A Reinforcement Learning Environment For Job-Shop Scheduling

Pierre Tassel, Martin Gebser, and Konstantin Schekotihin
Department of Artificial Intelligence and Cybersecurity, University of Klagenfurt

Introduction

- Scheduling is a **fundamental** task occurring in various automated systems applications
- We present an **efficient** environment to learn to solve job-shop scheduling

Contribution

- **Optimized** environment to solve the job-shop scheduling problem
- **Compact** yet **meaningful** state representation
- **Dense reward function,** correlated with the sparse make-span minimization objective

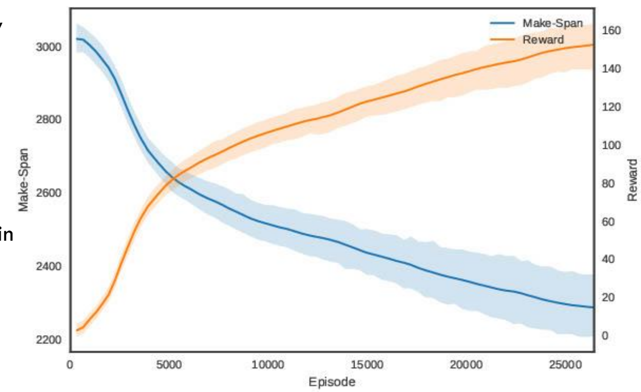
Evaluation setup

- Instances: Taillard's and Demirkol's **30 jobs** and **20 machines**
- Train using **PPO** algorithm for **10 minutes**
- Compare our approach against
  - the most widely used dispatching rules
  - a state-of-the-art CP solver

Results

Our **dense** reward is highly correlated with **the sparse objective function** used to evaluate a solution.

The dense reward tries to **maximize machine utilization,** avoiding gaps in the planning.



The **best solution make-span** for each approach per instance:

Dataset	Instance	Ours	FIFO	MWKR	(Zhang et al. 2020)	(Han and Yang 2020)	OR Tools	Upper Bound
Taillard	TA41	**2208**	2543	2632	2667	2450	2144	2005
	TA42	**2168**	2578	2401	2664	2351	2071	1937
	TA43	**2086**	2506	2385	2431	—	1967	1846
	TA44	**2261**	2555	2532	2714	—	2094	1979
	TA45	**2227**	2565	2431	2637	—	2032	2000
	TA46	**2349**	2617	2485	2776	—	2129	2004
	TA47	**2101**	2508	2301	2476	—	1952	1889
	TA48	**2267**	2541	2350	2490	—	2091	1941
	TA49	**2154**	2550	2474	2556	—	2089	1961
	TA50	**2216**	2531	2496	2628	—	2010	1923
	Average	**2203**	**2549**	**2449**	**2604**	**—**	**2058**	**1948**
Demirkol	DMU16	**4188**	4934	4550	4953	4414	3903	3751
	DMU17	**4274**	5014	4874	5379	—	3960	3814
	DMU18	**4326**	4936	4792	5100	—	4073	3844
	DMU19	**4195**	4902	4842	4889	—	3922	3764
	DMU20	**4074**	4539	4500	4859	—	3913	3703
	Average	**4211**	**4865**	**4712**	**5036**	**—**	**3954**	**3775**

Conclusions

- Our environment yields **excellent** performance compared with other **RL and non-RL approaches.**
- Having a **dense reward** correlated to the **objective sparse reward** will help future work to improve the agent's performance further, removing one of the obstacles to solving this problem with RL.
- This environment is also more complete than the previously proposed models as it allows the agent **not to schedule any operation** at a given time step.

Literature cited

Zhang et al, 2020, Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning.

Han et al, 2020. Research on Adaptive Job Shop Scheduling Problems Based on Dueling Double DQN.

Further information

Full approach code is available here: https://github.com/prosysscience/RL-Job-Shop-Scheduling

The gym environment is available as a pip package:

pip install JSSEnv