A Review of Doctoral Thesis
„Combinatorial algorithms for online problems: Semi-online scheduling on related machines“
by Mgr. Tomáš Ebenlendr

The thesis under review deals with one-by-one online scheduling problems with uniformly related machines. In particular, author suggests a novel deterministic algorithm for solving pre-emptive problems where the competitive ratio is given as a parameter. Then he shows how to compute optimal competitive ratio for various semi-online scheduling problems using linear programming techniques, namely when we know the sum of processing times or the maximal processing time, approximately know the optimal makespan, or the processing times are close to each other. For problems with a small number of machines the author also shows how to solve the linear program symbolically. Finally, a new lower bound for competitive ratio on deterministic algorithm solving the non pre-emptive problem is given.

I like the overall concept as well as the deterministic optimisation algorithm. The thesis shows that beyond the initial idea the student had to do a lot of technical work, though it seems that most of this work is mainly technical with limited innovation. The results were published at peer-reviewed international journals and conferences of good quality which justifies their importance. On the other side, I do not like the style of presentation of the results with a non-compact structure of thesis, missing motivation and explanation, and small mistakes that makes reading the text hard.

Doctoral thesis consists of three chapters that are only loosely connected. Introduction is not really an introduction to the area that motivates the research and explains the used notions to people not directly working on these specific problems. Some notions are re-introduced again at the beginning of Chapter 2, actually, whole section 2.1.3 seems to be the introduction of the topic again. Chapter 2 presents the main results and Chapter 3, though it is about a non-preemptive version of the problem, has no relation to techniques studied in Chapter 2 so it is not clear to me why this section is included. After Chapter 3, the text suddenly stops without any conclusion, summary of achieved results and description of possible future works. This is something that I have never seen before.

Scheduling is a very practically important area, but there is almost no motivation why the studied problems are of practical importance. There are several variants of semi-online scheduling problems described in the paper, but it would be also useful to see where the problems appear in practice. It is interesting to study the theoretical properties of abstract problems, but it is even more interesting to see that the abstractions are practically useful.

The style of text makes it hard to follow what the author wants to say. Definitions and propositions are sometimes placed in the text without telling the reader what the proposition/definition is going to say. Sometimes the used notation is even not introduced. To be more specific, the “introduction” of lemma 2.1.2 (p.15) says “The following lemma is due to Epstein and Sgall. We include the proof as it is helpful in understanding our results” so what is the lemma about? Only later the reader can realise that the lemma can actually be used to compute the lower bound for the competitive ratio. Moreover, the proof contains mistakes which again make it hard to understand it. If T_i is the time when *at most* i machines are running then how can be that at *least* i jobs are running at time T_i if

each machine can process at most one job at given time? What do you mean by schedule J_i? Is it the same as $J_{[i]}$?

When explaining algorithm RATIOSTRECH, there is again notation never introduced or explained before. What is R(M) and r(m) at page 19? It has never been used before. I also miss the definition of $W_k(t_j)$ which is used in the algorithm and later in the text. How is $W_k(t_j)$ computed? The critical parts of text explaining the algorithm (p. 19) are unclear to me. What do you mean by "We want to preserve the fastest virtual machine as fast as possible"? If we want to preserve something then we are not changing it so how can we preserve something as fast as possible? Similarly, how can we "optimise the speed of the fastest machine"? The speed is given so how can we "optimise" it? Honestly, I was not able to decode the following sentence at all "Again we want to save as large part of the fastest machine as possible, thus the job finishes exactly at given time being processed by either of these two virtual machines for each time $\tau < T$." There is a mistake in algorithm at line 10, there should be $W_{k+1}(t_j)$ at the beginning. What I really miss is the explanation how time t_j is found (of course I also miss the definition of $W_k(t_j)$ itself). This is never explained in the text.

Last but not least, there are no examples and only one figure (p. 23) explaining the underlying ideas. I suggest the author to pay much more attention to the reader by including some explanation of the work in addition to putting just the technical content to the thesis.

In summary, the thesis brings some interesting novel results that were published at international journals and conferences. The achieved results show that the student is capable of doing independent research. However, the presentation style should be improved significantly. I suggest that the work is accepted as Doctoral Thesis.

Doc. RNDr. Roman Barták, Ph.D.
Charles University in Prague
Faculty of Mathematics and Physics

Prague, May 12, 2011