Configuring *irace* using surrogate configuration benchmarks

Nguyen Dang[1], Leslie Pérez Cáceres[2], Thomas Stützle[2], Patrick De Causmaecker[1]

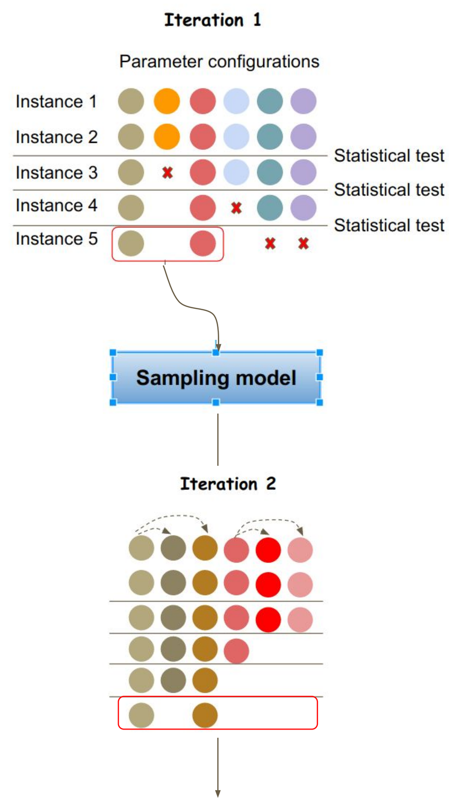
[1] *CODeS, imec - ITEC, KU Leuven, Belgium*

[2] *IRIDIA, Université Libre de Bruxelles (ULB), Belgium*

{nguyenthithanh.dang, patrick.decausmaecker}@kuleuven.be, {leslie.perez.caceres, stuetzle}@ulb.ac.be

1. *irace* parameters

Parameter	Type	Domain	Default value
N^{iter}	int	[1, 50]	$2 + log_2 n^{params}$
μ	int	[1, 20]	5
T^{first}	int	[4, 20]	5
N^{min}	int	[1, 20]	$2 + log_2 n^{params}$
test_type	categorical	{F-test, t-test, t-test-holm, t-test-bonferroni}	F-test
confidence_level	real	[0.5, 0.99]	0.95
enable_soft_restart	categorical	{true, false}	true
enable_elitist	categorical	{true, false}	true
elitist_instances	int	[1, 10]	1



2. Meta-tuning and algorithm configuration benchmarks

Use an automatic algorithm configurator to configure the parameters of a configurator

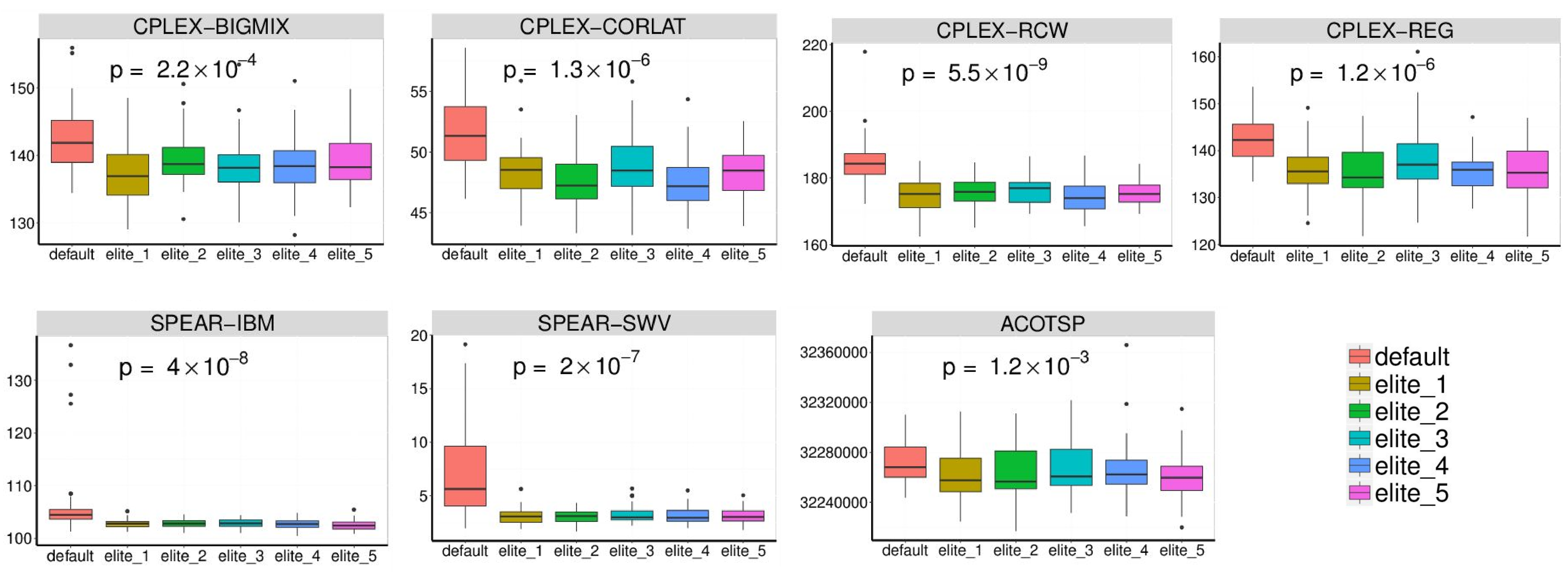
Algorithm configuration benchmarks for meta-tuning:

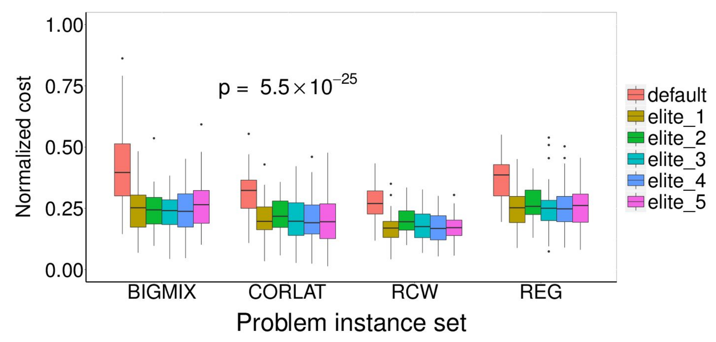
- Real benchmarks: extremely expensive
  - An irace run on SPEAR-IBM (budget: 5000 runs) → 2 CPU days
  - A meta-tuning on SPEAR-IBM (budget: 5000 irace runs) → 27.5 CPU years
- Artificial benchmark set
  - Unclear how to generate
  - Unclear how to match characteristics of real configuration tasks
- Surrogate benchmarks
  - A prediction model: configuration x instance → performance value
  - Build on real benchmark data

3. Empirical Performance Model (Hutter, Xu, Hoos & Leyton-Brown, 2014a)

- Random Forest regression model
- Training: 1000 random configurations x all problem instances

Algorithm	# Parameters (int/real/categorical/conditional)	Problem instance set (#instances)	Performance measure
CPLEX	76 (18/7/45/4)	BIGMIX (1000) REG (1000) CORLAT (1000) RCW (1000)	Running time
SPEAR	28 (4/12/10/9)	IBM (765) SWV (604)	Running time
Ant Colony Optimization	11 (4/4/3/5)	TSP (50)	Solution quality





Meta-tuning becomes computationally feasible

- An irace run on SPEAR-IBM (budget: 5000 runs) → 5 CPU minutes
- A meta-tuning on SPEAR-IBM (budget: 5000 irace runs) → 7.5 CPU days

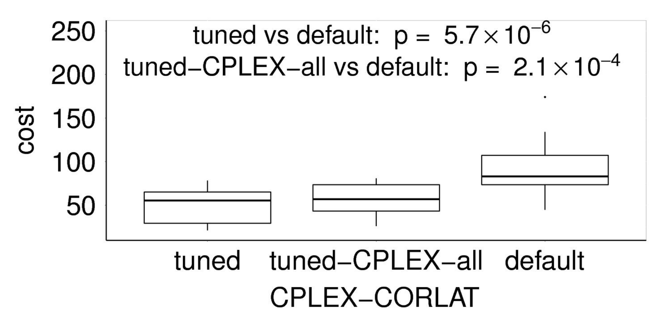
Consistency between surrogate and real benchmarks

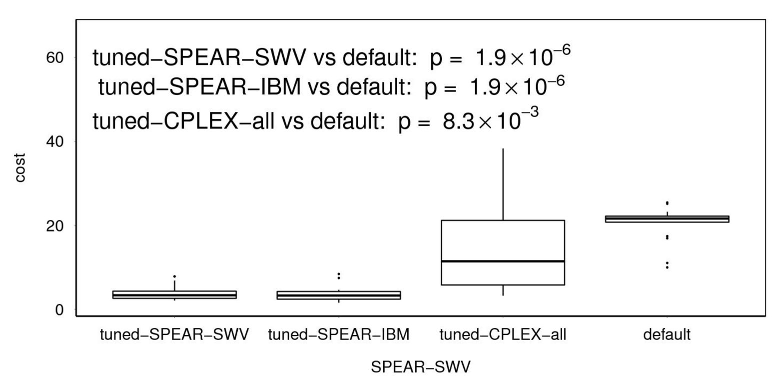
- Prediction accuracy
- Benchmark homogeneity (KendallW)
- Important parameters and pairwise interactions (fANOVA, Hutter et al 2014b)

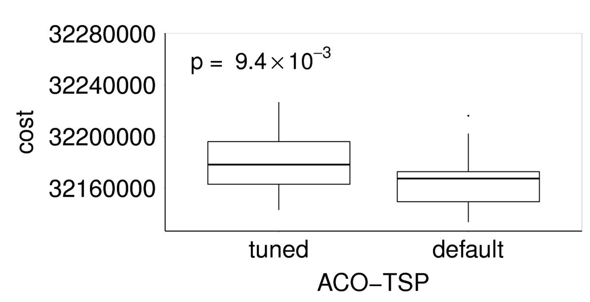
Post-analysis

- fANOVA (Hutter et al 2014b), ablation (Fawcett et al 2016), manual testing
- Most important parameters: N^{iter}, N^{min}, *elitist*
- Complex interactions between irace parameters
- Best irace versions exploit a more intense search than the default version.

4. Mixed results on real benchmarks







6. Ongoing works

- Try a new surrogate modelling method: Eggensperger, Lindauer, Hoos, Hutter & Leyton-Brown, 2017
  - Quantile Random Forest regression model + Hyper-parameter tuning
  - Training data: Trajectories of different automatic algorithm configurators + incumbents' performance + random data points
  - New benchmarks (*aclib2*)
- Collect more benchmarks
- Define quantitative characteristics of an algorithm configuration benchmark
- Meta-tuning on irace with capping
- Study relation between irace parameters and benchmark characteristics
- Study other state-of-the-art configurators

References

C. Fawcett and H. H. Hoos. Analysing differences between algorithm configurations through ablation. Journal of Heuristics, 22(4):431–458 (2016)

Dang, N., Pérez Cáceres, L, De Causmaecker, P. and Stützle, T. Configuring irace using surrogate configuration benchmarks. In: GECCO2017, pp. 243-250 (2017)

F. Hutter, L. Xu, H. H. Hoos, and K. Leyton-Brown. Algorithm runtime prediction: Methods & evaluation. Artificial Intelligence, 206:79–111 (2014a)

F. Hutter, H. H. Hoos, and K. Leyton-Brown. An efficient approach for assessing hyperparameter importance. In Proc. of the 31th Int. Conf. on Machine Learning, vol. 32, pp. 754–762 (2014b)

K. Eggensperger, F. Hutter, H. H. Hoos, and K. Leyton-Brown. Efficient benchmarking of hyperparameter optimizers via surrogates. In: AAAI, pp. 1114–1120. AAAI Press (2015)

Eggensperger, K., Lindauer, M., Hoos, H.H., Hutter, F. and Leyton-Brown, K. Efficient Benchmarking of Algorithm Configuration Procedures via Model-Based Surrogates. arXiv preprint arXiv:1703.10342 (2017)

M. López-Ibáñez, J. Dubois-Lacoste, L. Pérez Cáceres, T. Stützle, and M. Birattari. The irace package: Iterated racing for automatic algorithm configuration. Operations Research Perspectives, pp. 43–58 (2016)