A comparison of acceptance criteria for the ALNS metaheuristic

Alberto Santini, University of Bologna — a.santini@unibo.it

Stefan Røpke, Danish Technical University — ropke@dtu.dk

Lars Magnus Hvattum, Molde University College — hvattum@himolde.no

Introduction

The **Adaptive Large Neighbourhood Search** (ALNS) is a metaheuristic introduced by Røpke and Pisinger [10] and now widely applied to many diverse problems [9]. The idea behind ALNS is to explore the solution space using many large neighbourhoods. At each iteration we choose which one to explore, based on a score that reflects its past performance (the higher the score, the more likely is the neighbourhood to be chosen). Since the nighbourhoods are very large (e.g. exponential), they are not explored completely, but only sampled.

Algorithm 1 gives a general outline of the ALNS Framework. At each iteration, once we select the neighbourhood $N(x)$ to explore, we produce a new incumbent solution $x' \in N(x)$. A crucial part of the algorithm is determining whether to accept x' and make it become the *new current solution* $(x = x')$, or reject x' and explore another neighbourood of x. The rule to determine if x' is accepted or rejected is called **acceptance criterion**.

The aim of this work is to determine the impact of different choices of acceptance criteria to the overall ALNS performance. The question is a methodological one, as acceptance criteria can be seen as *black boxes* that do not need to know any detail of the problem being solved.

Algorithm 1: ALNS Framework

Input: Initial solution: x_0
Input: List of neighbourhoods: $\mathcal{N}$
Input: Neighbourhood scores: λ_N for $N \in \mathcal{N}$
Input: Acceptance parameters: p
Input: Objective to minimise: $f(\cdot)$

$x = x_0$ // Initialise current solution
$x^* = x_0$ // Initialise best solution
$i = 1$ // Iteration count
$\lambda_N = 1 \; \forall N \in \mathcal{N}$ // Initialise scores

while $i \leq K$ **do**
  Choose neighbourhood N with prob proportional to λ_N
  Pick $x' \in N(x)$
  if ***Accept new solution x' (using parameters p)*** **then**
    $x = x'$
  end
  if $f(x) < f(x^*)$ **then**
    $x^* = x$
  end
  Update scores λ
  Update acceptance parameters p
  $i = i + 1$
end
return x^*

Acceptance Criteria

At the extreme of the spectrum of acceptance criteria are **Hill climbing** (HC), which only accepts solutions that improve on the current one, and **Random Walk** (RW), which accepts any new solution.

Historically, however, ALNS used **Simulated Annealing** (SA) [6] as its acceptance criterion: the solution x' is accepted with a probability of $\exp((f(x)-f(x'))/T)$, where T is the acceptance criterion parameter, called *temperature*. The temperature starts at a high value T_{start} and decreases to T_{end} during the solution process. This decrease is exponential in the basic version of SA. In our work we also tried the following variations of SA:

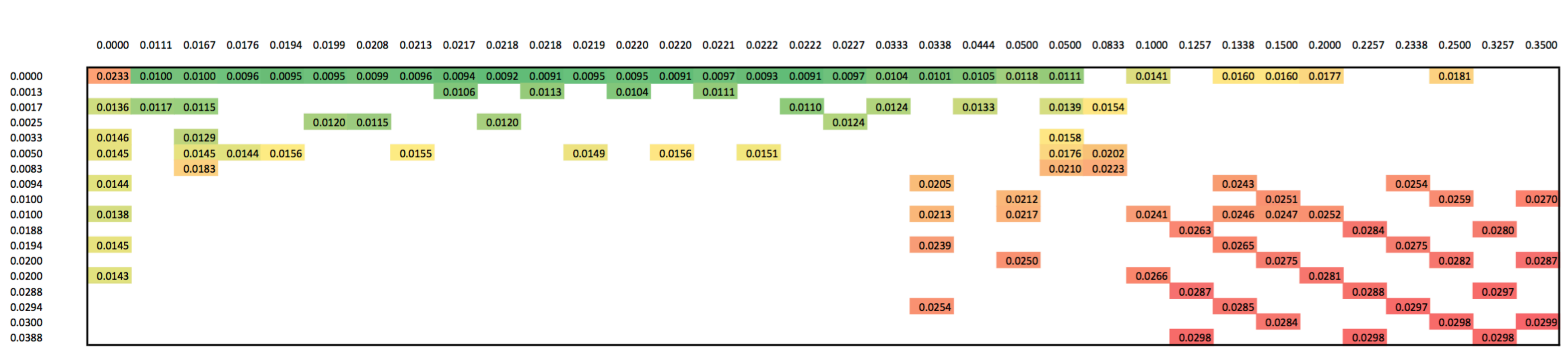
- **Linear SA**: the temperature decreases linearly from T_{start} to T_{end};
- **Linear SA 1par**: the temperature decreases linearly from T_{start} to 0;
- **SA with Reheating**: the temperature is periodically re-increased (see [2]);
- **Instance-scaled SA** (ISSA): the temperature is multiplied by a factor depending on the instance size (this is the criterion originally used by Røpke and Pisinger).

Two more criteria we tested are **Threshold Acceptance** (TA) [4] and **Record-to-Record Travel** (RRT) [3], which accept a new solution if the gap of $f(x')$ with $f(x)$ (for TA) or with $f(x^*)$ (for RRT) is lower than a certain threshold T. Analogously to SA, they also have explonential and *linear* variations, including the linear one with only *one parameter* (1par).

Other acceptance criteria we tested include: **Late-acceptance hill climbing** (LAHC) [1], **Great Deluge** (GD) [3], **Non-Linear Great Deluge** (NLGD) [7] and **Worse Accept** (WA). The latter was introduced by us and consists in accepting a worsening incumbent with a certain probability p which does not depend on $f(x')$ and which decreases during the solution process.

Parameter Tuning

We tested the impact of acceptance criteria on the ALNS, considering two problems: the Capacitated Minimum Spanning Tree Problem (**CMST**), and the Capacitated Vehicle Routing Problem (**CVRP**). For the latter, to remove any bias and focus on the impact of the acceptance criteria only, we also used a simpler Large-Neighbourhood-Search version of ALNS, in which we only have one neighbourhood N (**SimpleCVRP**).



We tuned the acceptance parameters p over a representative set of instances, performing 10 reruns for each instance and considering average values. The figure above, for example, shows the results of parameter tuning for the RRT acceptance criterion on the CVRP test bed. T_{start} is on the x-axis and T_{end} on the y-axis. The colour represents the quality of the solutions obtained with the parameter choice (the greener, the better) and the numbers are the deviation from the best known solution.

Results

We ran computational experiments with the parameters fixed by tuning, again performing 10 reruns for each instance. We ranked the methods according to the **gap** between the average and the best known solution (we also considered the gap between the best solution in the 10 reruns and the best known solution). The results are reported in the following tables, where the original criterion of [10] is highlighted.

CMST

Acceptance Criterion	Gap % (Avg)	Gap % (Best)
Lin. SA	0.399	0.108
ISSA	0.400	0.150
Lin. SA 1par	0.407	0.119
Exp. SA	0.409	0.127
Lin. TA 1par	0.418	0.119
Exp. SA Reheating	0.428	0.174
Lin. RRT	0.473	0.213
Lin. TA	0.474	0.120
Lin. RRT 1par	0.514	0.234
Lin. WA 1par	0.518	0.203
Exp. WA	0.552	0.181
Lin. WA	0.566	0.195
Exp. RRT	0.646	0.269
LAHC	0.655	0.244
GD	0.682	0.371
Exp. TA	0.759	0.315
NLGD	0.995	0.492
HC	2.226	1.215
RW	2.824	2.305

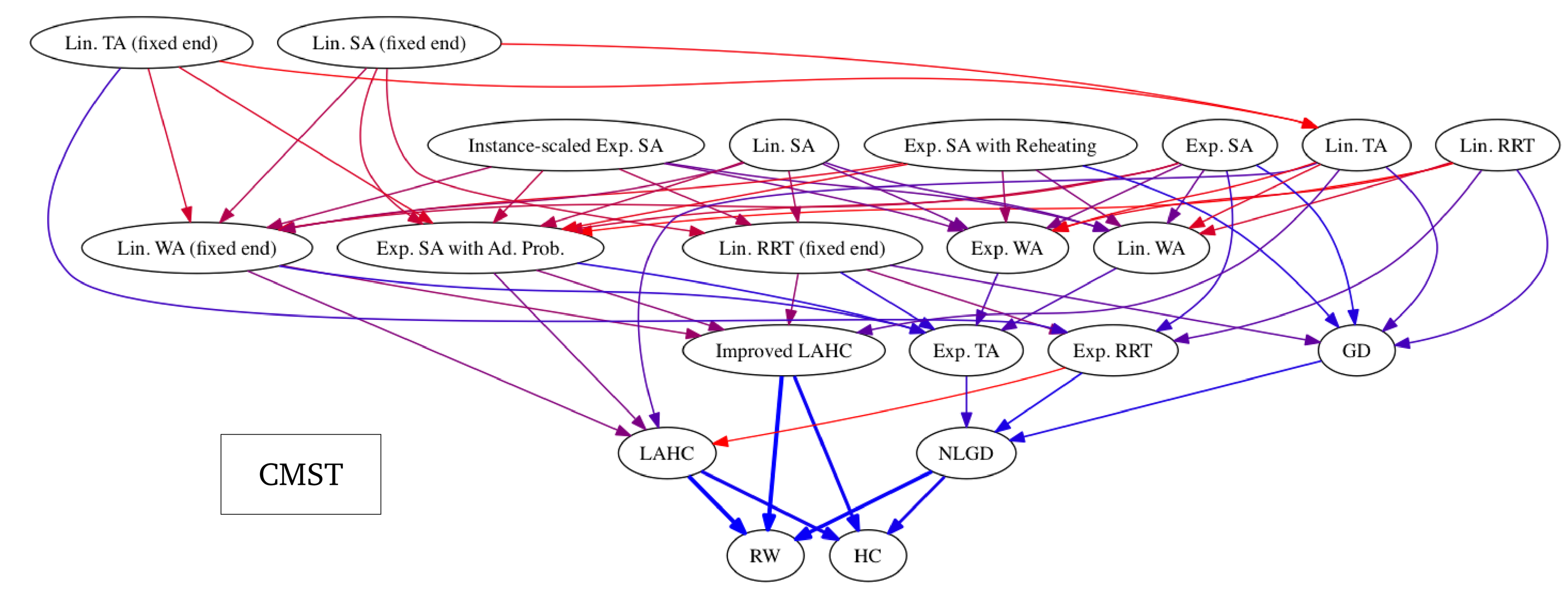
CVRP

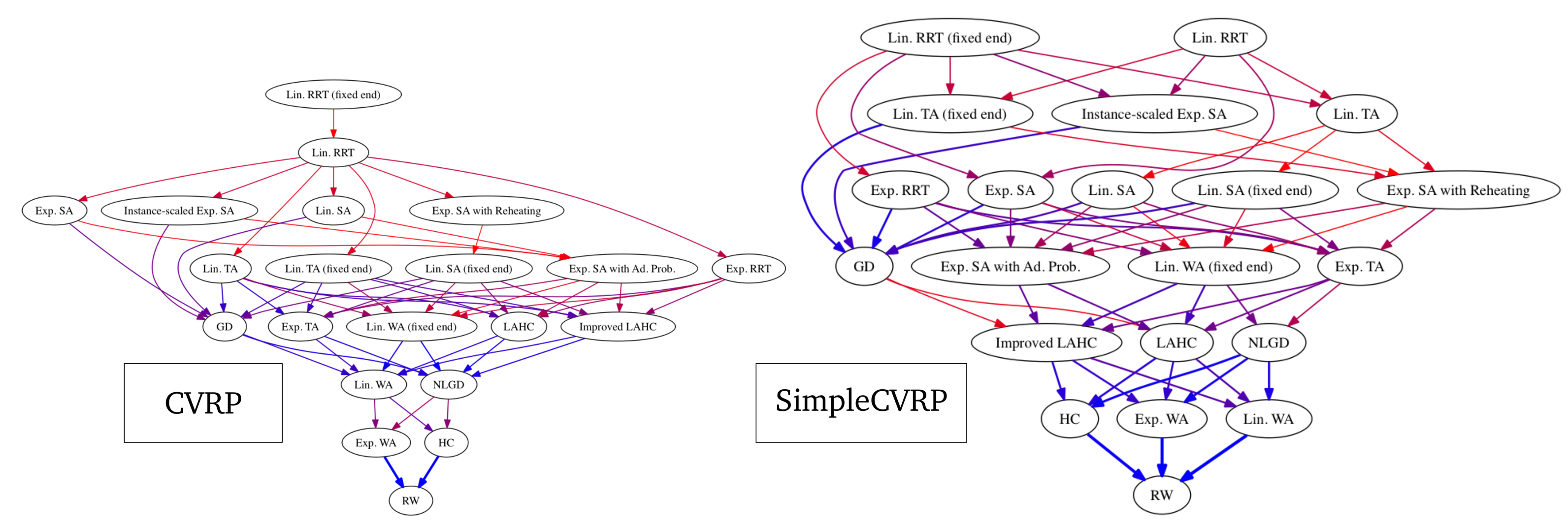
Acceptance Criterion	Gap % (Avg)	Gap % (Best)
Lin. RRT 1par	0.391	0.112
Lin. RRT	0.423	0.148
Lin. TA	0.497	0.179
Lin. TA 1par	0.511	0.197
Exp. SA Reheating	0.527	0.175
Lin. SA	0.527	0.167
Exp. SA	0.529	0.173
Lin. SA 1par	0.538	0.200
ISSA	0.542	0.159
Exp. RRT	0.551	0.126
Lin. WA 1par	0.661	0.301
LAHC	0.716	0.282
GD	0.726	0.463
Exp. TA	0.735	0.276
Lin. WA	0.963	0.496
NLGD	0.989	0.393
Exp. WA	1.147	0.510
HC	1.163	0.557
RW	2.583	2.226

SimpleCVRP

Acceptance Criterion	Gap % (Avg)	Gap % (Best)
Lin. RRT 1par	0.754	0.241
Lin. RRT	0.768	0.218
Exp. RRT	0.939	0.315
Lin. TA 1par	0.972	0.358
Lin. TA	0.973	0.328
Exp. SA	1.062	0.363
ISSA	1.076	0.399
Lin. SA 1par	1.086	0.443
Lin. SA	1.112	0.427
Exp. SA Reheating	1.150	0.445
Lin. WA 1par	1.270	0.580
Exp. TA	1.425	0.591
NLGD	1.695	0.713
GD	1.709	1.189
LAHC	1.870	0.986
Lin. WA	2.461	1.272
Exp. WA	2.516	1.312
HC	2.595	1.381
RW	3.946	3.340

To check that the differences in deviation have statistical significance, we performed pairwise Wilcoxon tests between the acceptance criteria. A negative Wilcoxon test (at 5% significance level) indicates that the results obtained using two acceptance criteria are drawn from the same distribution, and therefore any difference ultimately reduces to noise. If a criterion gives lower deviation than another, and the difference is significant, we say that the first criterion dominates the second. In the following graphs we draw an arc from dominant to dominated criteria (the thicker and bluer the arc, the greater is the difference in deviations).



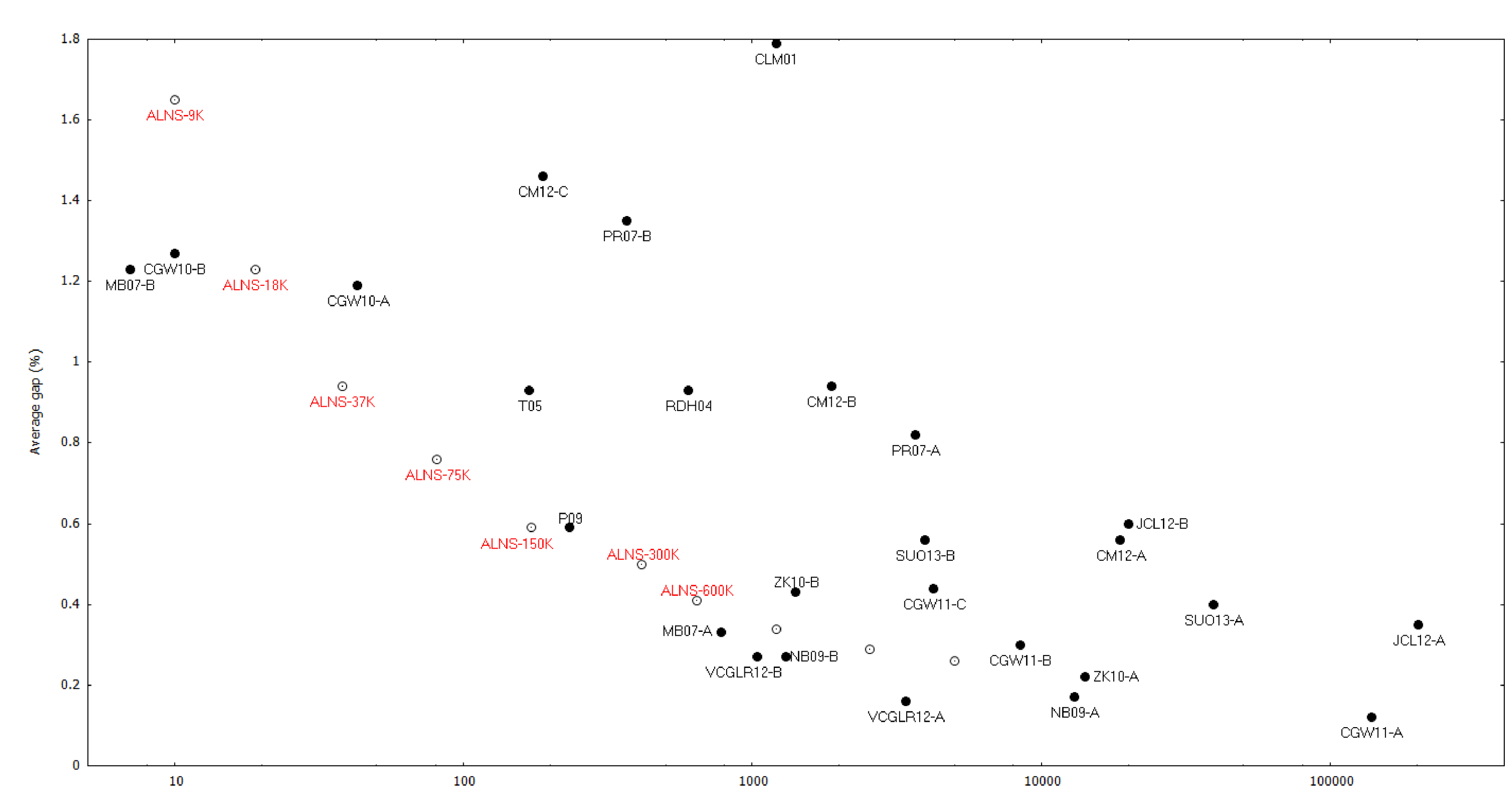


Conclusions

Key findings from our computational experiments show that:

- **Simulated Annealing**, the most widely used criterion, is good — but not necessarily the best;
- **Record-to-Record Travel** and **Threshold Acceptance** are very simple and give excellent results;
- **Linear** methods are no worse than exponential methods;
- And, in particular, linear methods with only **one parameter** give very good results, while being very easy to tune.

Furthermore, the results show that ALNS is still very competitive, especially in terms of speed. In the chart below, adapted from [8], we show results on the Golden CVRP instances [5] for the **Linear RRT 1par** acceptance criterion, comparing ALNS with other methods in the literature. The ALNS points are in red and correspond to different number of iterations (from 9'000 to 600'000); ALNS times have been multiplied by 8, as we ran a parallel version of the algorithm, with 8 concurrent threads.



References

[1] E.K. Burke and Y. Bykov. The late acceptance hill-climbing heuristic. *University of Stirling, Tech. Rep*, 2012.

[2] D. Connolly. General purpose simulated annealing. *Journal of the Operational Research Society*, pages 495–505, 1992.

[3] G. Dueck. New optimization heuristics: the great deluge algorithm and the record-to-record travel. *Journal of Computational Physics*, 104:86–92, 1993.

[4] G. Dueck and T. Scheuer. Threshold accepting: a general purpose optimization algorithm appearing superior to simulated annealing. *Journal of Computational Physics*, 90:161–175, 1990.

[5] B. L. Golden, E. A. Wasil, J. P. Kelly, and I. M. Chao. The impact of metaheuristics on solving the vehicle routing problem: algorithms, problem sets, and computational results. In T. Crainic and G. Laporte, editors, *Fleet management and logistics*, pages 33–56. Springer, 1998.

[6] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[7] D. Landa-Silva and J.H. Obit. Great deluge with non-linear decay rate for solving course timetabling problems. In *Intelligent Systems, 2008. IS'08. 4th International IEEE Conference*, volume 1, pages 8–11. IEEE, 2008.

[8] G. Laporte, S. Ropke, and T. Vidal. Heuristics for the vehicle routing problem. In Paolo Toth and Daniele Vigo, editors, *Vehicle Routing: Problems, Methods, and Applications*, chapter 4, pages 87–116. SIAM, 2nd edition, 2014.

[9] D. Pisinger and S. Ropke. Large neighborhood search. In *Handbook of Metaheuristics*, pages 399–419. Springer, 2010.

[10] S. Ropke and D. Pisinger. A unified heuristic for a large class of vehicle routing problems with backhauls. *European Journal of Operational Research*, 171:750–775, 2006.