Learning Heuristic Selection in Hyperheuristics for Examination Timetabling

Edmund K. Burke[†]

Mustafa Mısır[*]

Gabriela Ochoa[†]

Ender Özcan[†]

[†] *Automated Scheduling, Optimisation and Planning Research Group, School of Computer Science & Information Technology, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, United Kingdom*

{ekb, gxo, exo}@cs.nott.ac.uk

[*] *Yeditepe University, Department of Computer Engineering, Inonu Mahallesi, Kayisdagi Caddesi, Kadikoy, Istanbul 34755, Turkey*

mmisir@cse.yeditepe.edu.tr

Introduction

This paper will consider a non-traditional solution methodology for examination timetabling. Over the years, many approaches have been proposed for solving exam timetabling problems (Qu et al., 2008). Burke et al. (2003) describes hyperheuristics as “(meta-)heuristics to choose (meta-)heuristics to solve the problem at hand”. A hyperheuristic can be thought of as operating at a *high level* by utilizing problem independent information to guide the search process over the heuristic space formed by a set of *low level heuristics*. Low level heuristics of course, can be perturbative or constructive. Hyperheuristics with perturbative low level heuristics using a single configuration during the search draw upon an iterative methodology (Ozcan, Bilgin and Korkmaz, 2006; 2008). At each step, the most suitable heuristic is chosen using a heuristic selection method and a new state is generated after the application of the selected heuristic. This move is either accepted or rejected based on a move acceptance strategy. The iterations continue until a termination criterion is met. This paper presents a hyperheuristic of this type that embeds a learning mechanism for heuristic selection which is analyzed over a set of exam timetabling problems.

Nareyek (2004) experimented with a number of hyperheuristics combining reinforcement learning schemes and selection methods for choosing a heuristic to apply while accepting all moves. The learning mechanism employs a weight adaptation approach based on positive reinforcement as a rewarding mechanism and negative reinforcement as a punishment mechanism. At each step, the utility value of a heuristic gets updated depending on its performance. A comparison of different choices in reinforcement learning shows that selecting a heuristic having a maximal utility performs the best when additive adaptation for rewarding is used with root adaptation for punishment.

Kendall and Mohamad (2004) experimented with a hyperheuristic that used *simple random* (SR) heuristic selection and a *great deluge* (GD) acceptance criterion. A great deluge approach directly accepts improving moves, while non-improving moves are accepted if the objective value of the candidate solution is better or equal to an expected value, called the *threshold level* that changes at each step. The objective value of the first generated candidate solution is used as the initial level and the level is updated at a linear rate towards a final objective value.

Problem formulation

In this study, a hyperheuristic methodology which integrates reinforcement learning based on different adaptation schemes with a great deluge method is presented to address the exam timetabling problem presented in Bilgin, Ozcan and Korkmaz (2007). The problem formulation includes the following hard constraints:

- *Exam conflict*: A student can not take two exams at the same time.
- *Seating restriction*: The number of students seated for an exam cannot exceed the room capacity.

If a student has to take two exams in the same day, then a single time slot between them is preferable. The weighted average of constraint violations (c) for each constraint type is used to evaluate the quality of a candidate solution (T):

$$eval(T) = \frac{-1}{1 + \sum_{\forall c} violations(c)\, w_c} \qquad (1)$$

The Reinforcement Learning – Great Deluge Hyperheuristic

Ozcan, Bilgin and Korkmaz (2008) showed that combining a different heuristic selection method with a different move acceptance might yield an improved performance. All reinforcement learning heuristic selection mechanisms choose a heuristic with the maximal utility value for invocation. Also, they employ an additive adaptation for rewarding as a positive reinforcement scheme that increases the utility value u_i of the i^{th} heuristic by 1 in case of improvement. This rate is tested against three different negative reinforcement schemes for punishment where each one reduces u_i at a different rate whenever a worsening move occurs:

$$RL_1 : u_i = u_i - 1 \tag{2}$$

$$RL_2 : u_i = u_i / 2 \tag{3}$$

$$RL_3 : u_i = \sqrt{u_i} \tag{4}$$

The utility values are bounded arbitrarily in [1, 10×*number of heuristics*] and always rounded to the nearest integer as in the study of Nareyek (2004). RL_1, RL_2 and RL_3 employ subtractive (by 1), divisional (by 2) and root negative adaptation rates.

Four low level heuristics are implemented (Bilgin, Ozcan and Korkmaz, 2007). The first three heuristics are related (each) to a different problem constraint. A set of constraint based neighbourhoods are searched by each heuristic (Alkan and Ozcan, 2003). These heuristics attempt to reschedule the exam producing the highest violation to their associated constraints. The exam selected is rescheduled to the best available period, where the period is selected following a tournament strategy. The fourth heuristic randomly selects and reschedules an exam (following a uniform distribution).

Results

The Reinforcement Learning-Great Deluge hyperheuristics are evaluated over an arbitrarily selected subset of the Toronto benchmark data sets (Carter, Laporte and Lee, 1996) and compared to one of the top performing hyperheuristics, SR–GD in Bilgin, Ozcan and Korkmaz (2007). Table 1 summarizes the experimental results.

Embedding a learning mechanism might improve or worsen the performance of a hyperheuristic for a given problem. Determining the best adaptation rates seems to be a key issue in fully utilizing a reinforcement learning scheme within a hyperheuristic. Different adaptation rates might yield different performances. The results show that the RL_1 heuristic selection method delivers the best average performance when combined with the great deluge method as a hyperheuristic (Table 1).

Table 1 Performance comparison of reinforcement learning hyperheuristics with different negative adaptation rates. The rank of each approach for each problem instance is computed using its average best fitness over 50 trials

problem	exams	density	RL_1–GD	RL_2–GD	RL_3–GD	SR–GD
sta83 I	138	0.14	1	2	4	3
car92 I	543	0.14	1	4	3	2
kfu93	461	0.06	1	3	4	2
rye92	486	0.07	1	3	2	4
pur93 I	2419	0.03	2	3	4	1
		avr	1.20	3.00	3.40	2.40
		std	0.45	0.71	0.89	1.14

References

Alkan, A., Ozcan, E. (2003) Memetic Algorithms for Timetabling, Proc. of 2003 IEEE Congress on Evolutionary Computation, pp. 1796–1802

Bilgin, B., Ozcan, E., Korkmaz, E. E. (2007) An Experimental Study on Hyper-Heuristics and Exam Scheduling, PATAT2006, Springer-Verlag, selected papers, LNCS 3867:394–412

Burke, E., Kendall, G., Newall, J., Hart, E., Ross, P., Schulenburg, S. (2003) Hyper-heuristics: an Emerging Direction in Modern Search Technology, in Glover and Kochenberger (eds) Handbook of Metaheuristics, Kluwer International Publishing, vol. 57, pp. 457–474

Carter, M. W., Laporte, G., and Lee, S.T. (1996) Examination Timetabling: Algorithmic Strategies and Applications. Journal of the Operational Research Society 47:373–383

Kendall, G., Mohamad, M. (2004) Channel assignment in cellular communication using a great deluge hyper-heuristic. In: Proc. of the 12th IEEE Int. Conf. on Networks, v. 2, pp. 769–773

Nareyek, A. (2004) Choosing search heuristics by non-stationary reinforcement learning. In: Metaheuristics: Computer Decision Making, Kluwer Academic Publishers, pp. 523–544

Ozcan, E., Bilgin, B., Korkmaz, E. E. (2006) Hill Climbers and Mutational Heuristics in Hyperheuristics, LNCS 4193, Springer-Verlag, The 9th International Conference on Parallel Problem Solving From Nature, pp. 202-211.

Ozcan, E., Bilgin, B., Korkmaz, E. E. (2008) A comprehensive analysis of hyper-heuristics. Intelligent Data Analysis, 12:1, 3–23

Qu, R., Burke, E. K., McCollum, B., Merlot, L. T. G., Lee, S. Y. (2008) A survey of search methodologies and automated approaches for examination timetabling. Journal of Scheduling, to appear.