes how the standards can be implemented in classrooms and curricula. The DIMACS Associate Director for Education, Joseph Rosenstein (PM, RU), was one of the three principal authors of the "framework" and served as reviewer and chief editor of the entire document. The framework is filled with new and original examples and activities, appropriate for different grade levels, that were originally developed at DIMACS in the process of finding new ways to present discrete mathematics and other fields of mathematics to teachers, and required creative new approaches to engage teachers in problem-solving, the use of mathematics as a tool to make people think, etc. The "framework" and the 1997 DIMACS volume *Discrete Mathematics in the Schools* have begun to have a real impact on mathematics education locally, nationally, and internationally.

Through the various programs mentioned in this section, we have reached tens of 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 4.6.3, and the new Modules Series, described in Section 4.6.4.

Among the other undergraduate initiatives at DIMACS are the following. We continued our successful Research Experiences for Undergraduates (REU) program and made some innovative changes in it (as discussed below). We continued to host visitors from such four year colleges as Drew, Haverford, Lafayette, and Amherst. 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 American 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 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. At the end of the summer, students present their work in front of DIMACS members and visitors and write a short report on their work. This summer, we also received NSF funding to bring two students to DIMACS from our sister site DIMATIA at Charles University in Prague, Czech Republic and to send five of our students to DIMATIA for a three-week research program after they completed five weeks at DIMACS.

Our 1999 program consisted of 12 students from 10 schools. We continued our DIMACS REU speaker series of talks by distinguished scientists, including Jozsef Beck (PM, RU), Jeffry Kahn (PM, RU), Fred Roberts (PM, RU), Julia Abrahams (PM, RU), and Linda Lesniak (PM, Drew University). 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 1999 DIMACS REU program along with their mentor, project area, and home institution:

DIMACS "Regular" REU Participants

- Brian Chen
  - Mentor: Janos Komlos (PM, RU)
  - Project area: Graph Theory
  - Home insitution: Rutgers University
- Jean Griffin
  - Mentor: Wilma Olson, RU, Director of Center for Molecular Biophysics and Biophysical Chemistry
  - Project area: Computational Biology
  - Home institution: Rutgers University
- Nipha Saeteng
  - Mentor: William Craelius, RU, Electrical Engineering
  - Project area: Biomedical Engineering

- Home institution: Simmons College

- Zachary Saul
  - Mentor: Eric Allender (PM, RU)
  - Project area: Arithmetic Circuits and Branching Programs
  - Home institution: University of Idaho
- Edward Slavin
  - Mentor: Craig Nevill-Manning (PM, RU)
  - Project area: Computational Biology
  - Home institution: Lycoming College

DIMACS/DIMATIA REU Participants

- Jessica Kazimir
  - Mentor: Charles Weibel, RU, Math
  - Project area: Interior Methods in Linear Programming
  - Home institution: Montclair State University
- Carly Klivans
  - Mentor: Sven Dickinson (PM, RU)
  - Project area: Mobile Robot Navigation
  - Home institution: Cornell University
- David Kravitz
  - Mentor: Bill Steiger (PM, RU)
  - Project area: Convex Hull Computation
  - Home institution: University of Delaware
- Tanzy Love
  - Mentor: Peter Hammer (PM, RU)
  - Project area: Logical Data Analysis
  - Home institution: CSU Hayward
- Christopher Malon
  - Mentor: Sven Dickinson (PM, RU)
  - Project area: Mobile Robot Navigation
  - Home institution: University of Chicago
- Robert Samal
  - Mentor: Alex Samorodnitsky (PD)
  - Project area: Chromatic Polynomials in Random Graphs
  - Home institution: Charles University

- Jan Vondrak
  - Mentor: Michael Krivelevich (PD)
  - Project area: Extremal Problems in Color-Critical Graphs
  - Home institution: Charles University

While in Prague participating in the DIMACS/DIMATIA REU Program, the students attended the European Symposium on Algorithms and the Prague Mid-Summer Combinatorics Workshop. They also heard talks by several DIMATIA researchers, presenting their work and open problems. One open problem presented by Jaroslav Nešetřil was solved by Paul Dreyer (GS, RU), the coordinator of the REU, and a generalization of the problem was solved by REU student Christopher Malon. While in Prague, the three researchers co-wrote a paper which has been published as a technical report and has been submitted for publication to the journal *Discrete Mathematics*.

Although our students work on widely varying projects during the summer, several students collaborated with others on their projects and have continued this collaboration past the completion of the program. Edward Slavin is currently writing a simulator for the mobile robot navigation project of Professor Sven Dickinson and students Carly Klivans and Christopher Malon, and Brian Chen has undertaken an independent study at Rutgers to develop a simulator for the project of Professor Bill Craelius and student Nipha Saeteng on modelling the mechanics of prosthetic hands.

Most of the students have created webpages describing their projects. These can be seen at http://dimacs.rutgers.edu/REU. 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 1999 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. Another student, Tanzy Love, presented her work in a poster session at the Joint Statistical Meetings in Baltimore in August. The paper describing her work is contained in the conference proceedings.

The REU program is evaluated every year by questionnaires to the participating students. This year, every single participant who returned their evaluation rated the experience as a whole as 4 or 5 on a 5-point scale with 5 being the highest. The ratings reflected the overall level of satisfaction the students had with their success at research. 77% 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 definitely recommend the REU program to their peers, or already had, and the one remaining person responded "maybe". The international part of the program had similar positive ratings.

Here are a few sample comments from 1999 REU students:

"I've never attended a different REU, but I've heard about quite a few, and there is a big difference in the 'realness' of the work at the DIMACS REU. Our mentors give us problems that might not lead anywhere and are definitely not routine. It's really exciting."

"My project was a complete learning experience. From the onset, it was clearly fraught with problems which made learning accessible step by step. While the big picture was clearly visible, working out actual results was the important part of the exercise. I learned a great deal about graph theory, but even more about problem solving in general, and how to proceed with really big problems."

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, 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 two such postdocs are spending both the 1999-00 and 2000-01 academic years at DIMACS. We also agreed with NEC Research Institute to share a postdoc for the academic year 1999-00 and have plans to share a postdoc with IBM Research in 2000-01. 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.

During the period covered by this report, we have two "classes" of postdoctoral fellows, the ones here during 1998-99 and the ones here during 1999-00. A list of these fellows and their present affiliations follows:

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
  - Current Position: Member of Technical Staff, Bell Labs (no longer postdoctoral position)
- Gruia Calinescu
  - Large Scale Discrete Optimization Postdoc, Rutgers
  - Research Interests: algorithms, approximation algorithms
  - Current Position: faculty, Georgia Institute of Technology, College of Computing
- 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: Assistant Professor of Computer Science, Loyola University - Chicago
- 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: unknown

- Daniel Huson
  - Molecular Biology Postdoc, Princeton
  - Research Interests: combinatorial tiling theory and sequence analysis
  - Current Position: Research Computer Scientist, Celera Genomics Corporation
- 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: Assistant Professor of Computer Science, Tel Aviv University
- Owen Kelly
  - DIMACS-WINLAB postdoc on Simulations Project, Rutgers
  - Research Interests: statistical signal processing and applications of nonparametric models
  - Current Position: research staff, Nortel Wireless Networks
- Michael Krivelevich
  - DIMACS - Institute for Advanced Study Postdoc at Rutgers
  - Research Interests: probabilistic combinatorics, hypergraphs, Ramsey theory, extremal graph theory
  - Current Position: senior lecturer in Mathematics, Tel Aviv University
- Anuj Puri
  - Bell Labs DIMACS Postdoc
  - Research Interests: computer aided verification, real-time systems, hybrid systems, computer networks, transportation systems, optimization
  - Current Position: Research Associate, University of California, Berkeley
- Alex Samorodnitsky
  - DIMACS - Institute for Advanced Study Postdoc, Rutgers
  - Research Interests: combinatorics, especially extremal set theory and coding theory
  - Current Position: DIMACS - Institute for Advanced Study Postdoc, IAS
- Mark Smith
  - AT&T Labs DIMACS Postdoc
  - Research Interests: distributed algorithms, network protocols and formal verification techniques, the design of network protocols

- Current Position: INRIA; will become member of technical staff, Bell Labs, 8/00

- Elizabeth Sweedyk
  - Molecular Biology Postdoc, Rutgers
  - Research Interests: complexity theory and algorithm design in computational biology, phylogenetic tree construction and sequence assembly
  - Current Position: Clare Booth Luce Assistant Professor of Computer Science, Harvey Mudd College and Claremont Graduate School

- 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
  - Current Position: Assistant Professor of Computer Science, Columbia University

- Kasturi Varadarajan
  - Large Scale Discrete Optimization Postdoc, Rutgers
  - Research Interests: algorithms, data structures, computational and combinatorial geometry
  - Current Position: Assistant Professor of Computer Science, University of Iowa

- Francis Zane
  - Bell Labs DIMACS Postdoc
  - Research Interests: complexity theory, circuit complexity, satisfiability algorithms, combinatorics
  - Current Position: Bell Labs DIMACS postdoc

- Leonid Zosin
  - Networks Postdoc, Rutgers
  - Research Interests: approximation algorithms, combinatorial optimization, distributed computation, cryptography
  - Current Position: Postdoctoral researcher, NEC Research Institute

1999-2000 Postdocs

- Ee-chien Chang
  - joint DIMACS-NEC Postdoc
  - Research Interests: image-related problems which have the ingredients of discrete computation; matching and indexing of point sets extracted from images; pixel-based image matching for the purpose of image retrieval; image watermarking

- Mypongsu Choe
  - joint DIMACS-WINLAB postdoc on simulations project
  - Research Interests: modeling the global Internet; parallel discrete event simulation
- Raissa D'Souza
  - Bell Labs DIMACS postdoc
  - Research Interests: discrete models of physical systems
- Yuval Ishai
  - DIMACS - AT&T Labs postdoc, Networks postdoc
  - Research Interests: cryptography and privacy, complexity theory, computational learning theory
- Dieter van Melkebeek
  - DIMACS - Institute for Advanced Study Postdoc at Rutgers; Computational Intractability postdoc
  - Research Interests: computational intractability; pseudo-randomness and derandomization
- Ashwin Nayak
  - DIMACS - AT&T Labs postdoc
  - Research Interests: quantum information and quantum computing; computational biology
- Alex Samorodnitsky
  - DIMACS - Institute for Advanced Study Postdoc, at IAS
  - Research Interests: combinatorics, especially extremal set theory and coding theory
- Francis Zane
  - Bell Labs DIMACS Postdoc
  - Research Interests: complexity theory, circuit complexity, satisfiability algorithms, combinatorics
- Leonid Zosin
  - Networks Postdoc, Rutgers (through December 1999)
  - Research Interests: approximation algorithms, combinatorial optimization, distributed computation, cryptography
  - Current Position: Postdoctoral researcher, NEC Research Institute

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 travel or software awards to support the research of local students, small awards for summer or winter break salary to support their research, and jobs and projects for which graduate students are paid. The response to this program has been highly enthusiastic and includes projects proposed and mentored by researchers on the staffs of our industrial partners. We supported 42 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 supported graduate students are reported on in our discussion of research thrusts in Section 3.

DIMACS-Supported Graduate Students, February 1999 - January 2000

- Bernardo Abrego
  - Affiliation: RU, Mathematics
  - Research Thrust: discrete and computational geometry; graph coloring
  - Current Position: GS, RU
- Gabriela Alexe
  - Affiliation: RU, Operations Research
  - Research Thrust: Boolean geodetic functions; logical analysis of data; graph representation
  - Current Position: GS, RU
- Sorin Alexe
  - Affiliation: RU, Operations Research
  - Research Thrust: database modeling; stable sets in graphs
  - Current Position: GS, RU
- Yong Bai
  - Affiliation: RU, Electrical and Computer Engineering
  - Research Thrust: simulations of the global Internet
  - Current Position: GS, RU
- Tomas Chudlarsky
  - Affiliation: Charles University, Czech Republic, Mathematics
  - Research Thrust: perfect graphs; the P_4 structure of a graph
  - Current Position: GS, Charles University
- Bela Csaba
  - Affiliation: RU, Computer Science
  - Research Thrust: extremal graphs; deterministic k-servers

- Current Position: GS, RU

- Thomas Davoine
  - Affiliation: RU, Operations Research
  - Research Thrust: cuts in integer programming; semidefinite programming; graph coloring
  - Current Position: GS, RU
- Aynur Dayanik
  - Affiliation: RU, Computer Science
  - Research Thrust: learning
  - Current Position: GS, RU
- Erik Demaine
  - Affiliation: University of Waterloo
  - Research Thrust: cache oblivious data structures (data structures which approximate I/O performance even when details of the physical machine are not known)
  - Current Position: GS, University of Waterloo
- Atul Deshmukh
  - Affiliation: RU, Industrial and Systems Engineering
  - Research Thrust: artificial intelligence approach to forecasting for aircraft operators
  - Current Position: GS, RU
- Srikrishna Divakaran
  - Affiliation: RU, Computer Science
  - Research Thrust: on-line algorithms for design of scalable web servers
  - Current Position: GS, RU
- Paul Dreyer
  - Affiliation: RU, Mathematics
  - Research Thrust: domination in graphs; majority processes; computational biology
  - Current Position: GS, RU
- Adrian Dumitrescu
  - Affiliation: RU, Computer Science
  - Research Thrust: discrete and computational geometry; crossing numbers
  - Current Position: GS, RU
- Amr Elmasry
  - Affiliation: RU, Computer Science
  - Research Thrust: self adjusting data structures
  - Current Position: GS, RU

- Silvia Fernandez
  - Affiliation: RU, Mathematics
  - Research Thrust: discrete and computational geometry; graph coloring
  - Current Position: GS, RU
- Jarl Friis
  - Affiliation: University of Copenhagen, Institute of Computer Science
  - Research Thrust: semidefinite programming
  - Current Position: GS, University of Copenhagen
- Benjamin Gum
  - Affiliation: PU, Computer Science
  - Research Thrust: batched algorithms; algorithms to reduce hard disk read/write latency.
  - Current Position: GS, PU
- Tiina Heikkinen
  - Affiliation: RU, Operations Research
  - Research Thrust: applications of stochastic programming in communication networks; network economics; resource allocation in wireless networks
  - Current Position: GS, RU
- Gabriela Hristescu
  - Affiliation: RU, CS
  - Research Thrust: complex object feature extraction; structural similarity of proteins
  - Current Position: GS, RU
- Panayotis Kevrekidis
  - Affiliation: RU, Physics
  - Research Thrust: dynamics of lattice kinks; bifurcation of internal solitary wave modes from the essential spectrum
  - Current Position: GS, RU
- Ayman Khalfalah
  - Affiliation: RU, Computer Science
  - Research Thrust: private information retrieval; greedy edge coverings of graphs
  - Current Position: GS, RU
- Stefan Langerman
  - Affiliation: RU, Computer Science
  - Research Thrust: algorithms for hyperplane depth in the plane
  - Current Position: GS, RU

- Yuri Levin
  - Affiliation: RU, Operations Research
  - Research Thrust: directional Newton method for solving nonlinear equations in several variables; realizations of hypergraphs and connections to electronic circuit design
  - Current Position: GS, RU
- Hongbo Liu
  - Affiliation: RU, Electrical and Computer Engineering
  - Research Thrust: simulations of the Global Internet
  - Current Position: GS, RU
- Ying Liu
  - Affiliation: RU, Operations Research
  - Research Thrust: logical analysis of data; detecting structural information about data sets; moment problems with applications to derivative pricing
  - Current Position: GS, RU
- Sachin Lodha
  - Affiliation: RU, Computer Science
  - Research Thrust: private information retrieval; greedy edge coverings of graphs
  - Current Position: GS, RU
- Dean Lorenz
  - Affiliation: Technion, Electrical Engineering
  - Research Thrust: network and services management; improved routing when flow information is available
  - Current Position: GS, Technion
- Michael Malisoff
  - Affiliation: RU, Mathematics
  - Research Thrust: optimization and dynamic programming
  - Current Position: Research Scientist, Systems Science and Mathematics Department, Washington University in St. Louis
- Ryan Martin
  - Affiliation: RU, Mathematics
  - Research Thrust: graph coloring; uniform hypergraphs
  - Current Position: GS, RU
- Jana Maxova
  - Affiliation: Charles University, Czech Republic, Mathematics
  - Research Thrust: oriented perfect path double covers of graphs; finding long paths and long cycles in graphs with certain average or median degree

  - Current Position: GS, Charles University

- Susana Nascimento
  - Affiliation: Universidade Nova de Lisboa, Dept. Informatica,
  - Research Thrust: soft clustering: the multiple prototype fuzzy clustering model
  - Current Position: GS, Universidade Nova de Lisboa

- Mikhail Nediak
  - Affiliation: RU, Operations Research
  - Research Thrust: sequential optimal experimental designs and computerized adaptive educational testing
  - Current Position: GS, RU

- Ming Ouyang
  - Affiliation: RU, Computer Science
  - Research Thrust: algorithms for the satisfiability problem
  - Current Position: Postdoctoral Fellow, Department of Computer and Information Sciences, University of Delaware

- Unsal Ozdogru
  - Affiliation: RU, Operations Research
  - Research Thrust: analysis of two-valve fluid-flow systems with general distributed repair times
  - Current Position: GS, RU

- Murali Rangarajan
  - Affiliation: RU, Computer Science
  - Research Thrust: High performance distributed protocols on a cluster of PCs
  - Current Position: GS, RU

- Christian Rodriguez
  - Affiliation: RU, Computer Science
  - Research Thrust: algorithms for nearest neighbor searching in Hamming space
  - Current Position: GS, RU

- Ali Shoukoufandeh
  - Affiliation: RU, Computer Science
  - Research Thrust: extremal graph theory
  - Current Position: Assistant Professor of Mathematics and Computer Science, Drexel University

- Cliff Smyth
  - Affiliation: RU, Mathematics

  - Research Thrust: correlation inequalities and inequalities in percolation
  - Current Position: GS, RU

- Yueyue Song
  - Affiliation: RU, Computer Science
  - Research Thrust: performance study of a combined routing and flow control protocol on a ring-based WAN
  - Current Position: GS, RU

- Amy Stern
  - Affiliation: RU, Mathematics
  - Research Thrust: probabilistic methods and random graphs
  - Current Position: GS, RU

- Balakrishna Subramanian
  - Affiliation: RU, Physics
  - Research Thrust: Developing computational tools for analysis of proteins based on their electron density maps obtained from x-ray crystallographic experimentations
  - Current Position: GS, RU

- Sandra Sudarsky
  - Affiliation: Texas A&M, Computer Science
  - Research Thrust: large multidigraphs arising from telephone calling data; detecting cliques and quasi-cliques in massive sparse multidigraphs
  - Current Position: Visitor, Bell Labs

- Xiaodong Sun
  - Affiliation: RU, Computer Science
  - Research Thrust: extracting Internet topological information; small interpolation sets in learning theory and combinatorial group testing
  - Current Position: GS, RU

- Kostas Tsioutsiouliklis
  - Affiliation: PU, Computer Science
  - Research Thrust: data structures and dynamic graph algorithms
  - Current Position: GS, PU

- Tonguc Unluyurt
  - Affiliation: RU, Operations Research
  - Research Thrust: sequential testing of complex systems
  - Current Position: Assistant Professor of Operations Research, Sabanci University, Istanbul

- Pierangela Veneziani
  - Affiliation: RU, Operations Research
  - Research Thrust: Boolean probability bounding problems
  - Current Position: GS, RU
- Mark Watkins
  - Affiliation: University of Georgia, Mathematics
  - Research Thrust: analytic and algorithmic number theory, prime numbers
  - Current Position: GS, University of Georgia
- Xuerong Yong
  - Affiliation: Hong Kong University of Science and Technology, Computer Science
  - Research Thrust: The adjacency matrix and spectrum of a graph with negative third largest eigenvalue
  - Current Position: GS, Hong Kong University of Science and Technology
- Gang Yu
  - Affiliation: University of Georgia, Mathematics
  - Research Thrust: number theory
  - Current Position: GS, University of Georgia
- Hande Yurttan
  - Affiliation: PU, Civil Engineering
  - Research Thrust: alternative solution methods for semidefinite programming
  - Current Position: GS, PU
- Song Wang
  - Affiliation: RU, Computer Science
  - Research Thrust: simulations of the Global Internet
  - Current Position: GS, RU
- Yuanping Zhang
  - Affiliation: Hong Kong University of Science and Technology, Computer Science
  - Research Thrust: number of spanning trees in graphs and networks
  - Current Position: Faculty, Hong Kong University of Science and Technology
- Yi Zhao
  - Affiliation: RU, Mathematics
  - Research Thrust: packing trees into complete graphs
  - Current Position: GS, RU

The graduate students were asked to evaluate their experiences in DIMACS graduate student programs. In total, 87% 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:

"Yes, the experience introduced me to the approximation algorithms direction and to some prominent specialists in this area. .. I appreciate very much the DIMACS progams and I consider that the DIMACS fellowship I obtained helped me to obtain and to improve my research results." – Rutgers OR student.

"I found it a good chance to support my research, do some more work in the field and choosing my advisor." – Rutgers CS student.

DIMACS "support allowed me to exchange ideas with some of the leading experts in viscosity solution theory and numerical analysis, and I got a lot of ideas on how I can improve my thesis as a result of this. My DIMACS fellowships also made my work take new directions. Prior to my DIMACS fellowships, my work was very theoretical ... I'm much more interested in applications of viscosity solutions, now ..." - Rutgers Math student.

"Yes, I had the chance to meet world class scientists and improve my contact with some of them and start collaboratons with them. .. I had the chance to meet two scientists in a workshop sponsored by DIMACS, and had a chance to talk about my ideas and get invaluable feedbacks on my work. I started collaborations with one of the visitors." – Rutgers CS student.

"Through the Workshop on SDP (Large-Scale [Discrete] Optimization Special Year), I was able to listen to presentations by many distinguished researchers in the area, and I was very glad to meet some of them. I also got a chance to meet some students who were attending Princeton University's weekly colloquia in Operations Research." – Princeton Civil Engineering and Operations Research student.

"From the support from DIMACS, I have gotten introduced to the field of biomathematics and biophysics and hope to pursue a research career in this field." – Rutgers Physics student.

"I saw most of the great mathematicians in my field (Combinatorics) at DIMACS. It did give me a chance to know them better." – Rutgers Math. student.

"I have made great progress in my research while visiting at DIMACS. ... DIMACS is an ideal place to do research. " – Visiting Graduate student from Hong Kong University of Science and Technology

"My DIMACS research is really outside my field of specialization. The opportunity to work with Massive Data Sets has been very rewarding." – Visiting Graduate student from Texas A&M.

4.5 Involvement of Underrepresented Groups

A large percentage of the participants in center programs are US citizens. Of the 2117 visitors, workshop, and program participants at DIMACS over the course of the year, 1677 were from the US. US citizens were also involved in good numbers in our training programs involving undergraduates, graduate students, and postdocs. Of the total of 61 individuals involved in these programs, 18 were US citizens. We have been working to increase the percentage of US citizens in these training programs. The number is skewed by the large number of foreign local graduate students.

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 2117 visitors, workshop, and program participants at DIMACS over the course of the year, 417 were women. In the undergraduate, graduate, and postdoc training programs mentioned above, a total of 18 of the 61 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.

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. During academic year 1999-00, we are hosting two visitors from the University of Puerto Rico for semester-long visits and have opened discussions about joint programs. We continue to discuss joint programs with HBCU's Spelman College and Howard University and have opened discussions with HBCU Texas Southern University, which has agreed to be involved in our Leadership Program as a "franchise" site. We maintain contact with HBCU Clark Atlanta University through Fred Roberts' service on the External Advisory Board for the Minority Excellence Center (CREST Center) for Theoretical Studies in the Physical Sciences at Clark Atlanta University. Among the leaders of DIMACS programs are a number of leading African American and Hispanic researchers. This includes a member of our Executive Committee, a chief organizer of our Molecular Biology special year, a key figure in our Reconnect program, and a number of workshop organizers. In laying plans for continuation of our Reconnect Program, we have received commitments from three African American researchers, Earl Barnes (Georgia Tech), Nathaniel Dean (PM, Rice University), and William Massey (ATTL) to help us recruit minority participants for this and other DIMACS programs.

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 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 4.2.1.

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. The successor to this program, the DIMACS Connect Institute (DCI), will be funded by the National Science Foundation for three years starting in the summer of 2000, and we describe this new program briefly here as well.

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 '99) was held at the DIMACS Center on the Rutgers University campus July 19 - August 6, 1999. This year's topics were related to Graph Theory and Its Applications to Problems of Society. There were an overall total of 151 people registered for this program. Program attendees included 32 high school teachers; 30 resident researchers for week one; 42 resident researchers for week two; 38 resident researchers for week three; and 20 local participants (including undergraduate students, faculty and researchers) who attended one or more days of the entire program.

Linda Lesniak (PM, Drew University) was both overall program director and director of the research program. Barry Tesman (Dickinson College) 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 (education consultant, Pennsylvania), and Patrick Carney (teacher, The Montclair Kimberley Academy) were the lead teachers for the education program who assisted Tesman in running homework sessions, professional plan development, teacher implementation sessions and overall program planning. Additional classroom presentations were given by Moshe Rosenfeld (Pacific Lutheran University) and Andy Liu (University of Alberta).

The core of the education program was a series of daily lectures given by Barry Tesman. Beginning in 1998, DREI became a credit program with requirements including teacher professional plans of activities to be implemented during the subsequent academic year, with status updates/presentations at the follow-up sessions. Another successful part of the education program consisted of teacher presentations on a topic based upon the classroom lecture materials/All-Institute Activities 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 and discuss follow-up. Participants were introduced to The Math Forum, an online community funded in part by the National Science

Foundation and hosted by Swarthmore College. The teachers were also introduced to two graph theory software packages that they could utilize as a pedagogical tool, as well as a research tool.

The lead teachers played a central role in the program. They participated in program and schedule planning from the very first meeting. Lead teachers were able to do one "implementation" lesson for the entire group 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 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 and were, in fact, introduced early in the week to 2 or 3 researchers with special interest in their topic. The researchers were impressed and delighted to see these presentations.

The research program consisted of three week-long research conferences. The topic of week one was Matchings in Graphs and their Applications, chaired by Brian Alspach (Simon Fraser University); the topic of week two was Concepts of Domination in Graphs and Directed Graphs, chaired by K. Brooks Reid (California State University); the topic for week three was Graphs, Posets, and Algorithms, chaired by Tom Trotter (Arizona State University). 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. Brian Alspach (Simon Fraser University) will be collaborating with Dean Hoffman and Chris Rodger (Auburn University) and with Ed Ihrig (Arizona State). Richard Anstee (University of British Columbia) worked with Brian Alspach (Simon Fraser University) on a problem mentioned at DREI'99. Kenneth Bogart (Dartmouth College) had some very useful research conversations with Ann Trenk (Wellesley College) that helped each of them in their research. Stimulated by Bert Hartnell's (Saint Mary's University) and Doug Rall's (Furman University) talks, W. Edwin Clark (University of South Florida) started working on the Vizing Conjecture and has been able to establish it for many classes of graphs. Another paper for Clark was stimulated by interaction with participants and the activity of giving his talk and was also just submitted for publication (the paper finds circulant graphs with large minimum degree (around $n - \sqrt{n}$ where n = order) and domination number = 3). Besides catching up with the research of others, Peter Fishburn (PM, ATTL) began a collaboration of with Ann Trenk (Wellesley College) on a problem of interest to her, and with James Abello (PM, ATTL) on a problem of his. Jerrold Griggs (University of South Carolina) worked on a graph theory problem of Alexander Kelmans (V, University of Puerto Rico) and Dhruv Mubayi (Georgia Tech) after the workshop and Kelmans says he can now solve it. Long-term consequences can be expected to come from a poset problem Griggs gave at the workshop. Attila Sali (Hungarian Academy of Sciences), Glenn Hurlbert (Arizona State University), and Eva Czabarka (University of South Carolina) all have good ideas on it, and Griggs expects they all will continue thinking hard about it. Gregory Gutin (Brunel University) and Anders Yeo (University of Victoria) have started working on a new paper with a tentative title "Preserving Diameters of Orientions of Digraphs and the One-way and Gossip Problems". This was inspired by the talk of Fred Roberts (PM, RU) and some discussions during the workshop. Steve Hedetniemi (Clemson University) is collaborating on the following papers: "Domination Graphs of Undirected Graphs," with G. Fricke (Wright State University) and D. Knisely (Tennessee State University); "The k-domatic Number," with Renu Laskar (Clemson University), J. Dunbar (Converse College) and A.A. McRae (Appalachian State University); "The Red Square Problem," with E.O. Hare. John Mackey (Dartmouth College) will be working with Grant McMicken (High school teacher) on a packing problem relating to Hexaminoes. K. Brooks Reid (California State University) will probably start a new collaboration with Hedetniemi based on some conversations at DREI'99 on domination graphs of graphs.

The organizers heard many of the researchers say that they 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 four years now and each year we have made changes and improvements. DREI'99 was clearly the best DREI yet, and the evaluation shows it. Here are some excerpts from the preliminary draft of the evaluation. Results from the exit questionnaire indicate that participants continued to be very satisfied with their institute experience. A total of 87% of the high school teachers, 97% of the researchers/professors, and 90% of the graduate students rated their overall satisfaction with the institute as high to extremely high on a seven-point scale with extremely high the highest value. The very positive rating by the 1999 participants is comparable to previous years. There are some indications that cross-program interactions were more frequent in 1999 than the previous institute, evidence that the institute is reaching one of its important goals. Nearly all of the high school teachers and the researchers/professors reported they had at least daily interactions with those in their own program. However, 82% of the researchers/professors in 1999 reported they had at least daily interactions with high school teachers. This is a noticeable increase from 63% of the researchers/professors who reported this level of interactions with high school teachers in 1998. The percentage of high school teachers who reported interacting with researchers/professors at least daily remained high, 83% at least daily in 1999 compared to 86% in 1998. A higher percentage of high school teachers in 1999 than 1998 reported at least daily interactions with graduate students, 60% compared to 38%. There appeared to be more frequent interactions between high school teachers and researchers at a significant level of involvement than in 1998. A higher percentage of high school teachers in 1999 than in 1998 reported they engaged in collaboration or planning future joint work 2-to-3 times a week while at the institute-66% compared to 41% collaborated and 25% compared to 6% planned future work. High school teachers did not exchange information on education among themselves or with researchers/professors as frequently in 1999 as in 1998-74% compared to 95% in 1998 among themselves and 22% compared to 41% in 1998 with researchers/professors. High school teachers reported they exchanged information on mathematics in nearly the same frequency as in 1998-83% daily with other high school teachers and 50% daily with researchers/professors. Researchers/professors reported, in a higher percentage than in 1998, a daily exchange of information on education and mathematics with high school teachers-41% compared to 23% on education and 45% compared to 27% on mathematics. About the same percentage of researchers in 1999 as in 1998 reported they had at least 2 or 3 times a week collaborated or planned future work with high school teachers while at the institute-29% compared to 27% collaborated and 11% compared to 8% planned future work.

What the preliminary data from the evaluation of DREI'99 indicate is that the institute continued to meet the needs of the participants while increasing the already-high interactions of researchers/professors with high school teachers. In addition, the high school teachers engaged in more frequent collaboration with other high school teachers at the institute. About one-third of the researchers reported they had collaborated at least some with high school teachers and about 10% planned to continue such collaborations. We think this is a very dramatic success rate and evidence of true integration.

The following are some comments from the participants.

"Here at my university, I have a pretty good reputation as a teacher. That's for courses at the university level. At your conference, I was surrounded by many gifted teachers at the high-school level. Different worlds, yet a common goal. I certainly benefited from the interaction. Judging from many conversations I had there, I think they did too." – researcher, College of Sciences and Mathematics, Auburn University

"The contact with the teachers was also a very profitable one. I really wish I would keep in touch with some of these teachers who are so willing to do some research." – researcher, Department of Mathematical Sciences, Clemson University

"Overall, I think that the workshop was a great success and studying topics in graph theory by high school teachers is a great and useful idea." – researcher, Department of Mathematics and Statistics, Brunel University

"One aspect of the conference I liked was seeing a fair number of young teachers in this year's group. I think it is a good idea to make certain there are relatively new teachers each year." – researcher from Department of Mathematics, Simon Fraser University

"I liked the idea of having teachers and researchers together." – researcher, Department of Mathematics, Dartmouth College

"All in all I enjoyed the program very much. I also thought that the interaction between the 'researchers' and 'teachers' went quite well, much better than I would have anticipated." – researcher, Department of Mathematics, University of South Florida

"It was very interesting to find out what is going on in some of our high schools. Of course, the schools and teachers present represented above average schools. Nevertheless, as a college teacher I personally lead an isolated life with respect to high schools. Thus, this was very good. I expect to follow up some contacts I made with people from local high schools concerning our student teacher program and a student recruiting idea we are launching at Pace this fall." – researcher, Department of Computer Science and Mathematics, Pace University

"I really enjoyed interacting with the math teachers. I have started to look into trying to run something similar here for East Alabama - it is too early to tell if something will come of it, but I am hopeful." – researcher, Department of Discrete and Statistical Sciences, Auburn University "

"I would say that the DREI'99 Education Program opened my eyes to view a branch of mathematics that is not in the curriculum of my high school. Graph Theory is teachable to students in high school. The value for my students comes from exposing them to different branches of mathematics they may not have seen before." – High school teacher

"DREI '99 was simply the best educational experience that I've ever had in the field of mathematics... I've never had the opportunity to work with people who are at the cutting edge of mathematics before and I don't know of any other program that encourages or even provides for that. My students are already benefiting from my summer experiences." – High school teacher

"DREI 99 was among the best experiences of my life, and I have expressed this sentiment several times to those responsible in thanks and appreciation. I will never have the opportunity again to meet individuals as accomplished, intelligent, and interesting as those that I met at DREI. Not to mention of course the inspiration that I have to reform my Discrete Math class to incorporate all that I have learned. I am a better teacher and a more educated individual after this program. Thank you again." – High school teacher

4.6.2 DIMACS Connect Institute (DCI)

Starting in the summer of 2000, DIMACS will initiate a new program called the DIMACS Connect Institute (DCI) that is the successor to DREI. This will be a five year program, with the first three years supported by a new award from NSF and the last two years supported through a combination of funds from the DIMACS partners and new corporate supporters that we anticipate. The goals of DREI and DCI are similar. The major difference is that in DCI we aim to give teachers a true research experience that, in turn, can be reflected in the excitement they bring to their classroom teaching. The first two weeks of DCI are like DREI: research workshops, an educational program with lectures, and first research experiences through "scavenger hunts" requiring teachers to get information that is only available in library volumes, on the Internet, on MathSciNet, and through discussions with reseachers attending the research workshops associated with DCI. However, once the research workshops of the first two weeks are over, the teachers will remain at DIMACS for two more weeks. Here, they will work with researcher mentors on research projects, development of classroom implementation, and discussion of written materials. We will be preparing the teachers for significant in-class follow-up during the subsequent academic year.

During that academic year, teachers will organize projects in their schools, both in-classroom projects and out-of-classroom projects. We will arrange classroom visitations, utilizing teachers to visit other teachers and also Lead Teachers, mentors, and researchers to do some visiting, and we will provide written feedback after a visit. In May, there will be a one day conference at which teacher/student teams present their research. This is a major new innovation and puts more emphasis on student/teacher teaming.

Another major change is that our DCI teachers will return for two weeks in a second summer. Here, they will participate in a second research project, make presentations to the new group of teachers on their in-classroom and out-of-classroom experiences, and work on their written materials.

We are anticipating major support from corporate sponsors and in particular we are planning to work with such sponsors to pinpoint resource-poor school districts and minority teachers who could become involved.

Because of the success of the theme "Graph Theory and its Applications to Problems of Society" in our final two years of DREI, we will keep this theme. The topic lends itself beautifully to beginning research experiences for teachers and is ideal for in-classroom and out-of-classroom projects to students of varying abilities. In particular, we will emphasize location problems in one of the summer 2000 workshops. This is an ideal topic for DCI. The motivating problem involves finding the best place to locate a new facility such as a supermarket, parking lot, fire station, hospital, or garbage dump. This can be modeled using a graph with "users" at different locations in the graph (either at vertices or along edges). We wish to find a location that optimizes some criterion such as the sum of the distances to the users or the maximum distance to some user. The problem has close connections to voting and group decision making, since we can think of users as voters and the choice of a location as being voted on. Simple examples of graphs can be used to explore where to put locations. Real-world location problems can be researched by visits to local city officials or by investigating detailed local maps, and these problems can then be modeled using graph theory. While it is in general difficult to find the optimal solution to a given location problem under many of the widely studied criteria, these problems can be solvable for special classes of graphs that might arise in real applications and they can be explored for all graphs of a given type with a small number of vertices. Heuristic algorithms developed for a variety of location problem criteria can be explored for small examples. New heuristics can be developed. These are the types of things we would encourage our teachers to explore first. Another complication, an area of considerable current research interest, involves the difference in approach to problems where the facility is "friendly" (such as a supermarket) or "obnoxious" (such as a garbage dump). A natural extension brings in probabilities, such as the probability that a given user will need the facility, or usages and delays along different streets in the network, thus leading into facility location problems on weighted graphs. The axiomatic approaches to help us identify criteria we seek in a "good" solution to a voting problem are also relevant to location problems, and teachers can explore different axioms for good solutions to facility location problems as part of their research.

We are excited to be launching DCI and expect the same kinds of dramatic successes that we have had with DREI.

4.6.3 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 don't have 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 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 Modules Series. Participants are also encouraged to become involved in other DIMACS activities, such as workshops and the DIMACS Research and Education Institute (DREI) and DIMACS Connect Institute (DCI).

Reconnect Summer Program

The Reconnect '99 2-week conference in July, organized by Fred Roberts (PM, RU) and Rochelle Leibowitz (Wheaton College), had 23 participants from across the United States. The two topics for this year's conference were clustering and visibility in geometry. Both topics have seen rapid growth in research results in the last few years, have many real-world applications, are interdisciplinary, and can be brought into the undergraduate math and computer science classroom. Pierre Hansen (GERAD, University of Montreal) was the principal lecturer in clustering for both weeks, and Joseph O'Rourke (Smith) was the principal lecturer in visibility in geometry for week 1 and Ileana Streinu (Smith) for week 2. These lecturers were responsible for presenting the core material in their respective areas. In addition to the formal presentations by the lecturers, we encouraged our participants to be active rather than passive learners by scheduling time for "homework" problems and homework review, library reading, question-and-answer periods with the lecturers, and informal get-togethers with the lecturers and other researchers at DIMACS not directly connected to the conference. 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 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, and 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 lecturers were helpful in suggesting topics and offering advice to these groups. At the end of the first week, we had a session with previous participants of both Reconnect and DREI. These previous participants gave sound advice and answered many questions concerning the writing process.

During the second week of the conference, we kept to a similar model of having the lecturers present material, but we scheduled more time for writing and less for lectures. Again, time was scheduled for question-and-answer sessions and informal get-togethers with the lecturers. The lecturers also stayed around to help give advice and be the 'experts' for the writing groups. Writing group presentations were given during the last day of the conference. The conference ended with a charge to each group to polish its module and submit it to the DIMACS Educational Modules 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 consultant Regina C. Gray. The participants were very pleased with the organization of the conference, the clarity of the lectures, the topics covered and its relevance to the undergraduate classroom, and their exposure to the world of mathematical research. All of the respondents to the questionnaire at the end of the conference rated their overall satisfaction as 4 or 5 on a 5-point scale with 5 being highest. Last year's evaluation led to the suggestion that we provide notes and readings ahead of time. We did this for Reconnect'99 and it made a difference in participant satisfaction and readiness to do writing.

Here is a selection of comments from the participants.

Most Valuable Experiences at Reconnect'99:

"meeting the researchers and having extended interaction with them"

"seeing research at a level which allows nearly immediate application to real world problems"

"understanding the level of discourse my students need to attain to pursue graduate work"

"potential for publication of module ... to possibly lead to more in the future"

"the *wonderful* presenters"

"meeting peoplle that are excited about an area of research (our speakers) and finding out that many others have similar experiences as I do at my home institution (find themselves with little/no research time due to teaching and service obligations)"

Miscellaneous comments:

Reconnect "exceeded my expectations in reconnecting or creating academic interest in a new field."

"My goal was to jumpstart my sabbatical, meet some dedicated teachers, and refocus myself on content (rather than grading assignments) – I am so grateful for this experience! My expectations were exceeded."

"The lectures gave me new knowledge that I can pass on to my students. It also gave me a sense of being more connected to the Research Community."

"I feel reconnected. In fact I spoke to one presenter about devoting a semester's sabbatical to the subject."

"It makes me wonder how we get disconnected in the first place."

The modules are further along now than those from the 1998 Reconnect summer group were at the same time. Indeed, almost every one of the eight writing groups had a first draft of a module ready before leaving the conference. Editor Steve Maurer and the associate editors of the DIMACS Educational Modules Series are working closely with the participants in getting their modules ready for possible publication. One immediate consequence of Reconnect '99 is that one of the participants is competing with other Project NExT fellows to get a contract to write several chapters on graph theory in a book that a publishing company plans to produce. This participant is not a graph theorist but exposure to graph theory and its applications during Reconnect '99 has made her confident to compete. One of the organizers (Rochelle Leibowitz) will introduce one of the geometry modules on 'Art Gallery Theorems with Geometer's Sketchpad' in her math course for prospective elementary school teachers. Other Reconnect participants plan to introduce some of the modules, their own and others, into their classrooms.

Reconnect'98 has had continuing impact. One participant and a co-organizer of the 2-week conference ran 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 already introduced material that they learned this summer into their classrooms. A number of modules from Reconnect'98 participants are in the process of being improved for inclusion in the final modules series.

Plans are well underway for Reconnect '00. The principal speakers for the two-week conference will be Larry Gray (Univ. of Minnesota) on the topic probabilistic cellular automata with applications to such topics as traffic flow and Cathy McGeoch (Amherst) on experimental computing.

Reconnect Two 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. Participants attend a two-day workshop, work on modular classroom material, and return for a follow-up six months later.

The 1998 group finished their activites with a meeting on February 6, 1999, at Rutgers University. Fred Roberts (PM, RU) made a presentation entitled "From Genes to Archaeological Digs and From Traffic Lights to Childhood Development: The Many Applications of Interval Graphs." Participants also descussed final versions of their projects. These were bound into a single volume for distribution.

The 1999 workshops were organized by Jean Lane (Union County College) and Joseph Rosenstein (PM, RU). An advisory board with local and national members offered suggestions on content and publicity. The program was advertised in the Mathematical Association of America's "Focus" and in the Regional Newsletters for AMATYC (American Mathematical Association of Two-Year Colleges), as well as AMATYC's state affiliates newsletters. Lane and Charles Miller (Camden County College) screened the applications. Twelve applicants accepted invitations to join the Program and met May 22-23, 1999 at DIMACS, Rutgers University.

While New Jersey faculty predominated in the 1998 group, this year's group members were mostly out-of-staters. Presentation of the core material in discrete mathematics was done by Robert Hochberg (PM) and by L. Charles Biehl of the Charter School of Wilmington. In addition, Sven Dickinson (PM, RU) presented current research on "Object Representation and Recognition", and Bruno Simeone (V) from University of Rome spoke on "Political Districting: Traps, Criteria, Algorithms and Trade-Offs." Fred Roberts (PM, RU) gave a short presentation on "Writing Discrete Mathematics Modules" for the proposed DIMACS Modules Series.

As previously noted, last year's participants were mainly from New Jersey and, additionally, the entire group was larger. Therefore, groups of two, three and four formed to work on projects. The 1999 participants were more geographically diverse and preferred to work singly over the summer.

Participants reconvened on Oct. 23-24, 1999 to discuss their projects and obtain colleagues' comments and input. Jean Lane spoke briefly on Discrete Mathematics in the AMATYC Standards. Additional lectures were given by Adam Buchsbaum (PM, ATTL) on "Engineering the Compression of Massive Tables: an Experimental Approach" and by Dieter Van Melkebeek (PD) on "Solving Mazes: Deterministic and Randomized Strategies."

Final versions of the participants' projects are due Jan. 15, 2000 and will be bound together for distribution to interested parties.

The program was evaluated by a series of questionnaires. 83% of the participants felt that the material was applicable or very applicable to their classrooms. 100% of the respondents felt that the quality of the program was good or excellent. We asked the participants to summarize "in five words or less your reaction to the program." Here were a few of the comments:

"Fun, collegial, motivating."

"Good doing mathematics with others."

"Excellent ultra-compact frantically paced overview."

"Excellent opportunity to work with colleagues: focused collaboration."

"Very interesting, intellectually stimulating and informative!"

Here is a sample of a longer comment:

"Thank you and all of the excellent presenters for such a first class, academically informative and challenging 2-day workshop. I have not been exposed to so many discrete math topics, experts, examples, discussions, etc. before. This was a wonderful experience – we must spread the word to our colleagues and try to transmit as much of this information to our students as possible. I am particularly impressed by the high caliber background, marvelous teaching style and enthusiasm of all the presenters. ... Thank you! Thank you!"

4.6.4 Modules Series

DIMACS has commenced a modules 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.

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 Board has identified many current research topics which seem ripe for modules and has also identified researchers who may be receptive to writing modules, or serving as co-writers or advisors. The Modules Editorial Board is especially eager to arrange for collaborations between researchers, expositors, and teachers.

The first modules should come to publication in 2000, starting with a module on clustering and one on optimal nucleotide sequence alignment. So far, the best source of writers has been the Reconnect summer conferences from 1998 and 1999. Everyone from both programs has tried writing a module (in groups), and in the 1999 program the format was so organized that most of the 8 groups had first drafts of potential modules before they left. Stephen Maurer (PM, Swarthmore College), the Editor in Chief of the Module Series, attended for several days of both programs and gave advice and prompt feedback to get the groups started.

In addition to the work of the 1999 Reconnect conference, from which eight modules are underway, there are about 10 other modules well underway and others under discussion. Modules from the 1998 summer Reconnect group are on RNA string assembly; "carousel numbers" and DNA assembly; optimal sequence alignment; crossing numbers; clustering; and minimum spanning trees and clustering. In addition, there are some modules in progress from people who attended DREI and from attendees at the shorter 2-day Reconnect conferences in 1998 and 1999. People with more general connections with DIMACS, and even people who have simply obtained our brochure at conferences, have expressed an interest. For instance, we have drafts of modules from a researcher in Mexico (on Hamilton cycles) and from a high school teacher in Massachusetts (facility location). Both attended DREI in 1998.

5 Outreach and Knowledge Transfer

5.1 Industrial Interactions

DIMACS has four industrial partners: AT&T Labs-Research, Bell Labs (Lucent Technologies), NEC Research, and Telcordia Technologies (formerly 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. 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 7/1/98, the estimated value of their contributions are: Bell Labs $615,000; AT&T Labs - Research $515,000, NEC Research $250,000, Telcordia $400,000. These numbers include gifts of $100,000 from AT&T Research and $30,000 from NEC Research for shared postdocs. In the past, almost all of the contributions to DIMACS from our industrial partners have been earmarked for specific purposes. This year, as part of the transition of the center from an NSF STC center, we have approached our corporate partners for significant unrestricted contributions to the running of the center and three of them have responded with "membership gifts" of $40,000 to $50,000 each.

Special mention should also be made of the AT&T Foundation, which has been an especially important supporter of DIMACS' programs in education and at the interface between research and education. The Foundation has supported our Young Scholars Program for the past three years and just recently gave us another gift of $40,000 for that program.

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 Selection 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; NEC Research $160,000; Telcordia $200,000. These amounts are included in the amounts in the previous paragraph.

We have discussed or planned a number of joint activities with IBM Research, including joint workshops and a shared postdoc. The first joint workshop took place in August 1999 at IBM Research in Yorktown Heights and was a major culminating event in our Special Year on Large Scale Discrete Optimization with a major conference on the theory and practice of integer programming. The second joint workshop, on data mining in the Internet age, will take place at IBM Almaden in May 2000 as part of our new Special Focus in Next Generation Networks Technologies and Applications. IBM Research is also partnering with us in the upcoming Special Focus on Computational Molecular Biology.

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, SmithKline 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. Several companies have been extremely helpful in planning our forthcoming Special Focus on Computational Molecular Biology. In particular, researchers from Roche BioScience, SmithKline Beecham, and IBM are serving on the Advisory Committee for that activity, have agreed to organize workshops, and are in the process of raising corporate funds for that purpose. 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.

We anticipate significant industrial involvement in our upcoming Special Focus on Next Generation Networks Technologies and Applications. Three of the four major organizers are from industry, from AT&T Labs and Akamai Technologies. Our Advisory Committee represents a number of major companies from the information technology field and we have obtained promises of support for this activity from Microsoft and IBM, among other companies.

The year 1999 opened a new industrial partnership for DIMACS with the involvement of the Electric Power Research Institute (EPRI) in our workshop on Next Generation Unit Commitment Models. Not only did EPRI help support the workshop financially, but it is underwriting a large part of the cost of the conference volume by agreeing to purchase up to 1000 copies and distribute them to its industrial members. The workshop was exciting because of its mix of academic and industrial participants and the involvement of EPRI played an important role in bringing that about.

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 September 1999. 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, enhancing and expanding our many partnerships has remained a priority of DIMACS.

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 1998-99, we shared a postdoc at Rutgers, who is spending 1999-00 at IAS. A new DIMACS-IAS postdoc has joined us at Rutgers in 1999-00. We have agreed to continue the joint postdoc program for at least another three years. IAS has been a major partner in our special foci. The Special Focus on Discrete Probability was our first major joint event. A joint event was held at IAS during our Special Year on Large Scale Discrete Optimization. Avi Wigderson, a permanent part-time faculty member at IAS, has been a key organizer of our 1999-00 Special Year on Computational Intractability and is involved in arranging a follow-up year at IAS, with continuing activities at DIMACS. DIMACS is also a subcontractor to IAS for our DREI program that integrates research and education, as part of IAS' NSF grant for its Park City Mathematics Institute.

With the University of Pennsylvania, we have run a number of joint programs. In the past we had shared a postdoc in molecular biology and run several joint workshops. The cooperation continued during 1998-99 with a joint workshop with our sister STC at U. Penn, the Institute for Research in Cognitive Science, on "Logic and Cognitive Science."

Several of our activities in the special year on Large Scale Discrete Optimization have been jointly sponsored by Carnegie Mellon University and Georgia Tech. Faculty from those two universities are increasingly active in DIMACS activities. We also held a joint DIMACS - Georgia Tech workshop on the "Combinatorics of Statistical Mechanics Models" at Georgia Tech in April 1999. This was a follow-up to early events in our Special Focus on Discrete Probability.

At Rutgers, our simulations project has been carried out jointly with faculty at WINLAB, the Wireless Networking Laboratory, and WINLAB faculty are heavily involved in our research projects on simulations of telecommunications systems described in Section 3.10 and in planning for our Special Focus on Next Generation Networks Technologies and Applications, as members of the Advisory Committee and workshop organizers. WINLAB brings to this collaboration extensive practical experience in wireless and mobile computing topics, and faculty from WINLAB will be helping us with a major summer 2000 tutorial program on these topics. In addition, several postdocs, graduate students, and undergraduates at DIMACS have been mentored by WINLAB faculty.

In molecular biology and DNA computing, our special year activities have engaged us in serious collaborations with a number of centers 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 the Molecular Biology Department at Princeton has been one of the organizers of the DNA Computing Special Focus and of a number of its workshops. Professor Lee Silver from the Molecular Biology Department at Princeton hosted a DIMACS tutorial and this resulted in an interaction with DIMACS visitor Dannie Durand to produce 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, including the March 1999 workshop on the topic "Advances and Opportunities at the Biology/Math/Computation/Physical Sciences Interface." The Director of the Center, Wilma Olson, is one of the organizers of our new Special Focus on Computational Molecular Biology. She has also been mentoring in the DIMACS REU program for four summers. We have also partnered in workshops with the Program in Mathematics and Molecular Biology at Florida State University. PMMB was also a sponsor of the March 1999 workshop. 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.

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 served on the steering committee for our Massive Data Sets Special Year, and she organized one of the special year workshops. Victor Miller from IDA has served on the organizing committee for our January 2000 workshop on "Unusual Applications of Number Theory."

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 ran, as part of our Networks Special Year, a joint workshop with IMA on the topic "Scaling Phenomena in Communication Networks." Other joint programs are under discussion.

At the January 1999 and January 2000 American Mathematical Society national meetings, a number of the research centers in the mathematical sciences held and 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, the National Institute of Standards and Technology (NIST), and the Institute for Pure and Applied Mathematics at UCLA.

DIMACS has connections and joint programs with a number of professional societies. We have sponsored meetings for the NJ Chapter of the Mathematical Association of America and of AMATYC, the American Mathematical Association of Two Year Colleges. AMATYC is also a strong supporter of our "Reconnect" program. We have initiated plans to jointly sponsor a workshop with SIAM, the Society for Industrial and Applied Mathematics. IEEE, the Institute of Electronic and Electrical Engineers, has also been a collaborator on a workshop.

5.3 International Activities

DIMACS has been actively involved in 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 and Applications, 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, published in 1999, is entitled *Contemporary Trends in Discrete Mathematics*. During the present year, we ran a joint workshop on "Algebraic Methods and Arithmetic Circuits" at Rutgers in June 1999. These specialized workshops return to continuing research thrusts at DIMACS, in this case the Special Year on Logic and Algorithms. As part of the partnership with DIMATIA, we have had regular exchange of visitors, with an emphasis on young people. We have also launched 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. This is supported by a second grant from International Programs at NSF.

The Erdős Institute is part of the Hungarian Academy of Sciences in Budapest. We ran 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. We have 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. We have also discussed exchanges of visitors and students and have already hosted a BRICS graduate student on a long-term visit.

HKUST is the Hong Kong University of Science and Technology. We have started a collaboration with the Computer Science Department at HKUST which in 1998 resulted in a joint workshop. Mordecai Golin from HKUST spent a good part of his 1998-1999 sabbatical at DIMACS and we also hosted two HKUST graduate students during that year. Professor Golin was active in our Massive Data Sets Special Year and organized a workshop while he was at DIMACS.

We gave begun a collaboration with INRIA - Universite de Nice-Sophia Antipolis in France. Jean-Claude Bermond from Sophia-Antipolis visited DIMACS several times during 1999 and we also hosted Herve Bronimann from Sophia-Antipolis. We have agreed to plan several joint workshops and exchange students. Bermond is a key partner in our international efforts in connection with our Special Focus on Next Generation Networks Technologies and Applications. The Office of Naval Research, a sponsor of that special focus, asked us to involve an international partner, which very naturally was Sophia-Antipolis. We have also involved Sophia-Antipolis in future plans for other information technology projects.

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 last IMSI meeting.

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, 75% rated their experience at DIMACS 5 (the highest) out of a 5-point scale and 97% rated it 4 or 5. All but one said they would want to come back. (The one said he was too busy in the near future.) Here are a few selected comments from the visitors:

"I learned many new approaches and methods which I plan to use in my future work. As well

I found several new problems which I was unaware of where some methods I know could be used. ... The contacts with postdocs here at the center, as well as with senior scientists at ATT Research and Bell Labs were extremely helpful."

"I was able to reaffirm the direction I was taking with regards to the staff scheduling project I am working on. As well, I was able to continue my work in graph theory and with a new collaborator."

"My stay ... was a unique opportunity to make or to renew very useful contacts with many specialists in my field [large scale discrete optimization]... my stay has been, in particular, an opportunity to learn more about computational biology."

My visit "provided a great deal of additional background material that helped to broaden my understanding of the field of discrete optimization and the constraint manipulation algorithms used in this field. My prior research had involved a specific application of a specific algorithm. ... More importantly, DIMACS researchers helped me to find work that is related to my own work. I believe this will improve the quality of the work I am now doing, as well as make the completion of this work easier."

"During the last 4-5 years before my semester at DIMACS I spent almost no time on research. The biggest result of my stay was that I came away with the conviction that I am still capable of doing research." – visitor from 4-year college.

DIMACS asks all attendees at workshops to fill out a questionnaire. Of the workshop attendees during the past year, 52% rated the workshop as 5 (highest) on a 5-point scale, while 89% rated it 4 or 5 and 95% rated it 3 or higher. 93% of the attendees said that they made at least one new contact at the workshop and 51% of those said they expected to collaborate with their new contact, with 80% saying they expected to or could collaborate. Here are a few comments from workshop participants:

"Excellent specialized workshop with many leading researchers offering interesting ideas. Very good overall coverage of the area." – Workshop on Large Scale Discrete Optimization in Logistics

"Thank you for giving me the opportunity to participate in the workshop! It was very well organized, and the research topics were of the highest caliber." – Workshop on Advances and Opportunities at the Biology/Math/Computation/Physical Sciences Interface

"Thank you for this opportunity. I hope there will be an extension of the workshop in the near future." – Workshop on Logic and Cognitive Science

"I think the organizers have put together a nice program! I am flattered to have been invited." – Workshop on Discrete Mathematical Problems and Medical Applications

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 UltraSPARC/SPARC servers with over 60GB of storage capacity and 2GB of total RAM; 4 NT servers with 25 GB of storage and 500MB of RAM; 17 Sun workstations, a Silicon Graphics Indigo workstation with 384MB of RAM; 2 PC Labs with a total of 15 PCs; and an additional 8 PCs and 24 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 at our industrial partners' locations for collaborative research.

7 Administration and Management

7.1 Organization of the Center

The center has the following management:

Management

- Director: Fred S. Roberts, RU
- Co-Director: Robert Sedgewick, PU
- 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:

- Martin Farach-Colton, RU
- Peter Hammer, RU
- David Johnson, ATTL
- Jon Kettenring, TC
- Fred S. 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:

- Farid Alizadeh, RU

- Eric Allender, RU
- Endre Boros, RU
- 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
- Jon Kettenring, TC
- Janos Komlos, RU
- Clyde Monma, TC
- S. Muthukrishnan, ATTL
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael 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 from all or most of the DIMACS units. The committees are as follows:

Postdoctoral Selection Committee

- Sanjeev Arora, PU
- Joszef Beck, RU
- Enrico Bombieri, Institute for Advanced Study

- Martin Farach-Colton, RU
- Peter Hammer, RU
- David Johnson, ATTL
- Sanjeev Khanna, BL
- S. Muthukrishnan, ATTL
- Andy Ogielski, RU
- Sivaram Rajagopalan, TC
- Fred Roberts, RU
- Mike Saks, RU
- Bob Sedgewick, PU
- David Williamson, IBM Research

Projects Committee

- Farid Alizadeh, RU
- Ken Clarkson, BL
- Mike Fredman, RU
- Janos Komlos, RU
- S. Muthukrishnan, ATTL
- Rafi Ostrovsky, TC
- Fred Roberts, RU
- Mike Saks, RU
- Robert Sedgewick, PU
- Rebecca Wright, ATTL

Education Committee

- Bill Aiello, ATTL
- Adi Ben-Israel, RU
- Adam Buchsbaum, ATTL
- Giovanni DiCrescenzo, TC
- Sanjeev Khanna, BL
- Fred Roberts, RU
- Joseph Rosenstein, RU

- Michael Saks, RU
- Robert Sedgewick, PU

Graduate Student Selection Committee

- Sanjeev Arora, PU
- Endre Boros, RU
- Giovanni DiCrescenzo, TC
- Leonid Khachiyan, RU
- Sanjeev Khanna, BL
- Fred Roberts, RU
- Robert Sedgewick, PU
- Eduardo Sontag, RU
- Harold Stone, NEC
- Martin Strauss, ATTL

Visitor Committee

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

Computing Equipment Committee

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

Strategic Planning Committee

- Eric Allender, RU
- David Johnson, ATTL
- Fred Roberts, RU
- Joseph Rosenstein, RU
- Michael Saks, RU
- Robert Sedgewick, PU
- David Waltz, NEC

Budget Committee

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

DREI Steering Committee

- Chuck Biehl, Charter School of Wilmington
- Ethel Breuche, educational consultant, Pennsylvania
- Patrick Carney, Montclair Kimberley 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 S. Roberts, RU
- Joseph G. 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)
- Juris Hartmanis, Cornell University
- 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 Labs Research
- Dana Scott, Carnegie Mellon University
- Eva Tardos, Cornell University

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: "Language 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" 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," Editors: Doron A. Peled, Vaughan R. Pratt, and Gerard J. Holzmann.
Volume Thirty: "Parallel Algorithms: Third DIMACS Implementation Challenge," Editor: Sandeep N. Bhatt.
Volume Thirty One: "Descriptive Complexity and Finite Models," Editors: Neil Immerman and Phokion G. Kolaitis.
Volume Thirty Two: "The SPIN Verification System: The Second Workshop on the SPIN Verification System," 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.
Volume Forty Six: "Multichannel Optical Networks: Theory and Practice" Editors: Peng-Jun Wan, Ding-Zhu Du, Panos M. Pardalos.
Volume Forty Seven: "Mathematical Support for Molecular Biology" Editors: Martin Farach-Colton, Fred S. Roberts, Martin Vingron and Michael Waterman.
Volume Forty Eight: "DNA Based Computers III" Editors: Harvey Rubin and David Harlan Wood.
Volume Forty Nine: "Contemporary Trends in Discrete Mathematics: From DIMACS and DIMATIA to the Future" Editors: Ronald L. Graham, Kratochv, Jaroslav Nesetril and Fred S. Roberts.
Volume Fifty: "External Memory Algorithms" Editors: James M. Abello and Jeffrey Scott Vitter.

A.2 1999 Technical Reports

- **99-01** Stable Effectivity Functions and Perfect Graphs; Vladimir Gurvich and Endre Boros.
- **99-02** Scheduling to Minimize Average Stretch; Johannes E. Gehrke, S. Muthukrishnan, Rajmohan Rajaraman and Anthony Shaheen.
- **99-03** Massive Data Set Analysis in Seismic Explorations for Oil and Gas in Crystalline Basement Interval; Ilya Muchnik, Vadim Mottl and Vladimir Levyant.
- **99-04** Combinatorial Design of Multi-Ring Networks with Combined Routing and Flow Control; Avishai Wool and Bulent Yener.
- **99-05** A New Upper Bound on Codes Decodable into Size-2 Lists; A. Ashikhmin, A. Barg, and S. Litsyn.
- **99-06** A New Upper Bound on the Reliability Function of the Gaussian Channel; A. Ashikhmin, A. Barg, and S. Litsyn.
- **99-07** The Internet and Other Networks: Utilization Rates and Their Implications; Andrew Odlyzko.
- **99-08** The Economics of the Internet: Utility, Utilization, Pricing, and Quality of Service; Andrew Odlyzko.
- **99-09** Smart and Stupid Networks: Why the Internet Is Like Microsoft; Andrew Odlyzko.
- **99-10** Data Networks are Lightly Utilized, and Will Stay That Way; Andrew Odlyzko.
- **99-11** The Size and Growth Rate of the Internet; K. G. Coffman and Andrew Odlyzko.
- **99-12** Dynamic Behavior of Differential Pricing and Quality of Service Options for the Internet; Peter Fishburn and Andrew Odlyzko.
- **99-13** Optimal Orientations of Annular Networks; Jean-Claude Bermond, Johny Bond, Carole Martin, Aleksandar Pekec and Fred S. Roberts.
- **99-14** Symmetrization of Binary Random Variables; Abram Kagan, Colin L. Mallows, Larry A. Shepp, Robert J. Vanderbei and Yehuda Vardi.
- **99-15** Addition Chains as Test Trees and a Sequential Variant as the Huffman Problem; Julia Abrahams.
- **99-16** Characterizations of Consistent Marked Graphs; Fred S. Roberts and Shaoji Xu.
- **99-17** Note on Geometric Graphs, Geza Toth.
- **99-18** On Lower Bound of the Number of Channels for Deadlock-Free Whormhole Routing; Li Sheng and Jie Wu.
- **99-19** Deadlock-Free Routing in Irregular Networks Using Prefix Routing Algorithm; Jie Wu and Li Sheng.
- **99-20** Statistical Properties of Contact Maps; Michele Vendruscolo, Balakrishna Subramanian, Ido Kanter, Eytan Domany and Joel Lebowitz.

- **99-21** The Study of Three-body Interaction Hamiltonian on a Lattice; Balakrishna Subramanian and Joel Lebowitz.
- **99-22** Towards a Dynamic Optimal Alphabetic Tree and Cover Page; Ahmed A. Belal, Mohamed S. Selim and Shymaa M. Arafat.
- **99-23** On Some Tighter Inapproximability Results; Piotr Berman and Marek Karpinski.
- **99-24** Sequential Testing of Series-Parallel Systems of Small Depth; Endre Boros and Tonguc Unluyurt.
- **99-25** An Active Network Approach to Efficient Network Management; Danny Raz and Yuval Shavitt.
- **99-26** A Short Note on Linear Autarkies, q-Horn Formulas and the Complexity Index; Hans van Maaren.
- **99-27** Efficient Radio Wave Propagation for Wireless Radio Signal Prediction; Greg Perkins and Diane Souvaine.
- **99-28** On Parallel Edges in Cycles; E. Boros and V. Gurvich.
- **99-29** Camel Sequences and Their Applications; V. Gurvich and Li Sheng.
- **99-30** Fully Consistent Extensions of Partially Defined Boolean Functions with Missing Bits; Endre Boros, Toshihide Ibaraki and Kazuhisa Makino.
- **99-31** Recursive Generation of Partitionable Graphs; E. Boros, V. Gurvich and S. Hougardy.
- **99-32** Natural Language Understanding with the Generality Feedback; Boris Galitsky.
- **99-33** Millimix: Mixing in Small Batches; Markus Jakobsson and Ari Juels.
- **99-34** A Cost Upper Bound for Spanning Trees Connecting Faces of a Planar Subdivision; Greg Perkins and Diane Souvaine.
- **99-35** The Role Assignment Model Nearly Fits Most Social Networks; Aleksandar Pekec and Fred S. Roberts.
- **99-36** Randomized Complexity of Linear Arrangements and Polyhedra; Marek Karpinski.
- **99-37** New Bounds on Crossing Numbers; Janos Pach, Joel Spencer and Geza Toth.
- **99-38** Bifurcation of Internal Solitary Wave Modes from the Essential Spectrum; P.G. Kevrekidis and C.K.R.T. Jones.
- **99-39** A Combinatorial Algorithm for Pfaffians; Meena Mahajan, P. R. Subramanya and V. Vinay.
- **99-40** Camel Sequences and Quadratic Residues; V. Gurvich and Li Sheng.
- **99-41** Transformation of Regular Non-Dominated Coteries; Kazuhisa Makino and Tiko Kameda.
- **99-42** Modelling of Phase Separation in Alloys with Coherent Elastic Misfit; Peter Fratzl, Oliver Penrose and Joel L. Lebowitz.

- **99-43** A Minimax Game of Power Control in a Wireless Network under Incomplete Information; Tiina Heikkinen.
- **99-44** Packing Random Rectangles; E. G. Coffman, George S. Lueker, Joel Spencer, and Peter M. Winkler.
- **99-45** The Adjacency Matrix and Spectrum of a Graph with Negative Third Largest Eigenvalue; Paul A. Dreyer, Jr. and Xuerong Yong.
- **99-46** Universal H-colorable Graphs Without A Given Configuration; Paul A. Dreyer, Jr. , Christopher Malon, and Jaroslav Nesetril.
- **99-47** On the Maximum Number of Equilateral Triangles II; Bernard Abrego and Silvia Fernandez-Merchant.
- **99-48** Optimal Proof Systems and Sparse Sets; Harry Buhrman, Steve Fenner, Lance Fortnow and Dieter van Melkebeek.
- **99-49** On the Computational Power of Analog Neural Networks; Bhaskar DasGupta and Georg Schnitger.
- **99-50** Cluster-preserving Embedding of Proteins; Gabriela Hristescu and Martin Farach-Colton.
- **99-51** Dynamics of Lattice Kinks; P.G. Kevrekidis and M.I. Weinstein.
- **99-52** Improvements in Throughput Maximization for Real-Time Scheduling; Piotr Berman and Bhaskar DasGupta.
- **99-53** Ramsey-type Results for Unions of Comparability Graphs; Adrian Dumitrescu and Geza Toth.
- **99-54** On Polynomial Invariants of Codes, Matroids and Finite Interaction Models; Alexander Barg.
- **99-55** Apparent Motion in Spatial Time-varying Data: A Variational Approach to Pointwise Tracking of Coherent Structures in Computational Fluid Dynamics; Vadim Mottl, Alexander Blinov, Norman Zabusky and Ilya Muchnik.
- **99-56** An Approach of Information Theory to Multiple Sequence Comparison; Weiwu Fang, Fred S. Roberts and Zhengrong Ma.
- **99-57** An Optimal Algorithm for Hyperplane Depth in the Plane; Stefan Langerman and William Steiger.
- **99-58** Proof of a Conjecture of Bollobas and Eldridge for Graphs with Maximum Degree 3; Bela Csaba, Ali Shokoufandeh, Endre Szemeredi.
- **99-59** Note on the Work Function Algorithm; Bela Csaba.
- **99-60** Hardness of Approximation of the Loading Problem for Multi-layered Feedforward Neural Nets; Bhaskar DasGupta and Barbara Hammer.
- **99-61** A Geometric Model of Information Retrieval Systems; Myung Ho Kim.

- **99-62** Dual-Bounded Hypergraphs: Generating Partial and Multiple Transversals; E. Boros, V. Gurvich, L. Khachiyan and K. Makino.
- **99-63** On the Number of Vertices Belonging to all Maximum Stable Sets of a Graph by E. Boros, M.C. Golumbic and V.E. Levit.

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, The Johns Hopkins 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, Akamai Technologies
Research Areas: Network performance monitoring and management. Self-configuring systems for network security management. Models and algorithms for distributed computing and communications.

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, Telcordia Technologies (formerly Bellcore)
Research Areas: Telecommunications applications; interior point methods; stochastic optimization.

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

Chandra Chekuri, Bell Labs
Research Areas: Design and analysis of algorithms, with an emphasis on approximation, randomized, and online algorithms and their applications to various problems in combinatorial optimization, scheduling, databases, compilers, information retrieval and other applied areas.

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

Fan Chung, Univ. of California, San Diego
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, New Jersey Institute of Technology
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, Telcordia Technologies
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.

Giovanni DiCrescenzo, Telcordia Technologies
Research Areas: Cryptology and computational complexity, number theory, distributed computing and analysis of algorithms.

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.

Martin Eiger, Telcordia Technologies
Research Areas: Telecommunications network planning and engineering; optimization.

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.

Lance Jeremy Fortnow, NEC Research
Research Areas: Computational complexity theory; bounded rationality in game theory; cryptography; learning theory.

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.

Stefano Galli, Telcordia Technologies
Research Areas: Digital communications; channel equalization and coding; personal wireless communications; DSL technology.

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

Michael Garey, Bell Labs (retired)
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 (retired), Univ. of California, San Diego
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, Consultant
Research Areas: Combinatorics and graph theory.

Alan J. Hoffman, IBM Research
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, Telcordia Technologies
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: discrete mathematics, probability, geometry and computer science

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

Sampath Kannan, AT&T Labs - Research
Research Areas: Program reliability; algorithms and models for large data sets; wireless networking.

Narendra Karmarkar, Bell Labs
Research Areas: Linear programming.

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

Jon Kettenring, Telcordia Technologies
Research Areas: Multivariate statistics; cluster analysis; massive data problems.

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: Combinatorics and graph theory.

Keh-Wei Lih, AT&T Labs
Research Areas: Network design and optimization; mixed integer programming, algorithm design and development.

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

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

Zili Liu, Rutgers University
Research Areas: Visual perception: integrating human perceptual experiments (using computer graphics) and information-theoretic analysis, the "ideal observer" approach: using Bayesian statistics to determine what the best any visual system can do is, human or computer.

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.

Clyde Monma, Telcordia Technologies
Research Areas: Combinatorial optimization; applications to telecommunications.

S. (Muthu) Muthukrishnan, AT&T Labs - Research
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.

Craig Nevill-Manning, Rutgers University
Research Areas: Bioinformatics, inferring sequential structure, digital libraries, machine learning and data compression.

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, Telcordia Technologies
Research Areas: Cryptography; computer security; distributed protocols and algorithms.

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

Carl Pomerance, Bell Labs
Research Areas: Number theory and cryptography.

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

Sivaram Rajagopalan, Telcordia Technologies
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, Bell Labs
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.

Ronitt Rubinfeld, NEC Research
Research Areas: Program result checking; randomized algorithms; interactive proofs; computational learning theory.

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, Telcordia Technologies
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.

Aravind Srinivasan, Bell Labs
Research Areas: Telecommunications network design and routing; probabilistic methods in cs; randomized algorithms; approximation algorithms; combinatorial optimization; optimization and scheduling in information retrieval; scheduling theory and algorithms; distributed and parallel 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.

Alexander Stolyar, Bell Labs
Research Areas: Stochastic Processes; queues; high speed networks, wireless networks.

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; complexity; cryptography.

Mario Szegedy, Rutgers University
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.

Mikkel Thorup, AT&T Labs - Research
Research Areas: Combinatorial algorithms and data structures.

Robert J. Vanderbei, Princeton University
Research Areas: Optimization.

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

Suresh Venkatasubramanian, AT&T Labs - Research
Research Areas: theoretical computer science.

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.

Bulent Yener, Bell Labs
Research Areas: Data networks, multicasting; flow and access control protocols; routing; network design and optimization.

Peter Yianilos, Princeton University
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 1999 - January 2000

- **Tanya Alekseyevskaya (Gelfand)**; Rutgers University;
  Visit related to: General: Dates: 8/31/98 - 8/1/00
- **Noga Alon**; Tel-Aviv University;
  Visit related to: Computational Intractability: Dates: 10/6/99 - 10/18/99
- **Fred Annexstein**; University of Cincinnati;
  Visit related to: Networks: Dates: 1/1/2000 - 3/31/00
- **Reuven Bar-Yehuda**; Technion;
  Visit related to: General: Dates: 8/30/99 - 9/17/99
- **Richard Beigel**; University of Illinois at Chicago;
  Visit related to: Computational Intractability: Dates: 8/1/99 - 5/31/00
- **Iliana Bjorling-Sachs**; Lafayette College;
  Visit related to: General: Dates: 1/11/99 - 7/12/99
- **Maury Bramson**; University of Minnesota;
  Visit related to: General: Dates: 3/23/99 - 3/28/99
- **Lev Bregman**; Ben-Gurion University;
  Visit related to: Large Scale Discrete Optimization: Dates: 8/18/99 - 8/29/99
- **Herve Bronniman**; INRIA - Sophia-Antipolis;
  Visit related to: INRIA partnership: Dates: 10/1/98 - 12/31/99
- **Ondrej Cepek**; Charles University;
  Visit related to: Large Scale Discrete Optimization: Dates: 6/7/99 - 6/21/99
- **Tomas Chudlarsky**; Charles University;
  Visit related to: DIMATIA: Dates: 7/7/99 - 8/9/99
- **Italo Dejter**; University of Puerto Rico;
  Visit related to: General: Dates: 9/1/1999 - 1/31/00
- **Erik Demaine**; University of Waterloo;
  Visit related to: Networks: Dates: 1/23/2000 - 2/5/00
- **Guy Even**; Tel Aviv University;
  Visit related to: General: Dates: 8/1/99 - 8/31/99
- **Steve Freeland**; Cambridge University;
  Visit related to: DNA Computing: Dates: 11/30/98 - 4/30/99
- **Lixin Gao**; Smith College;
  Visit related to: Networks: Dates: 5/1/99 - 1/30/00
- **Fred Glover**; University of Colorado-Boulder;
  Visit related to: Large Scale Discrete Optimization: Dates: 4/28/99 - 5/1/99
- **Boris Goldengorin**; University of Groningen;
  Visit related to: Large Scale Discrete Optimization: Dates: 11/29/99 - 12/20/99

- **Mordecai Golin**; Hong Kong University of Science and Technology;
  Visit related to: Massive Data Sets: Dates: 9/1/98 - 6/18/99
- **Marty Golumbic**; Bar Ilan University;
  Visit related to: General/Networks: Dates: 8/31/99 - 9/9/99
- **Larry Gray**; University of Wisconsin;
  Visit related to: General: Dates: 6/1/99 - 6/5/99
- **Peter Hajnal**; University of Szeged;
  Visit related to: General: Dates: 5/11/99 - 6/6/99
- **Michael Huang**; Boston University;
  Visit related to: General: Dates: 1/3/99 - 1/9/99
- **Toshihide Ibaraki**; Kyoto University;
  Visit related to: Large Scale Discrete Optimization/Networks: Dates: 5/28/99 - 6/2/99
- **Marek Karpinski**; University of Bonn;
  Visit related to: Large Scale Discrete Optimization/Networks: Dates: 3/1/99 - 4/20/99
- **Meir Katchalski**; Technion;
  Visit related to: Discrete Probability: Dates: 12/9/98 - 3/28/99
- **James Kelly**; University of Colorado-Boulder;
  Visit related to: Large Scale Discrete Optimization: Dates: 4/28/99 - 4/30/99
- **Alexander Kelmans**; University of Puerto Rico;
  Visit related to: General: Dates: 8/1/99 - 1/31/00
- **Myung Ho Kim**; University of Auckland;
  Visit related to: General: Dates: 11/20/98 - 2/20/99
- **Suh-Ryung Kim**; Kyunghee University;
  Visit related to: General: Dates: 1/23/2000 - 1/29/00
- **Jan Kratochvil**; Charles University;
  Visit related to: DIMATIA Grant: Dates: 3/9/99 - 3/27/99
- **Charles Kurland**; Uppsala University;
  Visit related to: DNA Computing: Dates: 11/1/99 - 11/10/99
- **Manuel Laguna**; University of Colorado-Boulder;
  Visit related to: Large Scale Discrete Optimization: Dates: 4/28/99 - 4/30/99
- **Brenda Latka**; Lafayette College;
  Visit related to: General: Dates: 6/1/99 - 8/31/00
- **Heesang Lee**; Hankuk University of Foreign Studies;
  Visit related to: Networks and Large Scale Discrete Optimization: Dates: 8/1/99 - 1/31/00
- **Simon Litsyn**; Tel-Aviv University;
  Visit related to: Discrete Probability: Dates: 9/1/98 - 2/28/99
- **Meena Mahajan**; Institute of Mathematical Sciences;
  Visit related to: General: Dates: 5/24/99 - 8/8/99

- **Kazuhisa Makino**; Osaka University;
Visit related to: General: Dates: 4/22/99 - 6/22/99
- **Kirone Mallick**; Technion/Saclay;
Visit related to: General: Dates: 5/25/99 - 6/9/99
- **Jana Maxova**; Charles University;
Visit related to: DIMATIA partnership: Dates: 6/7/99 - 7/14/99
- **Angus McIntyre**; University of Edinburgh;
Visit related to: General: Dates: 9/10/99 - 10/10/99
- **Roy Meshulam**; Technion;
Visit related to: General: Dates: 9/1/99 - 9/30/99
- **Michel Minoux**; University Paris VI;
Visit related to: Large Scale Discrete Optimization/Networks: Dates: 5/5/99 - 5/14/99
- **Boris Mirkin**; German National Center of Cancer Research (DKFZ);
Visit related to: Molecular Biology: Dates: 9/5/99 - 11/5/99
- **Vadim Mottl'**; Tula State University;
Visit related to: Molecular Biology, Massive Data Sets: Dates: 12/16/98 - 2/27/99
- **Susana Nascimento**; University Nova de Lisboa;
Visit related to: Massive Data Sets, Molecular Biology: Dates: 9/16/99 - 10/16/99
- **Ashwin Nayak**; University of California at Berkeley;
Visit related to: Computational Intractability: Dates: 10/20/99 - 10/21/99
- **Arkadi Nemirovskii**; Technion;
Visit related to: Large Scale Discrete Optimization: Dates: 8/1/99 - 8/31/99
- **Jarik Nesetril**; Charles University;
Visit related to: DIMATIA Partnership: Dates: 9/19/99 - 10/1/99
- **Dmitrii Nogin**; Institute for Information Transmission Problems;
Visit related to: General: Dates: 11/1/99 - 11/28/99
- **Ramamohan (Mohan) Paturi**; University of California - San Diego;
Visit related to: Computational Intractability: Dates: 11/7/99 - 12/3/99
- **Oliver Penrose**; Heriot-Watt University;
Visit related to: General: Dates: 5/9/99 - 6/3/99
- **Gyorgy Petruska**; Eotvos University;
Visit related to: General: Dates: 5/3/99 - 6/27/99
- **Toni Pitassi**; University of Arizona;
Visit related to: Computational Intractability: Dates: 10/12/99 - 10/15/99
- **Yuval Rabani**; Technion;
Visit related to: General/Networks: Dates: 8/15/99 - 8/22/99
- **Jaikumar Radhakrishnan**; Tata Institute of Fundamental Research;
Visit related to: General: Dates: 5/11/99 - 6/4/99

- **Dieter Rautembach**; Lehrstuhl II fur Mathematik;
  Visit related to: Computational Intractability: Dates: 11/1/99 - 11/30/99
- **Ran Raz**; The Weizmann Institute of Science;
  Visit related to: Computational Intractability: Dates: 10/6/99 - 10/22/99
- **Alexander Razborov**; Steklov Mathematical Institute;
  Visit related to: Computational Intractability: Dates: 9/29/99 - 2/28/00
- **Bruce Reed**; University of Toronto;
  Visit related to: Computational Intractability: Dates: 11/1/99 - 11/30/99
- **Michael Roytberg**; Institute of Mathematical Problems of Biology of the Russian Academy of Sciences;
  Visit related to: General, Molecular Biology: Dates: 4/29/99 - 6/29/99, 11/15/99 - 12/11/99
- **Suleyman Cenk Sahinalp**; University of Warwick;
  Visit related to: Networks: Dates: 7/5/99 - 8/5/99, 1/1/00 - 1/31/00
- **Dimitriou Serpanos**; University of Crete;
  Visit related to: General: Dates: 9/1/99 - 1/31/00
- **Farhad Shahrokhi**; University of North Texas;
  Visit related to: General: Dates: 8/17/99 - 8/31/99
- **Amir Shpilka**; Hebrew University;
  Visit related to: General: Dates: 5/28/99 - 6/7/99
- **Bruno Simeone**; La Sapienza University;
  Visit related to: Large Scale Discrete Optimization: Dates: 1/1/99 - 6/30/99
- **Ewald Speckenmeyer**; Universitat zu Koln;
  Visit related to: Large Scale Discrete Optimization/Networks: Dates: 6/7/99 - 6/19/99
- **Tibor Szabo**; University of Illinois at Urbana-Champaign;
  Visit related to: General: Dates: 5/29/99 - 6/14/99
- **Gabor Tardos**; Eotvos Lorand Tudomanyegyetem;
  Visit related to: Computational Intractability: Dates: 10/23/99 - 11/19/99
- **Hans van Maaren**; Technical University of Delft;
  Visit related to: Large Scale Discrete Optimization: Dates: 4/19/99 - 4/29/99
- **Mark Watkins**; University of Georgia;
  Visit related to: General: Dates: 1/3/00 - 8/31/00
- **Xuerong Yong**; Xinjiang University;
  Visit related to: Massive Data Sets: Dates: 2/2/99 - 5/19/99
- **Gang Yu**; University of Georgia;
  Visit related to: General: Dates: 1/3/00 - 8/31/00
- **David Zuckerman**; University of California, Berkeley and the University of Texas at Austin;
  Visit related to: Computational Intractability: Dates: 10/12/99 - 10/22/99

C Awards and Honors

- Israel Gelfand (PM, RU): Chosen by Board of Governors of the New York Academy of Sciences as an Honorary Life Member. Election to honorary life membership is a distinction conferred on a limited number of scientists who made outstanding contributions to the advancement of science – only two are elected each year.
- Phil Gibbons (PM, BL): Yossi Levi Prize in Operations Research and Managerial Economics from Operations Research Society of Israel.
- C. Lee Giles (PM, NEC): New Jersey Internet Innovator Award.
- Jerry Goldin (PM, RU): Humboldt Research Prize from the Alexander von Humboldt Foundation.
- Peter Hammer (PM, RU): Euler Medal of the Institute of Combinatorics and its Applications.
- Joel Lebowitz (PM, RU): (1). AAAS Scientific Freedom and Responsibility Award; (2). Doctor Honoris Causa, Clark University.
- Michael Malisoff (GS, RU): Best Student Paper Award from IEEE Decision and Control Conference, December 1999.
- Carl Pomerance (PM, BL): Hedrick Lecturer, Mathematical Association of America.
- Fred Roberts (PM, RU): Distinguished Service Award of the Special Interest Group in Algorithms and Computing Theory (SIGACT) of the Association for Computing Machinery (ACM)
- Larry Shepp (PM, RU): Miller Fellow at Berkeley
- Peter Shor (PM, ATTL): (1). Dickson Prize in Science from Carnegie-Mellon University; (2). Godel Prize of European Association for Theoretical Computer Science (EATCS) and the Special Interest Group in Algorithms and Computing Theory (SIGACT) of the Association for Computing Machinery (ACM); (3). MacArthur Award
- Neil Sloane (PM, ATTL): AT&T Fellow
- Ken Steiglitz (PM, PU): IBM Partnership Award
- Harold Stone (PM, NEC): IEEE Computer Society Taylor Booth Award for contributions to education.
- David Waltz (PM, NEC): Elected a Fellow of the Association for Computing Machinery