

University of St Andrews

# A Preliminary Evaluation of Machine Learning in Algorithm Selection for Search Problems



Lars Kotthoff, Ian P. Gent, Ian Miguel  
larsko@cs.st-andrews.ac.uk

Fourth Annual Symposium on Combinatorial Search (SoCS 2011)

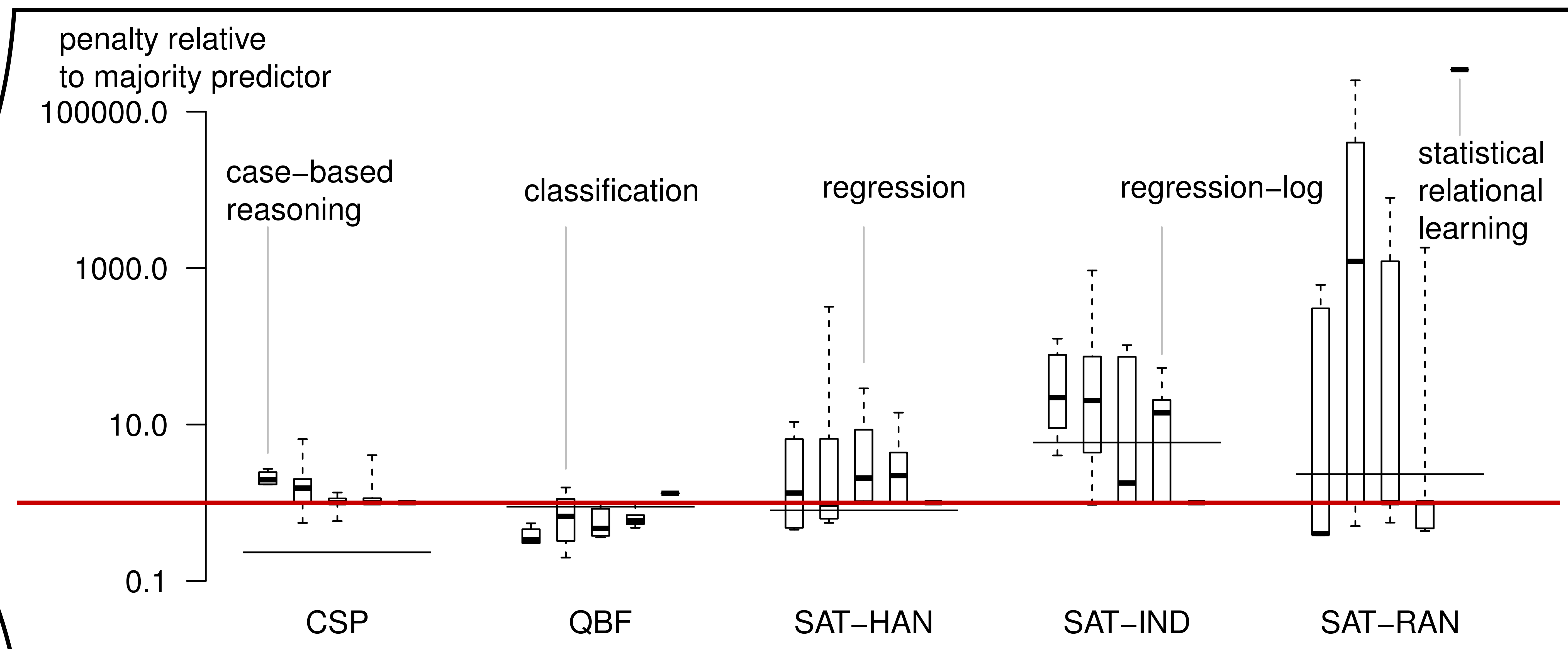
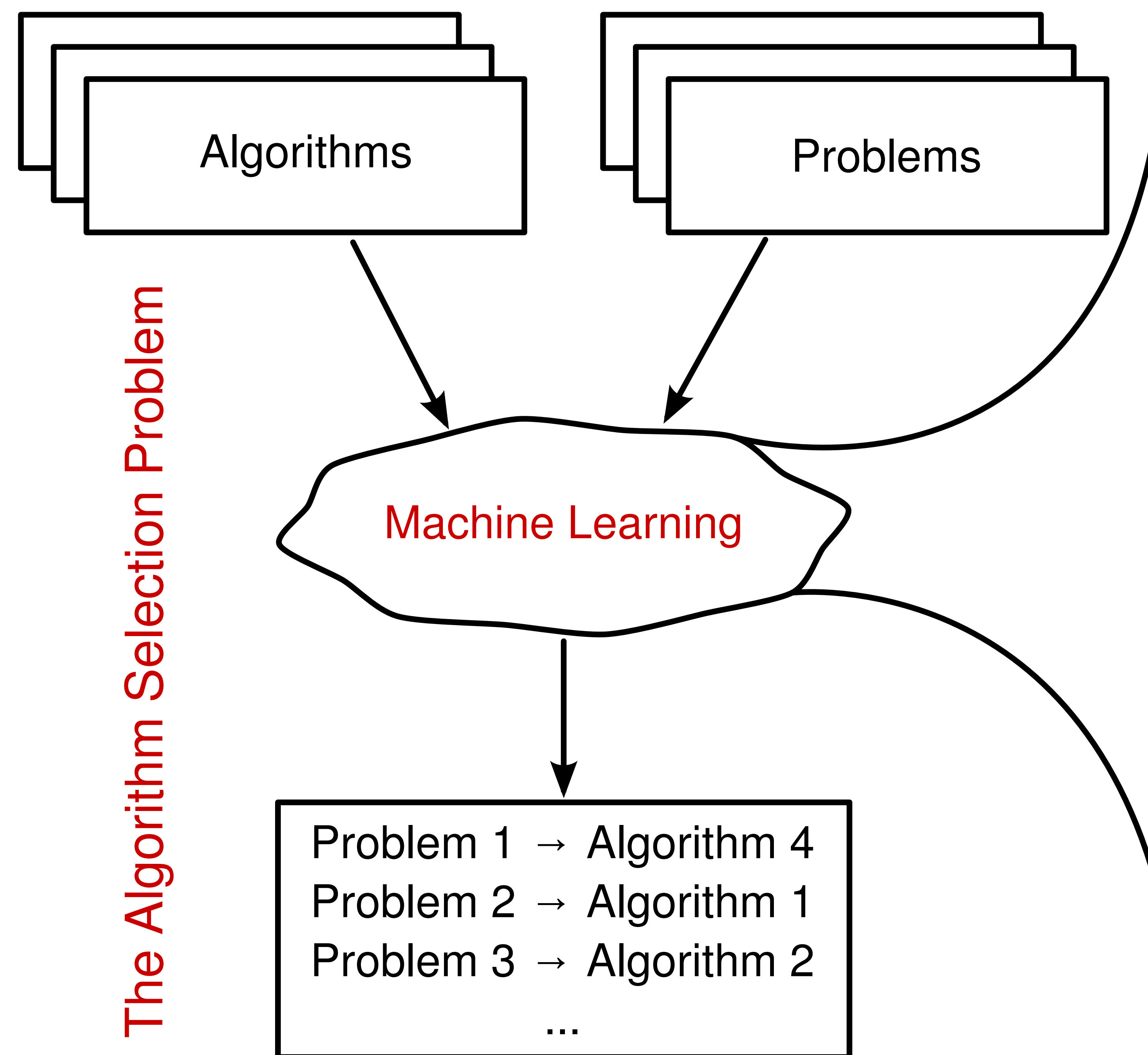
Task: given a set of algorithms/systems, select the most appropriate (fastest) one for the problem to solve

- ▶ use Machine Learning to solve this problem — but what kind of Machine Learning to use?

boxes show quartiles of the performance of the applicable kinds of Machine Learning

horizontal lines show the performance of the respective system from the literature

Machine Learning sorted by type  
data sets from Constraint Programming, Quantified Boolean Formula satisfaction and SAT



- ▶ No Free Lunch™ — no single kind of Machine Learning is the best
- ▶ **Statistical Relational Learning** applied to this problem for the first time, but no good results
- ▶ many sophisticated approaches perform worse than just using a single algorithm
- ▶ automatically selecting the most predictive problem features for Machine Learning improves the performance only marginally
- ▶ randomly deleting parts of the training data can improve performance
- ▶ **Support Vector Machines** predicting the log of the runtime of each algorithm perform well in general
- ▶ incentive for Machine Learning research to improve results in this area

many existing systems, but no comprehensive comparison of methods so far

- ▶ compare 31 different kinds of Machine Learning on 5 data sets from the literature and compare the performance to the respective existing systems

