Learning agents for the multi-mode project scheduling problem

Tony Wauters [a] Katja Verbeeck [a] Greet Vanden Berghe [a]
Patrick De Causmaecker [b]

[a] *CODeS, KaHo Sint-Lieven, Gebroeders Desmetstraat 1, B-9000 Gent, Belgium*
[b] *CODeS, Department of Computer Science, K.U. Leuven Campus Kortrijk, Etienne Sabbelaan 53, B-8500 Kortrijk, Belgium*

1 Introduction

The integration of machine learning methods into meta-heuristic search techniques is currently an important topic in the field of combinatorial optimisation. It is often referred to as *intelligent optimisation* [1]. Learning has the potential of helping the search process in several ways. Searching for instance, can become more generic when online learning and automation mechanisms tune the critical parameters of the search process themselves and avoid expensive manual finetuning. Some techniques also have the ability of learning to explain why, when and how the search is effective, resulting in better documented search. In this paper [6] we integrate multi-agent reinforcement learning and local search for building schedules for the multi-mode resource-constrained project scheduling problem (MRCPSP).

In the last few decades, the resource constrained project scheduling problem (RCPSP) has become an attractive subject in operational research. It considers scheduling a project's activities while respecting the resource requirements and the precedence relations between the activities. The academic problem has many relevant real world counterparts in the building and consultancy sector, for example, which results in a challenging list of instances with varying optimisation objectives. The MRCPSP is a generalized version of the RCPSP, where each activity can be performed in one out of a set of modes, with a specific activity duration and resource requirements (e.g. 2 people each with a shovel need 6 days to dig a pit, while 4 people each with a shovel and one additional wheelbarrow need only 2 days). The RCPSP is shown to be an NP-hard optimisation problem [2], thus so is the MRCPSP. Demeulemeester and Herrroelen [3] present a comprehensive research handbook on project scheduling.

The idea we propose here is to learn a good constructive heuristic that can be further refined by local search techniques for large scale scheduling problems. With the latter goal in mind, a network of distributed reinforcement learning agents was set up, in which agents cooperate to jointly learn a well performing heuristic. As we will show, our method generates results that are comparable to those obtained by the best-performing finetuned algorithms found in the literature, which confirms the above mentioned positive presumptions of using machine learning in search.

2 The multi-agent learning approach

The multi-agent learning approach proposed here is inspired by a commonly used activity network representation, namely the activity on node diagram (AON). Each activity is represented by a node in a graph, and the edges represent the precedence relations between the activities.

Our goal is to generate an activity list (AL) and a mapping from activities to modes, i.e. a mode assignment (MA), which can later be used to construct a schedule. We do this by using a learning agent for each activity. These agents make their decisions using 2 learning automata (LA) [5]. An LA is a simple decision making device based on reinforcement learning. One LA is used for choosing a successor activity

	J10	J12	J14	J16	J18	J20
(Jedrzejowicz and Ratajczak-Ropel, 2007)	0.72	0.73	0.79	0.81	0.95	1.80
(Jedrzejowicz and Ratajczak-Ropel, 2006)	0.36	0.50	0.62	0.75	0.75	0.75
(Bouleimen and Lecocq, 2003)	0.21	0.19	0.92	1.43	1.85	2.10
5000 schedules:						
(Jozefowska et al., 2001)	1.16	1.73	2.60	4.07	5.52	6.74
(Alcaraz et al., 2003)	0.24	0.73	1.00	1.12	1.43	1.91
(Jarboui et al., 2008)	0.03	0.09	0.36	0.44	0.89	1.10
(Ranjbar et al., 2009)	0.18	0.65	0.89	0.95	1.21	1.64
(Van Peteghem and Vanhoucke, 2009)	0.02	**0.07**	**0.20**	0.39	0.52	0.70
(Lova et al., 2009)	0.06	0.17	0.32	0.44	0.63	0.87
(Van Peteghem and Vanhoucke, 2010)	**0.01**	0.09	0.22	0.32	**0.42**	**0.57**
Multi-Agent Learning Approach RD	0.05	0.08	0.23	**0.30**	0.53	0.70

Figure 1: Comparison with other approaches for the MRCPSP - Average deviation from optimal(%)

to forward the control to, and one LA is used for choosing the mode in which the activity will operate. An additional dispatcher agent is introduced for forwarding the control to all agents.

3 Results and discussion

For testing the performance of the algorithm, we applied it to instances of the project scheduling problem library (PSPLIB) [4], which is available from the server of the University of Kiel (http://129.187.106.231/psplib/). Figure 1 shows the comparison of the multi-agent learning approach with other state-of-the-art approaches for the MRCPSP. The numbers denote the average deviation from the optimal makespan in percent. The best results are marked in bold.

When considering these results for the MRCPSP, we can conclude that the multi-agent approach performs very well when comparing it to the state of the art methods from the literature. We even reach the best result for one dataset (J16). We also tested the new approach on the largest single-mode RCPSP instances. The single-mode results reveal average performance when compared to the best algorithms reported in the literature. We can conclude that the power of our approach is its coupling between learning the activity order and learning the modes. Note that although our approach is distributed, it does not require mutual communication between the learning automata. The coupling of the LA happens through the common global reward signal.

We have shown that the combination of reinforcement learning and optimization is definitely useful and that this combination will certainly be further investigated in the future. One of the next topics will be to test the scalability of our local learning approach.

Acknowledgment: This research was carried out within the IWT SBO DiCoMAS project (IWT060837).

References

[1] R Battiti, M Brunato, and F Mascia. *Reactive Search and Intelligent Optimization*, volume 45 of *Operations research/Computer Science Interfaces*. Springer Verlag, 2008.

[2] J Blazewicz, J K Lenstra, and A H G Rinnooy Kan. Scheduling projects subject to resource constraints: Classification and complexity. *Discrete Applied Mathematics*, 5:11–24, 1983.

[3] E Demeulemeester and W Herroelen. *Project scheduling: a research handbook*, volume 49 of *International Series in Operations Research & Management Science*. Kluwer, 2002.

[4] R Kolisch and A Sprecher. Psplib - a project scheduling problem library. *European Journal of Operational Research*, 96:205–216, 1996.

[5] M A L Thathachar and P S Sastry. *Networks of Learning Automata Techniques for Online Stochastic Optimization*. Kluwer Academic Publishers, 2004.

[6] T Wauters, K Verbeeck, G Vanden Berghe, and P De Causmaecker. Learning agents for the multi-mode project scheduling problem. *Journal of operational research society*, 2010 (To appear).