# Learning to schedule from demonstrations: What we lose by only imitating the best?

Early success of Deep Reinforcement Learning (DRL) was rooted in arcade and board games, where expert behavior could be readily captured from top players. In these settings, demonstrations were used to bootstrap learning and accelerate policy convergence. In contrast, in combinatorial optimization problems, such as the Flexible Job-shop Scheduling Problem (FJSP), optimal demonstrations are costly to obtain. In this work, we build on a state-of-the-art DRL framework to investigate how the quality and diversity of demonstrations from FJSP solutions affect learning dynamics and policy generalization. We argue that representativity of the action space is more beneficial for pretraining than strict optimality. To that end, we consider an efficient Constraint Programming (CP) method and several composite heuristic rules as candidate experts. These were evaluated based on the final policy performance, the generalization to unseen instances, and the time required to gather expert FJSP solutions. Preliminary results show that agents pre-trained with diverse sub-optimal demonstrations converge faster to near-optimal policies than those trained solely on solver-based solutions. Moreover, combining CP and heuristic demonstrations leads to superior robustness to unseen instances. These findings suggest that diversity and representativeness in expert behavior may be more critical than optimality alone.

**Author:**   JESUS, Alexandre (Universidade de Coimbra)

**Co-authors:**   CORRÊA, Arthur (Departamento de Engenharia Mecânica, CEMMPRE, ARISE, Universidade de Coimbra, Coimbra, Portugal); SILVA, Cristóvão (Departamento de Engenharia Mecânica, CEMMPRE, ARISE, Universidade de Coimbra, Coimbra, Portugal); MONIZ, Samuel (Departamento de Engenharia Mecânica, CEMMPRE, ARISE, Universidade de Coimbra, Coimbra, Portugal)

**Presenter:**   JESUS, Alexandre (Universidade de Coimbra)

**Session Classification:**  Session 2.3 –Scheduling

**Track Classification:**  Session 2.3 –Scheduling