

AutoML: Algorithm Selection

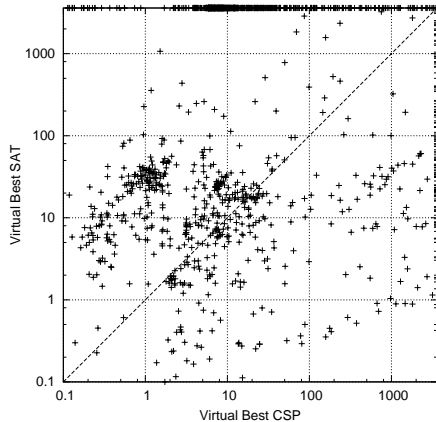
Bonus: Combinatorial Problems

Bernd Bischl Frank Hutter Lars Kotthoff
Marius Lindauer Joaquin Vanschoren

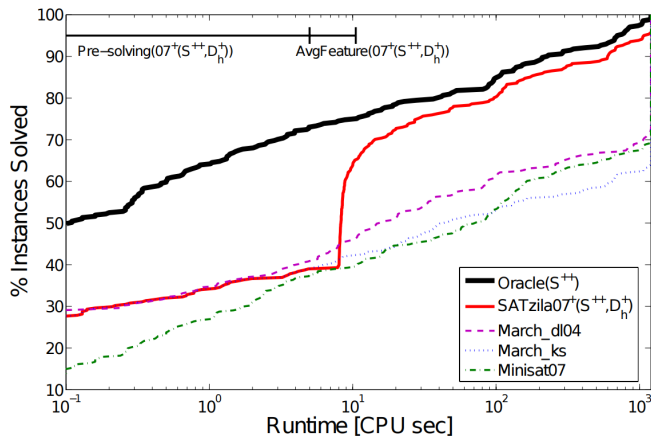
Motivation

- Algorithm Selection applied in many other domains
- success and performance improvements for combinatorial and optimization problems in AI dwarfs those in machine learning
- important application area of AI facilitating cross-disciplinary collaborations and advances

Motivation: Performance Differences [Barry et al. 2014] |



Motivation: Leveraging the Differences [Xu et al. 2008]



Algorithms [Huberman et al. 1997]

- constraint solvers
- search strategies
- modeling choices
- different types of consistency

Features

- number of variables, number of clauses/constraints/...
- ratios
- order of variables/values
- connectivity clause/constraints–variable graph or variable graph
- number of nodes/propagations within time limit
- estimate of search space size
- tightness of problem/constraints
- ...

Example System – SATzilla [Xu et al. 2008]

- portfolio of 7 SAT solvers, trained on 4811 problem instances
- syntactic (33) and probing features (15)
- ridge regression to predict log runtime for each solver, choose the solver with the best predicted performance
- later version uses random forests to predict better algorithm for each pair, aggregation through simple voting scheme
- pre-solving, feature computation time prediction, hierarchical model, selection of algorithms to include in portfolio based on overall performance
- won several competitions

- https://github.com/coseal/aslib_data
- SAT, CSP, QBF, ASP, MAXSAT, OR, ML...
- includes data used frequently in the literature that you may want to evaluate your approach on
- more scenarios in the pipeline
- <http://aslib.net>

autofolio <https://bitbucket.org/mlindauer/autofolio/>

LLAMA <https://bitbucket.org/lkotthoff/llama>

SATzilla <http://www.cs.ubc.ca/labs/beta/Projects/SATzilla/>

(Much) More Information [Kotthoff. 2014]

Comments? Suggestions? Corrections?
Lars.Kotthoff@univie.ac.at

Algorithm Selection Literature Summary

Last update 21 November 2018

click headings to sort
 click columns to expand

AssureSurvey

citation	dataset	features	predict what	predict how	predict where	portfolio	year
Largely 1983a, Largely 1983a	search	past performance	algorithm	hand-crafted and learned rules	offline and online	dynamic	1983
Calverton et al. 1981	planning	problem domain features, search statistics	search rules	explication based rule construction	online	dynamic	1981
Gratch and DeJong 1992	planning	problem domain features, search statistics	control rules	probabilistic rule construction	online	dynamic	1992
Smith and Smith 1992	software design	features of abstract representation	algorithms and data structures	circulated annealing	offline	static	1992
Aha 1992	machine learning	instance features	algorithm	learned rules	offline	static	1992
Broley 1993	machine learning	instance and algorithm features	algorithm	hand-crafted rules	offline	static	1993
Kamel et al. 1993	differential equations	past performance, instance features	algorithm	hand-crafted rules	offline	static	1993
Milson 1993a, Milson 1993a, Milson 1995	CSP	runtime performance	algorithm	hand-crafted and learned rules	offline	dynamic	1993
Calet 1984	software design	instance features	algorithms and data structures	frame-based knowledge base	offline	static	1984
Tsang et al. 1985	CSP	instance features	-	-	offline	static	1985
Brewer 1985	software design	runtime performance	algorithms, data structures and their parameters	statistical model	offline	static	1985
Wierwille et al. 1995, Joehi et al. 1995	differential equations	instance features	runtime performance	Bayesian belief propagation, neural nets	offline	static	1995
Bonnet et al. 1996	CSP	search statistics	switch algorithms?	hand-crafted rules	online	static, static order	1996
Allen and Hömmer 1998	SAT, CSP	problem	runtime performance	hand-crafted rules	online	static	1998
Sakkout et al. 1996	CSP	search statistics	switch algorithms?	hand-crafted rules	online	static	1996
Huelsenman et al. 1997	graph colouring	past performance	resource allocation	statistical model	offline	static	1997
Gomes and Selman 1997a, Gomes and Selman 1997a	CSP	problem size and past performance	algorithm	statistical model	offline	static	1997
Cook and Vareli 1997	parallel search	problem	set of search strategies	decision trees, Bayesian classifier, nearest neighbour, neural net	online	static	1997
Fink 1997, Fink 1998	planning	past performance	resource allocation	statistical model, regression	offline	static	1997
Lichten and Lemstra 1998	branch and bound	problem	runtime performance	hand-crafted rules	online	static	1998
Gomes et al. 1999	vehicle routing problem	runtime performance	algorithm	genetic algorithms	offline	static	1999
Hose et al. 1999	planning	instance features	resource allocation	linear regression	offline	static	1999
Tepstrima-Mann et al. 1999	scheduling	instance and search features	algorithm	genetic algorithms	offline	dynamic	1999
Wolpin et al. 2000	software design	instance features	data structures	nearest neighbour	offline	static	2000
Beck and Fox 2000	job shop scheduling	instance feature changes during search	algorithm scheduling policy	hand-crafted rules	online	static	2000
Braxell and Soares 2000	classification	past performance	tasking	distribution model	offline	static	2000
Lapoutsakis and Litzman 2000	order selection, sorting	instance features	estimating cost for each sub-problem	NECP	online	static	2000
Silko 2000	CSP	problem	cost of solving problem	statistical model	offline	static	2000
Plattinger et al. 2000	classification	instance features, problem	algorithm	9 different classifiers	offline	static	2000
Fuhrmann 2000	CSP	past performance	resource allocation	performance simulation for different allocations	offline	static	2000
Soares and Braxell 2000	machine learning	instance features	tasking	nearest neighbour	offline	static	2000
Gomes and Selman 2001	CSP, mixed integer programming	past performance	algorithm	statistical model	offline	dynamic	2001
Eysenck and Preuder 2001, Eysenck et al. 2002, Eysenck et al. 2008, Eysenck and Petrovic 2011	CSP	variable characteristics	algorithm	weights, hand-crafted rules	offline and online	dynamic	2001
Lapoutsakis and Litzman 2001	DPLL branching rules	instance features	estimating cost for each sub-problem	NECP	online	static	2001
Rosenfeld 2001	optimization	search statistics	expected utility of algorithm	reinforcement learning	offline and online	static	2001
Horvitz et al. 2001	CSP	instance and instance generator features, search statistics	runtime performance, model parameters	Bayesian model	offline and online	static	2001

<http://larskotthoff.github.io/assurvey/>