Review of: EJOR-D-15-00397 - “A generic feasibility-based heuristic for the container pre-marshalling problem”

May 6, 2015

The paper presents a generic approach to the Premarshalling problem, the problem of reorganizing the container terminal yard in such a way that containers in a bay will not need any further reshuffling during the retrieval phase.

The paper is written in a formal and rigorous way. The problem is well presented and the literature review allows to contextualize the work presented in this paper. The authors first formally define the premarshalling problem, then the general heuristic scheme is presented. Next, a specific implementation of the heuristic is used to carry our computational experiments on two benchmark sets and to compare the results obtained with two approaches from the literature.

While the approach to the problem is very rigorous and formal (and it is especially worth mentioning the attempt to formalize the problem through the definition of “state” and “state feasibility”), the paper suffers of a major drawback typical of other papers in the field. The pre-marshalling problem is tackled via a long list of heuristic rules and cases (and subcases). This heuristic-driven approach (as opposed to, *e.g.*, an approach based on a mathematical model) requires the use of a long list of symbols and makes the notation and the reading of the paper itself cumbersome and quite difficult. For example, it is easy to get lost (and to lose interest) around the algorithms presented in pages 13, 15, and 17 and the cases I1, I2, and I3 for the accomplishment of the internal tasks. The presentation should be improved, perhaps adding examples, to guide the reader through the implementation issues and the details of the proposed approach.

Some specific comments that should be addressed before the paper is accepted for publication are:

1. Improve and ease the presentation, especially Section 6. Maybe, including examples and pictures might help in achieving this goal. Similarly, while the precision and rigor of Algorithms 1-6 is appreciated, the author should rethink the way in which the ideas behind some of them (especially Algorithms 4 and 5) is currently presented.

2. Proposition 2 at page 9 on the existence of “tricky states”: The proposition states that a tricky states exists if the number of empty slots is less than

2(H-1). However, at lines 38 and 39, the authors are mentioning that what can cause the existence of a tricky state is the inability to reshuffle *unfixed* containers. In other words, could it happen that, despite $E < 2(H-1)$ being true, a tricky states does not exist (for example, in the case in which the number of unfixed containers in the remaining two stacks is less than E)? It seems to me that the existence of a tricky state must depend on the number of unfixed containers in the "extreme" state. Please, clarify this doubt.

3. Page 12, lines 49-52: The concept of "bottom tiers protection" is quite interesting. However, one might wonder what is the effect of different choices of parameters value, namely 2 and S. Would computational results and running times change significantly if these two parameters were modified? How?

4. The authors should present the running times of the proposed algorithm on both benchmark sets (if possible, along with the running time of the two benchmark heuristics).

5. Please, provide details of the implementation: Machine configuration (speed, ram, OS, etc.), language and compiler used, etc.

Minor comments:

1. Page 15, line 6 : The accomplishment for an internal task → The accomplishment *of* an internal task?

2. Page 15, line 6: "tack" → "task"

3. Why is the gap of Tables 2 and 3 measured with respect to the number of containers? Why is this gap preferred over a standard improvement measure (*e.g.*, gap = (gash-literature)/literature). For example, for the first row of table 2, to give a measure of the goodness of the GASH approach, I would use:

$$\gamma = \frac{LPVFH - GASH}{LPVFH} = \frac{11.25 - 11.28}{11.25} = -0.003$$

In this case, the γ value would indicate the net improvement, in percentage, (in this specific case, loss) of the proposed method with respect to the best result from the literature.