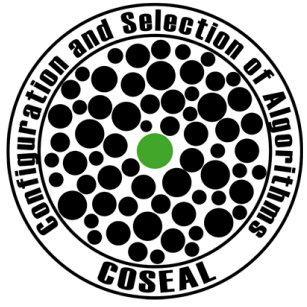
Update on Algorithm Selection Library (ASlib)

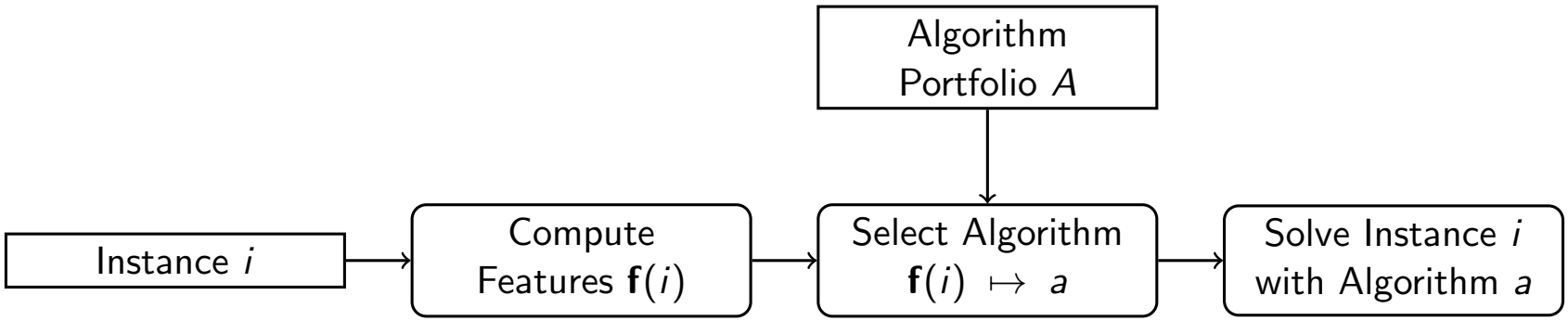
B. Bischl P. Kerschke L. Kotthoff Y. Malitsky A. Fréchette
H. Hoos F. Hutter K. Leyton-Brown K. Tierney J. Vanschoren

M. Lindauer

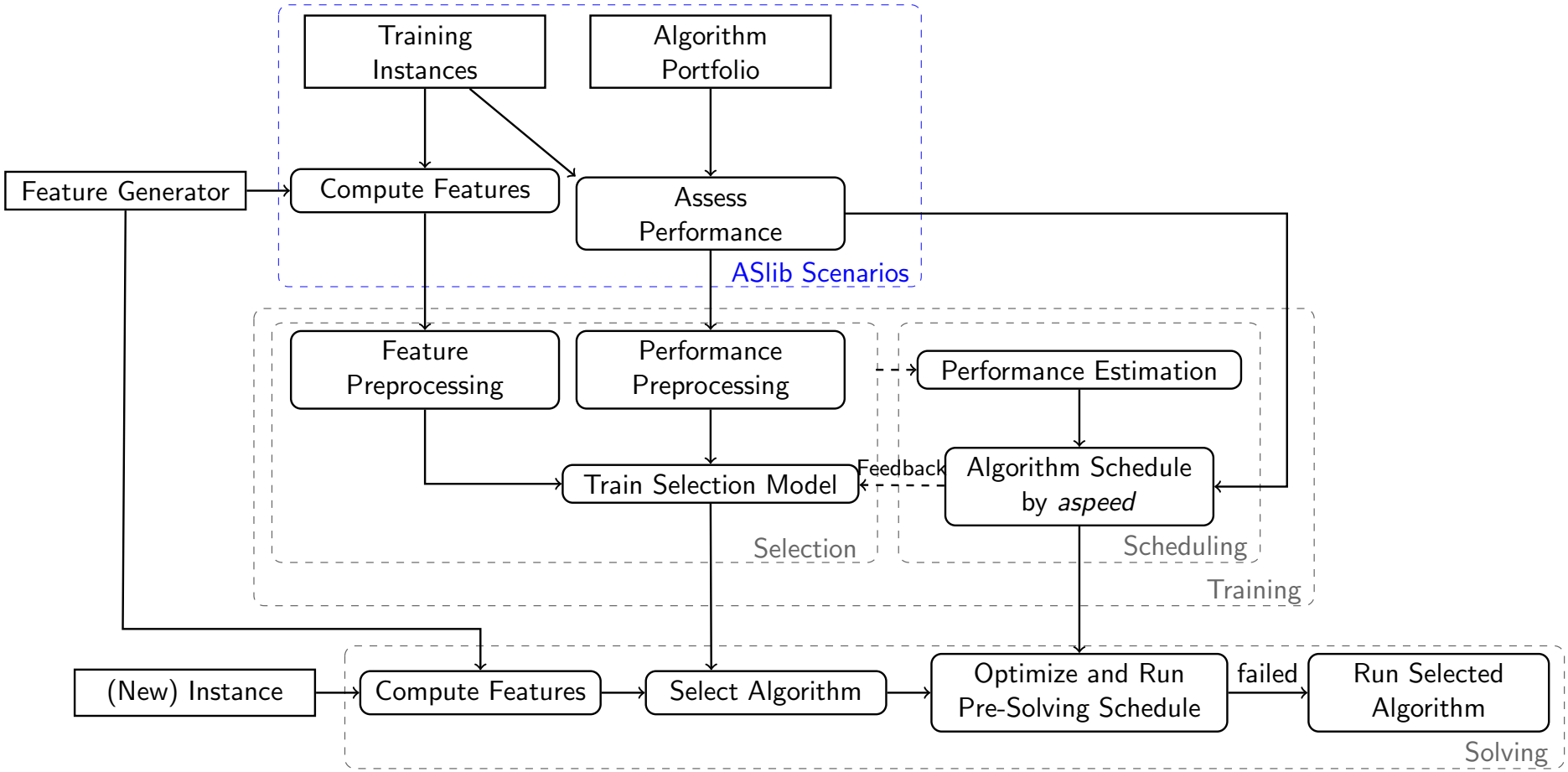
University of Freiburg



Algorithm Selection

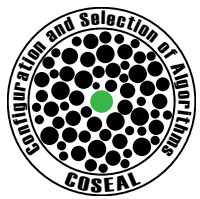


Algorithm Selection (Contd)



ASlib Scenarios

Scenario	Data Collection Time (CPU Days)
ASP-POTASSCO	25

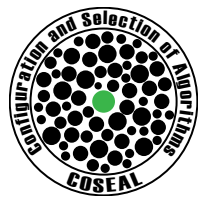


ASlib Scenarios

Scenario	Data Collection Time (CPU Days)
ASP-POTASSCO	25
PREMARSHALLING	28
CSP-2010	52
MAXSAT12-PMS	56

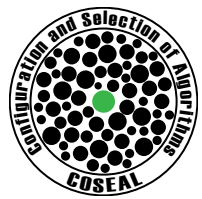
ASlib Scenarios

Scenario	Data Collection Time (CPU Days)
ASP-POTASSCO	25
PREMARSHALLING	28
CSP-2010	52
MAXSAT12-PMS	56
SAT11-INDU	128
SAT11-RAND	158
QBF-2011	163
SAT11-HAND	168



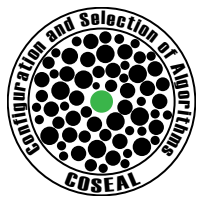
ASlib Scenarios

Scenario	Data Collection Time (CPU Days)
ASP-POTASSCO	25
PREMARSHALLING	28
CSP-2010	52
MAXSAT12-PMS	56
SAT11-INDU	128
SAT11-RAND	158
QBF-2011	163
SAT11-HAND	168
SAT12-HAND	234
SAT12-INDU	284



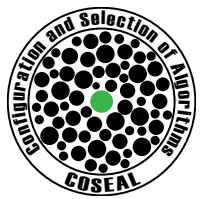
ASlib Scenarios

Scenario	Data Collection Time (CPU Days)
ASP-POTASSCO	25
PREMARSHALLING	28
CSP-2010	52
MAXSAT12-PMS	56
SAT11-INDU	128
SAT11-RAND	158
QBF-2011	163
SAT11-HAND	168
SAT12-HAND	234
SAT12-INDU	284
SAT12-ALL	415
SAT12-RAND	447
PROTEUS-2014	596



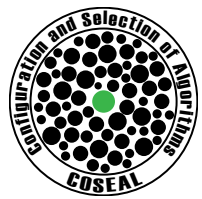
Issues with AS Papers

1. tedious and time-consuming task to collect data for more than $2-3$ scenarios – only few experimental results



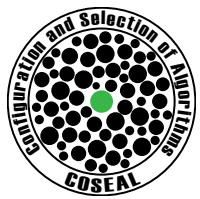
Issues with AS Papers

1. tedious and time-consuming task to collect data for more than $2-3$ scenarios – only few experimental results
2. everyone uses other scenarios – not comparable results
3. runtime are measured on different hardware – not comparable results



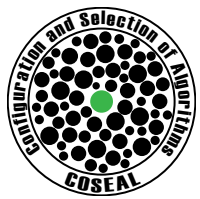
Issues with AS Papers

1. tedious and time-consuming task to collect data for more than $2-3$ scenarios – only few experimental results
2. everyone uses other scenarios – not comparable results
3. runtime are measured on different hardware – not comparable results
4. scenarios are not always publicly available – not repeatable results



Issues with AS Papers

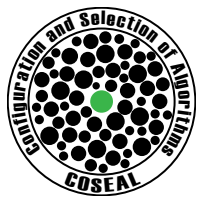
1. tedious and time-consuming task to collect data for more than $2-3$ scenarios – only few experimental results
2. everyone uses other scenarios – not comparable results
3. runtime are measured on different hardware – not comparable results
4. scenarios are not always publicly available – not repeatable results
5. beginners (and even experts) make mistakes – e.g., don't consider feature costs – invalid results



Goals of ASlib

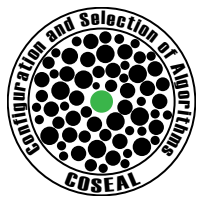
$\rightarrow$ provide a set of AS scenarios with all performance data and instance features with its associated costs

1. repeatable results
2. comparable results, e.g., independent of hardware
3. reduce the burden on AS developers to perform experiments



ASlib *v*.1.0 Scenarios

Scenario	$\lvert I \rvert$	$\lvert A \rvert$	$\#f$	$\varnothing t_f$	Ref.
ASP-POTASSCO	1294	11	138	1.3	[Hoos et al. 2014]



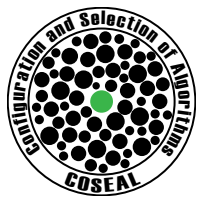
ASlib *v*.1.0 Scenarios

Scenario	$\lvert I \rvert$	$\lvert A \rvert$	$\#f$	$\varnothing t_f$	Ref.
ASP-POTASSCO	1294	11	138	1.3	[Hoos et al. 2014]
CSP-2010	2024	2	17	*n/a*	[Gent et al. 2010]
MAXSAT12-PMS	876	6	37	0.1	[Ansotegui et a. 2014]
PREMARSHALLING	527	4	16	*n/a*	[Tierney et al. 2014]
PROTEUS-2014	4021	22	198	6.4	[Hurley et al. 2014]
QBF-2011	1368	5	46	*n/a*	[Kotthoff et al. 2012]
SAT11-HAND	296	15	115	41.2	[Xu et al. 2012]
SAT11-INDU	300	18	115	135.3	[Xu et al. 2012]
SAT11-RAND	600	9	115	22.0	[Xu et al. 2012]
SAT12-ALL	1614	31	115	40.5	[Xu et al. 2012]
SAT12-HAND	767	31	115	39.0	[Xu et al. 2012]
SAT12-INDU	1167	31	115	80.9	[Xu et al. 2012]
SAT12-RAND	1362	31	115	9.0	[Xu et al. 2012]

Automated Exploratory Data Analysis (EDA)

`aslib.net` $\rightarrow$
`http://coseal.github.io/aslib-r/scenario-pages/index.html`

Scenario Name	Instances	Algorithms	Features	Type	Stoch. Feat.	Stoch. Alg.	Feature Costs
ASP-POTASSCO	1294	11	138	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
CSP-2010	2024	2	86	runtime	TRUE, max. reps = 1	FALSE, max. reps = 1	FALSE
MAXSAT12-PMS	876	6	37	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
PREMARSHALLING-ASTAR-2013	527	4	22	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	FALSE
PROTEUS-2014	4021	22	198	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
QBF-2011	1368	5	46	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	FALSE
SAT11-HAND	296	15	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT11-INDU	300	18	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT11-RAND	600	9	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT12-ALL	1614	31	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT12-HAND	767	31	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT12-INDU	1167	31	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE
SAT12-RAND	1362	31	115	runtime	FALSE, max. reps = 1	FALSE, max. reps = 1	TRUE



Automated Exploratory Data Analysis (EDA)

aslib.net →
http://coseal.github.io/aslib-r/scenario-pages/index.html

Scenario SAT12-ALL

```
## Scenario id                          : SAT12-ALL
## Performance measures                 : runtime
## Performance types                    : runtime
## Algorithm cutoff time                : 1200
## Algorithm cutoff mem                 : NA
## Feature cutoff time                  : 1200
## Feature cutoff mem                   : NA
## Nr. of instances                     : 1614
## Features (deterministic)       (115) : nvarsOrig, nclausesOrig, nvars, nclauses, reducedVars, re...
## Features (stochastic)                : -
## Feature repetitions                  : 1 - 1
## Feature costs                        : Yes
## Algo. (deterministic)          ( 31) : ebglucose, ebminisat, glucose2, glueminisat, lingeling, l...
## Algo. (stochastic)                   : -
## Algo. repetitions                    : 1 - 1
## Algo. runs (inst x algo x rep)       : 50034
## Feature steps                  ( 10) : Pre, Basic, KLB, CG, DIAMETER, cl, sp, ls_saps, ls_gsat, ...
## CV repetitions                       : 1
## CV folds                             : 10
```

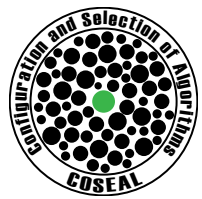
- README
- Download files
- Algorithm overview
- Feature overview
- Benchmark results
- Config
- Validator output

Back to scenario list

Automated Exploratory Data Analysis (EDA)

`aslib.net` $\rightarrow$
`http://coseal.github.io/aslib-r/scenario-pages/index.html`

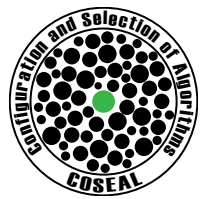
algo	model	succ	par10	mcp
baseline	vbs	0.988	241.318	0.000
baseline	singleBest	0.753	3079.886	302.509
baseline	singleBestByPar	0.753	3079.886	302.509
baseline	singleBestBySuccesses	0.753	3079.886	302.509
classif	rpart	0.791	2600.350	219.333
classif	randomForest	0.921	1060.722	77.848
classif	ksvm	0.904	1271.831	103.450
cluster	XMeans	0.749	3132.932	294.250
regr	lm	0.888	1473.018	128.754
regr	rpart	0.836	2072.328	176.903
regr	randomForest	0.931	943.202	66.367



On-Going Evaluation

aslib.net → https://docs.google.com/spreadsheets/...

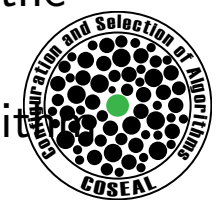
	Oracle	SBS	SNNAP	ISAC	aspeed	claspfolio	claspfolio+pre-solving	zilla	LLAMA/ASlib	SUNNY	state-of-the-art
Version			1.4	SNNAP 1.4	claspfolio 2.2	2.1	2.1	0.9.1b	0.8.1	1.0	
ASP	21.3	534.1	203.8	291.9	353.3	124.8	**115.5**	170.0	137.6	180.2	115.5
CSP	107.7	1,087.4	1,087.5	1,027.0	1,043.4	384.7	310.1	276.0	**247.7**	420.3	247.7
MAXSAT	40.7	2,111.6	895.0	786.4	769.2	264.0	629.6	**166.8**	322.6	307.7	166.8
PREMARSHALLII	227.6	7,002.9	9,042.1	5,880.8	**1,964.1**	2,513.8	2,395.7	3,179.1	5,546.5	2,221.5	1,964.1
PROTEUS-2014	26.3	10,756.3	4,058.7	3,328.0	1,444.7	1,729.5	1,909.7	2,050.3	2,267.8	**1,321.7**	1,321.7
QBF	95.9	9,172.3	7,386.2	3,813.5	1,883.5	1,068.0	1,072.9	1,245.2	1,053.7	**1,040.4**	1,040.4
SAT11-HAND	478.3	17,815.8	9,209.3	13,946.2	**4,976.1**	7,093.2	5,259.1	6,211.5	6,813.3	8,517.4	4,976.1
SAT11-INDU	419.9	8,985.6	6,632.6	8,461.2	8,507.8	7,851.2	**5,395.9**	8,048.8	5,650.2	7,329.9	5,395.9
SAT11-RAND	227.3	14,938.6	4,859.0	3,140.4	3,157.3	3,684.0	1,141.7	**877.5**	1,431.1	6,321.1	877.5
SAT12-ALL	93.7	2,967.9	1,427.5	2,989.3	2,694.8	1,694.5	1,027.2	876.9	**804.5**	1,297.2	804.5
SAT12-HAND	113.2	3,944.2	2,180.5	4,110.8	2,123.0	2,081.0	1,267.3	1,031.5	**886.3**	1,747.1	886.3
SAT12-INDU	88.1	1,360.6	789.0	1,409.5	1,645.4	1,027.2	816.1	839.7	**783.4**	1,354.2	783.4
SAT12-RAND	46.9	568.5	593.1	434.5	713.9	708.0	**425.5**	485.3	527.0	642.9	425.5



Papers using ASlib

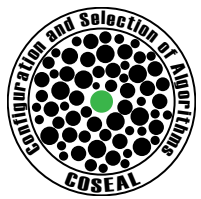
Bischl, B., Kerschke, P., Kotthoff, L., Lindauer, M., Malitsky, Y., et al (2015). Aslib: A benchmark library for algorithm selection. arXiv preprint 1506.02465. submitted to AIJ.

- Kotthoff, L. (2014). Towards an algorithm selection standard: data format and tools. ECAI.
- Hoos, H., Lindauer, M., & Schaub, T. (2014). claspfolio 2: Advances in algorithm selection for answer set programming. TPLP.
- Lindauer, M. Algorithm selection, scheduling and configuration of Boolean constraint solvers. PhD Thesis.
- Lindauer, M., Hoos, H., & Hutter, F. (2015). From Sequential Algorithm Selection to Parallel Portfolio Selection. LION.
- Kotthoff, L., Kerschke, P., Hoos, H., & Trautmann, H. (2015). Improving the State of the Art in Inexact TSP Solving Using Per-Instance Algorithm Selection. LION.
- Tierney, K., Malitsky, & Y. (2015). An Algorithm Selection Benchmark of the Container Pre-marshalling Problem. LION.
- Lindauer, M., Hoos, H., Hutter, F., & Schaub, T. (2015). Autofolio: Algorithm configuration for algorithm selection. JAIR.



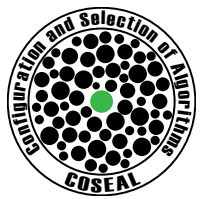
ICON AS Challenge 2015

- First challenge on algorithm selection
- Based on ASlib – fair comparison on an established benchmark set
- 8 participants



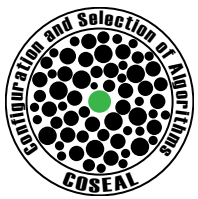
ICON AS Challenge 2015

- First challenge on algorithm selection
- Based on ASlib – fair comparison on an established benchmark set
- 8 participants
- more details in next talk by Lars Kotthoff



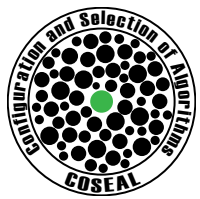
New in 2015

- New scenarios: COP-MZN-2013, CSP-MZN-2013



New in 2015

- New scenarios: COP-MZN-2013, CSP-MZN-2013
- `description.txt` is formatted in YAML – easier parsable, still human-readable

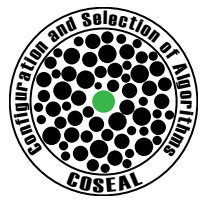


New in 2015

- New scenarios: COP-MZN-2013, CSP-MZN-2013
- `description.txt` is formatted in YAML – easier parsable, still human-readable
- Better understandable feature step definition
  - each feature steps specifies required steps
  - each feature steps specifies provided features
    – no set semantics anymore

Example (SAT12-ALL):

```
Basic:
      requires: Pre
      provides:
        - vars_clauses_ratio
        - POSNEG_RATIO_CLAUSE_mean
        - POSNEG_RATIO_CLAUSE_coeff_variation
        - POSNEG_RATIO_CLAUSE_min
```



www.aslib.net