Artificial Intelligence In Train Scheduling Problems

Kieran Molloy
17023997@stu.mmu.ac.uk

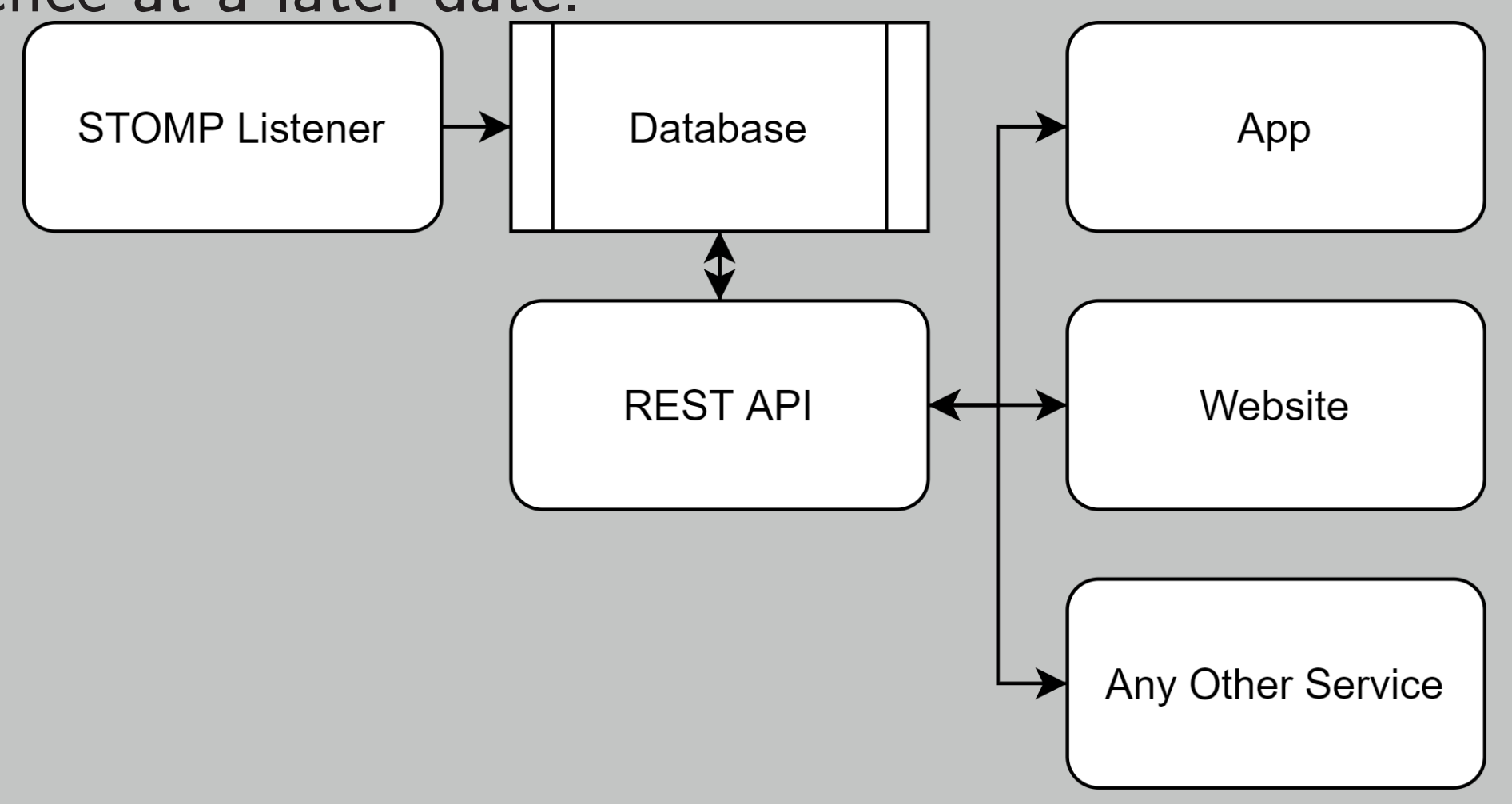
Department of Mathematics and Computing - Manchester Metropolitan University
Learn more at chopin.math.aca.mmu.ac.uk/kieran/

Abstract

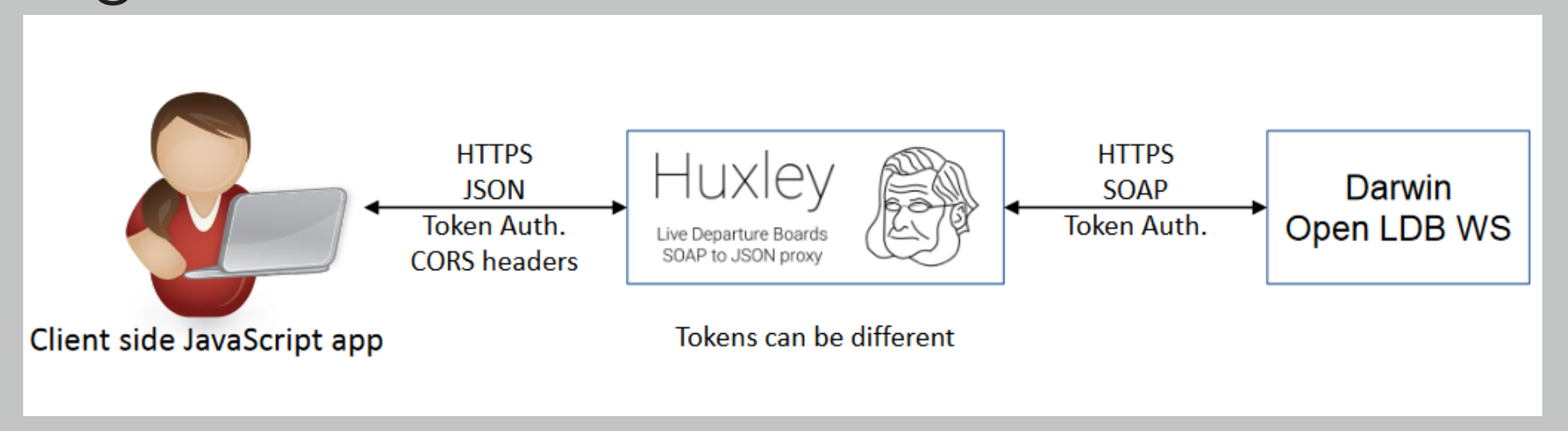
- We investigate if artificial intelligence can be used to independently solve, or aid a solution to the Train Scheduling Problem. We evaluate a Genetic Algorithm approach, in addition to a state-of-the-art Reinforcement Learning approach evaluated using the SBB Challenge.
- In the process, various Mixed Integer Programming formulations are given, based on ideas from Fischetti and Monaci, for the TSP and TRSP. The Reinforcement Learning algorithm is based on a Delayed Q-Learning system devised by Kakade.
- We conclude that Reinforcement Learning is outperformed by more direct heuristic methods, however if a pre-training system is devised performance would increase.
- We believe that the Reinforcement Learning system defined is the first of its kind, and provides a platform for further research.

Rail Data Feeds

- To curate a data set, a system is set up to record all live train movement in the UK and save this to a database for reference at a later date.



- The key components are STOMP Listener, which is setup to find the data stream and parse then save it in the database, and the API. The API has custom functions to calculate delay at specific berths and further functionality. A API endpoint documentation is available at the link in the header.
- The system previously is for Network Rail's official data stream, National Rail also provides Darwin. A similar, more detailed system, which runs all platform boards in stations. Using a system called Huxley, created by J.Singleton, this is also integrated into the API.



- The data collected can be used in later sections to inform learning algorithms.

Mixed Integer Programming TSP

- The Basic MILP TSP is extremely well studied, but can be summarised by the following objective function

$$\sum_{n=1}^{N} A^s_{n,M} - D^s_{n,0}$$

Example Constraints:

$$D^r_{n,m} \leq A^r_{n+1,m} - H \tag{1}$$

$$A^r_{n,m+1} \geq D^r_{n,m} + E_m \tag{2}$$

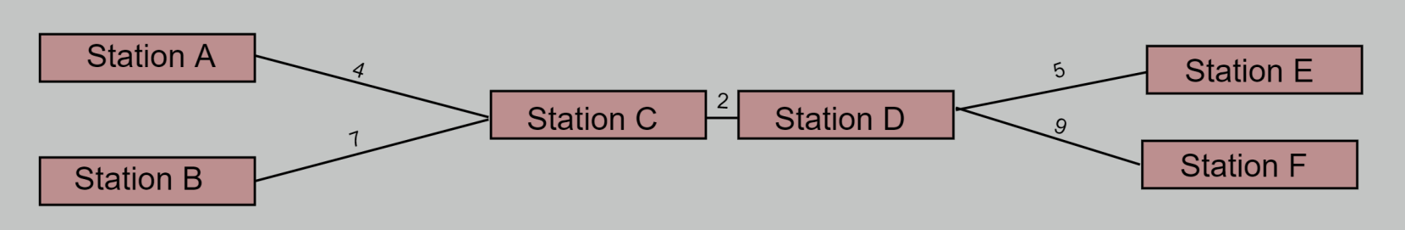
- The objective function for different TSP research papers varies, but generally it is always to minimise delay or waiting; whether that be for cost or environmental reasons.
- These constraints perform the following functions: (1) prevents multiple trains from occupying the same station, and enforces headway. (2) trains cannot arrive at a station faster than the minimum distance time.

Genetic Algorithm for TSP

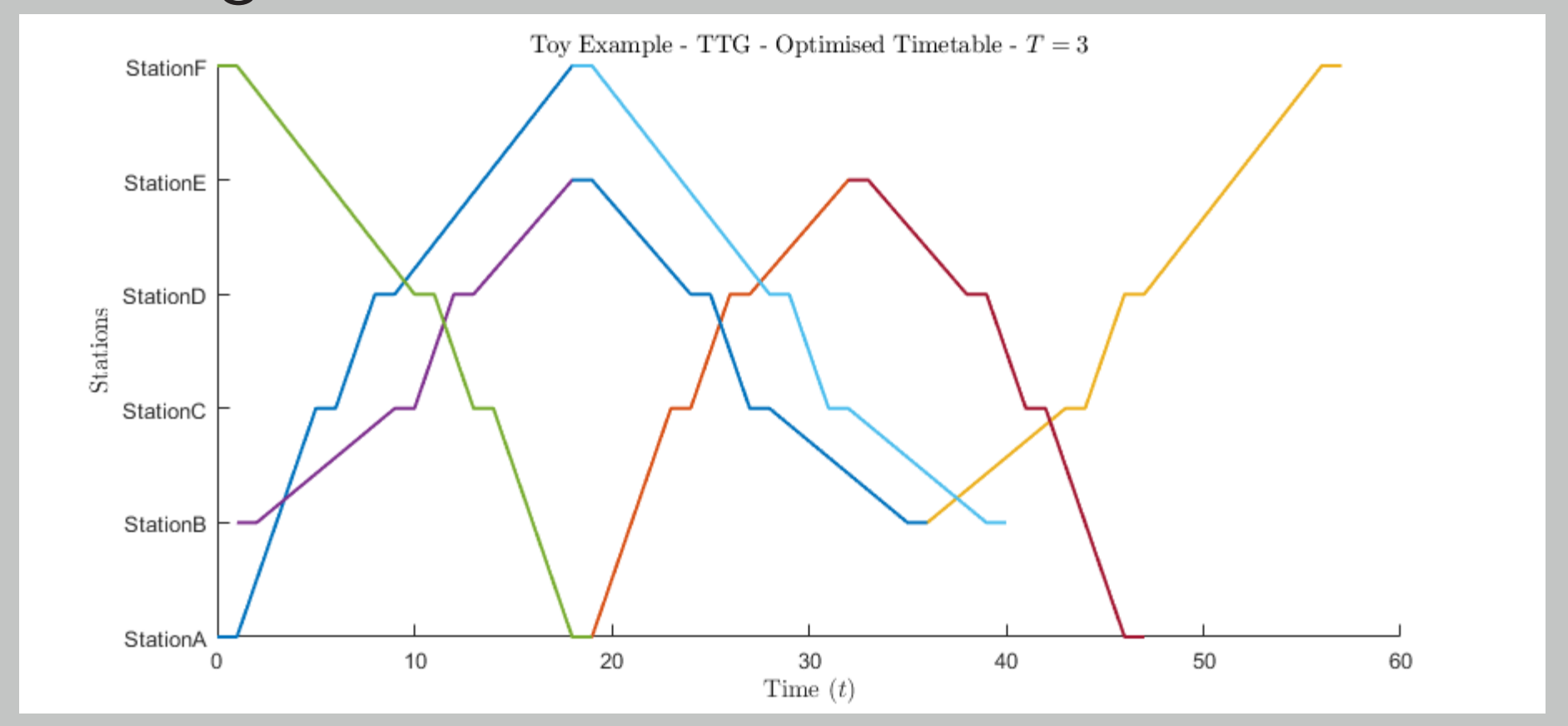
- A genetic algorithm is categorised by the functions it performs; generation, selection, crossover, mutation and fitness analysis. This section defines a framework for a solution to the Train Scheduling Problem

	toc	line	unit	origin	destination	intermediate
Bits	2	8	8	12	12	$12n$
Values	4	256	256	4096	4096	$n(2^{12})$

- The above table demonstrates the phenotype construction, this allows for full customisation of problems for all representations.



- The timegraph below represents a very basic scenario with 6 stations, 8 routes and 4 trains. It is presumed that bi-directional traffic is allowed. The genetic algorithm used basic heuristics to prevent blocking and attempt to maximise platform usage.



Markov Decision Processes

- Markov Decision Processes (MDP) comprises of Markov Reward Processes and Markov Processes. The reward-observation state relationship allows improvements.
- An MDP M is a $5x1$ array (S, A, T, R, γ) where,
  - S is the state space,
  - A is the action space,
  - T is a transition function $(T : S \times A \times S \to \mathbb{R})$,
  - R is a reward function $(S \times A \to \mathbb{R})$,
  - and γ is a discount factor $(0 \leq \gamma < 1)$.
- The optimal policy is denoted π^* and has value functions $V^*_M(s)$ and $Q^*_M(s,a)$.
- This logic is the basis of a reinforcement learning algorithm.

SBB Challenge

- The data sets provided by the CrowdAI SBB challenge, and framework provided by CrowdAI are massive resources for basing a study upon. The problem sets are defined in the table below

ID	Name	Trains	Routing	Difficulty
01	dummy	4	V.Few	V.Simple
02	a little less dummy	58	Few	Simple
03	FWA 0.125	143	Few	Simple
04	V1.02 FWA	148	Few	Medium
05	V1.02 FWA obs	149	Medium	Medium*
06	V1.20 FWA	365	High	V.Hard
07	V1.22 FWA	467	High	V.Hard
08	V1.30 FWA	133	V.High	V.Hard
09	ZUEZGCH 06001200	287	VV.High	VV.Hard

Table: Description of Problem Scenarios

Delayed Q-Learning

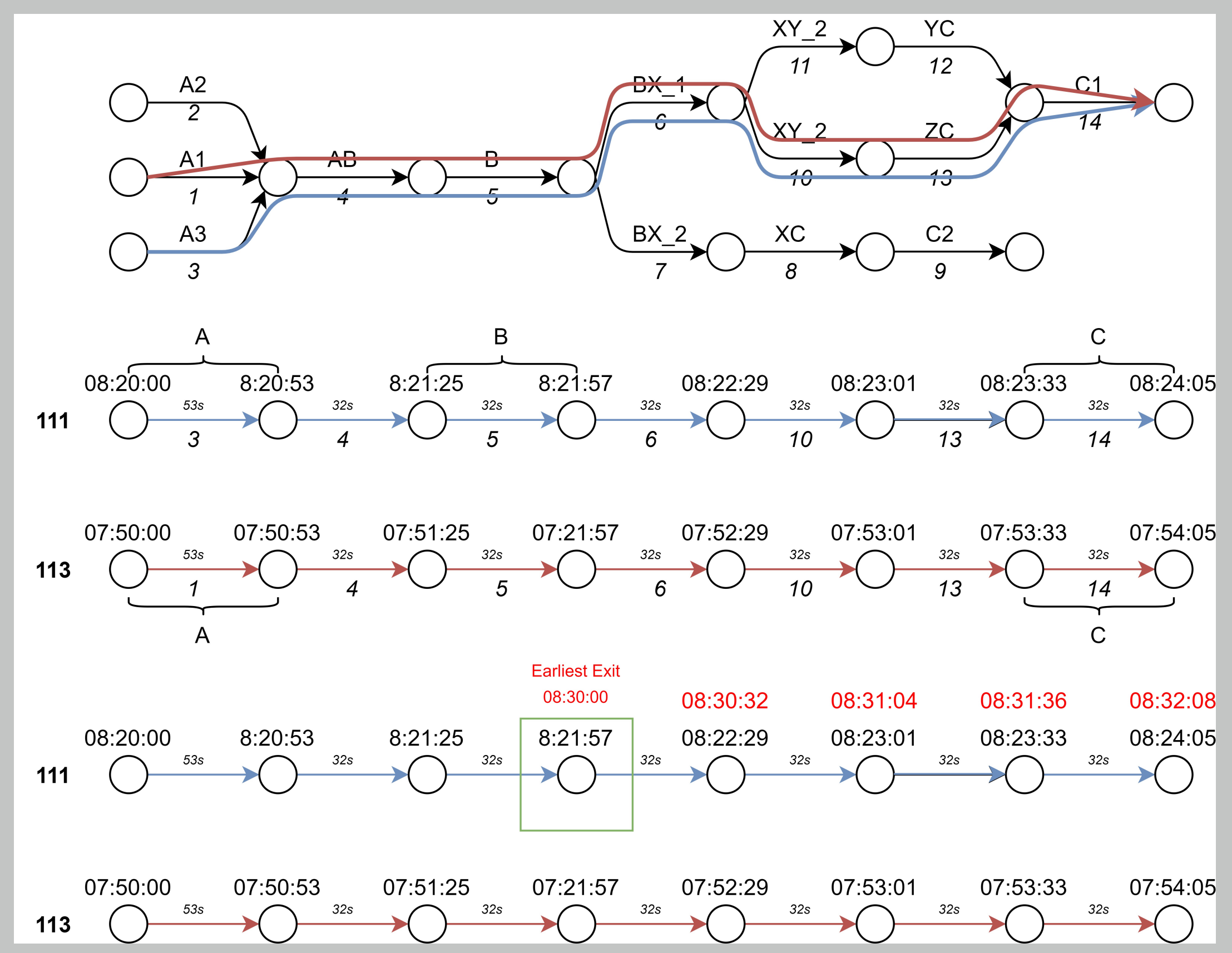
- Developed by Kakade, using a Learn Flag to indicate whether the learner is ready to make a change to the Q-Value estimates. The algorithm is PACMDP, and enhances the learning speed of a standard Q-Learning algorithm. The update function is as follows:

$$Q_{t+1}(s,a) = \frac{1}{m}\sum_{i=1}^{m}\left(r_{k_i} + \gamma V_{k_i}(s_{k_i})\right) + \epsilon_1$$

The following condition must be met to update:

$$Q_t(s,a) - \left(\sum_{i=1}^{m}\left(r_{k_i} + \gamma V_{k_i}(s_{k_i})\right)\right) \leq 2\epsilon_1$$

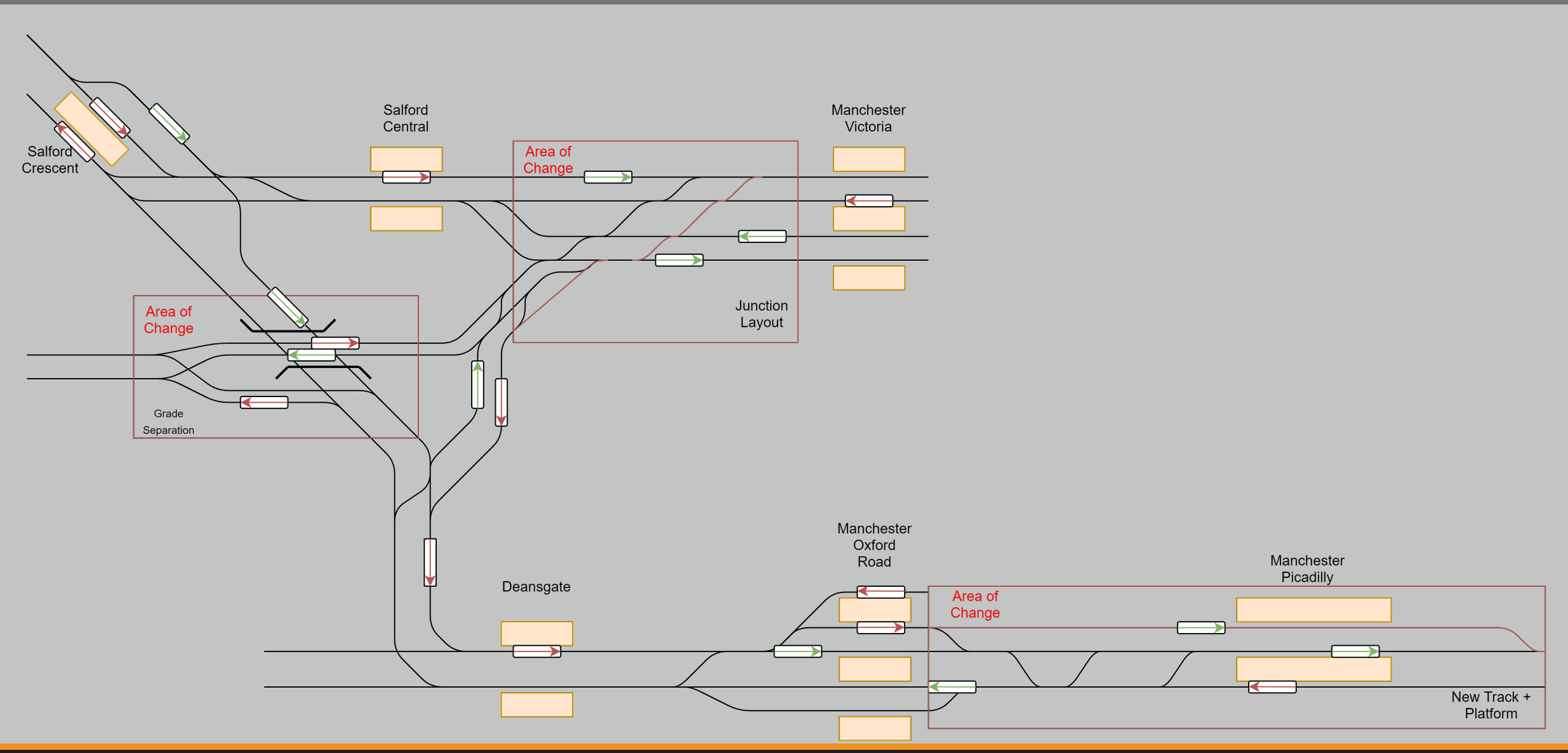
Reinforcement Learning TSP



- A Discrete Event Simulator is created in Python, handling train movement events from section to section. When a potential blockage is forecast the simulator looks at the Q-Value Table to determine the best action to take. The simulator will keep running until a score of 0 is achieved, or 30 minutes computation time.
- The figure demonstrates the sample data set, the 2 services 111 and 113. It shows the infeasiblity of 111, due to a earliest departure time constraint.

ID	System 01		System 02		System 03	
	Z	Time	Z	Time	Z	Time
01	0	5.452	0	4.296	0	7.590
02	0.43	87	0.43	72.087	0.43	77.042
03	0.86	45	0.96	198.273	0.11	200.234
04	24.94	430	24.94	439.121	024.94	435.945
05	n/a	1800	n/a	1800	n/a	1800
06	207.12	1355	207.12	1174.690	207.12	1750.154
07	456.60	1800	442.33	1800	467.10	1800
08	153.23	1800	122.70	1800	233.6	1800
09	138.95	1800	38.25	1800	43.95	1800

Case Study


Salford Crescent
Salford Central
Manchester Victoria
Area of Change
Junction Layout
Deansgate
Manchester Oxford Road
Manchester Piccadilly
New Track + Platform