Christoph Dürr
CNRS, CMM (*)
University of Chile
christoph.durr@CNRS.fr

Report on the PhD manuscript *Scheduling with uncertain processing times given by empirical distributions* by Antonín Novák

Scheduling is a wide range of optimization problems, where the goal is to plan the execution of tasks/jobs under some constraints, minimizing some cost function. In practice a main difficulty with such problems is that exact data is hardly available. This creates the need to study scheduling models, where only some distributional knowledge is available on the job processing times or other job characteristics. Antonín Novák's PhD studies exact those models and provides interesting results, which are meant to have a broad impact in the scientific community.

The manuscript contain 4 chapters, each is dedicated to a specific published work. The publications made it into high standard venues, 5 papers in European Journal of Operational Research, Annals of Operations Research, as well as conferences like the International Joint Conferences on Artificial Intelligence Organization for example.

Chapter 2 studies a very fundamental scheduling problem, where n jobs need to be assigned to m parallel identical machines, and ordered on each of the machines. The processing times are independent random variables. The goal is to minimize the expected total completion time of the jobs (or equivalently the expected average completion time). In this setting is enough to assign jobs to ranks, provided that at most m jobs are assigned to the same rank. The interpretation is that the rank 1 job, will complete last on its machines, the rank 2 job second last and so on. Hence the objective value is the sum over all the jobs of their rank multiplied with their processing time. If the random variables would follow a given distribution, one could just use linearity of expectation and work with the expected processing times. This becomes an easy and known problem, but might be unrealistic. Instead a set of distributions is given, and the goal is to optimize the expected objective value over the worst distribution from the set. Now the problem becomes interesting, but at the same time computationally intractable in practice. First the problem is formulated as minimization of a linear function with an additional robust term in the sense of the ℓ_2 norm. The form of the robust term influences the computational difficulty of the problem, and this is studied in this chapter. Improved upper bounds are given when jobs are independent. Then the problem is studied when there is dependency and is shown that it is negative correlation which makes the problem hard to solve. In addition is shown that the problem can be solved in polynomial time, using the robust term in the sense of the ℓ_1 norm. Multi-objective optimization is also considered in this chapter, focusing on expected quality and variance. An improved method to sample the Pareto set is proposed. The chapter ends with experiments, showing the superiority in running time of the proposed polynomial approximations, compared to a known second order cone integer programming formulation.

Chapter 3 concerns a scheduling problem with mixed criticality. In the considered scheduling model, one has to decide starting times of given jobs on a single machine, minimizing the makespan. Their exact processing time depends on a external discrete criticality variable, which can vary over time and is revealed only during the job execution time. The idea is that with higher criticality, some job processing times become longer, while some other jobs don't need to be processed anymore, which creates room for the longer jobs. For example when there are only two critical levels, jobs divide into low-critical jobs, which are not executed if the level becomes 2, and high-critical jobs, which have short processing time at level 1 and longer processing time at level 2. The difference in processing time can be seen as a sort of *capacity*. The idea of a good schedule is to assign the low-critical jobs to high-critical jobs, so that total processing time assigned to a high-critical job does not exceed its capacity, or as little as possible. Techniques to attack this problem have a packing problem type flavor. First Antonín shows that some greedy assignment is a 3/2 approximation when some inequality holds: no low-critical job size should not exceed a high-critical job capacity. Then they propose a mixed integer linear program formulation, which essentially specifies for each high-critical job, which low-critical jobs are assigned to it. It groups low jobs by their processing time, and specifies only how may jobs from each group are selected. A branch-and-price decomposition is proposed, based on a configuration linear program, and branching is done by interpreting fractional variables as probabilities. These techniques are extended to the case of three critical levels, where a schedule is obtained by grouping jobs in form of a tree, by assigning level 1 jobs to level 2 jobs, which in turn are assigned to level 3 jobs. The chapter ends with numerical experiments, on synthetic data, drawn from a distribution specified in this chapter. Most of the time the mixed integer programming formulation beats the branch-and-price approach. The conclusion contains research direction building on the proposed solutions.

In Chapter 4 instead Antonín studies a problem where the schedule is actually given to you, together with probability distributions for each job on the criticality level in which it will execute. Moreover in this schedule jobs can appear several times, augmenting the chance that they will eventually executed. So when the maximum criticality level is constant as well as the number of replications per job, then they can compute the expected number of executed jobs. However in general the problem is #P-hard. These are the main results presented in this chapter. For the hardness first it is shown than deciding if some given job has a positive probability to be scheduled is NP-hard, by a reduction from 3-SAT. The reduction uses a very pretty gadget, and can either work with constant criticality levels or constant job replications (but then reduces only from 2-SAT). As a result, computing the actually probability that a job is executed is #P-hard already when one of these parameters is constant. When both parameters are constant, a Bayesian network approach allows to compute the expected number of job executions efficiently. For this every job replica is modeled by a random variable, indicating its critical execution level, or indicating that its execution is rejected by a higher critical job, or that a previous replica has already been executed. Conditional probability tables relate the random variables with each other.

Finally Chapter 5 is devoted to parallel machine scheduling. It studies a similar model from the second chapter, namely the problem of scheduling jobs with uncertain processing time on identical parallel machines. But now a common deadline is given, and the goal is to maximize the probability that all jobs finish within the given time window. Job processing times follow a normal distribution with a given mean and variance. For this problem we have to partition the jobs, each set will be executed on one machine. This can be modeled by a linear program with binary decision variables, however the non-linear objective function makes it hard to solve. For this problem Antonín gives improved lower and upper bounds. The later uses the idea of contracting k parallel machines into a single machine running k times quicker. Also he provides a branch-and-price method to solve the problem and analyzes the underlying pricing problem. Finally he shows experimental results improving over a paper from 2012, and provides an exact algorithm for the special case of two parallel machines.

The manuscript wraps up with a conclusion, providing a roadmap for future work. In this manuscript Antonín Novák presented a selection of his main results obtained during the PhD, which give the impression of a round and complete work, involving quite a variety of different techniques. The contributions are strong, and the studied problems fundamental and important. Clearly, the author of the thesis proved to have an ability to perform research and to achieve scientific results. Antonín will do a brilliant academic career. In my opinion, this excellent work constitutes a valuable contribution to research, and I recommend awarding the author with a scientific degree Ph.D.

Christoph Dürr – Santiago, November 16th, 2022

(* From December 12th, 2022, Christoph Dürr will be affiliated with CNRS, Sorbonne University, Paris, France)