Evaluating anytime performance on NAS-Bench-101

Carlos Vieira
IMD, Universidade Federal do Rio Grande do Norte
Natal, RN, Brazil
carlosv@ufrn.edu.br

Leslie Pérez Cáceres
Escuela de Ingeniería Informática Pontificia Universidad Católica de Valparaíso
Valparaíso, Chile
leslie.perez@pucv.cl

Leonardo C. T. Bezerra
IMD, Universidade Federal do Rio Grande do Norte
Natal, RN, Brazil
leobezerra@imd.ufrn.br

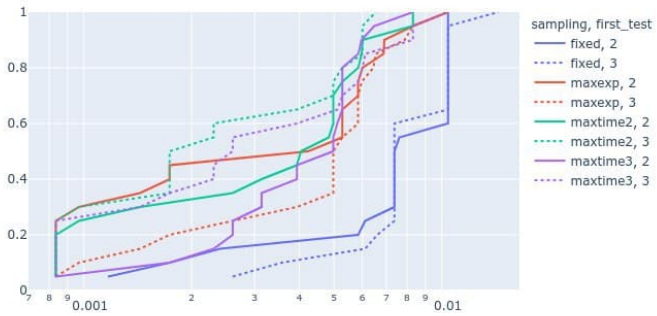

Fig. 1. Empirical cumulative distribution function (x-axis) of the final regret (y-axis) from 20 runs of each irace setup.

APPENDIX A
BENCHMARKING IRACE

Given that irace has not being applied to tackle NAS-Bench-101, we study its performance with different setups for a fair comparison with the other NAS-Bench-101 techniques. We note that when dealing with known benchmarks, ACs can benefit from tailored setups as they can exploit known characteristics of the benchmarks.

NOTE: I would suggest to omit the experiments with fixed number of configurations and also the ones that set the configuration budget as number of experiments, as they are not the best performing and it might simplify the analysis.

We performed 20 repetitions of irace setting 10 million TPU seconds as total configuration budget and evaluate the performance obtained by the following setup options:

Estimation budget Budget assigned for the initial estimation of the average execution time of an evaluation. We perform experiments with 2% and 0.2% of the total configuration time for budget estimation.

First test Number of evaluations required to perform the first elimination test in a race. We perform experiments with 2 and 3 evaluations for applying the initial test.

For further details of these setup options we refer to [1]. Figure 1 gives the ECDFs of the final quality obtained by the different setups of irace. The best overall performance is obtained by setting the estimation budget to 2% of the total configuration time and allowing 3 evaluations before the first elimination test (mean test regret 0.002837).

REFERENCES

[1] M. López-Ibánez, J. Dubois-Lacoste, L. P. Cáceres, M. Birattari, and T. Stützle, "The irace package: Iterated racing for automatic algorithm configuration," *Oper. Res. Perspect.*, vol. 3, pp. 43–58, 2016.