Flatland

Murphy

October 16, 2023

Contents

1 Introduction

1.1 Problem Background

The Flatland framework seeks to address the problem of automated train scheduling and rescheduling, a major challenge for modern railway systems. It does so by providing a simplified two-dimensional grid world environment to allow for fast experimentation of new approaches to this problem Mohanty et al. (2020).

1.2 Related Works

Scenarios in Flatland exist at the intersection of several well-explored problems. At its core, Flatland is a multi-agent pathfinding problem; several agents cooperate to complete their goals while managing limited resources within a shared, finite environment. Outside of pathfinding, Flatland is, in essence, a vehicle rescheduling problem—derived from the vehicle scheduling problem. An understanding of how each of these problems influences Flatland is a critical piece of finding ideal methods of producing solutions.

1.2.1 Multi-agent Pathfinding

Multi-agent pathfinding (MAPF) is a planning problem in which agents in a shared environment must find routes to their respective destinations without incurring collisions (Silver, 2005). MAPF has many applications, including in robotics, aviation, and vehicle routing (Standley, 2010). Many of the conflicts imposed on agents in MAPF problems are also imposed on agents within the Flatland framework, such as that they may not occupy the same cell at the same time step, or that two agents may not swap positions. These are to model collisions that would occur in real life.

Traditional grid environments seen in many MAPF problems, including in (Standley, 2010), have cells that are often four- or eight-connected; this means that an agent occupying one cell may move to any of its existing unoccupied neighboring cells. The Flatland framework is more restrictive in this sense, as agents are not free to move to any unoccupied neighboring cell, but rather may move to neighboring cells according to transitions governed by the track type of that cell and the orientation of the agent at that time step, all of which is further discussed in Section 3.1 and can be seen more clearly in Figure 1.2.1.

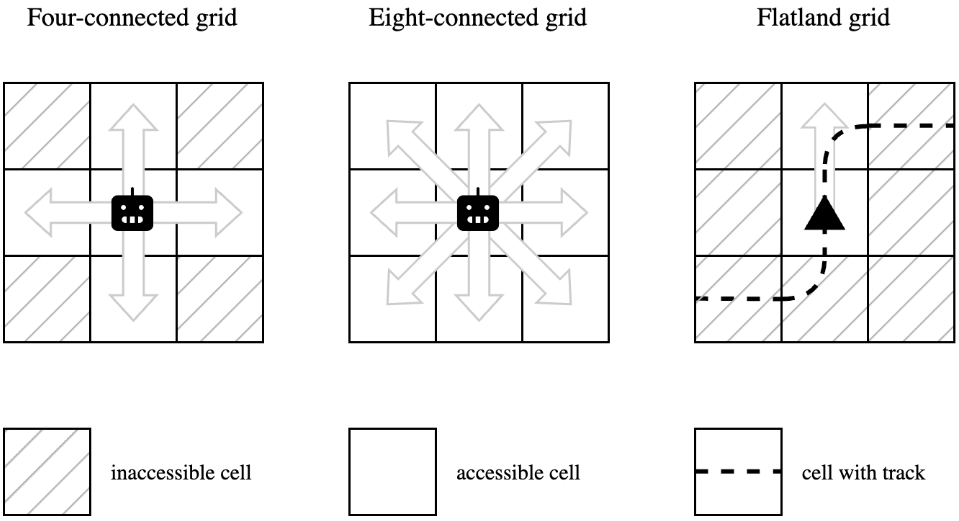

Figure 1: Examples of the degrees of freedom for agents in four-connected grid environments, eight-connected grid environments, and Flatland grid environments. The orientation of the agent in Flatland is necessary to determine which cells are accessible.

1.2.2 Vehicle Scheduling Problem

The vehicle scheduling problem (VSP) is a classical optimization problem in the field of operational planning of public transportation systems, consisting of assigning vehicles to trips, in such a way that each trip is associated to one vehicle and a cost function is minimized (Baita et al., 2000). This problem has been a topic of operational research for decades, but is not the primary focus of Flatland, as each instance already provides a set of agents with pairs of starting and ending positions. Since the agents come pre-assigned to the origin and destination, the importance is shifted from assigning vehicles to trips to actively determining their paths. However, since Flatland randomly inflicts breakdowns upon its fleet, the task effectively transforms into a vehicle rescheduling problem.

1.2.3 Vehicle Rescheduling Problem

The vehicle rescheduling problem (VRSP) is an extension of the VSP and arises when a previously-scheduled trip is disrupted due to interruptions such as a medical emergency or a vehicle breakdown (Li et al., 2007). Trips in the Flatland environment model this scenario by randomly assigning vehicle breakdowns, each of which stops a train in its current location for an unforeseen duration. Ideally, scheduled trips that are affected by a breakdown should be rescheduled in such a way that there are minimal impacts to the original plan. (Li et al., 2007) note the severity of this problem, in that it results not only in direct operational and delay costs for transit providers, but also in inconvenience for passengers. The lack of automated rescheduling policies, as well as algorithms that address this problem, highlights the importance of the role that Flatland plays in focusing its scenarios on the element of rescheduling disrupted trips.

2 Approaches

AIcrowd hosts a global competition in which participants submit crafted approaches to scenarios in the Flatland environment, which are scored and compared against one another. The variety of submissions highlights the diverse nature of approaches taken to solve the problem Flatland presents.

2.1 Reinforcement Learning

Reinforcement learning is a form of machine learning in which an agent with a goal is not told what steps to take, but rather discovers a path to it by determining which actions yield the greatest reward (Sutton and Barto, 2018). This differs from *supervised learning* in which agents are shown a set of labeled examples of desired outcomes so that it can generalize for unseen situations. This also differs from *unsupervised learning* in which agents seek to uncover structures or patterns within a set of unlabeled examples. A typical reinforcement learning approach in a Flatland scenario would include many iterations of agents relying on past actions that have yielded success and exploring new actions that aim to get them closer to the goal.

Although the competition is open to approaches from any domain, it has branded itself as a competition for multi-agent reinforcement learning. In the latest two iterations of the competition, separate leaderboards were posted for those pursuing the reinforcement learning track and those pursuing all other approaches. Despite this focus, the highest any participant has placed with a reinforcement learning solution has been seventh place.

2.2 Operations Research

Operations research is a broader field pertaining to the application of scientific methods to problems involving the operations of a system, such as by using the theories of probability, linear programming, queuing theory, or other methods (Gupta, 1992). The winners of the 2020 iteration of the competition were a team led by professors Daniel Harabor and Peter J. Stuckey of Monash University. **under construction** The team [utilized an conflict-based search (CBS) approach... it splits the search tasks into low-level pathfinding for individual agents and high-level conflict detection to prevent collisions for all agents] (Li et al., 2021).

In each the latest two iterations of the competition, teams with operations research-focused approaches have taken the top four spots. The winning team members from 2020 explain in (Li et al., 2021) why they believe reinforcement learning approaches have not produced the same level of success as other methods: **under construction**

- RL approaches need to predict future deadlocks, which seems difficult without directly reasoning about paths, as optimization approaches do
- as the density of trains grows, the number of situations that can lead to deadlocks increases non-linearly
- optimization approaches rely on global planners and have consistently outperformed RL approaches across all rounds of both the 2019 and 2020 Flat- land Challenges

2.3 Answer Set Programming

Answer set programming (ASP) is a form of declarative programming, based on the stable model semantics of logic programming, oriented toward difficult search problems (Lifschitz, 2019). ASP has wide-ranging applications in technology and science, with a particular inclination toward NP-hard[1]problems.

The pathfinding problem that is present in Flatland scenarios is, at its core, an NP-hard search problem. ASP presents itself as a particularly suitable approach to solving pathfinding problems in Flatland for several reasons:

[1]NP-hard represents a category of problems in computer science whose answers can neither be solved quickly nor verified quickly. *Quickly* is a relative term for the speed of an algorithm that can solve problems in polynomial time, a duration considered reasonably efficient.

- as a declarative language, encodings are constructed in terms of their goals, not their intermediate steps
- solutions are entirely transparent and can be clearly explained
- encodings can be easily adapted to changing conditions, such as by adding more agents or including additional constraints

under construction Text.

[Concerns or drawbacks about traditional machine learning methods compared to ASP]

3 Building Blocks

Flatland comprises a set of components that are fundamental to its ability to simulate real-world railway scenarios. In the forefront are the environment and the agents, which form the basis of modeling how trains traverse a physical network.

Further components include the scope of observation, the presence of breakdowns, and pre-determined timetables. However, these are not considered within the scope of this research.

3.1 Environment

An environment in Flatland is a grid that consists of cells that are either empty or contain tracks. A track may belong to one of nine types, and may be rotated in increments of 90º. Cells must be arranged in such a way that they form a cohesive network, even if the resulting network is simplistic. Each cell may only be occupied by a single agent at any given time step. ~~[Transitions are symmetrical]~~

3.1.1 Track Types and Transitions

Flatland documentation considers seven track types, however, two of them have alternative representations that this paper recognizes as individual types, bringing the total to nine. There is also a separate designation for an empty cell.

The various rail types can be grouped into two categories: non-switches and switches. Non-switches include tracks such as straight tracks, curved tracks, and dead ends. [They do not allow the course of the path to change.]

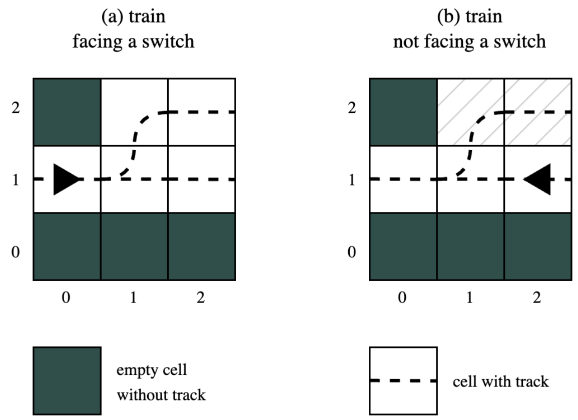

Figure 2: The switch on cell $(1,1)$ is only accessible from one side. In environment (a), the agent is facing east on cell $(0,1)$. When it reaches cell $(1,1)$, it will have the option to continue to the east or follow the switch to the north. In environment (b), the agent is facing west on cell $(2,1)$ and does not have access to the switch. Therefore, when it reaches cell $(1,1)$, it will only have the option to continue moving west.

Conversely, switches represent a decision point, as they force agents to make a choice. A switch will never present more than two options.

Each rail type determines the possible transitions for the agents, meaning which neighboring cells are accessible positions following a single move. Crucially, neighboring cells that are accessible from one location, may in some cases not be accessible had the train approached this same location from another direction. For this reason, recording both the position and orientation of each agents at every time step is necessary for determining its legal paths. A diagram explaining this quality is shown in Figure 3.1.1.

3.1.2 Points of Interest

under construction [Stations and cities]

3.2 Agents

Agents in Flatland represent the trains of a railway network. They can choose from a set of actions, are given starting locations and specified destinations, must avoid collisions and other conflicts with agents present throughout the environment, and have various speed profiles. The terms *agent* and *train* are used interchangeably.

3.2.1 Actions

Within the environment, agents follow a series of actions. Since Flatland is a discrete time simulation, the duration of each action conforms to a constant amount of time. At each time step, agents must choose from one of the following five actions:

1. `MOVE_FORWARD`: this action moves the train forward, provided there is a legal transition that allows this; this is also the appropriate action along curved track with no switch

2. `MOVE_LEFT`: this action moves the train along a switch to the left, provided there is a legal transition that allows this

3. `MOVE_RIGHT`: this action moves the train along a switch to the right, provided there is a legal transition that allows this

4. `STOP_MOVING`: this action stops the train, resulting in its remaining in the current cell

5. `DO_NOTHING`: this action compels the train to adhere to a continuation of its previous action; the one exception is that while stopped in a dead end, this reverses the direction of the train

So long as a valid action has been selected, the position and orientation of the agent will be updated in the following time step.

3.2.2 Starting and Ending Positions

Each agent is given a starting position and orientation, as well as a destination. The goal for a single agent is to traverse the environment by choosing valid actions along legal paths that lead the agent from its starting position to its destination.

For simplicity, trains are not recognized by the environment until they are actively underway. This prevents trains who have not yet begun their journeys from occupying space on the tracks and preventing the passage of other trains. Likewise, once a train reaches its destination, it is no longer recognized as actively being present in the environment. The track space therefore becomes free in the time step after the train has completed its journey.

3.2.3 Conflicts and Speed Profiles

Consistent with core tenets of any MAPF problem, any two trains must avoid a collision with each other. Agents, as with trains in real life, may not physically pass through one another. This prevents one train from overtaking a stopped train on the same set of tracks; it also prevents two trains facing each other from swapping positions.

Furthermore, agents in Flatland may have different speeds, depending on their train types, such as passenger trains or freight trains. This speed profile determines how quickly agents are capable of traversing the environment, and ultimately influences how they interact with one another. **under construction** [This means] that a train cannot travel faster than another train traveling immediately before it.

References

F. Baita, R. Pesenti, W. Ukovich, and D. Favaretto. A comparison of different solution approaches to the vehicle scheduling problem in a practical case. *Computers & Operations Research*, 27(13):1249–1269, 2000. ISSN 0305-0548. doi: https://doi.org/10.1016/S0305-0548(99)00073-8. URL `https://www.sciencedirect.com/science/article/pii/S0305054899000738`.

RK Gupta. *Operations research*. Krishna Prakashan Media, 1992.

Jiaoyang Li, Zhe Chen, Yi Zheng, Shao-Hung Chan, Daniel Harabor, Peter J. Stuckey, Hang Ma, and Sven Koenig. Scalable rail planning and replanning: Winning the 2020 flatland challenge. *Proceedings of the International Conference on Automated Planning and Scheduling*, 31(1):477–485, May 2021. doi: 10.1609/icaps.v31i1.15994. URL `https://ojs.aaai.org/index.php/ICAPS/article/view/15994`.

Jing-Quan Li, Pitu B. Mirchandani, and Denis Borenstein. The vehicle rescheduling problem: Model and algorithms. *Networks*, 50(3): 211–229, July 2007. doi: https://doi.org/10.1002/net.20199. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/net.20199`.

Vladimir Lifschitz. *Answer set programming*. Springer Heidelberg, 2019.

Sharada Mohanty, Erik Nygren, Florian Laurent, Manuel Schneider, Christian Scheller, Nilabha Bhattacharya, Jeremy Watson, Adrian Egli, Christian Eichenberger, Christian Baumberger, Gereon Vienken, Irene Sturm, Guillaume Sartoretti, and Giacomo Spigler. Flatland-rl : Multi-agent reinforcement learning on trains, December 2020.

David Silver. Cooperative pathfinding. In *Proceedings of the aaai conference on artificial intelligence and interactive digital entertainment*, volume 1, pages 117–122, 2005.

Trevor Standley. Finding optimal solutions to cooperative pathfinding problems. *Proceedings of the AAAI Conference on Artificial Intelligence*, 24(1):173–178, Jul. 2010. doi: 10.1609/aaai.v24i1.7564. URL `https://ojs.aaai.org/index.php/AAAI/article/view/7564`.

Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.