Combinatorial Algorithms Test Sets (CATS): The ACM/EATCS Platform for Experimental Research

Andrew V. Goldberg* Bernard M.E. Moret†

Abstract

We describe **CATS**, a recently started project, endorsed by the ACM and the EATCS, to develop a Web-based collection of test sets. These test sets, each specialized for a specific problem (such as network flow or satisfiability), are maintained by volunteers (associate editors), using contributions from researchers (authors). The purpose is to facilitate experimental research by standardizing common benchmarks, providing a mechanism for their evolution, and making them easily accessible and usable; and to identify significant open questions in the design of good test sets and the evaluation of performance of existing algorithms. The test sets should also facilitate algorithm selection for applications by characterizing subproblems and the behavior of competitive algorithms on these subproblems and encourage the development of high-quality implementations of advanced algorithms and data structures.

1 Background and Motivation

Experimental work is a necessary, and a growing, part of the process of algorithm development. As recently as fifteen years ago, however, experimental research was rare and not well regarded within the algorithm community. The lack of experimental work eventually resulted in the creation of many "paper layers" in the description of a new algorithm—layers of unimplemented previous results upon which the new result relied—and in the opening of a large gap between the research work in academic circles and the tools actually used in industry.

Over the last ten years, we have seen a return to the idea of testing theoretical ideas by implementing them, along with a welcome narrowing in the gap between academic research and industrial practice. What remains lacking in many areas, however, is a well-developed experimental methodology, comparable to the methodologies developed in the natural sciences. As a major step in that direction, the algorithm community needs an organized collection of data sets, both synthetic and derived from applications, with which to test its ideas. (Consider the speech recognition area: in the years since ARPA set up a research group whose main function has been to produce test sets and analyze and validate results from other groups, the area has seen very significant progress, following upon a prolonged period of stagnation.)

A well-organized collection of data sets stimulates research by providing a basis for assessment, by ensuring reproducibility, and by encouraging industry to provide real-world data sets. The resulting codes have well-understood practical performance and thus readily transfer to applications, in contrast to theoretical results of unknown practical significance. Such a collection greatly facilitates the choice of a particular algorithm or implementation for a specific application. This evolution will narrow the gap between theory and practice, in the process uncovering further algorithmic problems of interest in both applied and theoretical areas.

Such a collection must include both real-world and synthetic data: the former show what can be done in the context of an application and the latter enable researchers to assess detailed characteristics of the algorithms. Such a collection cannot be static: as hardware, algorithms, and applications evolve, typical or demanding problem instances will change. (This evolution is the reason why static collections, such as *The Stanford GraphBase*, while remaining good starting points, cannot fulfill the roles described above.)

The collection should also include pointers to papers describing relevant experimental work and to state-of-the-art codes.

2 Aims

CATS is a Web-based platform for experimental research in combinatorial optimization and discrete algorithms that provides data sets to form a collection of core experiments.

The primary aim of **CATS** it to facilitate experimental research by standardizing common benchmarks, providing a mechanism for their evolution, and making

*NEC Research Institute, 4 Independence Way, Princeton, NJ 08540, `avg@research.nj.nec.com`

†Department of Computer Science, University of New Mexico, Albuquerque, NM 87131, `moret@cs.unm.edu`

them easily accessible and usable. Because its maintainers are active researchers, they are able to identify significant open problems, whether in the area of test generation or of algorithmic design, and thus are able to suggest fruitful new problems.

Two other aims of **CATS** are to facilitate algorithm selection for applications and to encourage the development of high-quality implementations of advanced algorithms and data structures. Selection of an algorithm for a specific application has been a major obstacle to the use of research results in industry: faced with a large number of theoretical results, most couched in asymptotic terms, the application programmer has little guidance on how to choose the right algorithm to implement. By enabling each researcher to provide clearly documented behavior on a variety of test data, **CATS** will encourage the development of algorithms that are effective in practice and attractive to the application programmer.

3 Scope

CATS focuses on data sets to be used in testing combinatorial optimization and discrete algorithms. Thus it complements classical efforts in algorithmic research, along with efforts aimed at developing libraries of basic data structures and algorithms for use in combinatorial optimization (such as LEDA). **CATS** sites will point to other Web sites containing relevant theoretical results, implementation, and information about the current state of library development; we hope that these sites will in turn point to the **CATS** homepage.

CATS focuses initially on a few specific areas within combinatorial optimization, in particular: network flow, matching, shortest paths, minimum spanning trees, graph coloring, satisfiability, and the travelling salesperson problem, along with problems in computational geometry. Additional areas may be added at a later time, as long as such areas meet some basic criteria: (i) an initial collection of test data, used in quality research papers; (ii) a connection to some application problems that will allow for the inclusion of real-world data; and (iii) instances characterized by discrete data along with objectives characterized by combinatorial features.

4 Structure

CATS is composed of a collection of Web sites, each devoted to one specific problem. The sites are unified through a common "look-and-feel," so that researchers familiar with one site will find it easy to use the resources provided by another. One (or more) root site(s) points to all of the individual sites and includes search facilities; pointers to **CATS** (including pointers from the ACM and the EATCS) link to one of these root sites. The root site also includes links to papers on experimental methodology, to implementation tools, and to algorithm libraries.

Each site contains an organized collection of "core experiments"—data sets that, in the opinion of researchers, provide significant information about the characteristics of solution algorithms. Along with this core collection is a history of its evolution, including pointers to deprecated data sets. When experiments in the core collection are parameterized, suggested parameter values are included. All data sets and generators are freely available, with no restriction on copying and re-use. Each site includes a complete description of the format(s) used.

Each site also includes pointers to new data sets (proposed for inclusion in the core), recent articles of importance, and existing codes (public or commercial). When pertinent (e.g., when the best codes are commercial and the data is unpublished), a site includes benchmark results on most of the data sets in the core.

Each site is maintained by a volunteer, who is an active researcher in the area(s) represented at that site. Sites are set up using an existing pattern (see `www.jea.acm.org/CATS/`), so as to ease the task of volunteers and offer a consistent presentation. Maintenance involves adding references to significant new work, including new data sets of proven usefulness, removing deprecated data sets from the core collection, and injecting whatever expert domain knowledge (e.g., in characterizing subproblems) can prove useful.

5 Steering and Credits

The group of all site maintainers collectively forms the "Editorial Board" of **CATS**; the group periodically meets (electronically or otherwise) to review the state of **CATS** and discuss new initiatives. Site maintainers thus fulfill a role similar to that of an associate editor of a journal and should be given the same amount of credit for service to their profession. Similarly, authors who provide new generators or interesting data sets are credited formally by **CATS**, with the intent that such contributions be given the same type of credit as refereed publications by their academic peers.

6 History

CATS was endorsed by the ACM and the EATCS in early Summer 1998. The project was first announced (and its current URL, `www.jea.acm.org/CATS/`, made public) in August 1998 at the *Second Workshop on Algorithm Engineering* held in Saarbrücken (Germany).