The Second DIMACS International Algorithm Implementation Challenge: General Information

September 28, 1992

This document provides a general overview of the Challenge and gives information about where to find more material.

1 Introduction

In conjunction with its Special Year on Combinatorial Optimization, the Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) invites participation in an international Implementation Challenge to find effective optimization and approximation algorithms for the following NP-hard problems: **Maximum Clique**, **Graph Coloring**, and **Satisfiability**. The Implementation Challenge will take place between September 1992 and August 1993. Participants are invited to carry out research projects related to these problem areas and to present research papers at a DIMACS workshop to be held in Fall of 1993. A refereed workshop proceedings will be published by the AMS.

The purpose of this Challenge is to encourage high quality empirical research on difficult problems. The problems chosen are known to be difficult to solve in theory. How difficult are they to solve in practice? How can we best measure the practical difficulty of solution? If optimal solutions are unobtainable in practice, how close to optimal can we get in reasonable amounts of time? Do artificial problem generators generate instances with characteristics similar to those of real–world instances? What approaches are most robust for a wide variety of instances? How can special structure

in instances be exploited? These are some of the many questions that might be addressed within the Challenge.

The Challenge is intended to create several tangible benefits. One will be the refereed proceedings, that should provide a state–of–the–art overview of the field. Another will be a database of test instances, comparable to the traveling salesman problem testbed, TSPLIB. Other possible products include solution codes, problems generators, annotated bibliographies, and other items that are submitted by participants. Many of these will be accessible both during and after the Challenge through anonymous ftp and other electronic mail based methods.

The success of this Challenge will be based on the efforts of the participants. DIMACS's role will simply be to provide general guidance and coordination, to be a clearinghouse for instances, codes and other material, and to sponsor the eventual Workshop and Proceedings.

2 Description of Problems

Separate documents are, or will shortly be, available giving an overview and annotated bibliography for each of the three problems that are the subject of the Challenge. This section offers a brief description of each problem.

Maximum Clique Given an undirected graph, a *clique* of of the graph is a set of pairwise adjacent nodes. A *maximum clique* of a graph is a clique such that no other clique has more vertices. There are also weighted versions of the clique problem where either the nodes or the edges have weights and the objective is to find a clique that maximizes the total weight of the nodes in the clique or the edges connecting nodes of the clique. This problem is is equivalent to the *Independent Set Problem* on the complement of a graph, where an independent set is a set of pairwise non–adjacent nodes.

Graph Coloring Given a an undirected graph, a *coloring* of its nodes is an assignment label to each node such that adjacent nodes have different labels. A *minimum coloring* is a coloring that uses the fewest distinct labels. This value is called the chromatic number of a graph.

Satisfiability Satisifiability comes in many forms. Most generally, a satisfiability problem consists of a number of boolean variables x_i contained in some logical formula F. The problem is to determine if there is an assignment of values to the variables so that the formula evaluates to TRUE.

Often the formula is expected to be in a particular form. The most common form is *conjunctive normal form (CNF)*, where the formula consists of a number of *clauses*, each consisting of the conjunction of a number of variables or their negation. The formula is TRUE if and only if each clause evaluates to TRUE. This form is general, in the sense that every formula can be written in such a manner, though the form may be exponentially sized compared to alternative forms.

In addition to the pure satisfiability problem, there is also a optimality version of the CNF Satisfiability problem (find a truth assignment that satisfies the maximum possible number of clauses).

3 Participation

There are several ways to participate in the Challenge. In fact, almost any aspect of empirical research related to the three problems addressed would be appropriate for the Challenge. Below are a few possible directions.

1. Implement and evaluate algorithms. Participants may implement one or more algorithms (or perhaps alternative data structures or strategies for a single algorithm) for experimental evaluation.

DIMACS and the participants of the Challenge will determine a benchmark set of inputs for each problem; this will facilitate comparisons between approaches. Participants are invited to perform more extensive tests as desired. Guidelines and support tools for experimental study are also available from DIMACS.

These algorithms need not try to solve the instances to optimality. We are interested in heuristics and relaxations as well as optimal solution finding methods. We are also interested in determining the effectiveness of such general purpose heuristic techniques as simulated annealing, genetic algorithms, and tabu search applied towards these problems.

2. Create test inputs and input generators. Participants are also invited to submit test inputs or input generators for these problems.

- We invite generators of "random" inputs having some quantifiable combinatorial structure and which suggest a useful model of average-case performance.
- Also of interest are generators of inputs that are arguably "hard" for certain algorithms (causing loss of efficiency), and which can be used to distinguish algorithmic robustness with respect to input.
- Input instances from real applications are particularly encouraged. One natural issue is the relationship between artificially generated instances and real–world instances. Do algorithms work similarly on the two or are there differences?

Authors of test generators should be able to describe the kinds of instances generated and to argue why they are interesting. They should perform at least limited tests to determine whether their generators yield different answers than the standard benchmarks. With author's permission, submitted inputs and generators will be made available to Challenge participants; some may be included in the set of DIMACS benchmarks.

Generated inputs are expected to follow a standard format; see Section 5 for further information.

3. Test on an unusual architecture or environment. Participants with access to a machine not widely available are invited to adapt algorithms (or existing programs) for testing on that architecture. Implementations tuned for a particular architecture (Cray, Connection Machine, etc.) are welcome.

4. Exploit Special Structure One very active research area is in characterizing special cases of all these problems that can be solved effectively. The goal of this Challenge, however, is to examine the solution of general instances. As such, mechanisms for solving structured problems should concentrate either on addressing particularly well–motivated special cases or on using the special structure to attack the more general case. The latter might include recognizing that special structure

"often" occurs, using the special structure in a subroutine for the general problem, or other approaches.

4 Challenge Goals, Rules, and Specifications

A committee of DIMACS members will provide general direction for the Implementation Challenge. Committee members are Vašek Chvátal, Rutgers University, Bill Cook, Bell Communications Research, David Johnson, AT&T Bell Laboratories, Cathy McGeoch, Amherst College, Bob Tarjan, Princeton University, and Michael Trick, DIMACS Visiting Fellow/Carnegie Mellon University (Coordinator).

DIMACS can provide neither financial support for research projects nor machine cycles for the experiments, although some limited confirmatory tests may be carried out at DIMACS. The intent of the Challenge is to use DIMACS facilities (and the Internet), to provide a catalyst for experimental research on these problems, to promote communication among researchers in these areas, and to provide a clearing-house for exchange of implementations and input generators.

Important Dates. The Challenge begins September 1992 and continues through August 1993. The only firm deadline in the Challenge is **August 1, 1993**, by which time extended abstracts must be submitted for consideration for the workshop. Later deadlines will be set for submission of the full paper for the refereed proceedings.

The goal of this Challenge, however, is to use the period from September 1, 1992 through August 1, 1993 to provide feedback and interaction for ongoing research. It is during this period that the committee will be reviewing short abstracts and suggesting alternative approaches. It is also during this period that researchers will be able to coordinate and interact with others doing similar work. We therefore suggest that prospective participants do the following:

1. Register with the Challenge as soon as possible by following the instructions in Section 5. This will put the participant on a mailing list. This list will announce available documents as well as provide a forum about ongoing research.

2. Submit a two or three page abstract by **January 15, 1993** which will be reviewed by the committee.

3. Submit a progress report by **April 1, 1993** which will also be reviewed by the committee.

4. Submit an (up to ten page) extended abstract by **August 1, 1993** from which the workshop presenters will be chosen by the Committee.

Project proposals. At each review step, the committee will examine each project to identify duplicated research efforts and to suggest modifications as may seem appropriate.

Approximately one month before the workshop, the set of accepted extended abstracts in each problem domain will be circulated to those with accepted presentations. This will allow participants to prepare their workshop presentations with the related papers already in mind and will encourage further communication between participants. Shortly after the workshop, a DIMACS technical report will be prepared that consists of the (possibly revised) extended abstracts. Full papers will be due one or two months after the workshop for consideration in the refereed proceedings.

Support code, program development and testing tools, input generators, and input instances may be submitted and are welcome at any time. See Section 5 for submission information. General announcements of new files and test inputs available from DIMACS will be made periodically.

5 Accessing Other Documents

There are a number of other documents that are ready, or will be ready soon. These include

1. Definition of input formats. To achieve compatibility of input generators and test implementations, a standard input format will be defined for each problem. Translation programs for converting to and from other standard representations will be distributed as they become available.

2. Annotated bibliography and suggested research directions. Each of the three problems will have a separate review of the literature and suggestions for future research.

Announcement of these documents will be made to the mailing list. There are a number of ways to access these and any other documents.

Anonymous FTP. The easiest way to access Challenge documents is through anonymous FTP to DIMACS. To use this facility type the following command sequence. Human-generated commands appear in `typewriter` font; `name@address` refers to your full email address.

```
% ftp dimacs.rutgers.edu
Connected to dimacs.rutgers.edu.
220 dimacs.rutgers.edu FTP server (Version 6.15 Fri Apr 10 01:00:01 EDT
    1992) ready.
Name (local.system.name): anonymous
331 Guest login ok, send e-mail address as password.
Password: name@address
230 Guest login ok, access restrictions apply.
ftp> cd pub/challenge
250 CWD command successful.
```

Typing a `?` at the `ftp>` prompt will produce a list of commands. Typing `help commandname` will give a very short description of what the command does. Typing `remotehelp commandname` gives a description of command format. For further information, contact either your local support group or send mail to the Challenge address `challenge@dimacs.rutgers.edu` and we will send more detailed information.

The main directory contains several files of general interest. The file INDEX list the available files and describes the various subdirectories.

Electronic Mail. Electronic mail should be used to be placed on the mailing list. In addition, some participants may not have capabilities for anonymous FTP. In this case files and related information can be obtained by electronic mail. The following addresses may be used.

challenge@dimacs.rutgers.edu Requests to be put on the mailing list, project proposals, general questions and comments should be sent to this address.

challenge-submit@dimacs.rutgers.edu Submitted programs and data files may be mailed to this address. If a submission consists of more than one file, please bundle all files into a single mail message, preferably using the commonly available "shar" utility. Please also send a message to `challenge@dimacs.rutgers.edu` with a short description of your submission.

challenge-server@dimacs.rutgers.edu This is an automatic mail server, from which files can be requested. For help using it, send a message to it with the single line "help".