The ICON Challenge on Algorithm Selection

Lars Kotthoff, Barry Hurley, Barry O'Sullivan

Algorithm selection is of increasing practical relevance in a variety applications. Many approaches have been proposed in the literature, but their evaluations are often not comparable, making it hard to judge which approaches work best. The ICON Challenge on Algorithm Selection objectively evaluated many prominent approaches from the literature, making them directly comparable for the first time. The results show that there is still room for improvement, even for the very best approaches.

In many areas of AI decades of research have resulted in many different approaches to solving similar problems. These different approaches exhibit different performance characteristics on different problem types, and it is usually unclear when to choose which approach. This is known as the algorithm selection problem (Rice 1976).

The challenge of algorithm selection in practice has led to the development of various data-driven, automated solutions to it. Machine learning techniques are used to transparently select the most appropriate algorithm for the problem at hand (O'Mahony et al. 2008; Hurley et al. 2014; Xu et al. 2008). Such systems have demonstrated that significant performance improvements can be achieved over using just a single approach. The interested reader is referred to a recent survey for more information (Kotthoff 2014).

There are many different approaches to solving the algorithm selection problem. While the majority of these have been evaluated empirically in the literature, such evaluations often use different data sets, different performance measures, and different experimental setups. The results are not directly comparable and do not provide a clear picture of the state of the art.

The ICON Challenge on Algorithm Selection provided the first comprehensive, objective evaluation of several state-of-the-art approaches. Its results gave an overview of the state of the art at the time of the competition, highlighting the strengths and weaknesses of different approaches.

Challenge Setting

The challenge leveraged the ASlib benchmark library for algorithm selection (Bischl et al. 2016). We used thirteen scenarios, drawn from prominent publications, in release 1.0. They represent a number of important AI application areas, including SAT, CSP, QBF, ASP, and heuristic search.

Challenge participants were required to output a schedule of algorithms to run for each problem instance in a scenario. They were allowed to specify: (a) a list of scenarios to run on; (b) the problem features they wanted their submission to have access to (feature computation incurs a cost that may reduce overall performance; and (c) an algorithm to run as presolver for a small amount of time, thereby reducing the overhead of feature computation and the selection process on easy instances. Since the ASlib dataset is public and was available to contestants before they submitted their system, the submissions were trained on ten different bootstrap samples of a scenario and evaluated on the remaining data.

All submissions were required to provide the full source code, with instructions on how to run the system. The submissions, full details of the challenge, along with detailed results and all data used in the evaluation, are available at http://challenge.icon-fet.eu/.

Results and Discussion

The challenge received a total of eight submissions from four different groups of researchers comprising fifteen people. Participants were based in four different countries on two continents. In alphabetical order, the submitted systems were ASAP_kNN, ASAP_RF, autofolio, flexfolio-schedules, sunny, sunny-presolv, zilla, and zillafolio.

The overall winner of the ICON challenge was zilla, based on the prominent SATzilla (Xu et al. 2008) system. The ASAP_RF system received an honourable

Table 1: Submission ranking.

	System	Total score
1	zilla	0.36603
2	zillafolio	0.37021
3	autofolio	0.39083
4	ASAP_RF	0.41603
5	ASAP_kNN	0.42318
6	flexfolio-schedules	0.44251
7	sunny	0.48259
8	sunny-presolv	0.48488

mention as a new system that had not been described in the literature before and outperformed all other submissions on some of the ASlib scenarios.

Table 1 shows the final ranking. Scores were aggregated over all scenarios and samples, with 0 corresponding to perfect predictions where on each problem instance the optimal algorithm is chosen (oracle) and 1 corresponding to a static predictor that chooses the overall best algorithm on each problem instance (single best).

All submissions achieve significant performance improvements over always choosing a single algorithm. The scores of the top-ranked approaches are very close, and in practice all of them will likely achieve good performance.

Nevertheless, there is scope for improvement. Even the best approach is still far away from being a perfect predictor. Especially the industrial SAT scenarios turned out to be challenging, with many systems not even achieving the performance of the single best solver. Detailed results are presented in (Kotthoff 2015).

Acknowledgements

We would like to thank all the participants for taking the time to prepare submissions and their help in getting them to run; in alphabetical order: Alex Fréchette, Chris Cameron, David Bergdoll, Fabio Biselli, François Gonard, Frank Hutter, Holger Hoos, Jacopo Mauro, Kevin Leyton-Brown, Marc Schoenauer, Marius Lindauer, Michèle Sebag, Roberto Amadini, Tong Liu, and Torsten Schaub. We thank Barry Hurley for setting up and maintaining the submission website and Luc De Raedt, Siegfried Nijssen, Benjamin Negrevergne, Behrouz Babaki, Bernd Bischl, and Marius Lindauer for feedback on the design of the challenge.

This work was in part supported by EU FP7 FET project 284715 (ICON). We thank Microsoft Azure and the BETA lab at the University of British Columbia for the provision of computational resources.

About the authors

Lars Kotthoff is a postdoctoral researcher at the University of British Columbia, Vancouver, Canada, where he works on automated algorithm configuration and selection, and uses optimisation techniques for data mining and machine learning. He previously held a postdoctoral appointment at Insight Centre for Data Analytics, Cork, Ireland, and completed his PhD work at the University of St Andrews, Scotland.

Barry Hurley is a Ph.D. student in the Insight Centre for Data Analytics at University College Cork. His research considers the exploitation of machine learning for solving combinatorial decision and optimisation problems.

Barry O'Sullivan serves as Director of the Insight Centre for Data Analytics and Professor in the Department of Computer Science at University College Cork. Professor O'Sullivan was elected a Fellow of ECCAI, the European Coordinating Committee for Artificial Intelligence (ECCAI), and a Senior Member of AAAI, the Association for the Advancement of Artificial Intelligence, in 2012.

References

[Bischl et al. 2016] Bischl, B.; Kerschke, P.; Kotthoff, L.; Lindauer, M.; Malitsky, Y.; Frechtte, A.; Hoos, H. H.; Hutter, F.; Leyton-Brown, K.; Tierney, K.; and Vanschoren, J. 2016. ASlib: A Benchmark Library for Algorithm Selection. *Artificial Intelligence Journal (AIJ)*. In press.

[Hurley et al. 2014] Hurley, B.; Kotthoff, L.; Malitsky, Y.; and O'Sullivan, B. 2014. Proteus: A hierarchical portfolio of solvers and transformations. In Simonis, H., ed., *Integration of AI and OR Techniques in Constraint Programming - 11th International Conference, CPAIOR 2014, Cork, Ireland, May 19-23, 2014. Proceedings*, volume 8451 of *Lecture Notes in Computer Science*, 301--317. Springer.

[Kotthoff 2014] Kotthoff, L. 2014. Algorithm selection for combinatorial search problems: A survey. *AI Magazine* 35(3):48--60.

[Kotthoff 2015] Kotthoff, L. 2015. ICON challenge on algorithm selection. *CoRR* abs/1511.04326.

[O'Mahony et al. 2008] O'Mahony, E.; Hebrard, E.; Holland, A.; Nugent, C.; and O'Sullivan, B. 2008. Using case-based reasoning in an algorithm portfolio for constraint solving. In *Proceedings of the 19th Irish Conference on Artificial Intelligence and Cognitive Science*.

[Rice 1976] Rice, J. R. 1976. The algorithm selection problem. *Advances in Computers* 15:65--118.

[Xu et al. 2008] Xu, L.; Hutter, F.; Hoos, H. H.; and Leyton-Brown, K. 2008. SATzilla: Portfolio-based Algorithm Selection for SAT. *J. Artif. Intell. Res. (JAIR)* 32:565--606.