ABSTRACT
Transportation systems fundamentally impact human well-being, productivity, and sustainability. It is thus crucial to address the disproportional benefits that their design can lead to. In my research, I explore the trade-off between efficiency and fairness in transport network design. I argue that Multi-Agent Reinforcement Learning frameworks can be used to study the dynamics of mobility and simulate the impact of transportation network design on alleviating urban inequalities.

KEYWORDS
Multi-Agent Reinforcement Learning; Transportation Networks; Socially-aware Artificial Intelligence

1. INTRODUCTION
Mobility is the foundation of the modern, fast-paced urban life [6]. However, not everyone is able to enjoy travelling to places quickly and cheaply, even in modern rich countries. According to a recent study, approximately one million people in Canada [1] are in high risk of suffering from mobility-related poverty.

Mobility inequality is not a natural phenomenon, but rather the outcome of centuries-long injustice in urban planning. Modern cities have been designed according to the traditional economic imperatives: efficiency and growth [6]. Recently, transportation researchers have been extracting new insights and creating indices to measure mobility inequality. Popular works focus on accessibility, with metrics such as the number of reachable opportunities [9], affordability of reaching them [7] and combinations of those into more complex indices [10]. While these works set important foundations for measuring mobility inequality, their capacity is limited in assessing the status quo. Yet, if we want to achieve the United Nations goal of providing affordable and accessible urban systems to everyone, we need tools that not only measure, but learn the dynamics of mobility and generate alternative urban designs.

Artificial intelligence can be utilized to create these tools. Transportation Network Design (TNDP) is an NP-hard optimization problem, where a transport line is created with the goal to optimize for some definition of utility (usually satisfied mobility demand) [3]. The TNDP can be modelled as a sequential decision-making process, where at each time-step a location is selected to place a station on. Following this framework, a line is generated by taking an action, receiving a reward and adapting based on it. Reinforcement Learning can be used to tackle this formulation and offers better results over linear optimization and heuristics methods, without the need to specify many constraints, or reduce the solution space [11]. However, these models have ignored the emergence of disproportional benefits from the new lines, and as I show later on, can give unfair advantages to the most-privileged groups. Applying the insights of the transportation community to this problem is crucial for building better cities, but has been largely absent from relevant research [8].

Urban mobility is a dynamic process, thus, even if an seemingly fair network is designed, it is not guaranteed to alleviate inequalities. Populations tend to adapt to changes in city environments, leading to unintended consequences. Modelling this behavior and simulating adaptation to changing environments is therefore crucial for taking better long-term decisions. Multi-agent reinforcement learning has recently proven to be a suitable tool to study adaptation to changing environments under notions of fairness, both in simple games [4] and in complex economic simulations [12]. No previous work has thus far applied these concepts to network-based environments of urban mobility.

My research thus aims to advance learning tools that offer desirable compromises between efficiency and equity when designing transportation systems. It is conducted in collaboration with the Municipality of Amsterdam. To the best of my knowledge, no other work has been studying the inter-disciplinary problem in inclusive urban systems design with multi-agent reinforcement learning.

2. PROGRESS SO FAR
In my research so far, I have worked toward exploring mobility inequality, defining fairness in transportation network design and creating a reinforcement learning agent that designs networks under different fairness goals.

2.1 Fairness in Transportation
There are multiple fairness criteria in transportation; I outline the most common ones [2]:

- Utilitarianism: maximize total benefits.
where the objective is defined as the total captured travel demand of the created line, expressed as a function $U_{od}$ of the estimated Origin-Destination (OD) matrix [3], where $\{U_{od}\}_{ij}$ represents traveling magnitude from location $i$ to $j$. The optimization objective is to find the transportation graph $G(N, E)$, such that:

$$\max U_{od}(G(N, E))$$

$$\text{s.t. } \text{cost}(G) \leq B \tag{1}$$

Recent work has shown that a policy gradient Deep RL model outperforms other heuristic approaches, when optimizing for total captured demand (origin-destination flows) [11]. However, as I show in Figure 1, this approach (red color) leads to big disparities between different groups. Specifically, when applied to the city of Amsterdam, with groups defined as quintiles of the average house price index, the areas in the most disadvantaged cohort benefit the least from the generated line, amplifying already-present disparities. This leads to unfairness in the added benefits, as they are neither equalised nor benefit the groups mostly in need.

The traditional optimization objective ignores how the benefits of the newly designed line are distributed between groups. To address this I introduce the group-based satisfied mobility flow. I define a set $A$, which represents $d$ different groups based on socio-economic indicators, such as income, development index, or education. Each cell $h \in H^{n \times m}$ of the environment is then associated with a group $a \in A$ and the optimization goal is a welfare function that incorporates the needs of the different groups. I focus on defining fair RL policies for the TNDP problem, aiming to reduce disparities in captured demand, according to the criteria outlined in section 2.1. This is achieved by modifying the reward function, without the need of introducing new constraints. I propose reward functions that optimize for the distribution of benefits, using the Generalized Gini Index (GGI) of the captured demand. As shown in Figure 1 (striped purple bars), the proposed method generates a metro line that achieves near-equality in added benefits. As expected, this comes at the expense of total utility, with our model performing 20% worse. Part of this work has been presented at the IJCAI 2022 Data Science and Optimisation workshop 2 and the full results in two cities (Amsterdam and Xi’an) will be presented as an extended abstract at AAMAS 2023.

2994

### 2.2 Fair Transportation Network Design

A public transport network is defined as a spatial graph $G(N, E)$, where nodes $n \in N$ are stations around the city and edges $e \in E$ are the transport lines. A city is represented as a two-dimensional grid environment $H^{n \times m}$. The TNDP is an NP-hard optimization problem where the objective is defined as the total captured travel demand of the created line, expressed as a function $U_{od}$ of the estimated Origin-Destination (OD) matrix [3], where $\{U_{od}\}_{ij}$ represents traveling magnitude from location $i$ to $j$. The optimization objective is to find the transportation graph $G(N, E)$, such that:

$$\max U_{od}(G(N, E))$$

$$\text{s.t. } \text{cost}(G) \leq B \tag{1}$$

Recent work has shown that a policy gradient Deep RL model outperforms other heuristic approaches, when optimizing for total captured demand (origin-destination flows) [11]. However, as I show in Figure 1, this approach (red color) leads to big disparities between different groups. Specifically, when applied to the city of Amsterdam, with groups defined as quintiles of the average house price index, the areas in the most disadvantaged cohort benefit the least from the generated line, amplifying already-present disparities. This leads to unfairness in the added benefits, as they are neither equalised nor benefit the groups mostly in need.

The traditional optimization objective ignores how the benefits of the newly designed line are distributed between groups. To address this I introduce the group-based satisfied mobility flow. I define a set $A$, which represents $d$ different groups based on socio-economic indicators, such as income, development index, or education. Each cell $h \in H^{n \times m}$ of the environment is then associated with a group $a \in A$ and the optimization goal is a welfare function that incorporates the needs of the different groups. I focus on defining fair RL policies for the TNDP problem, aiming to reduce disparities in captured demand, according to the criteria outlined in section 2.1. This is achieved by modifying the reward function, without the need of introducing new constraints. I propose reward functions that optimize for the distribution of benefits, using the Generalized Gini Index (GGI) of the captured demand. As shown in Figure 1 (striped purple bars), the proposed method generates a metro line that achieves near-equality in added benefits. As expected, this comes at the expense of total utility, with our model performing 20% worse. Part of this work has been presented at the IJCAI 2022 Data Science and Optimisation workshop 2 and the full results in two cities (Amsterdam and Xi’an) will be presented as an extended abstract at AAMAS 2023.

### 3 Future Work

In the future, I plan to build on the previous results using multi-objective reinforcement learning, and to implement a multi-agent framework to simulate mobility under alternative designs.

**Designing lines with multiple objectives.** In Section 2, I model TNDP as a single objective optimization problem, engineering rewards to achieve the best compromise between efficiency and fairness. However, since the solution space is large, it is not guaranteed that the proposed solutions are pareto optimal. The satisfied mobility demand of different groups could be modelled as objectives competing with each other, making pareto-optimal solutions important. I am currently formulating the problem using multiple objectives, to find solutions without limiting the agent to a small subset of linear scalarizations [5].

**Simulating population adaptation in multi-agent environments.** A fair public transport network could mitigate mobility inequality, but it would do so considering current mobility preferences. Would such gains hold in time, given how citizens adapt to changes in the transportation network? To answer this, we need a way to model human behavior and simulate scenarios with different designs, as well as incorporate different conflicting objectives of individuals. Inspired by recent works on economics-based multi-agent simulations [12], I plan to create a two-level deep reinforcement learning framework, where a planner-agent will create and extend transportation networks (like in Section 2), and multiple citizen-agents will learn to adapt their mobility to these changes. I believe this will help to better understand urban inequalities and provide new insights towards reducing them.

2https://sites.google.com/view/ijcai2022dso
REFERENCES


