Democratizing Traffic in Smart Cities

Marcin Korecki[a,*], Damian Dailisan[a,*], Joshua Yang[a], Dirk Helbing[a,b]

[a]*Computational Social Science, ETH Zürich, Stampfenbachstr. 48, Zürich, 8092, Switzerland*
[b]*Complexity Science Hub Vienna, Josefstädter Str. 39, Vienna, 1080, Austria*

Abstract

To improve the performance of systems, optimization has been the prevailing approach in the past. However, the approach faces challenges when multiple goals shall be simultaneously achieved. For illustration, we study a multi-agent system, where agents have a plurality of different, and mutually inconsistent goals. Taking decisions based on suitable voting procedures turns out to lead to favorable solutions, which perform highly for several goals rather than optimally for one goal and poorly for others. This opens up new opportunities for the management or even self-governance of complex systems that require the consideration and achievement of multiple goals, such as many systems involving humans. Here, we present results for traffic flows in urban street networks, which suggest that "democratizing traffic" would be a promising alternative to centralized control of traffic flows. Interestingly, the cumulative proportional voting method can often successfully include minority interests without sacrificing overall system performance.

Keywords: Traffic Control, Reinforcement Learning, Social Participation, Voting, Smart Cities

1. Introduction

As the research into Artificial Intelligence (AI) progresses, so does the range of fields to which such methods are applied. Many real-world systems that used to be managed by humans are now increasingly taken over by computer programs and machine learning (ML) approaches, illustrating the apparent victory of AI technologies.

*Corresponding author
Email addresses: mkorecki@ethz.edu (Marcin Korecki), ddailisan@ethz.edu (Damian Dailisan)

However, as this transition happens, can we be sure that these new approaches are always delivering better solutions? To answer this, it is of utmost importance to ensure that the goals pursued by AI systems, which control increasingly many crucial systems in our societies, are aligned with the goals of said societies and their individual members. Currently, such an alignment is often not there or even avoided, as it appears to be difficult to make everyone happy or to reach a consensus about the overall goal. In many cases, decisions about the goal are, therefore, taken by a "social engineer", who invents some software- and data-based solution and imposes it on the social system of interest.

However, what if there was a method that would not need a consensus on one goal, and could handle people with a plurality of different goals? Moreover, what if that method would perform highly with regard to many goals? Then, such a method could serve the different interests of many and, thereby, overcome a lot of the conflicts occurring in complex societies today. In fact, we will show that the application of suitable voting methods can offer just that. Accordingly, we propose to democratize the management of complex systems involving multiple goals, particularly of social systems involving humans.

For illustration, we will consider traffic flows in urban street networks, which were initially uncontrolled (or controlled only by fixed signs) [1], then controlled by human traffic signal operators [2], later replaced by cyclical signaling schemes [3], then supplanted by algorithmic systems [4], which are now increasingly being replaced by AI methods (usually based on Deep Reinforcement Learning) [5]. Our contribution shows how AI-based control can be further improved, by giving people decision-making power in such a way that multiple objectives can be reached. This, however, does not exclude an implementation involving Internet of Things technology, if people prefer.

Note that traffic systems are a good example of complex systems involving humans. Such systems often show large variability in their behavior, are hardly predictable, and are difficult to control. Moreover, it is not a priori clear what goal or "objective" the AI (or any other control approach) should optimize. Some users may prefer the system to prioritize sustainability and, thus, minimize emissions. Others may only care about travel time, i.e. getting to their destination as fast as possible. Some may care instead about the number of stops, traffic flows, densities, or occupancies. Still, others may prefer to divide their preferences among several goals, i.e. to optimize an index weighting several goals. Unfortunately, however, perhaps due to the additional challenges and inherent limitations of multi-objective optimization [6], most traffic control methods attempt to focus on just one goal.

We note that such a one-dimensional, "utilitarian" approach, especially when applied to complex systems such as pluralistic societies, is often considered to lead potentially to some kind of technological totalitarianism, also known as "data dictatorship" [7]. Such concerns have also been raised in connection with some "smart cities" approaches. In fact, the algorithmic approaches applied in many larger cities around the world to control traffic and other urban infrastructures and services give their users little or no control over the objectives prioritized. At the same time, they are often based on closed-source proprietary solutions, making it impossible for citizens (and politicians) to even know what is going on, why, and how.

Nevertheless, even though "black box" AI poses certain risks, they often offer novel solutions and a new way forward. If powerful data-based learning and optimizing capabilities of modern-day Machine Learning (ML) approaches could be combined with direct user engagement, smart cities could be efficient, sustainable, and democratic. This is why this paper proposes a multi-objective AI system, where the AI learns how to optimize for a certain objective, while the users of the system are enabled to vote on what objective is to be chosen as the basis for a particular decision at a given place and time. By using advanced input schemes and aggregation schemes (which together define a particular voting method), the system can arrive at a collective (local) solution that incorporates potentially contradictory objectives based on the relative preferences for them. We use a traffic simulation environment to test our proposed method in pretty realistic urban settings. Moreover, we compare the effects of different input and aggregation schemes (i.e. various voting methods) on the functioning of our method, and we test it for different preference landscapes of the system's users.

While voting schemes have previously been applied in traffic settings [8, 9], these did

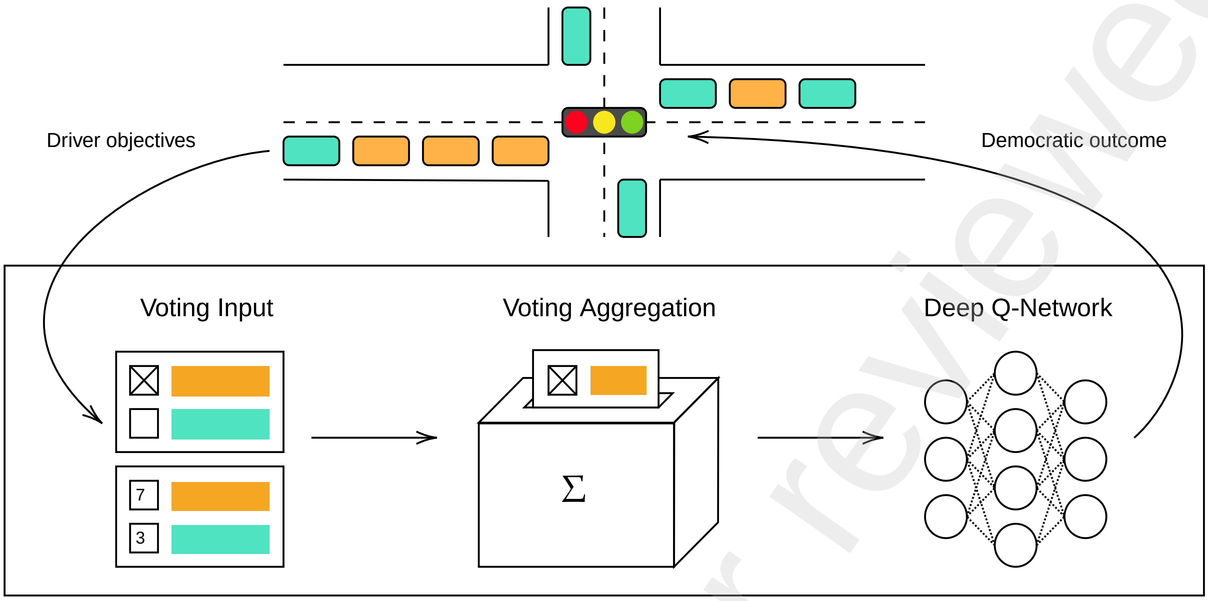


Figure 1: Visualisation of our proposed system. Drivers approaching an intersection engage in a voting round, where they express their preferences regarding the objective to be pursued by the intersection controlling the traffic lights. The voting input can be in the form of binary or cumulative choice. The votes are aggregated using a selected aggregation method (for example proportional or majority). The votes are then used as weights for the Q-values generated by Deep Q-Networks trained on the objectives. Finally, the collective objective is used to determine the phase of the traffic light to be set at the intersection.

not take a multi-objective approach. The novelty of our proposed method is that it allows for the democratization of traffic (and other complex systems such as smart cities), namely by combining AI approaches with insights from social choice theory. Thereby, it makes a major step towards "trustworthy AI, in line with our shared democratic values", as recently demanded [10].

Such democratization of previously technocratic systems is expected to have significant positive effects on social systems, as it can better handle pluralistic settings, where multiple goals are pursued. Among other things, it would likely increase trust and the sense of agency of its users, which could further translate into increased social engagement and have positive reverberations in other social systems besides traffic [11].

2. Background Literature

In this section, we will introduce the background literature relevant to our topic. We will focus on introducing prior research into urban traffic and its mitigation and control (especially ML based). We also present our perspective on social choice theory, the methodologies of which will be incorporated in this work.

2.1. Urban Traffic

Traffic flows in urban road networks are a widespread, real-world example of complex dynamic systems [12]. Such systems are often inherently difficult to control, hard to predict, and prone to phase shifts [13], potentially also a rapid degradation of system performance. An example of a negative property emerging from complexity is congestion, which occurs due to the many direct and indirect interactions between individual driver-vehicle units [14]. The main goal of urban traffic management is usually to limit the outbreaks of congestion, minimize their duration, and mitigate their consequences. Two popular methods to achieve these goals are traffic signal control and road pricing.

2.1.1. Traffic Signal Control

Traffic signal control is a direct (and arguably the most common) method of controlling the flow of vehicles at intersections. Historically, the flows were controlled by humans. Nowadays some cities rely on automatized control cycles, while others have modernized their systems for control by a range of "smart" algorithms.

One example of such an intelligently designed control algorithm is the self-organizing traffic signal control, which relies on local interactions between the neighboring intersections [15, 16]. Another example is a method that similarly relies on self-organization, but improves on the previous one by accounting for short-term traffic anticipation and traffic physics [4].

Recently, significant attention has been paid to methods that rely on machine learning (ML) and, specifically, Reinforcement Learning (RL). Such implementations model the intersection as an "RL agent". Its *state* is the representation of the situation at the intersection and the *actions* are the phases that the traffic lights can take (i.e. which vehicle flows get a green signal). The key specification of the agent is its reward (function), which relates to a goal to be achieved (or optimized).

It has been shown that different specifications of the reward may lead to diametrically different results [17]. Mis-specification of the reward may also lead to solutions, which are highly unfair across the possible flows [18]. Whatever the choice may be, the reward (function) is currently chosen by the designer ("social engineer") of the RL agents. It is not in control of the local government and cannot be influenced by drivers.

Some examples of RL approaches to traffic include methods that rely on an analytic exploration heuristic, which allows for faster learning [19], or on methods, which rely on meta-learning [20] or pre-training [21]. The aim is to increase the agent's adaptability to various conditions. Nevertheless, some concerns have been raised regarding the resilience of RL solutions to disruptions and their performance compared to previously existing, non-learning algorithms [22].

The novelty of the present work is the introduction of a pluralistic framework supporting multiple goals, which enables the democratization of control decisions using voting methods. While we illustrate the suitability of the approach for the case traffic signal control, generalizations to other systems are easily possible.

2.1.2. Road Pricing

Road pricing is an indirect way of diminishing congestion by discouraging drivers through pricing from using infrastructure. While in this work we focus on traffic signal control, road pricing is relevant to social aspects of urban traffic.

One issue of road or congestion pricing is that it affects the different "socio-economic classes" in different ways, thereby creating a risk that some members of society may be "priced out". To some extent this can be addressed through appropriate mechanism design [23] and/or bidding or credit systems [24]. However, concerns remain. For a review of the issue of fairness in road pricing, the reader is referred to Ref. [25].

Our present work focuses on mechanism design, which is based on different voting aggregation and input schemes. This approach is fair, but at the same time gives some amount of direct control to the users. Current road pricing schemes, in contrast, are top-down approaches, which *reduce* the level of control that users of the traffic system can exert.

2.2. Value Alignment

The research field on Value Alignment (VA) is often considered together with AI systems and follows some already established lines of research on ethical and social consequences of automation and technology [26]. Currently, the field has a large focus on the risks posed by a hypothetical, powerful, misaligned Artificial General Intelligence (AGI) [27–29]. While this position gives a strong argument for the importance of Value Alignment, it has also been criticized [30]. The study of VA involves multiple disciplines and brings insights from the study of ethics, sociology, economics, and AI [31]. The philosophical arguments are summarized by Ref. [32]. The technical side of the field addresses the specific causes of misalignment, such as reward hacking [33]. Some helpful impossibility and uncertainty theorems have also been provided, using the work performed by ethicists as a basis [34].

2.3. Social Choice Theory

Social Choice Theory (SCT) provides an analytical lens to understand the complex process of aggregating individual preferences into collective decisions [35]. This field is highly relevant to the value alignment problem, especially when multiple preferences are aggregated into a cohesive output for a given system. SCT has consistently been identified as a key element in addressing this problem, with several works investigating value alignment through the lens of social choice impossibility theorems, like Gibbard's theorem [36].

Rooted in interest in the construction of social order during the French Revolution, pioneers like Jean-Charles De Borda and Marquis De Condorcet laid the groundwork for understanding democratic decision-making [37]. In the recent decade, emerging sub-fields like Computational Social Choice extend to computational tasks in voting contexts, incorporating elements of Social Choice Theory, theoretical computer science, and multi-agent system analysis [38].

Research has illustrated the relevance of SCT for understanding value alignment. Approaches like the use of voting-based systems have been proposed for ethical decision-making in AI, based on large datasets of preferences.

Voting methods, such as approval voting, have been suggested to guide group decision-making using AI systems [39]. These insights have been employed to design AI systems with decision-making capabilities allowing them to identify and aggregate potentially conflicting preferences [40].

In democratic societies, the legitimacy of decision-making processes heavily relies on the effective integration of these preferences [41]. Recent research underscores the context-dependent nature of the perceived legitimacy of different voting methods within a democratic setting [11], highlighting the need for careful design of voting methods to handle the Value Alignment problem in various real-world contexts well.

Our work builds upon this work, by integrating a voting layer directly into the decision-making procedure of a reinforcement learning (RL) agent. Echoing the insights of Baum [40], our approach prioritizes the explicit identification and aggregation of potentially divergent preferences, thereby avoiding an over-reliance on the RL agent to independently manage these contradictions. Specifically, we favor a co-learning process, where the RL agents are guided by collective democratic inputs.

Key theorems in SCT have implications for these contexts. For binary choices, such as when there are only two policy alternatives, May's theorem posits that a binary voting procedure can be seen as fair if it fulfills the conditions of anonymity and positive responsiveness [42].

To handle different intensity levels of preferences, cumulative voting allows voters to assign scores to each of the objectives [43]. Aggregation methods like majority aggregation, often justified by the Condorcet Jury Theorem, select the objective with the most votes [44]. In comparison, proportional aggregation ensures a representation of diverse preferences [45].

Arrow's Impossibility Theorem suggests that designing a social welfare function satisfying a set of conditions when aggregating preferences from three or more alternatives is generally impossible [46]. However, this is not the approach taken in our paper, as we are not trying to determine one social welfare function applicable to all, but rather focus on local contexts, preferences, and needs.

In summary, SCT provides a rich theoretical background for understanding and addressing value alignment problems. By effectively incorporating individual preferences, it becomes possible to create systems that deliver optimal or high performance, while respecting a broad spectrum of values and objectives.

3. Methods

In this section, we introduce our methodology. We describe the details of the methods employed and the simulation experiments run. Our approach implements RL traffic signal agents that each control a single intersection. The design builds on previous work [18], where simple voting methods have been tested for a toy example (single intersection). Here, we considerably extend this work by comparing various input and aggregation methods and studying their effectiveness in scenarios involving many intersections. Furthermore, we study the functionality under a variety of preference landscapes.

3.1. Traffic Control Agent

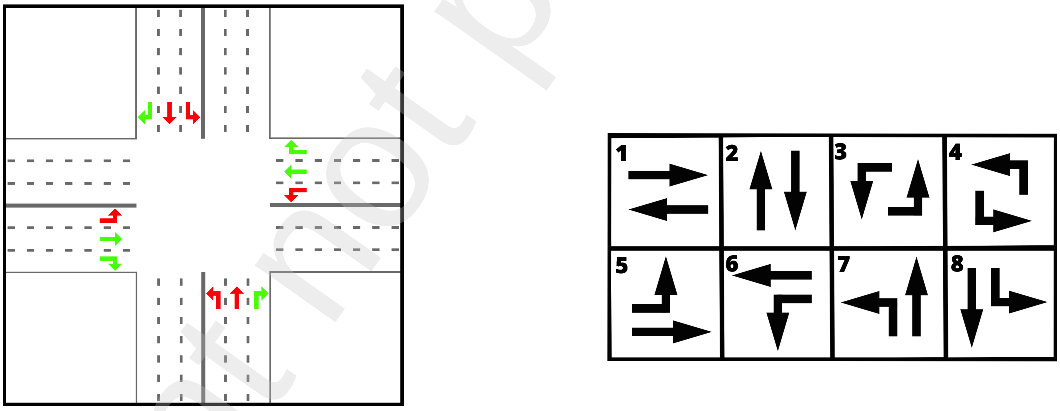


Figure 2: Left: Representation of an intersection with 8 flows (not counting the right turns, which are assumed to be given “conditional green” always). The arrows correspond to the possible flows (green arrow: the flow is given green; red arrow: the flow is given red. The green phase depicted in this figure corresponds to phase 1 in the figure on the right. Right: possible phases (actions) available at intersections with 8 flows (as illustrated on the left).

In this paper, we consider RL traffic signal agents (short: “RL traffic agents” or “RL agents”), which control the phases at their associated intersections (see Figure 2). The *state* of the agents represents the situation at their respective intersections. The *actions* correspond to the phases, and the *reward* is the objective that we wish to optimize. The agents take action every 10 seconds. An agent is given by the following variables:

- **State:** occupancy of each incoming flow (the ratio of the road length covered by vehicles), occupancy of each outgoing flow, and (one hot-encoded) current phase.
- **Actions:** phases as indicated in Figure 2.
- **Reward:** in this work we study two different kinds of rewards. One reward minimizes the number of stops (Equation 1), which corresponds to a sustainability criterion (as most emissions are generated during stopping, starting the engine, and accelerating). The other minimizes the wait times (Equation 2), which could be seen as a criterion for driver comfort. That is, we consider:

$$r_{\text{stops}} = -(\text{number of stops in the past } t_{\text{act}} \text{ steps}) \tag{1}$$

$$r_{\text{wait}} = -(\text{duration of stops in the past } t_{\text{act}} \text{ steps}) \tag{2}$$

3.2. Pluralistic Framework to Consider Multiple Objectives

The key feature of our pluralistic framework is its ability to incorporate multiple goals/objectives. This is achieved by training separate Deep Q-Networks (DQN) for each objective. Each DQN, in a given state, outputs Q-values for each of the n possible actions. These Q-values are interpretable and represent the sum of discounted future rewards. Typically the action with the highest Q-value is chosen. Consider a DQN for objective A that outputs the Q-values for n actions, Q_1^A through Q_n^A. One could also consider the relative quantities, so if $Q_i^A \gg Q_j^A$ (where i denotes the highest Q-value and j denotes the second highest), there is a clear preference for one action over the other, while if $Q_i^A \approx Q_j^A$ neither action is strongly preferred.

Another observation is that the relative Q-values could also be compared across different DQNs (objectives). One might want to compare the Q_i^A to Q_i^B (where B is another objective) in order to gauge the effects of taking a given action on different objectives.

Note that the Q-values by two different DQNs might not be scaled in the same way. However, if we only care about the relative values (that is how much better action 1 is compared to another action n with respect to a given objective) we can normalize the values $\{Q_1^A$ through $Q_n^A\}$. In this work, we use a softmax normalization in connection with objective A. This yields a set of $\{q_1^A$ through $q_n^A\}$ such that $\sum_n q_n^A = 1$. If we do the same for objective B, we are able to compare q_i^A with q_i^B. In our simulation experiments, the DQN outputs Q-values for $n = 8$ actions (one for each phase in Figure 2).

3.3. Voting System

The key aspect of the system we study is that the users (in this case car drivers) are able to affect the decisions taken by the control agents. Thus we need a way of polling and aggregating drivers' preferences. In a given decision step (corresponding to a given agent's action frequency, which is set to 10 seconds, here), all drivers of cars on incoming road sections related to an intersection's control agent are polled for their preferences. We test and compare several input and aggregation methods, which we explain below.

3.3.1. Input Methods

Binary Voting: This method implies a straightforward choice between two options. For instance, a driver could vote for either minimizing their wait time at traffic lights or reducing their expected number of stops. This can be represented mathematically as $v_i \in \{0, 1\}$, where v_i is the vote of driver i.

Cumulative Voting: This method is more nuanced and allows for greater expression of preference intensity. Drivers can distribute 10 points across different objectives based on their preference for each. For example, a driver could give 7 points to reducing wait time, and 3 points to minimizing the number of stops. This system is able to reflect more complex preference landscapes and takes into account that drivers might not have a single objective, but rather a mix of priorities. With $s_{ij} \geq 0$, where s_{ij} is the score of driver i for objective j, and $\sum_j s_{ij} = 1$ for each i, one can normalize across objectives (which can be done in the above example by dividing the given points by 10).

3.3.2. Aggregation Methods

Majority Aggregation: This method operates based on a winner-takes-all principle, where the objective with the majority of votes takes precedence in decision-making. For instance, if most drivers in a binary voting system vote to reduce their wait time at traffic signals, the system will optimize traffic management predominantly for that objective, regardless of the number of votes other objectives may have received. This methodology aligns the traffic system optimization with the most preferred goal among users, potentially sacrificing the lesser-supported objectives in the process.

Proportional Aggregation: Unlike majority aggregation, proportional aggregation seeks to represent all objectives proportional to their voting support. This method doesn't solely focus on the objective with the most votes, but distributes focus across several objectives according to the number of votes each has received. For example, if reducing wait times receives 70% of the total points, minimizing stops receives 30%, the system will optimize traffic management with the corresponding distribution, prioritizing wait times the most, but also taking into account the importance of reducing stops. In this way, the system respects the nuanced preference landscape among users and provides a more balanced and inclusive approach.

3.3.3. Representation of Preferences

In this work, we make certain assumptions regarding the behavior of the traffic users, who participate in the voting. Namely, we assume that the users are rational, but do not misrepresent their preferences strategically. That is, they are always assumed to vote in a way that most closely fits their preferences. We would like to mention, however, that some voting schemes are prone to manipulation through misrepresentation of preferences by voters [47]. This particularly applies to rank-based input methods with three or more options. In this work, however, we only deal with situations, where voters choose among two options. Accordingly, this manipulation risk does not occur, here.

3.4. Integration Layer

The integration layer allows to combine the output values produced by the DQN and the weights computed from the users' preferences. The combination follows Equation 3, where for action n a q'_n is computed. Such a q' is computed for every action and the action with the highest q' is selected.

$$q'_n = w_A q_n^A + w_B q_n^B \tag{3}$$

This approach allows for incorporating the users' preferences as well as the insights gained through interpreting the Q-values of the DQN. Practically, in an example with two possible actions, when the users have a strong preference (e.g. $w_A \gg w_B$), the action under objective A (e.g. $q_i^A > q_{n \neq i}^A$) will be selected unless $q_{n \neq i}^B \gg q_i^B$ and $q_i^A \approx q_{n \neq i}^A$. Thus, the system follows the users' preferences unless the relative difference between the expected rewards within one objective is large and within the other small. Conversely, for a weak preference ($w_A \approx w_B$), the system prioritizes the objective that would lose the most from not being followed. The entire system is visualized in Figure 3 for a case with two possible actions. The system, however, can be easily extended to incorporate an arbitrary number of actions.

3.5. Voting Scenarios

Below we list the scenarios that we explore in our computer simulation experiments.

1. **Bipolar Preference Distribution:** In this scenario, drivers are divided evenly into two distinct groups. Half allocate approximately 90% of their points to 'stops', indicating a strong preference to avoid them. The remaining half distributes roughly the same proportion of points, but favors avoiding 'waits'. This reflects a pronounced bipolarity in drivers' preferences.
2. **Committed Minority-Indifferent Majority:** In this scenario a small but adamant minority, namely 20% of the drivers, allocates around 90% of their points to 'stops', demonstrating a strong preference. However, the majority (80%) is largely indifferent. They show only a slight preference to avoid 'waits', assigning roughly 60% of their

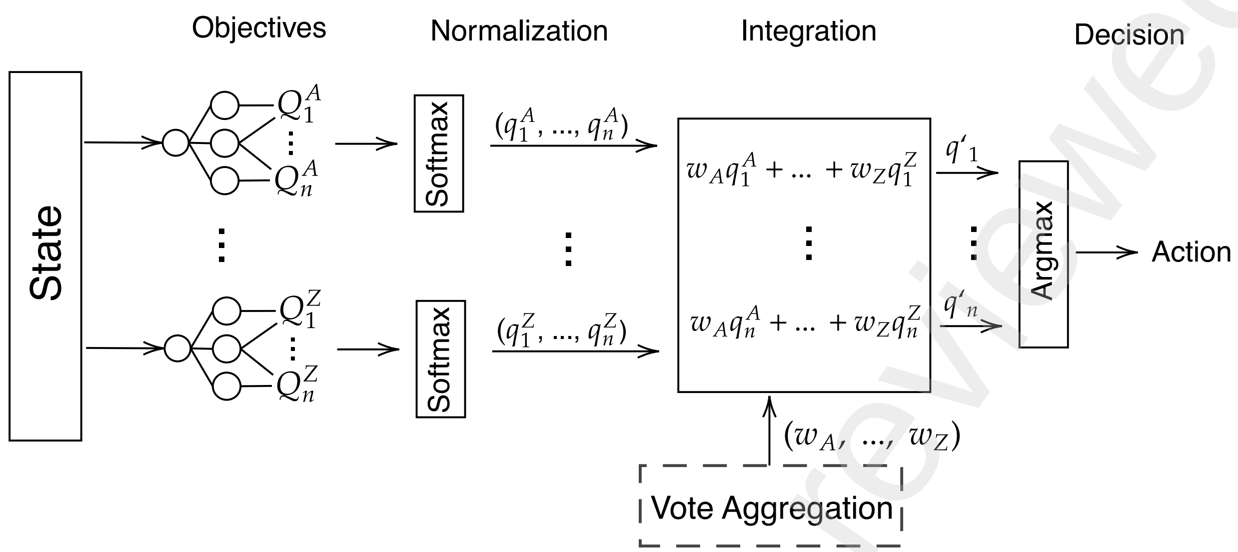


Figure 3: Schematic representation of the multi-objective voting system. A state s is fed through two DQNs, whose resulting Q-values are normalized and integrated with the votes. An argmax of the resulting q' values is taken and the corresponding action is chosen.

points to this. However, considering the extent of polarization, the aggregated preference is neutral. This neutrality is mathematically represented as $(0.2 \times 0.1) + (0.8 \times 0.6) = 0.5$, indicating a balanced preference for 'stops' and 'waits' at an aggregate level, despite the large individual variability. In the text and figures this scenario is referred to as "committed minority".

3. **Random Distribution:** In this instance, the allocation of points between 'stops' and 'waits' is purely random for each driver. The distribution follows a multinomial pattern, yielding a 50% probability each to cast a vote for 'stops' or 'waits'. This represents a state of complete randomness, where drivers do not seem to care about what will happen.

For all voting scenarios, driver point allocations are subject to a random Gaussian spread of 0.5 of the 10 points. Each of these scenarios provides insights into how voter preferences can impact collective decisions in different ways. Studying these scenarios can reveal interesting insights into how preferences influence decision outcomes and the dynamics of collective choice.

Note that voting decisions may vary from one intersection to another, as the composition of drivers on the incoming road sections may be different. Hence, each control agent responds to the respective local traffic situation and the corresponding local distribution of preferences.

3.6. Simulation Experiments

Here we provide further details of the experiments we conduct in our simulation environment. The traffic simulator we employ is Cityflow [48]. Our main reason for using this software is that it is open source, boasts higher performance than many popular alternatives, and has been extensively used in publications relating to intersection control and machine learning.

We run our experiments on two simple real-world-inspired traffic networks visualized in Figure 4. The first scenario is a grid, while the second one is a corridor. Thus, each represents a different traffic situation. In each traffic scenario (which we refer to as Hangzhou and NY16 from now on), we run three experiments corresponding to the voting scenarios described in subsection 3.5. Furthermore, in each of these, we test three combinations of voting and aggregation methods, namely:

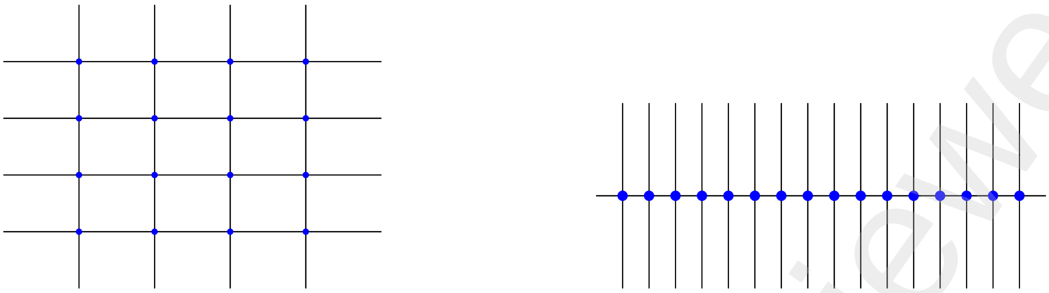
(i) binary voting using majority aggregation,

(ii) binary voting using proportional aggregation, and

(iii) cumulative voting using proportional aggregation.

A single run of a traffic scenario corresponds to 3600 seconds of real time. For each combination of a traffic scenario with a voting scenario, we generate 50 runs and report their averages.

To assess the effectiveness of our proposed method in different traffic and preference conditions as well as the appropriateness of various voting methods, we report the average performance of the traffic system in terms of wait times, number of stops and speeds. Moreover, we report the distribution of the value alignment of the voting population for each voting method and traffic scenario.



(a) Hangzhou scenario road network.

(b) NY16 scenario inspired by the Upper East Side in New York.

Figure 4: Illustration of the road networks underlying the traffic scenarios studied in this paper.

3.6.1. Value Alignment

Once votes are cast and the points have been integrated to decide the next action of the traffic control agent, the resulting action $a^r = \text{argmax}(q')$ of the traffic controller can be compared against the supposed actions of the pure controllers $a^k = \text{argmax}(q^k)$ for each objective k. With "pure controller" we mean one that optimizes only for one reward function in all its actions. When votes for an objective result in an action that is the same for the integrated controller and the pure controller, the points cast by a voter achieved the desired result and, thus, the controller can be said to be aligned with the voter's preferences. More formally, we can sum the value alignment across all objectives and determine the total alignment,

$$\text{total alignment}_i = \frac{\sum_j \delta(a_n^r, a_n^k) w_i^j}{\sum_j w_i^j}, \quad \text{where } j \in \{A, B\}. \tag{4}$$

Here, w are the voting points allocated to an objective j, n denotes the voting round, and i the voter. Votes can only be aligned if the integrated action a_n^r matches the pure action a_n^k. The value alignment is determined each round at each intersection.

4. Results

In the following section, we present the results of our simulation experiments.

4.1. System Performance

In Figure 5 we report the performance of the system for both traffic scenarios under the pure control methods. By pure control method we mean a traffic controller based on

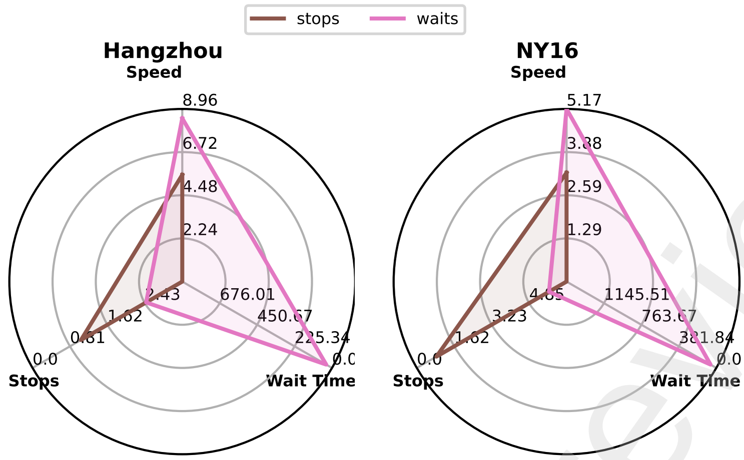


Figure 5: Radar plots for both traffic scenarios with three evaluation metrics (stops (seconds), wait times (seconds), and speeds (meters per second)) for the pure preference cases (controller following only one objective). Larger triangle areas can be associated with a more desirable performance across all metrics.

a DQN that only ever optimizes one metric. We report the results of two controllers—one optimizing the number of stops, and the other wait times. Such pure controllers are expected to perform very well in terms of the metric that they optimize, without any guarantees of good performance for the other metrics. Indeed we can see the waits-optimizing controller achieving very short wait times (∼35 seconds in Hangzhou and ∼58 seconds in NY16) and rather bad results in terms of stops in both scenarios. Conversely, the stops controller achieves very good number of stops (close to 0) and very high wait times. We also note here that the speed is correlated with both optimizing stops and wait times (with a somehow stronger correlation for wait times).

Let us refer to Figure 5 once again to highlight how to read our figures. A method that is performing well across all our metrics will be represented by a large, approximately equilateral, triangle. A weaker method that satisfies all objectives to a lesser extent, would be represented by a triangle with a smaller area but still equilateral. If a given method achieves good results on one objective and weak results on the other it will be represented by much smaller triangles (much like the triangles in Figure 5). Visually, the overall area of the triangle indicates a represented method's overall performance, while the shape of the triangle indicates how balanced the performance is with respect to different objectives.

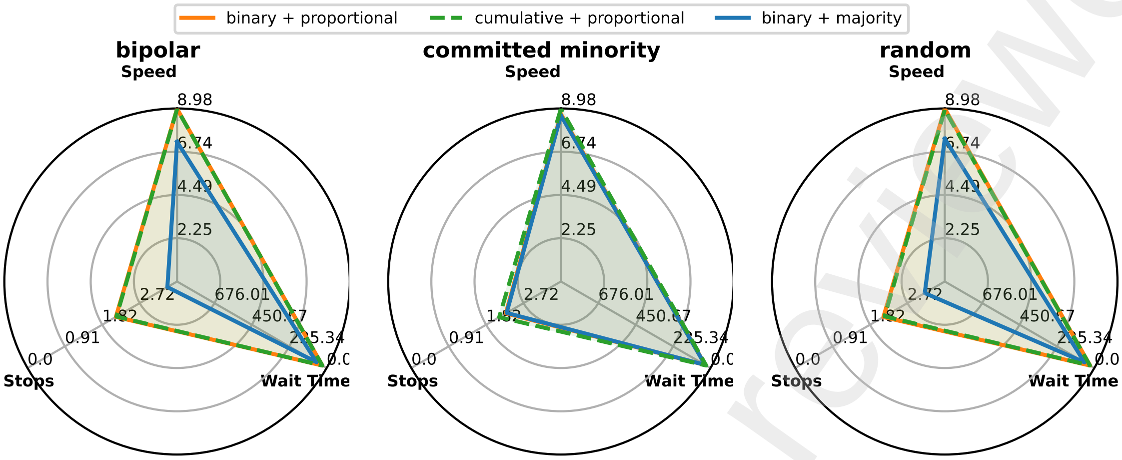


Figure 6: Radar plots for the Hangzhou scenario with three evaluation metrics (stops (seconds), wait times (seconds), and speeds (meters per second)) for three different voting distributions. Larger triangle areas can be associated with a more desirable performance across all metrics.

We report the results of the system performance by means of averages of three different metrics: Wait Times (seconds), Stops (number of stops), and Speed (meters per second) for both traffic scenarios, i.e. Hangzhou (Figure 6) and NY16 (Figure 7).

The columns of Figure 6 and Figure 7 represent the different voting scenarios (bipolar, committed minority, and random) and the different colored triangles correspond to the different input/aggregation methods. As seen in Figure 6 the binary proportional and cumulative proportional methods perform on the same level in the bipolar and random scenarios (the triangles are overlapped). The binary majority method performs significantly worse in the bipolar and random scenarios in all three metrics. In the committed minority scenario, the three voting methods perform at a similar level, with the cumulative proportional method slightly outperforming the other two.

Based on Figure 7, we note that the cumulative proportional method outperforms the other two in the committed minority voting scenario. In that scenario, the two binary methods reach very low wait times, but make sacrifices in terms of a larger number of stops, while the cumulative proportional method achieves better performance in terms of stops,

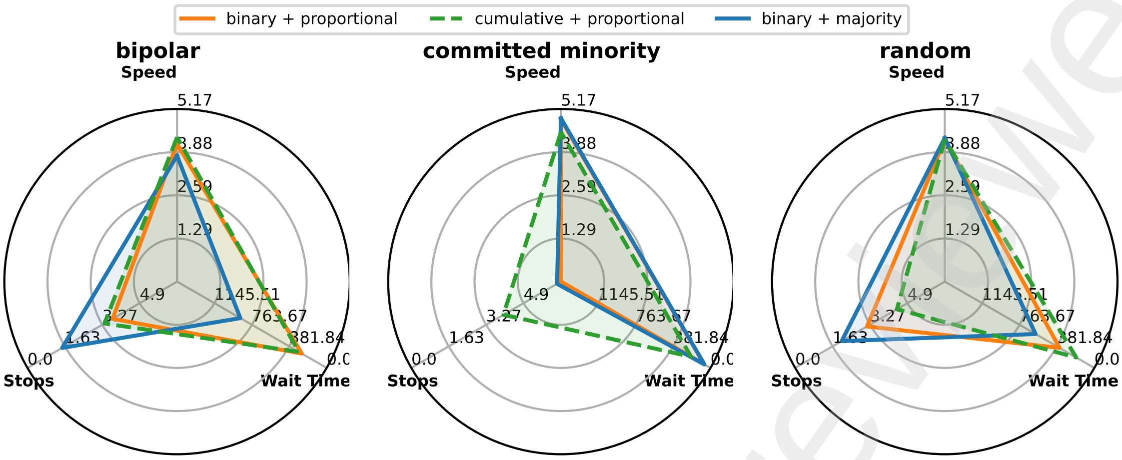


Figure 7: Radar plots for the NY16 scenario with three evaluation metrics (stops (seconds), wait times (seconds), and speeds (meters per second)) for three different voting distributions. Larger triangle areas can be associated with a more desirable performance across the multiple (three) performance metrics.

while maintaining a good result in terms of wait times. In the bipolar scenario, cumulative proportional and binary proportional methods perform at a similar level—slightly worse than the binary majority method in terms of the number of stops, but much better in terms of wait times. For the random scenario, the binary proportional and cumulative proportional methods perform at a comparable level, while the binary majority method prioritizes the stops at significant costs in terms of wait times, leading to a worse performance overall.

We also provide an alternative visualization of the multi-objective performance in Figure 8. Here, we rescaled the lower and upper limits of the radar plot dimensions to obtain objective score indices ranging from 0 to 1. Thus, summing up over these objective score indices across all of the evaluation metrics yields a scalar metric that indicates the overall multi-objective performance of each voting scenario. Overall, Figure 8 shows that, with proportional voting, we can achieve multi-objective scores higher than just choosing a pure controller optimized solely for waits in a grid-like (Hangzhou) traffic scenario. In the arterial (NY16) traffic scenario, employing proportional voting achieves a total objective score that is basically on par with the pure controller. Moreover, all voting-based results largely outperform the pure stops scenario.

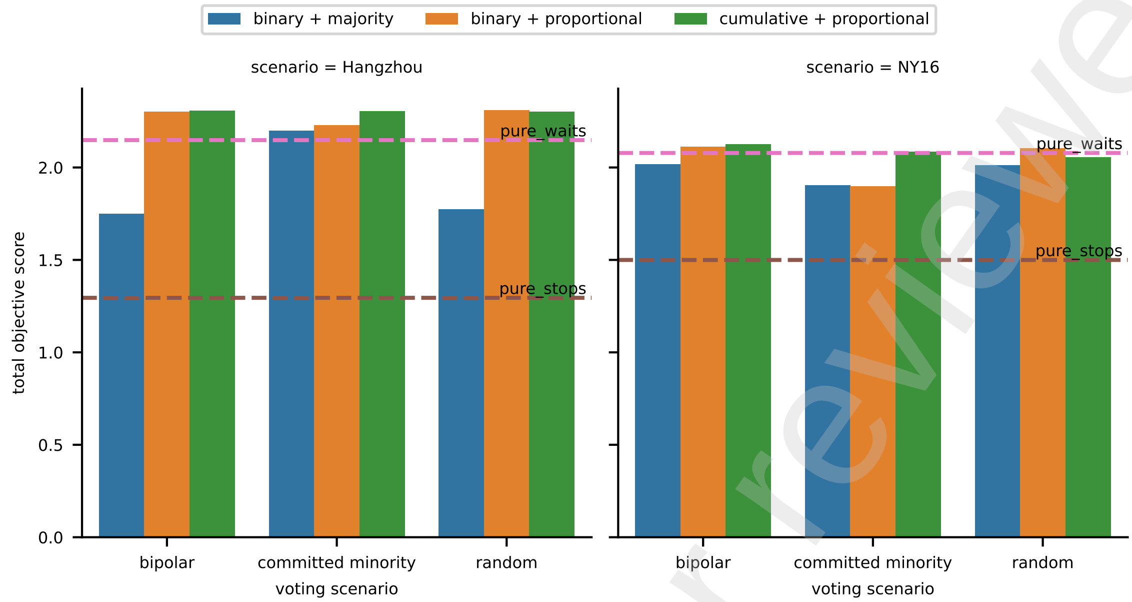


Figure 8: Total score across multiple objectives expressed as a scalar index ranging from 0 to 3. We present results for the grid (Hangzhou) and arterial road (NY16) scenarios. The dashed lines on the axis reflect the multi-objective scores of the pure objective controllers. Cumulative proportional voting yields results that are better than or at par with the purely wait-oriented controller across multiple scenarios, and much better than the purely stops-oriented controller.

4.2. Value Alignment

The results in terms of the Value Alignment of the system's actions with the values of the system's users following the Equation 4 are presented in this section.

In Figure 9, we present the value alignment results for the Hangzhou traffic scenario. The rows represent the different voting methods, while the columns represent the different preference distributions of the voters. We note that the binary proportional and cumulative proportional methods appear to lead to highly similar value alignment distributions in all three voting preference conditions. In the bipolar scenario, they produce a multi-modal distribution, where the populations around the two modes are of equal size. The binary majority method, on the other hand, induces a uni-modal value alignment distribution with a higher average value alignment than the other two methods. In the random scenario, the binary majority leads to a higher average value alignment with a uni-modal distribution having a higher peak as compared to the more spread-out distributions induced by cumulative proportional and binary proportional methods. In the committed minority condition, all three methods induce a similar value alignment distribution.

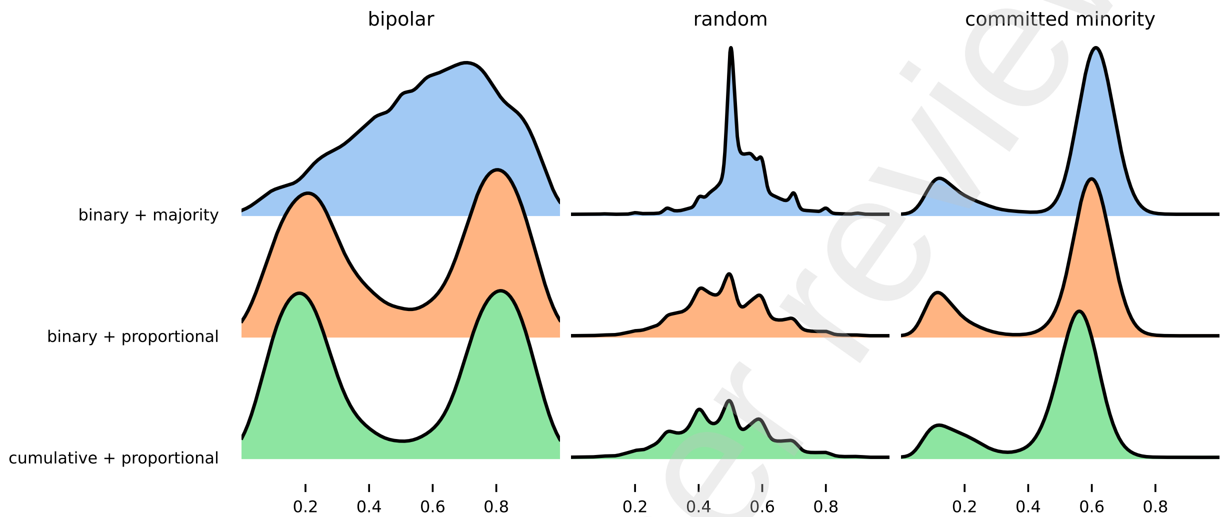


Figure 9: Distribution of the value alignment for drivers in the Hangzhou traffic scenario for the three voting preference distributions (columns) and three input/aggregation methods (rows).

In Figure 10, we present the value alignment results for the NY16 scenario. Here the value alignment is very similar across the voting methods for the bipolar and random scenarios. However, for the committed minority scenario the cumulative proportional method leads to a higher value alignment on average. The two binary methods produce a multi-modal distribution, where a minority suffers of poor value alignment and a majority enjoys a high value alignment. The value alignment for the cumulative proportional method is uni-modal with small bumps at the low and high ends of the distribution. In the NY16 case, the cumulative proportional method enables minorities to protect their interests, fostering a more balanced and inclusive outcome.

5. Discussion

Here we discuss the results reported in the previous section.

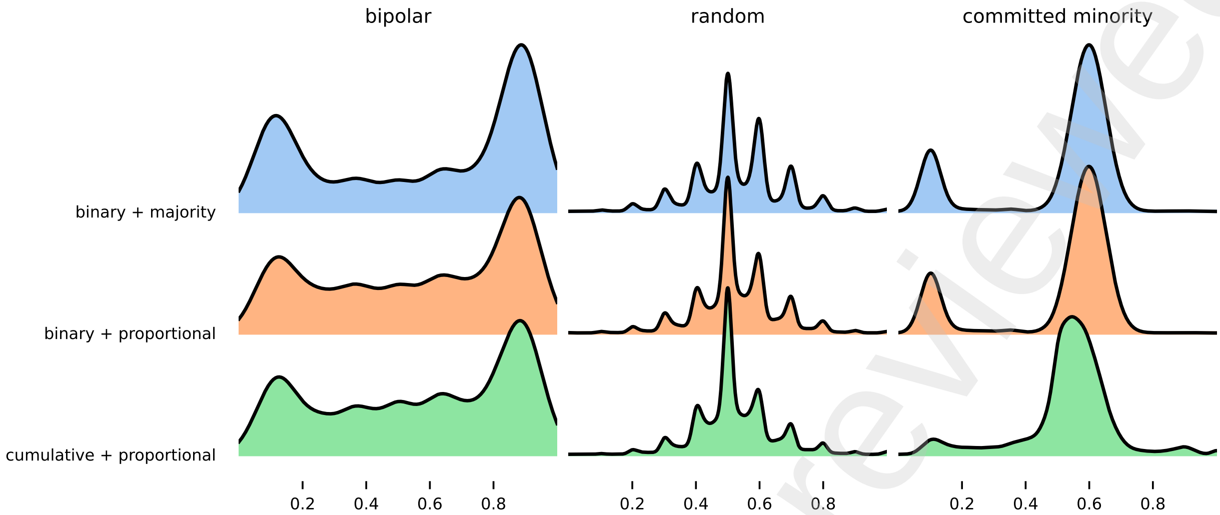


Figure 10: Distribution of the value alignment for drivers in the NY16 traffic scenario for the three voting preference distributions (columns) and three input/aggregation methods (rows).

In terms of the system's performance, we note that the cumulative proportional method and the binary proportional method perform the best in the Hangzhou scenario. The cumulative proportional method performs best in the NY16 scenario, outperforming binary proportional in the committed minority condition and performing on a similar level in the bipolar and random scenarios. Thus, altogether we can state that the cumulative proportional voting method appears to lead to the best performance in terms of the three metrics we use to measure it, followed closely by the binary proportional method. The binary majority method tends to perform worse than the two methods using proportional aggregation. This is explained by the fact that the majority voting does not allow taking the relative Q-values into account, as visualized in Figure 3. Thus, only the prevailing preferences of the voters are taken into account, which in some cases might lead to highly sub-optimal decisions in terms of one or both of the objectives. This suggests design choices for our system that are capable of taking into account not only user preferences, but also the consequences of expected actions on the overall system performance.

We also note that in the committed minority condition of the NY16 scenario, the cumulative proportional method is able to include the minority position (which has a strong preference for the number of stops). The two other methods fail to do this, as indicated by their lower results in terms of the number of stops. Overall, this indicates the potential of the cumulative method to account for minority positions, leading to more inclusive outcomes. Conversely, in the NY16 bipolar condition, the binary majority method ignores half of the population with a 50–50 preference split and prioritizes only the other half with a strong preference for stops. This clearly shows the limitations of majority aggregation.

Regarding the Value Alignment, the voting aggregation method seems to have a more significant influence on it than the voting input method. This is due to the fact that the binary aggregation method discounts for the influence of the Q-values in the decision process entirely. This points our attention to the potential tension between the Q-values induced decisions and value alignment. Indeed the inclusion of Q-values in the decision process (allowed by the proportional aggregation mechanism) can lead to situations, where an action that would have been indicated by the voting method alone is not chosen (thereby leading to a lower alignment) . While this may lower the alignment, it leads to higher system performance. This tension should nevertheless be taken into account when designing such systems. As we have shown, with the cumulative proportional method, one is able to achieve a decent Value Alignment in most cases, while not sacrificing system performance overall—getting the best of both worlds so to speak.

6. Conclusions

In this work, we have extended traffic signal control by voting methods, which we have applied to larger action spaces and more realistic, bigger traffic scenarios. This approach empowers drivers to influence how traffic lights control intersections, based on the local preferences of the drivers using such intersections. Thereby, humans are put back into the equations and algorithms used to manage smart societies (which have been increasingly automated recently, often removing agency of people). With our approach, humans get back decision-making power, in a way that benefits systemic performance.

Interestingly, the cumulative proportional method can often successfully include minority interests without sacrificing system performance. In contrast, majority aggregation, which is currently used in most democratic parliaments, is particularly prone to producing a tension between user preferences and system performance. We have taken care of this tension with our system design, which attempts to balance short-term preferences of users with long-term consequences of following them. While our computational results are based on DQNs, the underlying voting mechanism(s) should, in principle, be extendable to any sort of control algorithm that associates scores (instead of Q-values) to a finite set of actions.

We believe that our methodology offers a strong alternative to technocratic top-down approaches currently used to control traffic systems in many cities. We expect that our democratic, user-driven, bottom-up approach presented here can be transferred to many other complex systems with a few adjustments. It should be easy to generalize this novel approach to applications beyond traffic, especially to smart cities and other socio-economic systems, where limited resources need to be shared.

Overall, our impression is that the state of the world could be improved, if foundational principles of democracies, namely suitable voting methods, were considered more often in the design of AI systems. This is particularly relevant when complex systems shall be managed in ways that can achieve multiple goals simultaneously. We propose to explore this approach further in future case studies and careful real-life experiments or laboratory experiments (potentially using VR technology), with a focus also on user experience.

In many settings, it is important to acknowledge that people are different, among others for genetic and cultural reasons, and due to different personal histories. The diversity resulting from this has many advantages for societies, for example, in terms of division of labor, but also as a driver of innovation [49] and economic wealth [50], furthermore as an enabling factor of collective intelligence [51–53] and basis of societal resilience [54]. Such diversity comes with diverse individual goals, which are the basis of modern, pluralistic societies. Those societies thrive if they manage to meet the goals of many people, such that they can contribute with their individual talents, engagement, and ideas.

In fact, our presented work demonstrates the power of a novel approach, which allows one to satisfy many, even potentially incompatible goals with a high level of performance. It is an approach, which allows one to manage or even self-govern complex systems, and can be well combined with artificial intelligence methods. For this reason, it makes significant advances towards "inclusive artificial intelligence (AI) governance and interoperability to achieve our common vision and goal of trustworthy AI, in line with our shared democratic values," as it was recently demanded [10].

Glossary

Term	Definition
ML	Machine Learning
RL	Reinforcement Learning
AI	Artificial Intelligence
DQN	Deep Q-Networks
VA	Value Alignment
SCT	Social Choice Theory
State	Vector of observations passed to the DQN
Action	Chosen phase to be activated by the DQN
Reward	Quantity that signals (dis)incentives to the RL agent
Binary Voting	Voting scheme where only one option is chosen by voters
Cumulative Voting	Voting scheme where voters can distribute points among options
Majority Aggregation	Aggregation scheme where the option with the most points "wins"
Proportional Aggregation	Aggregation scheme where the outcome of the vote considers point allocation

Acknowledgments

This work has benefited from funding by the European Union's Horizon 2020 research and innovation programme through the DIT4TraM project (Grant Agreement no. 953783). It was further inspired by the ERC Advanced Investigator Grant on Co-Evolving City Life and an SNF project on Digital Democracies within the NFP 77 "Digital Transformation". Last but not least, MK was partially supported by an ETH Zurich Doc.Mobility Fellowship.

References

[1] E. A. Mueller, Aspects of the history of traffic signals, IEEE Transactions on Vehicular Technology 19 (1) (1970) 6–17. doi:10.1109/T-VT.1970.23426.

[2] A. Hamilton, B. Waterson, T. Cherrett, A. Robinson, I. Snell, The evolution of urban traffic control: changing policy and technology, Transportation planning and technology 36 (1) (2013) 24–43. doi:10.1080/03081060.2012.745318.

[3] T. Urbanik, A. Tanaka, B. Lozner, E. Lindstrom, K. Lee, S. Quayle, S. Beaird, S. Tsoi, P. Ryus, D. Gettman, et al., Signal timing manual, Vol. 1, Transportation Research Board Washington, DC, 2015.

[4] S. Lämmer, D. Helbing, Self-control of traffic lights and vehicle flows in urban road networks, Journal of Statistical Mechanics: Theory and Experiment 2008 (4) (2008) 1–36. doi:10.1088/1742-5468/2008/04/P04019.

[5] M. Korecki, Deep reinforcement meta-learning and self-organization in complex systems: Applications to traffic signal control, Entropy 25 (7). doi:10.3390/e25070982.

[6] C. Carissimo, M. Korecki, Limits of optimization, Minds and Machines (2023) 1–21doi:10.1007/s11023-023-09633-1.

[7] D. Helbing, B. S. Frey, G. Gigerenzer, E. Hafen, M. Hagner, Y. Hofstetter, J. v. d. Hoven, R. V. Zicari, A. Zwitter, IT-revolution: Digitale Demokratie Statt Datendiktatur (Dec 2015).
URL https://www.spektrum.de/news/wie-algorithmen-und-big-data-unsere-zukunft-bestimmen/1375933

[8] S. L. Dennisen, J. P. Müller, Agent-based voting architecture for traffic applications, in: Multi-agent System Technologies, Springer International Publishing, 2015, pp. 200–217. doi:10.1007/978-3-319-27343-3_11.

[9] R. Hendrych, Democratic traffic lights, Ph.D. thesis, University of Dublin (2009).

[10] G7 Hiroshima Leaders, G7 hiroshima leaders' communiqué (2023).
URL https://www.g7hiroshima.go.jp/documents/pdf/Leaders_Communique_01_en.pdf

[11] C. I. Hausladen, R. Hänggli, D. Helbing, R. Kunz, J. Wang, E. Pournaras, On the legitimacy of voting methods, Available at SSRN 4372245doi:10.2139/ssrn.4372245.

[12] C. Gershenson, Self-organizing urban transportation systems, Complexity theories of cities have come of age (2012) 269–279doi:10.1007/s11023-023-09633-1.

[13] G. Caldarelli, E. Arcaute, M. Barthelemy, M. Batty, C. Gershenson, D. Helbing, S. Mancuso, Y. Moreno, J. Ramasco, C. Rozenblat, et al., The role of complexity for digital twins of cities, Nature Computational Science 3 (2023) 1–8. doi:10.1038/s43588-023-00431-4.

[14] E. Manley, T. Cheng, Understanding road congestion as an emergent property of traffic networks, in:

14th WMSCI, Citeseer, 2010.

[15] C. Gershenson, Self-organizing traffic lights, Complex Systems 16.

[16] D. Helbing, S. Lämmer, J.-P. Lebacque, Self-organized control of irregular or perturbed network traffic, Optimal Control and Dynamic Games: Applications in Finance, Management Science and Economics (2005) 239–274doi:10.1007/0-387-25805-1_15.

[17] G. Zheng, X. Zang, N. Xu, H. Wei, Z. Yu, V. Gayah, K. Xu, Z. Li, Diagnosing reinforcement learning for traffic signal control (2019). doi:10.48550/arXiv.1905.04716.

[18] M. Korecki, D. Dailisan, C. Carissimo, Dynamic value alignment through preference aggregation of multiple objectives, Submitted to Transactions on Machine Learning Research (Under review).
URL https://openreview.net/forum?id=CsFQKrfGvX

[19] M. Korecki, D. Helbing, Analytically guided reinforcement learning for green it and fluent traffic, IEEE Access 10 (2022) 96348–96358. doi:10.1109/ACCESS.2022.3204057.

[20] X. Zang, H. Yao, G. Zheng, N. Xu, K. Xu, Z. Li, Metalight: Value-based meta-reinforcement learning for traffic signal control, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, 2020, pp. 1153–1160. doi:10.1609/aaai.v34i01.5467.

[21] M. Korecki, Adaptability and sustainability of machine learning approaches to traffic signal control, Scientific Reports 12 (1) (2022) 16681. doi:10.1038/s41598-022-21125-3.

[22] M. Korecki, D. Dailisan, D. Helbing, How well do reinforcement learning approaches cope with disruptions? the case of traffic signal control, IEEE Access 11. doi:10.1109/ACCESS.2023.3266644.

[23] D. Jalota, J. Lazarus, A. Bayen, M. Pavone, Credit-based congestion pricing: Equilibrium properties and optimal scheme design (2022). doi:10.48550/arXiv.2210.15907.

[24] E. Elokda, C. Cenedese, K. Zhang, J. Lygeros, F. Dörfler, Carma: Fair and efficient bottleneck congestion management with karma (2022). doi:10.48550/arXiv.2208.07113.

[25] D. Levinson, Equity effects of road pricing: A review, Transport Reviews 30 (1) (2010) 33–57. doi:10.1080/01441640903189304.

[26] N. Wiener, Some moral and technical consequences of automation: As machines learn they may develop unforeseen strategies at rates that baffle their programmers., Science 131 (3410) (1960) 1355–1358. doi:10.1126/science.131.3410.1355.

[27] S. Russell, Provably beneficial artificial intelligence, in: 27th International Conference on Intelligent User Interfaces, 2022, pp. 3–3. doi:10.1145/3490099.3519388.

[28] S. Russell, Artificial intelligence and the problem of control, Perspectives on Digital Humanism (2022) 19doi:10.1007/978-3-030-86144-5_3.

[29] D. Hadfield-Menell, A. Dragan, P. Abbeel, S. Russell, The off-switch game, in: Workshops at the Thirty-First AAAI Conference on Artificial Intelligence, 2017. doi:10.48550/arXiv.1611.08219.

[30] T. Walsh, The singularity may never be near, AI Magazine 38 (3) (2017) 58–62.

[31] D. Hadfield-Menell, G. K. Hadfield, Incomplete contracting and ai alignment, in: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 2019, pp. 417–422. doi:10.1145/3306618.3314250.

[32] I. Gabriel, Artificial intelligence, values, and alignment, Minds and Machines 30 (3) (2020) 411–437. doi:10.1007/s11023-020-09539-2.

[33] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, D. Mané, Concrete problems in ai safety (2016). doi:10.48550/arXiv.1606.06565.

[34] P. Eckersley, Impossibility and Uncertainty Theorems in AI Value Alignment (or why your AGI should not have a utility function) (2019). doi:10.48550/arXiv.1901.00064.

[35] A. Sen, Chapter 1 – INTRODUCTION, in: A. Sen (Ed.), Collective Choice and Social Welfare, Vol. 11 of Advanced Textbooks in Economics, North-Holland, Amsterdam, 1979, pp. 1–6. doi:10.1016/B978-0-444-85127-7.50007-1.

[36] A. Fickinger, S. Zhuang, A. Critch, D. Hadfield-Menell, S. Russell, Multi-principal assistance games: Definition and collegial mechanisms (2020). doi:10.48550/arXiv.2012.14536.

[37] A. Sen, FROM SOCIAL CHOICE TO DEVELOPMENT, THE INFLUENCE OF NICOLAS DE CONDORCET AND JEAN-CHARLES DE BORDA, Revue Tiers Monde 198 (2) (2009) 263–267. doi:10.3917/rtm.198.0263.

[38] F. Brandt, V. Conitzer, U. Endriss, J. Lang, A. D. Procaccia, Handbook of computational social choice, Cambridge University Press, 2016.

[39] M. Prasad, Social choice and the value alignment problem, Artificial intelligence safety and security (2018) 291–314.

[40] S. D. Baum, Social choice ethics in artificial intelligence, AI & SOCIETY 35 (1) (2020) 165–176. doi:10.1007/s00146-017-0760-1.

[41] D. Helbing, S. Mahajan, R. Hänggli, A. Musso, C. I. Hausladen, C. Carissimo, D. Carpentras, E. Stockinger, J. Argota Sánchez-Vaquerizo, J. Yang, et al., Democracy by design: Perspectives for digitally assisted, participatory upgrades of society, Participatory Upgrades of Society (November 2, 2022) 71. doi:10.1016/j.jocs.2023.102061.

[42] K. O. May, A set of independent necessary and sufficient conditions for simple majority decision, Econometrica: Journal of the Econometric Society 20 (4) (1952) 680–684. doi:10.2307/1907651.

[43] S. Bhagat, J. A. Brickley, Cumulative voting: The value of minority shareholder voting rights, The Journal of Law and Economics 27 (2) (1984) 339–365.
URL http://www.jstor.org/stable/725580

[44] N. d. Condorcet, Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité

des voix, Cambridge Library Collection - Mathematics, Cambridge University Press, 2014. doi:10.1017/CBO9781139923972.

[45] H. Ashworth, T. Ashworth, Proportional Representation Applied To Party Government: A New Electoral System, Project Gutenberg, Project Gutenberg., 2004.
URL https://books.google.ch/books?id=Zp4DxQEACAAJ

[46] K. J. Arrow, Social Choice and Individual Values, Yale University Press, 1951.
URL http://www.jstor.org/stable/j.ctt1nqb90

[47] A. Gibbard, Manipulation of voting schemes: a general result, Econometrica: journal of the Econometric Society 41 (4) (1973) 587–601. doi:10.2307/1914083.

[48] H. Zhang, Y. Ding, W. Zhang, S. Feng, Y. Zhu, Y. Yu, Z. Li, C. Liu, Z. Zhou, H. Jin, CityFlow: A multi-agent reinforcement learning environment for large scale city traffic scenario, in: The World Wide Web Conference, 2019. doi:10.1145/3308558.3314139.

[49] M. P. Feldman, D. B. Audretsch, Innovation in cities:: Science-based diversity, specialization and localized competition, European economic review 43 (2) (1999) 409–429. doi:10.1016/S0014-2921(98)00047-6.

[50] L. M. Bettencourt, H. Samaniego, H. Youn, Professional diversity and the productivity of cities, Scientific reports 4 (1) (2014) 5393. doi:10.1038/srep05393.

[51] J. Lorenz, H. Rauhut, F. Schweitzer, D. Helbing, How social influence can undermine the wisdom of crowd effect, Proceedings of the national academy of sciences 108 (22) (2011) 9020–9025. doi:10.1073/pnas.1008636108.

[52] L. Hong, S. E. Page, Groups of diverse problem solvers can outperform groups of high-ability problem solvers, Proceedings of the National Academy of Sciences 101 (46) (2004) 16385–16389. doi:10.1073/pnas.0403723101.

[53] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, T. W. Malone, Evidence for a collective intelligence factor in the performance of human groups, science 330 (6004) (2010) 686–688. doi:10.1126/science.1193147.

[54] S. Mahajan, C. I. Hausladen, J. A. Sánchez-Vaquerizo, M. Korecki, D. Helbing, Participatory resilience: Surviving, recovering and improving together, Sustainable Cities and Society 83 (2022) 103942. doi:10.1016/j.scs.2022.103942.