COMPUTER SCIENCE

Introducing Deep Reinforcement Learning Framework for Intelligent Baseball Scheduling

Peiyuan Xu, Antwan Clark, Anton Dahbura

Johns Hopkins University | Whiting School of Engineering | Baltimore, MD

Design Day 2023

Abstract

Scheduling sports matches is still a very challenging problem. The goal of sports scheduling is not only to create fair and balanced baseball schedules but also to satisfy various kaleidoscopic constraints. While advancements have been made through optimization approaches such as semi-definite programming, hybrid approaches, and heuristic methods, these approaches are limited to fixed scenarios and do not completely consider real-time scheduling demands. Motivated by these limitations, our work consists of the following:

1) Developed a custom constraint library for the required and desired constraints
2) Proposed a reinforcement learning framework between the baseball league schedule and the RL agents to create satisfying baseball scheduling schedules, which is based off a Markov Decision Process.

Our simulative experiments with specific constraints of the 2022 High A Central season suggests that the RL agents were able to achieved consistently high schedule satisfaction score (over 70%) on average after training with Deep Q Networks.

Introduction

Scheduling for baseball leagues need to respect various constraints organized into the following categories:

1. *Required (or immovable) constraints* – Constraints that must be satisfied to produce workable schedules

2. *Desired (or movable) constraints* – Constraints that don't need to be satisfied to produce workable schedules

Each of these categories include other types of constraints that include 1) uniqueness of home play; 2) number of games played in a specific region; 3) how opponents are scheduled; and 4) schedule fairness

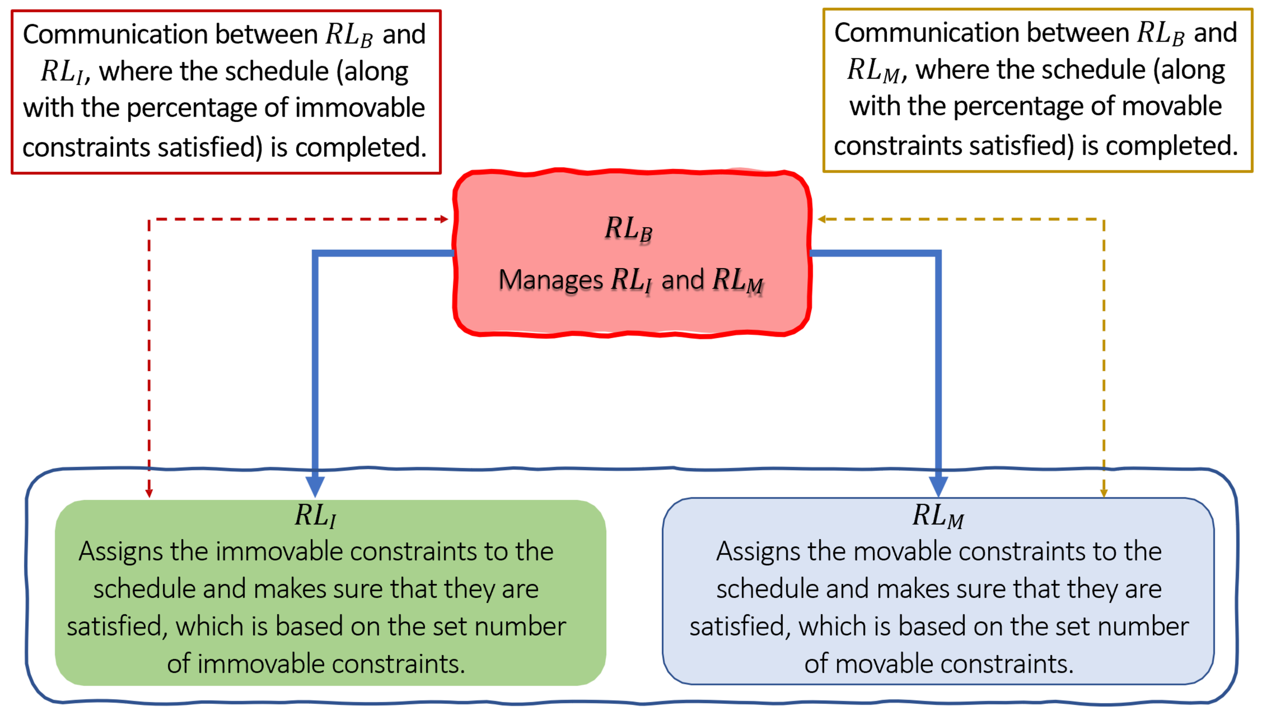
Previous advancements have been made through combinatorial optimization approaches such as semi-definite programming, hybrid approaches, and heuristics have been used. But they approaches do not completely consider the real-time scheduling demands and changes that are typically league-specific.

Theoretical Development

Our process is based on the following Markov Decision Process (MDP):

- The state space $\boldsymbol{S}$ includes the number and type of constraints added and satisfied to the baseball schedule
- The action space $\boldsymbol{A}$ includes the number and type of constraints added, removed, modified, or substituted to the baseball schedule.
- The rewards $\boldsymbol{R} = \{-\mathbf{2}, -\mathbf{1}, \mathbf{0}, +\mathbf{0.5}, +\mathbf{1}\}$ are assigned based on how the agents modify and make corrections

Reinforcement Learning Framework



Deep Reinforcement Learning Framework

Testing Strategy

The goal of our testing strategy is to robustly test the proposed RL framework consisting of the following objectives.

1. To ensure that each respective subordinate RL agent is doing its proposed job in terms of application the immovable and movable constraints.

2. To ensure that the subordinate RL_I and RL_M agents are effectively interacting with the RL_B agent while generating effective schedules

We test this on different combinations of immovable and movable constraints from the 2022 High A Central Baseball League requirements as shown in the table below

Scenario	No. Immovable Constraints	No. Movable Constraints	Number of Constraint Combinations	Number of Constraint Combinations with at least one unsatisfactory constraint
I	2	6	147	126
II	3	6	245	231
III	4	6	245	238
IV	5	3	735	735
V	5	5	441	147
VI	6	2	147	441
VII	6	6	49	49
VIII	7	7	1	1

Results

Scenario	No. Immovable Constraints	No. Movable Constraints	Average Immovable CSP achieved	Average Movable CSP achieved
I	2	6	0.612	0.156
II	3	6	0.499	0.171
III	4	6	0.368	0.157
IV	5	3	0.277	0.150
V	5	5	0.268	0.154
VI	6	2	0.281	0.157
VII	6	6	0.235	0.146
VIII	7	7	0.429	0.0

Table I: Testing Results for the Sampling Actions from Markov Decision Process

Scenario	No. Immovable Constraints	No. Movable Constraints	Average Schedule Score	Average Immovable CSP achieved	Average Movable CSP achieved
I	2	6	0.825	0.850	0.765
II	3	6	0.807	0.826	0.763
III	4	6	0.756	0.747	0.777
IV	5	3	0.803	0.796	0.819
V	5	5	0.836	0.796	0.819
VI	6	2	0.781	0.829	0.854
VII	6	6	0.812	0.833	0.761
VIII	7	7	0.5	0.714	0.0

Table II: Testing Results for the Deep Q Network

These tables present simulative results when testing our proposed RL framework. It is important to note that after training with Deep Q Network, the agents can achieve consistently high schedule satisfaction score over 70%. It's also worth noting that the consistently high average immovable and movable CSP values attained by the RL_I and RL_M agents respectively, that significantly outperforms the immovable and movable CSP values achieved by random sampling actions from the Markov Decision Process without training. This suggests that the agents are learning to modify and make corrections to account for infeasibility through enough experience.

Conclusion

In this work, we proposed a deep reinforcement learning framework for intelligent baseball scheduling where experiments with the High A Central League demonstrate promise in our approach. We plan to further improve the framework to help the agents better adapt to new scenarios while also improving the computational efficiency in scheduling more complex baseball leagues and processes.