A Benchmark Framework for a Computational Argumentation Competition

Federico CERUTTI [a], Nir OREN [a], Hannes STRASS [b], Matthias THIMM [c,1], and Mauro VALLATI [d]

[a] *Department of Computing Science, University of Aberdeen, UK*
[b] *Computer Science Institute, Leipzig University, Germany*
[c] *Institute for Web Science and Technologies, University of Koblenz-Landau, Germany*
[d] *School of Computing and Engineering, University of Huddersfield, UK*

Abstract. We introduce probo, a general benchmark framework for comparing abstract argumentation solvers. probo is intended to act as the core of an argumentation competition intended to run in 2015.

Keywords. Argumentation, argumentation semantics, argumentation competition

1. Introduction

Computational models of argumentation are an active research discipline within Artificial Intelligence, having grown since beginning of the 90s [1]. While still quite young when compared to research areas such as SAT solving and Logic Programming, there has been serious interest in the application of argumentation due to its intuitive approach and similarities to human reasoning. However, most computational argumentation research is still theoretical in nature, concerning itself with aspects such as abstract argumentation, semantics, structured argumentation, and dialogues, to name just a few. There are, however, already several implemented domain-independent argumentation tools such as ASPARTIX, CEGARTIX, ConArg, Dung-O-Matic, and Tweety, aimed at identifying extensions of abstract argument systems. With the exception of [2], no organized and thorough comparison of these solvers has been performed. While the need for a practical evaluation of argumentation research via a systematic approach to benchmarking has been recognized before [3], we take this idea one step further and propose an organized competition on computational models of argumentation. Other research communities such as SAT solving, Logic Programming, and Planning have shown that conducting competitions on problem solvers nurtures research and rapid development of effective algorithms and implementations. These competitions serve as a comparison platform for state-of-the-art research and have been proven to accelerate the evolution of a field from theoretical discussions to the development of applicable solutions. We argue that the time has come for a first competition on computational models of argumentation.

[1] Corresponding author

In this extended abstract we present probo (lat. "judge" or "to show"), a general benchmark framework for comparing solvers of abstract argumentation problems. This benchmark framework has been developed to serve as the comparison platform for a competition on computational models of argumentation which is currently being organized and scheduled for the year 2015. The aim of this framework is to easily compare different implementations for solving argumentation problems in terms of *correctness* and *performance*. The framework provides a very general interface (command line and Java) that can be implemented by a solver in order to be assessed by the framework. Currently, probo focuses on abstract argumentation but extensions to structured argumentation frameworks is foreseen for the future.

2. Implementation Overview

probo is written in Java and publicly available at SourceForge.[2] It provides a well-documented and stable interface that can be used by developers of solvers for argumentation problems. Currently, probo supports comparison of solvers for different computational problems (e. g. enumerating extensions, deciding skeptical/credulous inference), with a variety of semantics (conflict-free, admissible, complete, preferred, stable, semi-stable, grounded, cf2, ideal, naive, stage), and different file formats (e. g. trivial graph format, ASPARTIX). Solvers are asked to solve problems on different argumentation graphs that have been either artificially generated using different randomization models, generated from other more general problems (e. g. graphs that encode search problems), or application-oriented problem instances. Correctness of solutions is verified by the Tweety solver [4], which directly implements the semantical definitions of the different types of extensions but is slow in practice.

3. Discussion

Argumentation is a highly promising field in Artificial Intelligence and we believe that it is high time to go beyond theoretical discussions and into implementations and applications. With our contribution and the idea of a competition on computational argumentation we aim at providing a common platform for comparing and evaluating practical approaches within the field.

While probo is still under development, we aim to run a first competition on computational models of argumentation in 2015. A description of the current command line interface used by probo can been found in our repository. We invite developers of abstract argumentation solvers to comment on the framework and prepare for the competition.

References

[1] T. J. M. Bench-Capon and Paul E. Dunne. Argumentation in Artificial Intelligence. *Artificial Intelligence*, 171:619–641, 2007.

[2] S. Bistarelli, F. Rossi, and F. Santini. A First Comparison of Abstract Argumentation Systems: A Computational Perspective. In *Proceedings CILC'13*, 2013.

[3] W. Dvořák, S.A. Gaggl, S. Szeider, and S. Woltran. Benchmark libraries for argumentation. In Sascha Ossowski, editor, *Agreement Technologies*, pages 389–393. Springer, 2013.

[4] M. Thimm. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In *Proceedings KR'14*, 2014.

[2] `https://sourceforge.net/projects/probo/`