CSPLib: a status report

Toby Walsh
APES research group
University of York
United Kingdom
tw@cs.york.ac.uk

Over the last two or so years, I have been collecting together a problem library for the constraint satisfaction and optimization community (see www.csplib.org). There were many motivations for developing such a library. One of the main motivations was to provide a common set of problems on which different groups can quickly benchmark their algorithms. However, there were other motivations including focusing the research community away from random problems onto more realistic problems, and focusing the research community onto modelling. More recently, I have taken on the job of Benchmark Editor in Chief for the Constraints Journal. It is our hope that problems submitted to the Benchmark corner of this journal will be archived electronically in CSPLib.

Other areas have benchmark libraries so it seemed about time that constraints did too. In theorem proving, for example, the TPTP (Thousands of Problems for Theorem Provers) library was started by Suttner and Sutcliffe in 1993, and has since proved highly successful and influential. It is a valued resource in the theorem proving community, and there is an annual system competition based on problems drawn at random from TPTP. The library contains over four thousand different problems, in 28 different domains, as well as comprehensive references on their source and a variety of problem generators. A software tool converts the problems into input formats used by the major theorem proving systems. In addition, the library provides guidelines on how to evaluate theorem proving systems.

When we launched CSPLib, we realized that for it to be successful, the research community must become active users of the library and contributors to it. If CSPLIB is not used and contributed to, the time spent in setting it up will have been wasted. However this is one aspect of the library beyond the control of the maintainers. We hoped that researchers in constraints will come to view CSPLIB as a one stop shop for finding constraints benchmarks, and disseminating new benchmarks. So how have we done? It appears that the CSPLib library has been of some use to community with other 5000 different visitors in its first year or so of existence. However, it has not been as successful as I would have liked. In particular, the number of people contributing problems has been very modest. Most of the problems in the library have come from members of my immediate research group. Why is this so? It could be that CSPLib is not a good idea. I would hope this is not the case, so what other reasons could there be for a low submission rate?

First, submission could be too difficult. I do not believe this is the case as we will take a plain text problem description and turn it into a HTML entry. We could provide a web form but this is no easier than emailing csplib@cs.strath.ac.uk? Second, submission could offer too little reward. We could increase this reward by seeking suitable supporting funds and offering renumeration. We could also offer non-monetary rewards (perhaps CSPLib t-shirts: "I contributed. Have you?", or a CSPLib dinner at the CP conference, ...). I welcome any other suggestions you might have. Third, people might have problems that they could submit but are unwilling to for reasons of confidentiality. It is hard to know how to tackle this problem other than to offer to anonymize data. Again I welcome any other suggestions you might have for tackling this problem.

I have found CSPLib useful for my own research. A number of my recent papers have used problems that are from CSPLib. I hope others will find useful problems there. And if you come across any other problems, do send them in!