Market Interfaces for Electric Vehicle Charging

Sebastian Stein   ss2@ecs.soton.ac.uk
Enrico H. Gerding   eg@ecs.soton.ac.uk
Adrian Nedea   adrian.nedea@gmail.com
University of Southampton, Southampton, United Kingdom

Avi Rosenfeld   rosenfa@jct.ac.il
Jerusalem College of Technology, Jerusalem, Israel

Nicholas R. Jennings   n.jennings@imperial.ac.uk
Imperial College, London, United Kingdom
King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

We consider settings where owners of electric vehicles (EVs) participate in a market mechanism to charge their vehicles. Existing work on such mechanisms has typically assumed that participants are fully rational and can report their preferences accurately via some interface to the mechanism or to a software agent participating on their behalf. However, this may not be reasonable in settings with non-expert human end-users.Thus, our overarching aim in this paper is to determine experimentally if a fully expressive market interface that enables accurate preference reports is suitable for the EV charging domain, or, alternatively, if a simpler, restricted interface that reduces the space of possible options is preferable. In doing this, we measure the performance of an interface both in terms of how it helps participants maximise their utility and how it affects deliberation time. Our secondary objective is to contrast two different types of restricted interfaces that vary in how they restrict the space of preferences that can be reported. To enable this analysis, we develop a novel game that replicates key features of an abstract EV charging scenario. In two experiments with over 300 users, we show that restricting the users' preferences significantly reduces the time they spend deliberating (by up to half in some cases). An extensive usability survey confirms that this restriction is furthermore associated with a lower perceived cognitive burden on the users. More surprisingly, at the same time, using restricted interfaces leads to an increase in the users' performance compared to the fully expressive interface (by up to 70%). We also show that some restricted interfaces have the desirable effect of reducing the energy consumption of their users by up to 20% while achieving the same utility as other interfaces. Finally, we find that a reinforcement learning agent displays similar performance trends to human users, enabling a novel methodology for evaluating market interfaces.

1. Introduction

There is a growing awareness that our use of energy is causing unsustainable stress on the environment (Stern, 2006). To mitigate this, governments are increasingly advocating the widespread adoption of electric vehicles (EVs), coupled with a switch to renewable energy sources (Royal Academy of Engineering, 2010). However, this transition poses considerable challenges. A significant uptake in EVs will place unprecedented strains on the existing electricity infrastructure (Ipakchi & Albuyeh, 2009). At the same time, renewable energy is

typically intermittent, creating a potential mismatch between supply and demand (Lund, 2007).

One promising approach for dealing with these emerging challenges is the use of market mechanisms, which enable the efficient allocation of scarce resources in multi-agent systems (Ramchurn, Vytelingum, Rogers, & Jennings, 2012; Samadi, Mohsenian-Rad, Schober, & Wong, 2012). Specifically, previous work has proposed auction-like mechanisms for scheduling the charging of EVs (Mets, D'hulst, & Develder, 2012; Robu, Gerding, Stein, Parkes, Rogers, & Jennings, 2013; Hayakawa, Gerding, Stein, & Shiga, 2015). These achieve a high efficiency because they take the individual preferences of drivers (i.e., their availability and willingness to pay) into account when allocating a limited supply of electricity. Other work relies on real-time price signals to incentivise autonomous charging agents to shift or curtail their consumption when supply is low (Ramchurn, Vytelingum, Rogers, & Jennings, 2011; Flath, Ilg, Gottwalt, Schmeck, & Weinhardt, 2014).

However, such approaches assume that the human end-users have perfect knowledge of their preferences, i.e., they can reason accurately about the value of different amounts of electricity, considering all possible, often uncertain future opportunities of using it. This is often not realistic (Simon, 1972; Kahneman, 2000). Moreover, providing the complete preferences is tedious and the associated cost could outweigh the benefits of doing so (Larson & Sandholm, 2005). Specifically, in the electric vehicle charging setting we consider here, this would require users to reason about the potential journeys they may wish to complete in the future, the value of these journeys, and their inherent uncertainty (which may be due to external events such as the weather, decisions by friends, family and work colleagues, or even emergencies).

To address this challenge, there has been some work on auctions with *restricted reporting*, i.e., where bidders do not report their full preferences, but rather choose from a restricted set of possible bids. Such auctions can lead to equilibria with certain desirable properties, including higher revenue for the auctioneer or better computational tractability (Milgrom, 2010; Dütting, Fischer, & Parkes, 2011; Dütting, Henzinger, & Starnberger, 2013). Specifically, some work has analysed settings where the messaging space is reduced to a small set of discrete options. For example, Bergemann, Shen, Xu, and Yeh (2012) describe how to select these options to maximise either social welfare or revenue, and Blumrosen and Feldman (2013) characterise the associated loss of efficiency. Other work, e.g., by Sandholm and Boutilier (2006) and Baarslag and Gerding (2015), considers how complex preferences can be elicited through an incremental query process. However, these approaches all assume rational agents and do not evaluate the auctions with non-expert human bidders.

Another strand of work explicitly considers non-expert market participants. For example, research on hidden market design has looked at building simple interfaces that hide the rules and pricing mechanisms of a complex underlying market (Seuken, Jain, Tan, & Czerwinski, 2010b; Seuken, Charles, Chickering, & Puri, 2010a). However, that work cannot be applied directly to EV charging, as it considers an exchange market for computational storage without financial payments. Related to this, Seuken, Parkes, Horvitz, Jain, Czerwinski, and Tan (2012) investigated market user interface design. That work focuses on simplifying complex market interactions by asking users to select from a discrete set of options. However, it considers a setting with posted prices (as opposed to more complex auctions)

and is significantly simpler than the EV setting. Crucially, neither work provides a direct comparison to fully expressive interfaces, which leaves their relative benefits unclear.

To address these limitations, we conduct the first study with human participants that specifically considers how to design market interfaces for the EV charging setting and that directly compares fully expressive and restricted interfaces. The key research questions we seek to answer are whether fully expressive or restricted interfaces are preferable in this setting (in terms of maximising the participants' utility, and reducing their deliberation time and perceived cognitive burden) and which types of restricted interfaces work best (i.e., a small set of discrete options or a large but lower-dimensional reporting space). In so doing, we make several novel contributions:

- First, we formalise the EV charging setting to capture several real-world challenges that give rise to complex preferences and we design two restricted interfaces for reporting preferences in this setting: one that reduces the dimensionality of the reporting space (but retains an infinite number of options) and one that restricts the reporting space to a discrete set of options.

- To evaluate these interfaces with real users, we then develop a game that serves as an abstract representation of the EV charging setting. Using this game, we experimentally compare the restricted interfaces to a fully expressive interface in two large user studies involving a total of over 300 participants hired on Amazon Mechanical Turk, showing that:

  - restricted interfaces significantly alleviate the participants' cognitive burden. We measure this through a range of metrics, including the time they take to interact with the market (halving their deliberation time in some cases) and a number of usability metrics collected through a survey (showing that restricted interfaces lead to lower perceived mental demand, less frustration and higher ease of use).

  - restricted interfaces lead to a significantly better performance than the fully expressive interface (up to 70% in some cases). This difference in performance is most pronounced for new users.

  - some interfaces lead to the same performance for their users, but induce different behaviours. Specifically, a restricted mechanism that offers users the option to set a maximum charge amount leads to a 20% reduction in energy consumption in some settings (while otherwise achieving the same performance as another restricted mechanism without this option).

  - a reinforcement learning agent displays similar general behaviour trends as the human players in our experiments. This potentially paves the way for a new approach for optimising market interfaces using reinforcement learning agents to predict the responses of human users, thereby reducing the need for large-scale user trials.

The remainder of this paper is structured as follows. In Section 2, we discuss related work, focussing on market and mechanism design, as well as human-computer interaction. In Section 3, we then formalise the EV charging problem, including how we model the market

mechanism and the driver's travel opportunities. Given this formalisation, we discuss how to solve it in Section 4, using either an optimal model-based algorithm or a model-free reinforcement learning approach. These solutions serve as benchmarks against which we later compare human behaviour. In Section 5, we then formalise the restricted market interfaces we consider in this paper and in Section 6, we describe Bid2Charge, the game we use to evaluate our market interfaces. We discuss our experimental setup and results in Section 7 and finally conclude in Section 8.

2. Related Work

There is already a significant body of literature that deals specifically with the problem of charging EVs within the constraints of a limited electricity distribution infrastructure. Some address this by using a central planner to coordinate the charging of EVs (Clement-Nyns, Haesen, & Driesen, 2010; Petrou, Quiros-Tortos, & Ochoa, 2015), while others use decentralised control algorithms (Ahn, Li, & Peng, 2011). However, those approaches assume that EV drivers are cooperative and provide accurate information about their charging requirements. This neglects the fact that drivers are typically self-interested agents that may strategise about the information they reveal (if this is to their own benefit) and that need to be motivated through appropriate incentives. To address this, recent work on EV charging has looked at market-based approaches, which either incentivise truthful reporting (Robu et al., 2013; Hayakawa et al., 2015) or balance supply and demand through real-time prices (Ramchurn et al., 2011; Flath et al., 2014).

However, these market mechanisms have an inherent complexity that must be addressed for non-experts to use them effectively. Thus, in Section 2.1, we discuss existing work that has looked at restricting the complexity of markets. Then, we look more generally at the related area of designing human-computer interaction mechanisms in complex environments (Section 2.2). That work motivates the guiding hypothesis of this paper that restricted interfaces, which intentionally limit the information that can be reported by users, are beneficial in real-world markets with non-expert participants.

2.1 Markets and Mechanism Design

Many markets in practice use mechanisms which are not fully expressive, i.e., where bidders are not able to fully express their entire utility function (Milgrom, 2010). A well-known example is the auction market for selling radio frequency spectrum licences for mobile, radio and television. Preferences in such settings are complex since there are complementarities (i.e., a set of multiple items may be more valuable than its constituent parts individually) as well as interdependencies (i.e., an agent's utility depends on the private information known by other agents), but the auctions used, mostly variants of simultaneous ascending clock auctions, often do not package bids and consequently do not allow complex preferences to be fully expressed (Cramton, 2010). This is because it would simply be too costly or even infeasible for bidders to fully specify all the dependencies, and indeed this is not always needed to obtain efficient outcomes (Ausubel, 2004; Hartline & Roughgarden, 2009). Another example is online advertising. In sponsored search (where ads appear alongside search results, also called position or keyword auctions), several ads are auctioned simultaneously by bidding on (combinations of) keywords, and top ads are more valuable. However, the

mainstream mechanism, called the generalised second-price (GSP) auction, does not allow bidders to express different valuations for different slots, resulting in inefficiencies (Benisch, Sadeh, & Sandholm, 2008b). Similarly, in display or banner advertising, multiple slots on the same webpage are auctioned off simultaneously but, typically, in separate ad auctions, preventing the expression of combinatorial preferences (such as substitutes and complements). The main advantage is that these mechanisms are not only simpler to understand, but they are also computationally simpler. This computational simplification is critical in online settings where winners need to be determined within milliseconds. While many papers exist that analyse these applications, the focus is often on the performance (in terms of revenue and/or efficiency) of the specific auction rather than explicitly comparing auctions based on their expressiveness, which is what we do in this paper.

Several papers do explicitly consider the effect of expressiveness of the auction design. Specifically, Blumrosen and Feldman (2006, 2013) study the level of discretisation of the action space (i.e., the bids) in single-parameter domains and its effect on revenue. A more general approach is taken by Benisch, Sadeh, and Sandholm (2008a). In their work, the authors note that, in general, expressiveness cannot be measured by the dimensionality of the agents' reports since any finite-dimensional report can be mapped to a single-dimensional one without loss of expressiveness. Instead, they propose new measures of expressiveness based on impact and show that more expressive mechanisms are more efficient. On the other hand, Milgrom (2010) shows that a restricted message space can lead to improved revenue by excluding low-revenue equilibria. This work is extended by Dütting et al. (2011), who consider restricted message spaces for both sponsored search and combinatorial auctions, and consider settings with both complete and incomplete information about other agents' valuations. These papers (Milgrom, 2010; Dütting et al., 2011) focus on so-called *tight* simplifications, where no new equilibria are introduced. In contrast, Dütting et al. (2013) consider a range of classes of non-fully expressive preference representations in combinatorial auctions, from additive to subadditive representations, but where the actual preferences are more general. They analyse the widely-used Vickrey-Clarke-Groves (VCG) mechanism, which is no longer truthful in such settings. They show that computing pure Nash equilibria is NP-hard when the representation is additive but the preference domain is subadditive. Finally, they compute bounds on the price of anarchy (i.e., the proportion between the outcome maximising social welfare and the social welfare of the worst equilibrium outcome) and show that welfare loss w.r.t. to the worst-case Nash equilibrium increases with expressiveness.

All of the work discussed so far analyses the auctions theoretically. However, given that often multiple equilibria exist and people do not often behave as predicted, especially in complex settings (Kahneman, 2000; Rosenfeld & Kraus, 2012), it is not clear how humans will behave when faced with different trade-offs between simplicity and expressiveness. The closest work that analyses different market mechanisms with human trials is by Seuken et al. (2010b, 2012). However, the markets analysed in those papers do not consist of auctions but rather a fixed set of packages offered to the user, and the aim is to present those choices as to maximise efficiency. Crucially, other than varying the number of choices (from 3 to 6), the complexity of the interfaces remains the same. This is different from our work, where we analyse three types of designs, in which the restrictions in terms of expressiveness, and the complexity for the user, differ substantially.

2.2 Human-Computer Interaction

While auctions may be an effective mechanism to allocate energy, to date there has been only limited research on how people could participate in such energy auctions. However, there has been some relevant work on how to design appropriate interfaces for more general home energy management applications. Specifically, a number of papers have focused on the visualisation of the amount of energy devices use, while implicitly attempting to advocate conversations through better user awareness of their consumption (Pierce & Paulos, 2012; Bartram, 2015). Other work has begun to consider how to design interfaces that explicitly promote sustainability by persuading a user to change their behaviour to be more energy efficient (Kluckner, Weiss, Schrammel, & Tscheligi, 2013; Coskun, Zimmerman, & Erbug, 2015). There is also parallel work that considers specifically how to design the in-car interface for an EV to better inform drivers about the current state of the vehicle as it is being driven (Strömberg, Andersson, Almgren, Ericsson, Karlsson, & Nåbo, 2011; Jung, Sirkin, Gür, & Steinert, 2015). This work has highlighted new challenges, e.g., about clearly presenting electrical concepts to non-expert users (Strömberg et al., 2011) or about informing them about the inherent uncertainty of an EV's range (Jung et al., 2015). These are issues that are relevant to our application, but they do not deal with the complexity of an EV charging market. Indeed, as has been pointed out by Pierce and Paulos (2012), there has been little research on what type of interface is best suited for proper engagement with emerging energy systems, such as the energy markets we consider for EVs in this paper.

The key difficulty in designing interfaces for complex markets is deciding what information should be presented to the user. Generally, one might assume that presenting the user with more information is always advantageous, as supported by Endsley's idea of Situation Awareness (SA) (Endsley, 1995). SA quantifies the user's ability to identify, process, and comprehend critical elements of information. Three levels of SA exist: perception, comprehension and projection. The lowest level, perception, assumes the user perceives, but does not fully understand the important elements in the environment. Comprehension indicates that the user not only sees the information, but also comprehends its significance. The third level assumes that user also can use the information to effectively plan her next activities. Based on SA, one might assume that interfaces that present the most information are always advantageous as they present the user with enough information to properly plan optimal activities.

These three SA levels loosely correspond to the different interface choices we evaluate in this paper. The most restricted interface we present has only three static choices, allowing the user's actions to depend only on broad perceptions but without the need for deep comprehension or projection. At the other extreme, the fully expressive interface we consider allows the user to fully control and plan their actions optimally, requiring both comprehension and detailed forward planning. Between these extremes is a middle ground, reflecting the second SA level, which requires the user to comprehend the meaning of specific bids but does not necessitate full consideration of all possible outcomes.

However, previous work has raised questions both as to the general importance of SA (Flach, 1995), and if creating interfaces that provide more information is useful (Seuken et al., 2010b, 2012; Elmalech, Sarne, Rosenfeld, & Erez, 2015b; Elmalech, Sarne, & Grosz, 2015a). Similar to our work, Onal, Schaffer, O'Donovan, Marusich, Yu, Gonzalez, and

Höllerer (2014) consider how much information should be presented to a user in an abstract trust game based on the iterated diner's dilemma. They found that the complexity of interfaces had a significant effect on the main performance metric, with more expressive interfaces allowing users to perform significantly better. Conversely, there is growing work that questions if people should be given all information as doing so may actually hurt their performance. Elmalech et al. (2015b) demonstrate that at times people should not be given optimal advice as their bounded rationality prevents them from properly processing and accepting it. This is related to our approach in this work, where we contrast fully expressive interfaces, which theoretically allow people to achieve optimal performance, with simpler interfaces that ostensibly entail lower levels of SA and that do not allow users to act optimally.

Another relevant strand of work within human-computer interaction is the automatic design of user interfaces (Hodes, Katz, Servan-Schreiber, & Rowe, 1997; Nichols, Myers, Higgins, Hughes, Harris, Rosenfeld, & Pignol, 2002; Gajos, Weld, & Wobbrock, 2010). These approaches compose interface components based on high-level descriptions and constraints, in some cases optimising the interfaces for a particular user (Gajos et al., 2010). In a similar manner, multivariate landing page optimisation (MVLPO) is widely used by organisations to optimise the placement of content within Web interfaces (Ash, 2011). These approaches are complementary to our own — while we are interested in the fundamental types of market interfaces that work well in an EV charging setting, their respective components and layouts could be further optimised using such approaches.

3. The EV Charging Problem

In this section, we present an abstract model of the EV charging problem. Our aim is to capture the key challenges that are inherent in the domain, while still retaining a succinct and general model.[1] More specifically, we capture the following challenges that are found in realistic settings:

- First, electricity has **no intrinsic value**. Its value instead depends on how it is utilised to complete journeys.

- Second, the problem is of an inherently **uncertain** nature. On the one hand, this is a key feature of the market itself, as supply and demand may fluctuate significantly over time. On the other hand, there may be considerable uncertainty over what journeys the driver needs to complete in the future.

- Third, there are **complementarities**, i.e., the driver's preferences are typically highly nonlinear over the *quantity* of electricity they receive. For example, a driver may require a minimum overnight charge of 10 kWh to drive to work the next day. Receiving any less has no value.

---

1. A succinct and general model here provides us with a basic framework that focuses on the main challenges and that is easy to explain to participants in our user trials. It also constitutes a solid basis for future work exploring other challenges in the EV domain, such as range anxiety or en-route charging.

Given this, we consider a general setting where EV drivers participate in a market mechanism to charge their vehicles. We abstract away from the particular market[2] and assume that a driver simply reports her preferences for charging to an autonomous agent at regular intervals (e.g., once a day) through an appropriate market interface. In this section and the next, we will assume that the market interface is fully expressive, i.e., that it allows the driver to report from the full set of possible preferences. Later, in Section 5, we will then describe restricted interfaces. Given the preferences, the agent then participates in the market on its owner's behalf and procures electricity for the vehicle while it is plugged in (e.g., overnight). At the end of the charging period, the driver can use the vehicle to complete journeys and derives value from doing so.

More formally, the problem consists of a sequence of n *days*, $D = \{1, 2, \ldots, n\}$.[3] An EV starts with a given initial *state of charge (SOC)* $s_1 \in [0, s_{\max}]$ (in kWh). Then, at the start of each day d, the driver reports her charging preferences to an autonomous agent, which procures electricity from the market to charge the EV.

3.1 Electricity Market

We assume electricity is sold in discrete, unit-sized quantities (we use 1 kWh, w.l.o.g.). To participate in the market for day d, the EV driver reports her preferences for each quantity of electricity to her charging agent via a suitable market interface. This is done in the form of a *preference report vector* $w_d = [w_{d,1}, w_{d,2}, \ldots, w_{d,s_{\max}-s_d}]$, which indicates the driver's maximum willingness to pay for a given charge on that day. Here we assume that the report vector is bounded by the capacity of the battery, i.e., no valuations are reported for units that would exceed the battery capacity. This structure allows complementarities to be expressed — for example, a preference vector $w_d = [0, 4, 5]$ indicates that the driver is not willing to pay anything for receiving 1 kWh (e.g., because this is too little to complete any journeys), she will pay up to \$4 for 2 kWh and up to \$5 for 3 kWh.

Given this preference vector, the charging agent then participates in the market and obtains $x_d(w_d) \leq |w_d|$ units of electricity at a price of $p_d(w_d)$. These are uncertain, depending on the day's market conditions (e.g., supply of renewables and fluctuations in aggregate demand), but we assume that the agent maximises the driver's utility in expectation, such that truthful reporting is optimal, i.e., $\forall w_d, \hat{w}_d : \mathbb{E}[w_{d,x_d(w_d)} - p_d(w_d)] \geq \mathbb{E}[w_{d,x_d(\hat{w}_d)} - p_d(\hat{w}_d)]$, where w_d is the driver's true willingness to pay. In practice, this can be achieved by participating in an incentive compatible mechanism (Robu et al., 2013; Hayakawa et al., 2015) or by acting strategically on the owner's behalf, e.g., when optimising charging decisions given price predictions for real-time pricing (Ramchurn et al., 2011). Throughout the paper, we will also assume that future market conditions are independent from a driver's previous reports, i.e., x_{d+x} and p_{d+x} are independent from w_d for any $x \geq 1$. This is reasonable in markets with many participants and where supply and demand depend significantly on

---

2. For example market mechanisms, see the work of Ramchurn et al. (2011), Robu et al. (2013) and Hayakawa et al. (2015).

3. We focus on a finite horizon here for practical considerations. First, it enables the efficient calculation of the optimal solution, and, second, our experimental trials are necessarily limited by time. In practice, n can be an arbitrarily large horizon (e.g., over a month or a year) and extensions to infinite settings with discounted future rewards are straight-forward.

external factors, such as the availability of renewables or the need to recharge vehicles for daily journeys.

3.2 EV Utilisation

Given the outcome of the market interaction, the new SOC of the EV is now $s'_d = s_d + x_d(w_d)$. This can then be used by the EV driver to complete journeys. Specifically, every day, there is a set of *potentially available* journeys, $j_d \subseteq J$, where $J = \{1, 2, \ldots\}$ is the set of all journeys. These represent possible journeys the driver may wish to make during the day (e.g., driving to work or to the supermarket). To model realistic scenarios, at the time of reporting preferences, there may be uncertainty about which journeys will be available (e.g., because they depend on the weather or because some journeys will only be necessary in exceptional circumstances). To reflect this uncertainty, a journey is defined by a probability $r_j \in [0, 1]$ (we assume these are independent, although this can be easily relaxed, as discussed briefly in Appendix A) and a value $v_j \in \mathbb{R}$. The value of a journey reflects its importance to the driver and may represent the inconvenience cost incurred when missing the journey, or even the cost of alternative transport (e.g., to travel to work). The sets of potential journeys j_d are known in advance for all days, but the set of *actually available* journeys on day d, denoted by $j'_d \subseteq j_d$, only becomes known after charging is complete on that day.

Given the set of journeys j'_d, the driver now chooses which subset $a_d \subseteq j'_d$ of these to complete. Doing so reduces the EV's SOC, as given by an appropriate cost function $\gamma : 2^J \to \mathbb{R}_{\geq 0}$ (journeys that exceed the SOC cannot be completed). Furthermore, the driver receives the total value of these journeys. Thus, her total utility is the difference between the overall value derived from journeys over time horizon D, and the total costs incurred. In the following subsection, we conclude this section with an illustrative example of the model.

3.3 Illustrative Example

To illustrate the EV charging problem, we here briefly describe a scenario that highlights the key features of our model. Specifically, Figure 1 shows three potential destinations that a driver may wish to travel to on a given day and that are at different distances to her home. The first is a trip to the office, journey 1, which is completely predictable, $r_1 = 1$, and carries a high value, $v_1 = \$20$. The second, journey 2, is a visit to the gym, which is of a medium value, $v_2 = \$5$, and less likely to be available, $r_2 = 0.5$ (this uncertainty may be because the driver is only able to go the gym if they finish work early). Finally, the third, journey 3, is an emergency trip to an important customer. This is unlikely ($r_3 = 0.05$), but carries a high value ($v_3 = \$200$). The costs for individual journeys are $\gamma(\{1\}) = 5$ kWh, $\gamma(\{2\}) = 4$ kWh and $\gamma(\{3\}) = 7$ kWh, which are derived from the distances of the respective destinations to the driver's home location (Figure 1 shows one-way journey costs on the edges between locations). The gym is near the office, so the two journeys can be combined and $\gamma(\{1, 2\}) = 6$ kWh, but the customer is in the opposite direction, so adding journey 3 will always use another 7 kWh (e.g., $\gamma(\{1, 2, 3\}) = 13$ kWh).
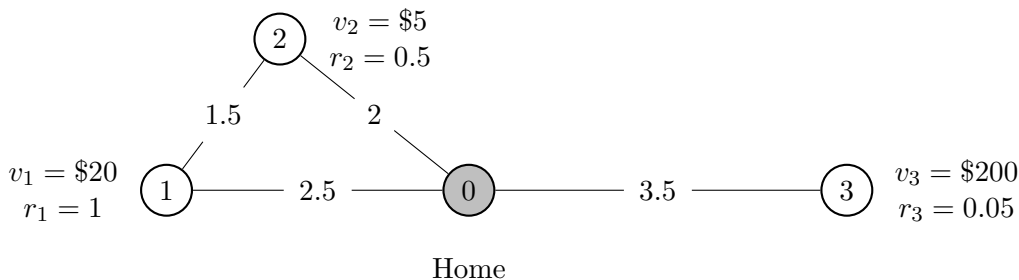
Figure 1: Illustrative example with three potential journeys (1,2,3). Node 0 denotes the user's home location. Edges are labelled with the one-way travel costs (in kWh) and the journey values (v_i) and their probabilities of being available (r_i) are shown.
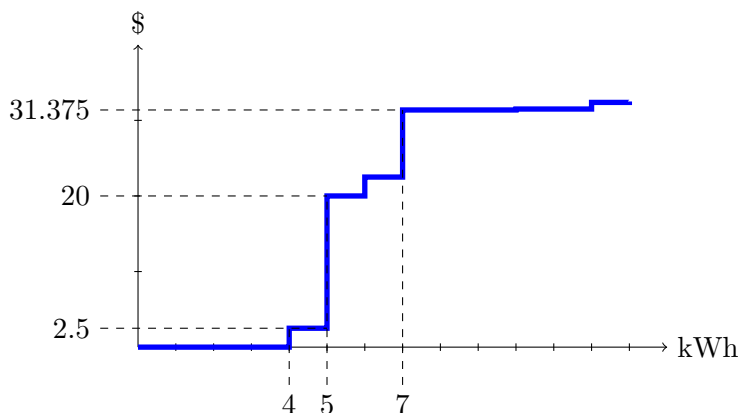

Figure 2: Driver's true willingness to pay, w_d, for the example in Figure 1.

Assuming there is only a single day,[4] and that the battery is initially empty ($s_1 = 0$), it is now possible to calculate the driver's true willingness to pay for each amount of electricity given the costs, probabilities and values of the potentially available journeys. For this particular example, this is $w_1 = [0, 0, 0, 2.5, 20, 22.5, 31.375, 31.375, 31.375, 31.5, 31.5, 32.375, 32.5]$, as shown in Figure 2. Here, $w_{1,x} = 0$ for $x < 4$, because the driver has an insufficient charge to complete any journeys, $w_{1,4} = 2.5$, because she has just sufficient charge to complete journey 2 for a value of $5, but the journey's probability of appearing is only $r_2 = 0.5$. Next, $w_{1,5} = 20$, because the driver now has enough charge to reach the office, which is a better choice than the gym, yielding a value of $20, and it is definitely available. The next significant increase in the driver's willingness to pay happens for $w_{1,7} = 31.375$, when she has sufficient charge to complete either journey 3 (if it is available) or both journeys 1 and 2.

---

4. This is a simplifying assumption to keep the example brief. The optimal solution discussed in the next section takes into account future days.

This example illustrates that even a simple setting can lead to complex and complementary valuations for electricity that the user needs to report via a market interface. In the following section, we will discuss two approaches for computing these valuations computationally. One approach is optimal and one is a learning approach that may be a better model of human behaviour.

4. Modelling EV Driver Strategies

In this section, we consider the potential behaviour of an EV driver. Although we are primarily interested in how real users behave in the EV charging problem, these will constitute important benchmarks to which we can compare the users' performance when interacting with specific market interfaces. The first benchmark is an optimal strategy (Section 4.1), which provides an upper bound for how well a fully rational participant could perform. However, as it assumes a perfect model of all potential market allocations and prices, as well as future journeys, we also present a second benchmark that starts with no prior model and learns a good strategy only through repeated interactions (Section 4.2). To keep the exposition brief, the strategies are outlined here at a high level, but the full details can be found in Appendices A and B.

4.1 Optimal Solution

To optimise her utility, an EV driver needs to report a preference vector w_d on every day d that expresses the respective expected utility for receiving each possible quantity of charge. This is a complex problem, as it needs to take into account the (uncertain) availability of journeys not only on the current day, but also on future days, as well as likely future market conditions, as the battery of the EV allows the driver to store surplus electricity.

To derive the optimal strategy for a perfectly rational driver, we will assume that the driver has knowledge about the expected prices $p_d(w_d)$ and the distribution of allocations $x_d(w_d)$ for a given preference vector. With this, we can model the problem as a Markov Decision Process (MDP) (Howard, 1960), the full details of which are in Appendix A. This MDP is a single-agent MDP, as we assume throughout this paper that the overall market can be modelled probabilistically without dependence on the agent's individual strategy (which in practice could be achieved by using an appropriate incentive compatible market mechanism, as argued earlier). The solution to such an MDP is a *policy* that selects both appropriate preference vectors to report to the charging agent and sets of journeys to complete, depending on the state (the current day d, the current SOC and the journeys available on a given day).

Solving this MDP optimally is NP-hard in general (as the journey selection generalises the Knapsack problem), but problems of realistic sizes can still be solved quickly (e.g., the settings considered in Section 7 are solved within a few seconds). This is because we consider a limited time horizon here, allowing the use of backwards induction and dynamic programming. Furthermore, the possible journeys that a driver will seriously consider on a given day will be small, perhaps in the order of a dozen or fewer. We also discretise the state space of the SOC to include all reachable states, given the cost function γ. Finally, as the charging agent participates in the market optimally, the optimal preference vector is

simply the driver's true valuation for each level of charge, and this is obtained readily from the MDP solution.

4.2 Reinforcement Learning Agent

Clearly, the optimal solution makes some assumptions that are unlikely to hold in practice (such as knowledge of the distributions of $p_d(w_d)$ and $x_d(w_d)$, and all potential future journeys). Hence, a second approach for solving the EV driver's decision problem is to design a reinforcement learning agent (Sutton & Barto, 1998) (note this agent is different from the agent described in Section 3 that interacts with the market). This approach does not require an explicit model of future market interactions or available journeys, but rather learns the optimal policy from repeatedly interacting with the environment and observing realised costs and rewards.

Specifically, we employ the widely-used Q-Learning algorithm (Watkins & Dayan, 1992) with an ϵ-greedy exploration strategy (see Appendix B for details). We hypothesise that this learning approach may be a good approximation of how real users interact with the system, by trying some preference reports, observing how the system responds and then making small adjustments to their strategies (rather than reasoning about the optimal strategy). In fact, reinforcement learning has been used as a computational model to explain learning and decision-making behaviours in animals (Montague, Hyman, & Cohen, 2004).

5. Restricted Market Interfaces

In practice, the drivers solving the above optimisation problem will be humans, and, as we argued, they may not have the time or capacity to act optimally. Neither of the solution approaches discussed in the previous section are feasible for completely automating people's decision processes — the optimal solution requires full information about all possible future journeys, while reinforcement learning requires a long training phase until it performs well (potentially thousands of days, as will be shown in Section 7).

Instead, we here focus on simplifying the market interface for drivers. In particular, we use a range of interfaces that intentionally restrict the reporting space for the user (but without changing the underlying market mechanism). Knowing which interface to present to users can alleviate the cognitive burden, as the user has to consider fewer options.

Given this, we denote the full space of possible reports in a fully expressive interface by $W = \mathbb{R}_{\geq 0}^{s_{\max} - s_d}$. In the following, we will use two approaches for restricting this space by providing the user with an alternative set of possible reports, W', which maps to W through a function $\omega : W' \to W$, to determine the corresponding report that is used by the charging agent. Note that here W, ω and possibly W' depend on s_d to ensure that only valid preference reports are included (which do not exceed the capacity of the battery), but we leave this implicit for clarity. Both approaches rely on significantly reducing the dimensionality of the decision space, while still retaining the ability for users to express a range of valuations.
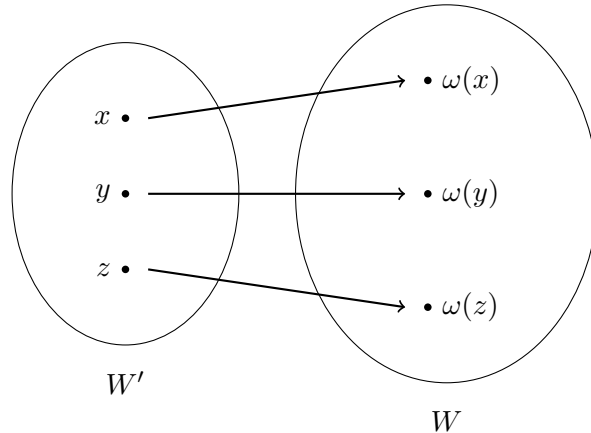
Figure 3: The user in a restricted market interface is given a restricted set of reports (in this example $W' = \{x, y, z\}$), which map to elements of the (usually infinitely large) fully expressive space W via a function ω.

5.1 Single Marginal Value with Quantity (SMV)

The first restriction, SMV, reduces the driver's possible reports to a single marginal value $m_d \in \mathbb{R}_{\geq 0}$ and a maximum quantity $q_d \in \{1, 2, \ldots, s_{\max} - s_d\}$ she wishes to acquire. Here, m_d expresses the value she gains for receiving each *additional* unit of electricity up to a total quantity of q_d. Thus, the space for this restriction is $W^{\text{SMV}} = \mathbb{R}_{\geq 0} \times \mathbb{N}$ and we denote a report by $w_d^{\text{SMV}} = (m_d, q_d)$. The corresponding mapping function is $\omega^{\text{SMV}}(m_d, q_d) = [m_d, 2m_d, 3m_d, \ldots, q_d m_d]$.

The rationale behind providing this restriction is that it reduces the dimensionality of the problem to two dimensions, which may be significantly easier for a user to understand and solve. Furthermore, providing a single value in the auction is closer to the decision space in a single-item auction, which the user may be familiar with. The additional quantity parameter deals with the multi-unit aspect of the auction and allows the user to restrict their risk of receiving too many items at a high price (e.g., if the user urgently requires 2 kWh of charge, is willing to pay a high price for this, but does not want to win more than that).

However, the disadvantage of this approach is that the space of possible reports is still infinitely large. While q_d is discrete and restricted by the battery size, m_d may take on arbitrary values. In practice, the latter could be restricted to discrete increments, e.g., $0.01, and it is bounded by the values of the bidder's tasks, but this still represents a large decision space with at least hundreds of feasible bids (assuming low expected task values in the order of $1). Thus, as bidders may find it easier to select from a smaller set of options, we consider this case in the next mechanism.

5.2 Finite Set of Alternatives

In the second restriction (FINITE), which has been used in related work (Bergemann et al., 2012; Seuken et al., 2012; Blumrosen & Feldman, 2013), we select a finite subset of f

alternatives from the full report space, $\{\alpha_1, \alpha_2, \ldots, \alpha_f\} \subset W$. The corresponding restricted report space is then $W^{\text{FINITE}} = \{1, 2, \ldots, f\}$, such that a report $w_i^{\text{FINITE}} = x$ expresses the user's choice of alternative α_x, i.e., $\omega^{\text{FINITE}}(x) = \alpha_x$.

The advantage of this restriction is that the driver has to consider a very small number of alternatives. In practice, f can be chosen to trade off the cognitive burden on the user with the expressivity of the space (typically, we envisage f to be a handful or less). The alternatives could be chosen to represent a cross section of the full report space, could be manually selected by domain experts or could even be selected by an autonomous agent that adjusts these alternatives to a particular user.

Note that both the optimal solution and the reinforcement learning approach can be adapted for the restricted interfaces through appropriate discretisation. The details are in Appendices A and B.

6. The Bid2Charge Testbed

To test how human participants interact with our interfaces, and to empirically determine which one works best, we designed a web-based game called Bid2Charge.[5] This replicates the EV charging setting, allowing players to repeatedly participate in an auction to charge their simulated EVs and then complete journeys with it. There are several motivations for framing this as a game. First, this is a low-cost way of gathering data from large numbers of users in a short space of time — in contrast, a real field trial would require expensive cars, charging stations and data would have to be collected over long periods of time. Second, the game offers a controlled setting, allowing us to modify the experimental parameters (such as the tasks available, the levels of uncertainty and the interfaces that are used). Third, we can record the users' actions and compare these to the optimal decisions of a rational agent. The last two features in particular are impossible to achieve in a field trial, where many parameters are beyond our control and where a ground truth about the available options to the drivers may not be available.

In the following subsection, we describe the game in more detail.

6.1 Bid2Charge Overview

In Bid2Charge, the player takes the role of an EV delivery van driver. This provides an intuitive explanation to players of what journeys represent (in the game, journeys are referred to as delivery tasks and result in a certain payment), what the objective of the game is (maximise overall profit) and what the uncertainty means (delivery tasks may or may not come up on a given day). We use an incentive compatible auction based on the well-known VCG mechanism (Nisan, Roughgarden, Tardos, & Vazirani, 2007) as the market mechanism in this game, and so the player's reports in each of the market interfaces are framed as bids for this auction.[6] Since strategic behaviour is not the focus of this work, there is no interaction between multiple players of the game. Instead, auction prices are

---

5. The game can be accessed at `http://www.bid2charge.com/jair` and the source code is available at `https://github.com/soton-agents/bid2charge`.
6. VCG fulfils the conditions for optimality in Section 3.1. It is dominant strategy incentive compatible, as long as prices on future days remain independent from the agent's bids (as we assume throughout this paper).
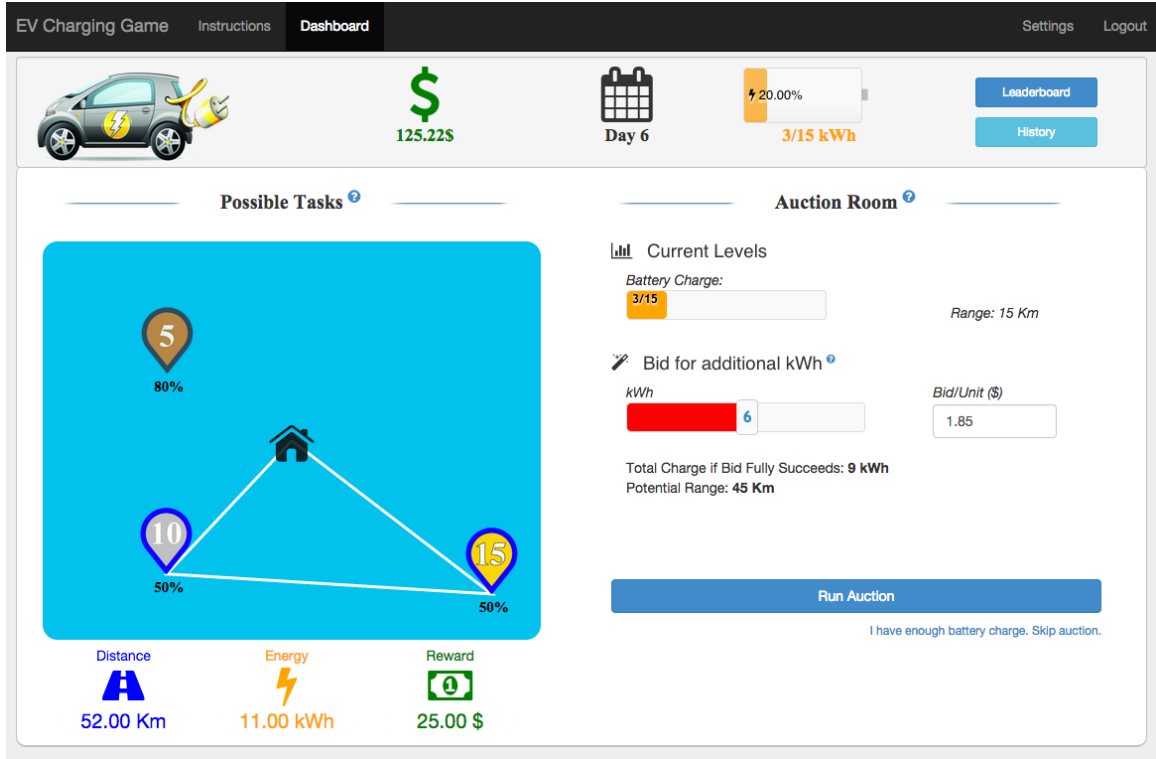
Figure 4: Main game view.

determined according to pre-defined probability distributions (which will be discussed in more detail in Section 7).

When starting the game, the player is first presented with the main interface, a task planning view and an auction view.

6.2 Main Interface and Task Planning View

Figure 4 shows the main game screen. At the top, there are some general statistics, showing the player's accumulated profit, the current day and current SOC. Below that, on the left, there is a task planning view. This provides the user with information about j_d, i.e., the tasks that are potentially available on the current day. Both the value, v_j, and the realisation probability, r_j, are shown for each task j. Furthermore, the user can select subsets of tasks, $j'_d \subseteq j_d$ to inspect both the total value $(\sum_{j \in j'_d} v_j)$ and the total cost $(\gamma(j'_d))$ if those tasks are completed $(\gamma(j'_d)$ is calculated based on the Euclidean distance of the shortest route past all tasks in j'_d, as described in more detail in Section 7.2). Note that interacting with this view does not affect the game — it simply provides the player with information about the available tasks. In Figure 4 the user has selected the \$10 and \$15 tasks and is informed that this will require 11 kWh in total for a reward of \$25.

After interacting with the task planning view, the player can submit their bid in the auction view.

189

Figure 5: The FINITE interface (showing three options).
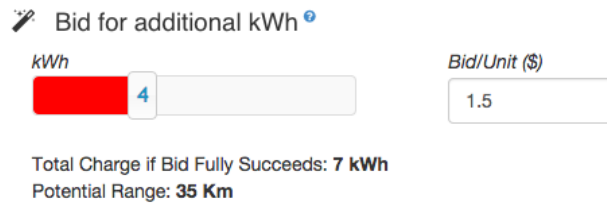
Figure 6: The SMV interface (showing a bid with $m_d = 1.5, q_d = 4$). The EV in this example already has a charge of 3 kWh.
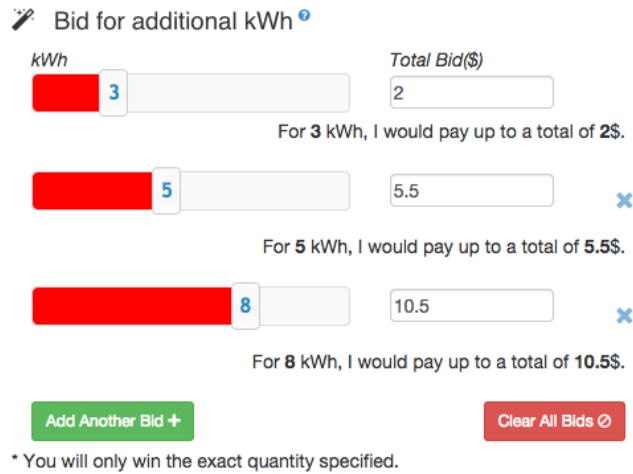
Figure 7: The Fully-Expressive interface (showing bid $w_d = [0, 0, 2, 2, 5.5, 5.5, 5.5, 10.5]$).
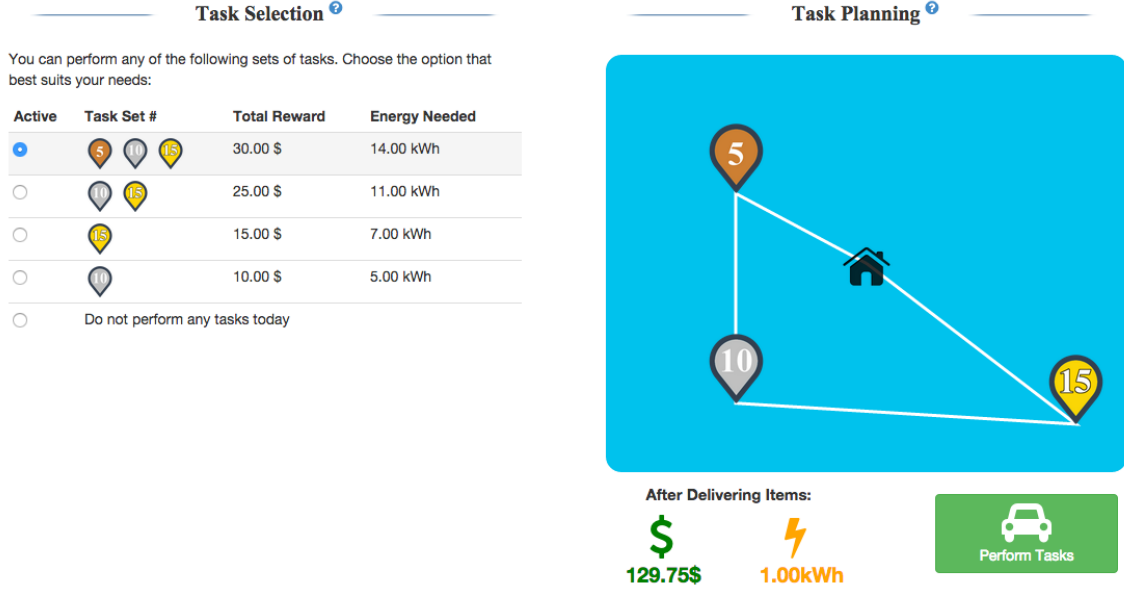
Figure 8: Task view with actually available tasks.

6.3 Auction View

The auction view, which is to the right of the task planning view, allows the player to submit their bid w_d for the current day. The view supports all three interfaces discussed in this paper, as shown in Figures 5 – 7: FINITE, SMV and a fully expressive interface. Only one of these is shown to each user, depending on which experimental treatment they have been assigned to. Here, the candidates for FINITE in Figure 5 can be customised by the game administrator, but this particular view shows $[1, 2, 3, \ldots, s_{\max} - s_d]$, $[2, 4, 6, \ldots, s_{\max} - s_d]$ and $[4, 8, 12, \ldots, s_{\max} - s_d]$, which we also use for our experiments. The view for SMV in Figure 6 shows a user entering the bid $m_d = 1.5, q_d = 4$. Finally, the view for the fully expressive interface in Figure 7 shows a user entering $w_d = [0, 0, 2, 2, 5.5, 5.5, 5.5, 10.5]$. The user always has the option to skip the auction. On pressing the "Run Auction" button, they are informed of the outcome and then taken to a task view.

6.4 Task View

In the task view, representing the EV utilisation phase and shown in Figure 8, the user is presented with the realisation of tasks for the current day and can select their desired set of tasks to complete. The options for this are given as a table with corresponding rewards and costs (ordered by decreasing rewards). Any dominated or infeasible solutions are automatically removed, i.e., where another set offers a higher total reward for the same or lower cost, or where the total cost exceeds the current SOC.

In the following, we will discuss how we used the Bid2Charge testbed to conduct a large experimental evaluation of the market interfaces discussed in Section 5.

7. Experimental Evaluation

The purpose of our experiments is to investigate how real human users interact with the two restricted interfaces proposed in this paper, as well as the fully expressive interface, and whether the choice of a particular interface influences the performance of users. Our approach is to evaluate this through a randomised, controlled experiment, where we allocate market interfaces randomly to users, in order to exclude self-selection bias. Participants were given instructions only for the particular interface they were assigned to and were not told that others existed. In carrying out our experiments, we were guided by two main hypotheses:

Hypothesis 1 *Players using more expressive interfaces achieve the same or a higher profit than those using restricted interfaces.*

Hypothesis 2 *Players using more expressive interfaces spend more time on the game than those using more restricted interfaces.*

The first hypothesis is based on the fact that the more expressive interfaces offer more possible reports to the players. Experimental evidence also suggests that market user interfaces with more options lead to the same or better performance than those with fewer options (Seuken et al., 2012). However, more expressivity may also incur a higher cognitive burden, which is expressed by Hypothesis 2.

We tested these hypotheses through one initial experiment, where we asked participants to play through a single game consisting of 30 days. To further explore how players learn and improve over time, we carried out a second experiment, where we asked another set of participants to play through three identical 10-day games in sequence (although journeys and their probabilities were identical, their realisations were not necessarily the same). In addition to validating the results from the first experiment, the main hypothesis we tested in the second setting was:

Hypothesis 3 *Players using more complex interfaces improve their profit more significantly over repeated plays of the game than those using more restricted interfaces.*

This hypothesis reflects the fact that the more complex interfaces have a larger decision space, which may be initially overwhelming for users and may require some learning and experimentation to achieve their full performance. Conversely, the simpler interfaces offer fewer options, and thus we envisage that users will already perform well on the first play, with less improvement on subsequent iterations.

Furthermore, this more sequential setting allowed us to compare the human learning trends to those of the reinforcement learning agent we proposed in Section 4.2.

7.1 Player Recruitment

We recruited players from Amazon Mechanical Turk, a platform that allows requesters to advertise tasks to a large audience of online workers.[7] Mechanical Turk is increasingly being

---

7. http://www.mturk.com/

used to recruit participants for scientific studies (Mason & Suri, 2012), and these have been shown to yield valid results that are comparable to more traditional methods of subject recruitment (Paolacci, Chandler, & Ipeirotis, 2010).

For our experiment, we asked workers to first read a set of instructions explaining the rules and objectives of the game, then they had to give consent to participating in a study. To ensure players understood the game, they also had to correctly answer a short multiple-choice questionnaire. The instructions and questionnaire can be seen in full in Appendix C. This was followed by the game and then a brief final survey to gain feedback about their game experience as well as some demographic information. This final survey can be found in Appendix D.

We told workers the task would take about 25 minutes, they would receive a base payment of $2.50 and then a bonus based on the profit they made in the game. Specifically, in the first experiment, we offered $0.02 for each $1.00 earned in the game, while (due to budget constraints) in the second experiment we offered $0.01 for each $1.00 earned in each of the three games. In both experiments, this was capped at $3.00.[8] This bonus was chosen to ensure that the incentives of the player were aligned with the objective of the game. We assigned participants to interfaces using a block randomisation scheme, to achieve a balanced spread of participants across the interfaces. In the first experiment, 130 workers played the game; while in the second experiment 189 played three games each. Table 1 shows the full demographic data we collected from the players. This indicates that most players were young, well-educated, owned a car, but had not driven an electric vehicle (including hybrid) before. Due to restrictions imposed by Amazon, the players were predominantly US residents.

7.2 Experimental Parameters

To simulate the auction, we determine the marginal prices of units,[9] denoted by $p_{d,x}$ for the xth unit, using random distributions. These are identical for all interfaces, on all days and for all participants (but the prices were sampled independently each time). Specifically, each price was determined by first setting $p_{d,0} = 0$, and then iteratively determining each $p_{d,x}$ as $p_{d,x} = p_{d,x-1} + \epsilon_x$, where ϵ_x was drawn from a uniform distribution $\mathcal{U}(0.2x - 0.2, 0.4x + 0.6)$. The rising mean of this distribution ensures that marginal prices for each unit generally increase. Given these prices, we then set the allocation to $x_d(w_d) = \text{argmax}_x w_{d,x} - \sum_i^x p_{d,i}$. This is consistent with a VCG mechanism, where the prices represent the externality imposed on other agents and the allocation maximises both the agent's utility and the social welfare.[10] Furthermore, the distribution we chose here lead to a realistic and interesting setting, as increasingly higher bids are needed to secure a larger number of units, as would be expected in an efficient VCG mechanism that allocates to the highest bidders first. However, as in VCG, marginal prices can occasionally decrease due to the complementarities in valuations (e.g., once a high-value agent that requires many units simultaneously is outbid, marginal prices may drop).

---

8. This maximum was set to limit our potential spend. Only three players managed to reach this.
9. Each unit here corresponds to 1 kWh. The maximum capacity of a player's EV is 15 kWh.
10. However, we do not simulate the full VCG mechanism, as this is computationally challenging in general.

		Experiment 1	**Experiment 2**
Dates:		14–15th November 2014	11–14th September 2015
Total Participants:		130	189
Treatment:	FINITE (1,2,4)	42 (32.3%)	52 (27.5%)
	FINITE (1,2,3)	–	49 (25.9%)
	SMV	45 (34.6%)	48 (25.4%)
	Fully-Expressive	43 (33.1%)	40 (21.2%)
Gender:	Male	62 (47.7%)	104 (55%)
	Female	68 (52.3%)	85 (45%)
Age:	18–24	20 (15.4%)	26 (13.8%)
	25–34	57 (43.8%)	77 (40.7%)
	35–44	22 (16.9%)	52 (27.5%)
	45–54	17 (13.1%)	23 (12.2%)
	55–64	11 (8.5%)	10 (5.3%)
	65+	3 (2.3%)	1 (0.5%)
Education:	None	–	1 (0.5%)
	Primary school	1 (0.8%)	–
	Middle school	2 (1.5%)	1 (0.5%)
	High school	35 (26.9%)	73 (38.6%)
	University	92 (70.8%)	114 (60.3%)
Car Owner:	Yes	105 (80.8%)	154 (81.5%)
	No	25 (19.2%)	35 (18.5%)
Used Hybrid/EV:	Yes	14 (10.8%)	25 (13.2%)
	No	116 (89.2%)	164 (86.8%)
Country:	United States	106 (81.5%)	172 (91%)
	India	18 (13.8%)	16 (8.5%)
	United Kingdom	2 (1.5%)	–
	Brazil	1 (0.8%)	–
	Canada	1 (0.8%)	–
	Mexico	1 (0.8%)	–
	Sri Lanka	1 (0.8%)	1 (0.5%)

Table 1: Demographics of participants

The game in the first experiment was played for 30 simulated days, and we varied the number of tasks every 1–4 days (with between 1–6 tasks available every day). We did not give players information about tasks on future days (only the total number of days to play), and there was no prior information about the distribution of auction prices and allocation probabilities. This is reasonable because in real-world settings these would also be highly uncertain, and because we did not want to overwhelm players with a large amount of information. In the second experiment, a 10-day game was played three times by each player, enabling some learning.

A map of all tasks used throughout our experiments and their (x, y) locations on a 35km $\times$ 35km map are shown in Figure 9 and Table 2. This abstract representation corresponds to the types of maps that are shown to a user each day (see, for example, Figure 4, where tasks B, E and I are shown). Location A represents the player's starting location, which is depicted by a house in Figure 4 and to which players need to return by the end of the day. Each unit on the map corresponds to 1km and the distance between two locations corresponds to the Euclidean distance between them, i.e., $\delta((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ km. To compute the total battery cost of completing a given set of tasks a_d (i.e., to compute the cost function $\gamma(a_d)$), we select the path with the shortest total distance that originates from location A, visits all task locations in a_d and then returns to location A. Denoting the length of that shortest path by $\Delta(a_d)$, we then compute the cost as $\gamma(a_d) = \lceil \lfloor \Delta(a_d) \rfloor \cdot 0.2 \frac{\text{kWh}}{\text{km}} \rceil$.[11] Here, the average consumption of 0.2 kWh per km travelled is a reasonable figure representing current EVs.[12] As an example, applying this to the path shown in Figure 9, the distance from A to E is about 12.374 km, from E to I it is 22.817 km and from I to A it is 17.5 km. Thus, the total distance is 52.691 km, which is truncated to 52 km. The total required consumption is $\lceil 52 \text{ km} \cdot 0.2 \frac{\text{kWh}}{\text{km}} \rceil = \lceil 10.4 \text{ kWh} \rceil = 11$ kWh, as shown in the figure.

Finally, Tables 3 and 4 show the set of potentially available journeys for each day in the first and second experiment, respectively. Here, each journey is first specified by the corresponding location in Figure 9 and a tuple (v_j, r_j) denoting the journey's value und probability of being available. For example, day 6 in the first experiment is associated with three journeys E, I and B, with respective values 10, 15 and 5, as well as probabilities 0.5, 0.5 and 0.8. These correspond to the tasks shown to the user in Figure 4.

7.3 Benchmarks

To establish upper and lower bounds for the possible performance of players, we compare them to a number of benchmarks:

- **Optimal**: This is the optimal strategy assuming a fully expressive interface. We also show two variants, **Optimal (SMV)** and **Optimal (FINITE)**, for the restricted interfaces. All of these are obtained by solving the MDP, as described in Appendix A.

- **QL(λ)**: This is a reinforcement learning agent that has played the game for a number (λ) of episodes. Here, each episode is a full run of the game (30 days in the first experiment and 10 days in the second experiment) and, in order to measure the

---

11. We round the values here to ensure that only integers rather than fractions are shown to the player.
12. Such as the Nissan Leaf (Davis, Alexander, & Duvall, 2013).
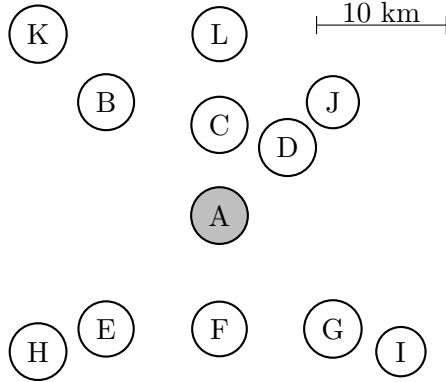
Figure 9: Map of task locations.

	x	y
A	17.50	17.50
B	8.75	26.25
C	17.50	24.50
D	22.75	22.75
E	8.75	8.75
F	17.50	8.75
G	26.25	8.75
H	3.50	7.00
I	31.50	7.00
J	26.25	26.25
K	3.50	31.50
L	17.50	31.50

Table 2: (x, y) locations.

Days	Potentially available journeys					
1, 2	C(5,1.00)					
3	C(5,1.00)	D(5,0.80)				
4, 5	C(5,1.00)	D(5,0.80)	G(10,0.75)			
6, 7, 8	E(10,0.50)	I(15,0.50)	B(5,0.80)			
9, 10, 11	B(5,0.50)	E(10,1.00)				
12, 13	B(5,0.50)	I(15,0.50)				
14, 15	B(5,0.50)	I(15,0.50)	H(15,0.50)			
16, 17	H(15,1.00)	B(5,0.50)				
18, 19, 20	D(5,0.40)	B(10,0.25)	F(5,0.25)	E(10,0.25)	I(15,0.20)	
21, 22	B(5,0.50)	F(5,0.50)				
23, 24, 25, 26	I(15,0.40)	H(15,0.40)	B(10,0.60)	F(5,0.80)	G(10,0.50)	C(5,0.50)
27, 28	B(5,0.50)	F(5,0.50)				
29, 30	B(5,0.50)	E(10,1.00)				

Table 3: Potentially available journeys (i.e., tasks) in the first experiment. Each journey is represented by the letter of its corresponding location and a tuple (v_j, r_j), representing the journey's value and probability, respectively.

performance, we temporarily set the ϵ parameter to 0 (i.e., the agent switches to pure exploitation). We also test the performance of the reinforcement learning agent on the restricted interfaces, denoted by **QL(SMV,λ)** and **QL(FINITE,λ)**.

- **RandomGreedy**: This is a benchmark that places a random bid (chosen from the FINITE options) and then greedily chooses the highest-value tasks until it runs out of battery.

- **MaxGreedy**: This places a bid that is high enough to fully charge the EV each day and then greedily chooses the highest-value tasks until it runs out of battery.

Days	Potentially available journeys			
1	B(5,1.00)			
2	B(5,1.00)	B(5,0.75)		
3	G(5,1.00)	C(10,0.25)		
4, 5	E(15,0.75)	I(15,0.50)		
6	C(10,0.50)	B(5,0.50)	E(15,0.20)	
7	G(5,1.00)	C(10,0.25)		
8, 9	C(10,0.50)	D(5,1.00)	E(15,0.50)	I(15,0.25)
10	C(10,0.30)	E(15,0.50)		

Table 4: Potentially available journeys (i.e., tasks) in the second experiment.

The last two strategies represent simple baseline approaches that a worker could employ to complete the task with as little effort as possible, which is a possibility on Mechanical Turk (Mason & Suri, 2012) and which we here use as a lower bound on performance. We present the average performance of all benchmarks over 1000 trials, except for the reinforcement learning agent, for which we average over 250 trials, due to the significantly longer runtime required for training.

7.4 Results of the First Experiment

We first consider the overall performance in terms of the overall profit achieved, as this is the main objective of the game and to verify Hypothesis 1. Figure 10 shows this for the three interfaces with human participants (in red, plain), for the three optimal policies (in green, hatched), the two baseline benchmarks (in blue, finely hatched) and the reinforcement learning agent after 100 and 25000 iterations of the game, which are representative of early and late stages of learning (in orange, dotted). All results are shown with 95% confidence intervals. Focusing first on the performance of the optimal (in green), it is interesting to note here that there is little difference between the optimal performance in the restricted setting (in particular for SMV) and the fully expressive setting. This is encouraging, showing that, despite the severe restrictions, there is little loss in the utility players could, in theory, achieve. Note that this is a general trend we observed across a wide range of possible settings, not just the particular setting chosen for the experiment here. This is likely because units of electricity can be retained for future days and so winning each one generates some utility for the driver. In turn, this means that the SMV and FINITE interfaces (both of which attach uniform marginal values to each unit of electricity) are reasonable here.

Next, considering the performance of human players, these are generally situated between the optimal and the baseline benchmarks. Unsurprisingly, the humans perform significantly worse than a rational agent, given that the problem is highly complex due to its inherent stochasticity and combinatorial nature. However, when comparing the human players to the baseline benchmarks, there is a marked improvement. This provides some evidence that the participants are putting effort into the game rather than selecting strategies that require the least effort.

When comparing the performance of the human players using different market interfaces, several interesting trends emerge. First, we note that the choice of mechanism seems to
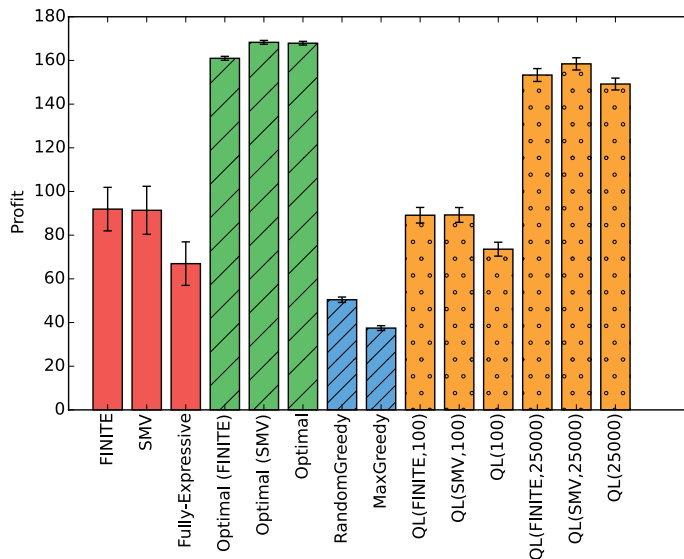
Figure 10: Average profit in the first experiment.

have a significant influence on performance.[13] Counter-intuitively, the players using the fully expressive interface achieve the overall lowest performance with an average profit of only $66.96, while players with SMV achieve an average profit of $91.38 and players with FINITE achieve an average $91.93.[14] This constitutes an improvement of over 35% compared to the fully expressive interface. This means our Hypothesis 1 needs to be rejected. This poor performance of the fully expressive interface is likely due to its substantial cognitive burden, as users are faced with a complex decision problem (we will return to the issue of cognitive burden in Section 7.6, where we discuss the post-experiment feedback we collected).

Considering the performance of the reinforcement learning agent, several interesting trends emerge. First, the fully-expressive interface consistently performs worse than the other two interfaces (with the same amount of learning). This is because the extremely large reporting space takes longer to explore than in the other two interfaces. Furthermore, after 100 learning episodes, the performance of the reinforcement learning agent resembles that of human players, indicating that it may be useful for predicting human behaviour.

Although there is no significant difference in profit between FINITE and SMV for human players, interesting trends emerge when considering the actual money that was spent on acquiring electricity and how much was gained from completing tasks. This is shown in Figure 11a. Here, players using FINITE spent an average $208.44 acquiring electricity and achieved an average reward of $300.37. In contrast, players using SMV spent only

---

13. This is confirmed by ANOVA with $p = 0.001$.

14. A post-hoc Bonferroni test confirms that there is a significant difference between the fully expressive mechanism and each of the other two (with $p = 0.004$). There is no significant difference between the performance of users with FINITE and SMV.
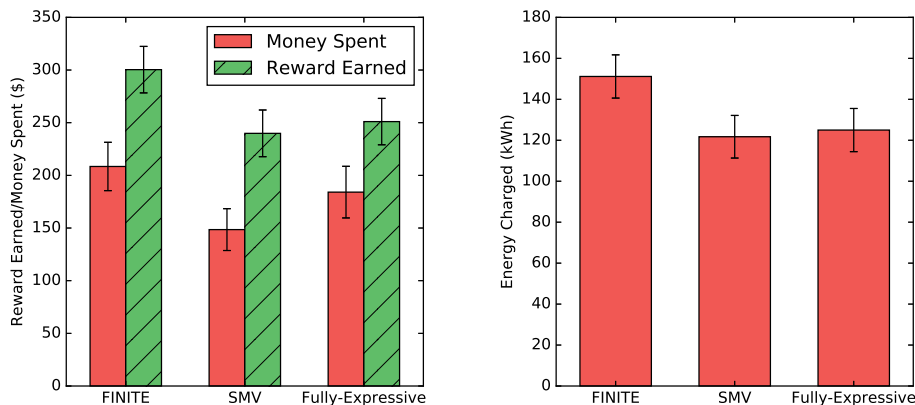
(a) Average spend and reward (in $). (b) Average energy charged (in kWh).

Figure 11: Differences in player performance in the first experiment.

an average of \$148.50 and earned \$239.89. The differences here are highly significant.[15] Thus, while the profit for both is similar, FINITE induces a very different behaviour in the players — they spend significantly more on acquiring electricity and are then able to use this to complete a larger number of tasks. This trend is confirmed when considering the average amount of electricity that is won by the participants throughout the game, shown in Figure 11b — players using the FINITE interface charged an average of 151.15 kWh, while players using SMV charged only 121.71 kWh on average, or about 20% less.[16]

This difference in observed behaviours could be due to several factors. One possible explanation is that SMV requires users to explicitly set a maximum number of units, thereby focusing them on this parameter and implicitly suggesting they restrict this demand. Another possibility is that the good strategies in both mechanisms are just fundamentally different. In both cases, this is an interesting result, showing that significant changes in behaviour can be caused by the right choice of mechanism. Especially in the energy domain, low overall consumption may be a particularly desirable goal and SMV encourages this.

Next, to investigate Hypothesis 2, Figure 12 shows the time that players spent on the auction and task screens. Participants using the fully expressive interface spent the longest time on the auction (689 seconds on average), while participants using FINITE spent the least amount of time on the auction (502 seconds on average). This supports the hypothesis.[17] Note that the difference in the time spent on journey selection is not significant here.

---

15. ANOVA confirms this for both metrics with $p \leq 0.002$. Post-hoc Bonferroni tests confirm differences between FINITE and SMV with $p = 0.001$.

16. ANOVA and post-hoc Bonferroni tests confirm significance of this difference with $p \leq 0.001$.

17. ANOVA ($p = 0.026$) and a Bonferroni test confirm a difference between the expressive and FINITE ($p = 0.021$.)
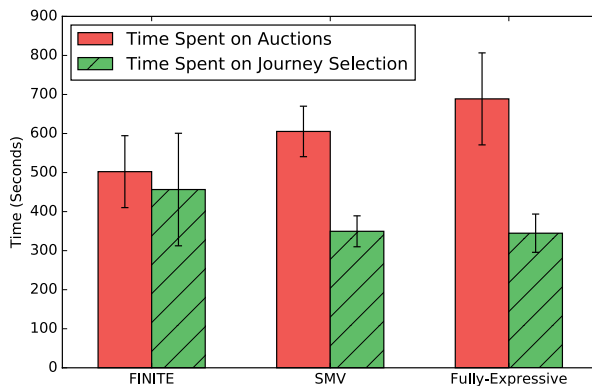
Figure 12: Average time spent by players on auctions and journeys.
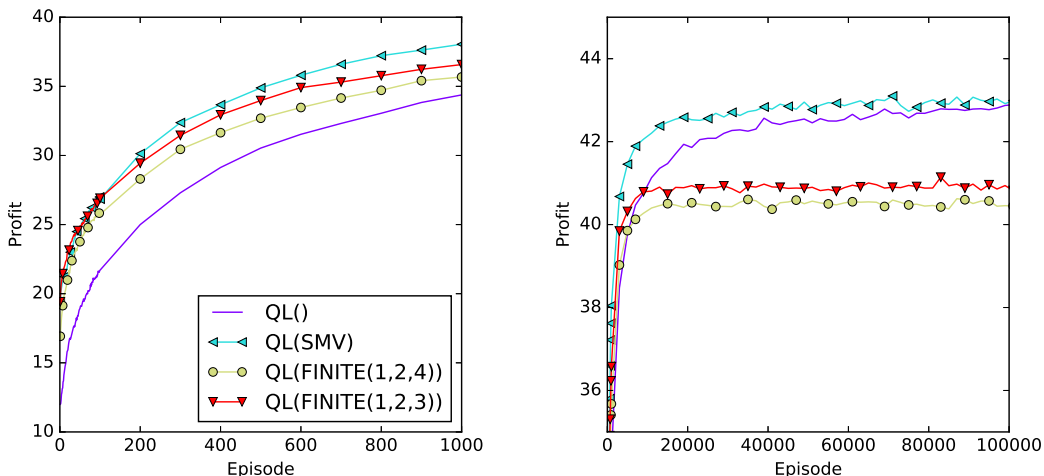
(a) Profit over first 1000 learning episodes.

(b) Profit over first 100000 learning episodes.

Figure 13: Average profit over number of learning episodes obtained by reinforcement learning agent.

7.5 Results of the Second Experiment

To further examine whether our proposed reinforcement learning agent can be applied to optimising market user interfaces, we ran the agent in the setting of the second experiment and identified a small change to the FINITE strategy that led to an improvement in the agent's performance. Specifically, we decreased the marginal valuation reported in the third alternative from \$4/kWh to \$3/kWh. We used both alternatives in the second experiment, and in the following, we will refer to interfaces FINITE(1,2,3) and FINITE(1,2,4) to distinguish between these two. Figure 13 shows the performance of the reinforcement learning agent over time using the various interfaces.
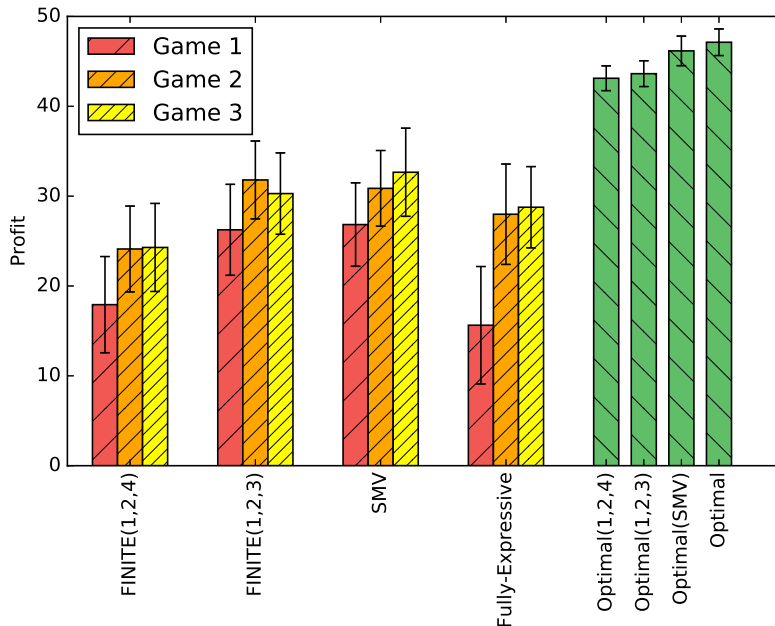
Figure 14: Average profit during each of the three games in the second experiment, along with benchmarks.

Figure 14 shows the overall performance results from the second experiment, grouped by the three games each player completed. First, it is clear that there is a strong learning effect — on all treatments, the players perform better over time. On the first game, there is also a very clear performance difference between the interfaces. Fully expressive performs worse (average profit \$15.6) while SMV performs best (average profit \$26.8).[18] However, participants on the fully expressive interface then start to perform significantly better on subsequent games, supporting Hypothesis 3. This is likely because they begin to exploit the higher expressivity of the interface. At the same time, they are still taking significantly longer (approximately twice as long) to deliberate during the auction phase (taking 373, 266 and 227 seconds for the three games, compared to 178, 123 and 116 of FINITE(1,2,3), as shown in Figure 15), and they still do not outperform the restricted interfaces.

Finally, comparing the results with the predictions of the reinforcement learning agent, several broad trends are confirmed. First, although the specific profits and the absolute differences between the interfaces vary between the human players and the reinforcement learning agent, the relative performance trends between the interfaces are similar. Specifically, FINITE(1,2,3) indeed outperforms FINITE(1,2,4), as predicted by the agent, indicating that our approach of optimising the interface based on the agent's response is sensible. This is interesting, because there is no discernible difference in the *optimal* performance of the two interfaces (see Figure 14). Other trends indicated by the agent are also confirmed: SMV consistently performs well throughout, while the fully expressive interface initially

---

18. ANOVA ($p = 0.006$) and a Bonferroni test find a significant difference ($p = 0.036$) between these two.
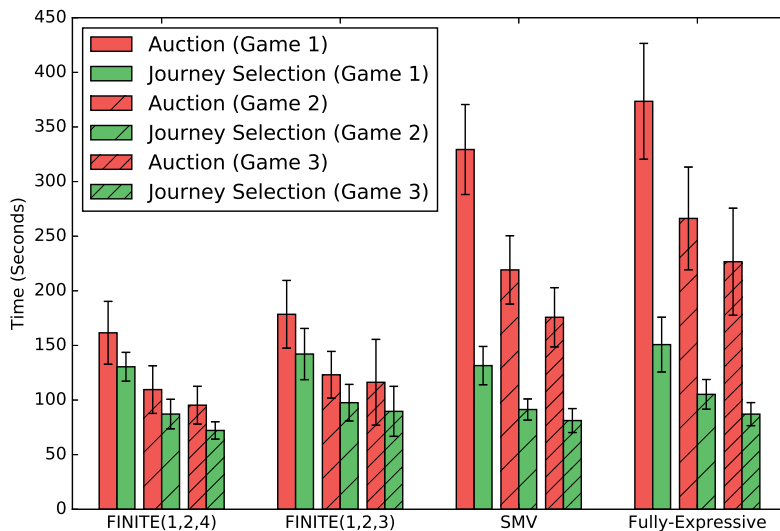
Figure 15: Average time taken for auctions and journey selection during each of the three games in the second experiment.

performs poorly, but then catches up with the others as more learning and exploration take place.

In order to compare the behaviour of human players and the reinforcement learning agent in more detail, Figure 16 shows the proportion of reports that are chosen by both when using the FINITE(1,2,4) and FINITE(1,2,3) interfaces, and after various stages of learning. Specifically, we recorded the proportions after 100, 1000 and 20,000 learning episodes for the agent, and in games 1, 2 and 3 for the human players. For comparison, the optimal agent is also shown. Again, although the absolute values differ, the reinforcement learning agent is arguably closer in its behaviour to the human players than to the optimal strategy. Specifically, the optimal strategy rarely chooses the highest report (in only 2-3% of cases), but both the reinforcement learning agent and the human players frequently pick this option (in between 11–26%, of cases, depending on the stage of learning). As such a high report is typically avoided by the optimal agent (indicating that the marginal value of electricity is usually below $3 per unit), this offers a plausible explanation for why both approaches achieve a lower performance when using FINITE(1,2,4) than when using FINITE(1,2,3): occasionally reporting a value of $4 is significantly further from the optimal than $3. Interestingly, both the agent and the human players reduce the frequency of choosing the high option with more learning, as might be expected.

To conclude this section, we briefly note that there is no significant difference between the energy used by SMV and the two FINITE variants, as was observed in the first experiment. This may be due to the particular setting used in the second experiment, which included fewer tasks (2.4 tasks per day on average rather than 3.0) and so presented fewer opportunities to spend large amounts of energy for a correspondingly larger reward.

(a) FINITE(1,2,4). (b) FINITE(1,2,3).
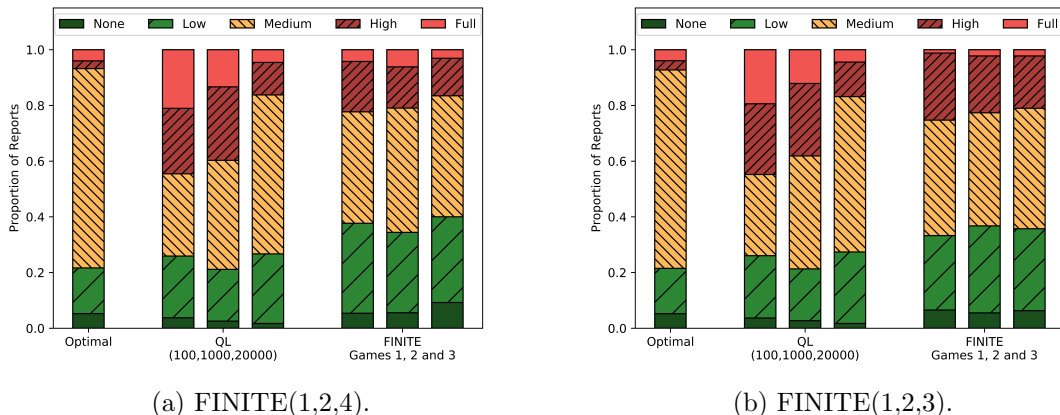
Figure 16: Average proportion of reports chosen by the optimal, the reinforcement learning agent and human players with the FINITE interface. Here, (from bottom to top for each bar) *None* means the auction was skipped, *Low*, *Medium*, and *High* correspond to the three possible bids in each interface, and *Full* means the battery was full and therefore no report was necessary.

7.6 User Feedback

In addition to measuring the users' performance during the game, we conducted a survey at the end of the game during the second experiment. Related to Hypothesis 2, the purpose of this survey was to explore in more depth how users of the restricted and fully-expressive interfaces perceived the respective cognitive burden associated with their interfaces. Specifically, the overarching hypothesis guiding this survey was that users of the more complex interfaces perceive a significantly higher cognitive burden than those on the more restricted interfaces.

In more detail, the survey was based on the mental demand, performance, effort, and frustration categories from the NASA-TLX scale (Hart & Stavenland, 1988). Specifically, we asked each participant to what extent they agreed with the following statements, corresponding to the four NASA-TLX categories mentioned above:

S1 : I found the auction system unnecessarily complex.

S2 : I was successful in accomplishing what I was asked to do.

S3 : I had to work hard to achieve my level of performance.

S4 : I felt irritated, stressed or annoyed during the game.

As one of the motivations behind this study was to determine the impact of market user interface design within potential applications, we also included three questions from the System Usability Scale (SUS) (Brooke, 1996), as well as a general question about using auctions for EV charging (S5):

S5 : I would be happy to use this type of auction system for charging a real EV.

S6 : I thought the auction system was easy to use.

S7 : I needed to learn a lot of things before I could get going with this game.

S8 : I felt very confident in playing the game.

We asked participants to rate their agreement with these statements on a five-point Likert scale with responses corresponding to "Strongly Agree", "Slightly Agree", "Neutral", "Slightly Disagree" and "Strongly Disagree". We grouped the FINITE variants into a single category, as the interfaces are similar and deliberation times (shown in Figure 15) were comparable.
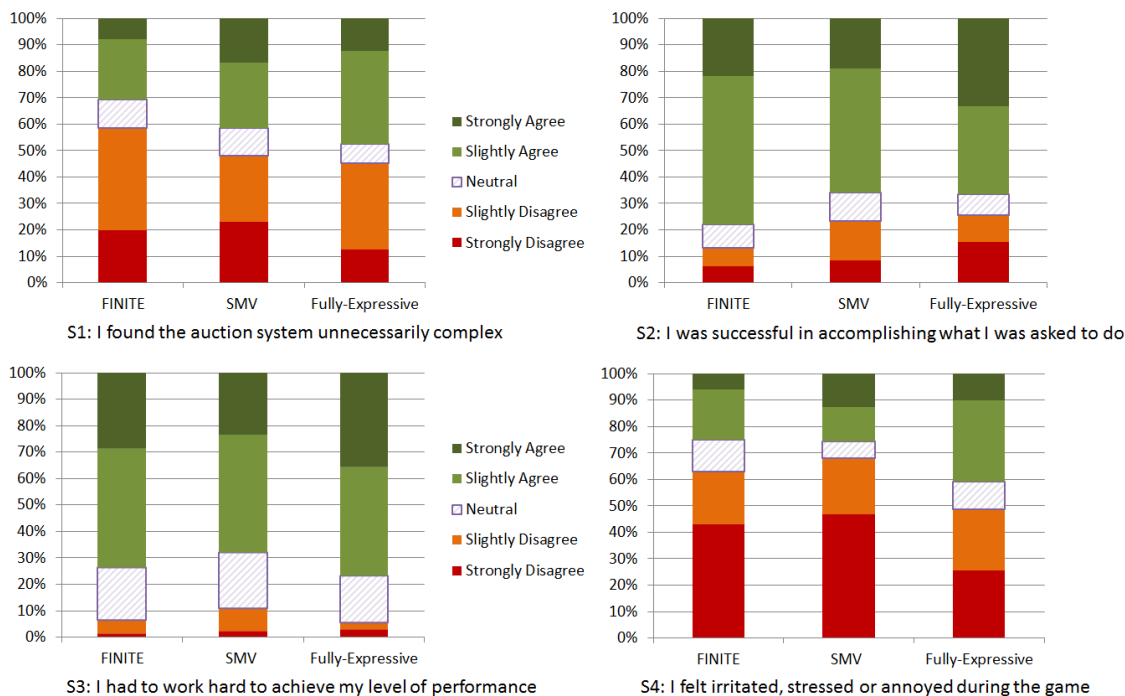
Figure 17: Distributions of responses to statements from the NASA-TLX scale.

Figure 17 visually presents the distributions for the mental demand, performance, effort, and frustration categories from the NASA-TLX scale. As is evident in the results, users found the Fully-Expressive interface was significantly more complex (S1) than the more limited FINITE interface[19] and that users felt that they were more successful in accomplishing their task (S2) when using the FINITE interface compared to the more complex SMV interface.[20] Similarly, the Fully-Expressive interfaces was more likely to produce feedback that agreed with the statement that users felt irritated, stressed or annoyed (S4) during the game.[21] There were no significant differences in how hard users felt they needed to work to

---

19. This is confirmed by pairwise t-tests with $p = 0.04$. Note that, unless indicated otherwise, all tests in this section are pairwise t-tests on the numerical values of the Likert scale (where 5 is "Strongly Agree" and 1 is "Strongly Disagree").

20. Borderline significance with $p = 0.09$.

21. Significant with $p = 0.03$ between FINITE and Fully-Expressive and $p = 0.05$ between FINITE and SMV.
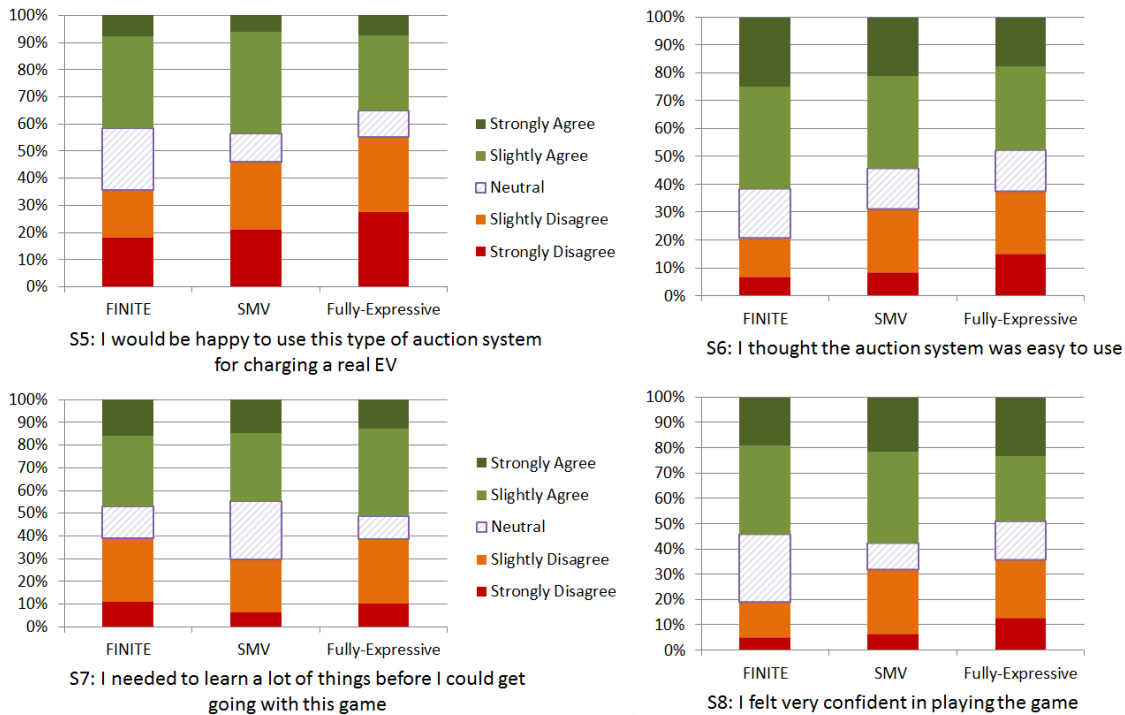
Figure 18: Distributions of responses to statements from the System Usability Scale scale.

achieve their performance (S3). Nonetheless, these results overall support our hypothesis that higher interface complexity leads to a higher perceived cognitive burden. This holds particularly when contrasting FINITE and Fully-Expressive, where differences are typically the most pronounced.

Next, Figure 18 visually presents the distributions for the four remaining statements with the trend again being similar. The Fully-Expressive interface yields more negative feedback than the FINITE interface for the statement S5, "I would be happy to use this type of auction system for charging a real EV".[22] Similarly, for statement S6 ("I thought the auction system was easy to use"), agreement was significantly higher overall for FINITE compared to that of Fully-Expressive.[23] FINITE was also less likely to receive negative responses to statement S8 ("I felt very confident in playing the game").[24]

In conclusion, the restricted interfaces, and FINITE in particular, are generally more usable and lead to a lower perceived cognitive burden than the Fully-Expressive interface. This is aligned with our earlier finding that more complex interfaces lead to longer deliberation times, which we used as an objective indicator for cognitive burden.

---

22. Although a t-test does not find a significant overall difference, a chi-square test on the proportion of negative feedback (score 1 or 2) is borderline significant with $p = 0.07$.
23. The difference between FINITE and Fully-Expressive is significant at $p = 0.02$, but the difference between FINITE and SMV is not statistically significant with $p = 0.11$.
24. Again a chi-square test on the proportion of negative feedback is borderline significant with $p = 0.06$.

8. Conclusions and Future Work

As appropriate market mechanisms can efficiently allocate scarce resources to competing consumers, they hold considerable promise in addressing emerging challenges in the energy domain. However, in many realistic settings, including the EV charging problem, there are non-expert participants. In this paper, we studied how the cognitive burden on such participants can be alleviated by using restricted market interfaces.

This paper's extensive evaluation used a framework called Bid2Charge that allowed us to compare people's behaviour in a simulated EV charging setting while using a variety of interfaces. As such, our work is the first comprehensive study of a wide spectrum of market interfaces, ranging from fully expressive to parameterised (SMV) to very limited (FINITE). In carrying out this study, we found that using restricted market interfaces has several key advantages over standard fully expressive approaches.

First, participants spend significantly less time deliberating (half the time in some settings), which indicates a lower cognitive burden. This observation was confirmed by a detailed survey, where participants using more restricted interfaces responded more positively to questions about the system's usability.

Second, despite taking less time, these participants tend to perform better than those using a fully expressive, theoretically optimal interface. This performance gap is particularly pronounced during the first interactions, and then gradually closes with experience.

Third, particular types of interfaces induce different behaviours in participants, while achieving the same utility. This could be a promising tool for nudging people towards particularly desirable behaviours, such as energy conservation, and we show that the parameterised SMV interface can reduce energy consumption by up to 20%.

Last, we found that a reinforcement learning agent was able to predict broad trends in the relative performance of alternative interfaces. As such, we believe it constitutes a valuable tool for the evaluation and optimisation of market user interfaces.

There are some limitations to the empirical study conducted in this paper. Our subjects were recruited on Mechanical Turk and interacted with the various market interfaces in one sustained interaction (where days were simulated in a matter of minutes). In reality, users would interact with the interfaces only once or twice a day, but over much longer periods of time. This will affect how they engage with the interface, both in the short and in the long term, and this will likely raise further interesting issues, e.g., how to retain engagement and how to assist users in learning the interface, especially when feedback is significantly delayed. Moreover, the stakes in a real setting will be higher, as missed journeys will result in significant inconvenience to the user, and so real users may behave in a more risk-averse manner than in our game. Hence, in future work, we plan to evaluate these potential differences in behaviour through more realistic trials with real vehicles.

Here, one of the key findings that needs to be validated in practice is the difference in energy consumption between FINITE and SMV. Due to the higher stakes in missing journeys, this result may not be as pronounced in a real setting. However, it constitutes an interesting point of departure for investigating how users may be nudged to alter their behaviour slightly through an appropriate interface. Given this, it may be promising to apply multi-objective optimisation methods to trade off various goals, such as the driver's utility, energy usage and deliberation time.

Moreover, we plan to build on the framework discussed in this paper and develop new optimised interfaces that adapt to users. This adaptation may include setting appropriate parameters for the interfaces, for example to optimise the options available in the FINITE interface based on a particular user's driving patterns or to switch from more restricted to more complex interfaces as users become familiar with the market mechanics.

Building on the reinforcement learning agent used in this paper, we will also explore how an intelligent agent can support a human market participant. This could take place on a flexible spectrum of autonomy, ranging from simple advice on what bids to place (based on historical data) to bidding completely automatically on the human's behalf, and switching between these modes depending on the agent's confidence and the user's private information about their requirements.

Finally, it is important to note that while Bid2Charge was evaluated using EV market mechanisms, it is a general framework that is easily adapted to other markets. Along these lines, we plan to study restricted interfaces and learning agents in other settings where market mechanisms hold promise, including on-demand mobility systems, cloud computing or demand response in the Smart Grid.

Acknowledgments

Appendix A. Optimal Solution

Here, we present the MDP formulation of the EV charging problem, describe an agent's policies and present the optimal solution.

A.1 MDP Formulation

We formalise the EV charging problem from Section 3 as a tuple $(\Sigma, \Sigma', J, j_{d \in D}, r_{j \in J}, v_{j \in J}, \gamma, X_{d \in D}, P_{d \in D})$, where:

- $\Sigma = D \times [0, s_{\max}]$ is the set of states *before* the market interaction, and $\Sigma' = D \times [0, s_{\max}] \times 2^J$ is the set of states *after* the market interaction (but before selecting from the available journeys). Both include the current day, d, and state of charge, s_d or s_d', while a state $\sigma' \in \Sigma'$ also includes the set of available journeys j_d'.

- J, j_d, r_j, v_j and γ describe the journeys, their potential availability per day, realisation probabilities, values and journey costs, as defined in Section 3.

- $X_d : (W \times \{0, 1, \ldots, s_{\max}\}) \to [0, 1]$ is the probability distribution of x_d, i.e., $X_d(w_d, x)$ is the probability of obtaining x units when reporting w_d during a market interaction.

- $P_d : W \to \mathbb{R}$ is the expected price of a market interaction given a report, i.e., $P_d(w_d) = \mathbb{E}\left[p_d(w_d)\right]$.

The actions, transition probabilities and rewards of this MDP are fully described by the tuple above. Specifically:

- The available **actions** depend on the current state as follows. For $\sigma \in \Sigma$, the actions are the reports in W (restricted by the battery capacity as outlined in Section 3.1). For $\sigma' = (d, s_d', j_d') \in \Sigma'$, the set of actions is $\{a \in 2^{j_d'} \mid \gamma(a) \le s_d'\}$, i.e., all subsets of the available journeys that can be completed with the current state of charge.

- **Transition probabilities** from a state σ to a state σ' are determined by the chosen report w_d and X_d (to determine s_d') and by j_d and r_j (to determine j_d'). Transitions from a state σ' to σ are deterministically given by the chosen action a_d and cost function γ.

- **Rewards** are incurred when transitioning from one state to the next, and they correspond to the prices paid for electricity (negative) and the values derived from completing journeys (positive). Specifically, the former is given by $-P_d(w_d)$, while the latter is $\sum_{j \in a_d} v_j$.

A.2 Agent Policy

An agent's policy determines which preference vector to report for the market interaction, and which journeys to complete, given the current state. Thus, it is described by a tuple (π, π'), where $\pi : \Sigma \to W$ determines the reports and $\pi' : \Sigma' \to 2^J$ the journeys.

Given this, we can now define the expected utility of a policy (π, π') using a value function for each type of state. For state $\sigma = (d, s_d) \in \Sigma$:

$$V(d, s_d, \pi, \pi') = -P_d(\pi(\sigma)) + \sum_{i=0}^{s_{\max} - s_d} \left[X_d(\pi(\sigma), i) \sum_{j_d' \subseteq j_d} R(j_d') \; V'(d, s_d + i, j_d', \pi, \pi') \right], \quad (1)$$

where $R(j_d') = \prod_{j \in j_d'} r_j \prod_{j \in j_d \setminus j_d} (1 - r_j)$ is the probability that j_d' is the set of journeys available.[25]

For state $\sigma' = (d, s_d', j_d') \in \Sigma'$, with $d < n$:

$$V'(d, s_d', j_d', \pi, \pi') = \left[\sum_{j \in \pi'(\sigma')} v_j \right] + V(d + 1, s_d' - \gamma(\pi'(\sigma')), \pi, \pi'), \quad (2)$$

while for the last day, $V'(n, s_n', j_n', \pi, \pi') = \sum_{j \in \pi'(\sigma)} v_j$.

---

25. Note that while this definition of $R(j_d')$ assumes independence between journeys, it could be redefined to account for correlated journey probabilities without changing the solution approach or its computational complexity.

A.3 Optimal Solution

Given the above definition, the optimal policy (π^*, π'^*) is simply:

$$(\pi^*, \pi'^*) = \underset{(\pi, \pi')}{\operatorname{argmax}} V(1, s_{\text{initial}}, \pi, \pi'), \qquad (3)$$

where s_{initial} is the initial state of charge. As discussed in Section 4.1, this can be found using dynamic programming with backwards induction and by recognising that the agent's best strategy is to bid its true valuation for each level of electricity. This valuation is given directly by V', such that:

$$\pi^*(d, s_d) = \left[V'(d, s_d + 1, \pi^*, \pi'^*), V'(d, s_d + 2, \pi^*, \pi'^*), \ldots, V'(d, s_{\text{max}}, \pi^*, \pi'^*) \right] \qquad (4)$$

The full algorithm for computing the optimal policy is shown in Algorithm 1. In order to obtain optimal policies for the restricted interfaces, we can replace line 12 with the following:

13: $\qquad w^* = \operatorname{argmax}_{w \in W'} -P_d(\omega(w)) + \sum_{i=0}^{s_{\text{max}}-s} \Big[X_d(\omega(w), i) \cdot$

$\qquad \qquad \sum_{j'_d \subseteq j_d} R(j'_d) \, V'(d, s + i, j'_d, \pi, \pi') \Big]$

14: $\qquad \pi^*(d, s) = \omega(w^*)$

For SMV, this requires a discretisation of the marginal value. In our experiments, we use $m_d \in \{\$0, \$0.01, \$0.02, \ldots, \$5\}$, which is sufficiently fine-grained to lead to a near-optimal performance. Furthermore, in our experiments we approximate X_d and P_d using Monte Carlo simulations with 1000 trials for each possible action.

The run-time of the optimal solution is reasonable for problems of realistic sizes. For example, computing the optimal policy for the first experiment in Section 7 using a Python implementation of the above algorithm took 4.75 seconds on an Apple MacBook Pro with a 3.3 GHz Intel Core i7 CPU and 16GB RAM (including the Monte Carlo simulation of X_d and P_d). For the SMV interface, we conducted the Monte Carlo simulation of the discretised action space first, which took 77.39 seconds. Solving the corresponding MDP was then completed in 8.59 seconds. For the FINITE interfaces (with three options), the simulation took 0.45 seconds and solving the MDPs took 0.03 seconds. In the second experiment, solving the MDPs took about a third of the time (reflecting the shorter time horizon).

Appendix B. Reinforcement Learning Agent

Algorithm 2 shows the reinforcement learning algorithm we use as a more realistic benchmark than the optimal. This algorithm does not need knowledge of the underlying MDP and instead uses a Q-Learning approach to gradually learn the value of making particular reports in a given state (expressed using a $Q(\sigma, w)$ function). It takes three parameters: an exploration probability ϵ, a learning rate parameter α, and the set of possible reports W. In the experiments, we set $\epsilon = 0.2$ and $\alpha = 0.1$ (our results are not particularly sensitive to this choice). W is determined by the relevant interface.

In more detail, for each day d of each episode e, the agent first senses the current state σ (line 6). It then selects a report to submit using the PICKREPORT function. This

---

Algorithm 1 Optimal Solution

1: **procedure** FINDOPTIMAL
2:     Initialise π^*, π'^*, V and V' to be empty
3:     Initialise Γ to hold feasible charge levels given γ
4:     **for** $d \in \{n, n-1, n-2, \ldots, 1\}$ **do**
5:         **for** $s \in \Gamma$ **do**
6:             **for** $j_d' \subseteq j_d$ **do**
7:                 **if** $d = n$ **then**
8:                     $\pi'^*(d, s, j_d') = \text{argmax}_{a \subseteq j_d' \mid \gamma(a) \leq s} \sum_{j \in a} v_j$
9:                 **else**
10:                     $\pi'^*(d, s, j_d') = \text{argmax}_{a \subseteq j_d' \mid \gamma(a) \leq s} \sum_{j \in a} v_j + V(d+1, s - \gamma(a), \pi^*, \pi'^*)$
11:             Calculate $V'(d, s, j_d', \pi^*, \pi'^*)$ using Equation 2
12:             Calculate $\pi^*(d, s, \pi^*, \pi'^*)$ using Equation 4
13:             Calculate $V(d, s, \pi^*, \pi'^*)$ using Equation 1
14:     **return** (π^*, π'^*)

---

function depends on the interface the agent is using. Algorithm 3 shows this for the Fully-Expressive interface, while Algorithm 4 is used for FINITE and SMV. Both approaches use an ϵ parameter that balances exploration (choosing a random, potentially untested report) with exploitation (choosing the best-performing report). As discussed in Section 4.2, the Fully-Expressive uses a local search technique to generate new reports during exploration. Here, the MODIFYRANDOMELEMENT(w) function takes a report w, sets one randomly chosen element w_i to a random number (in the experiments, this is on the interval $[0, 6i]$), and then adjusts the other elements to ensure the vector is non-decreasing, by decreasing elements before i and increasing elements after i as necessary.

The agent then submits its chosen report, observes the incurred cost in the market and the new state (lines 8–10). It then picks the best set of journeys to complete, given the Q-values of states on the following day (lines 11–14). These state transitions are deterministic, given the costs of journeys, so we do not learn separate Q-values for journey choices. However, note that most Q-values for the next states will be zero initially, so the agent will start by greedily completing journeys to maximise its immediate reward.

Next, the agent updates the Q-function for its chosen report based on the overall profit achieved during the day (lines 15–20). Finally, with a small probability ϵ, the agent picks a new random set of journeys, which is again used for exploration; otherwise, it executes its chosen set a_d (lines 21–23).

Algorithm 2 Reinforcement Learning Agent

1: **procedure** QLearningAgent(ϵ, α, W)
2: $\quad \hat{Q} \leftarrow \{\}$ $\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ State/action pairs that have been tried
3: $\quad \forall \sigma \in \Sigma, w \in W : Q(\sigma, w) \leftarrow 0$ $\qquad\qquad\qquad\qquad\qquad$ ▷ Initialise Q-function
4: $\quad$ **for** $e \in \{1, 2, \ldots\}$ **do** $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ Episodes
5: $\qquad$ **for** $d \in \{1, 2, \ldots, n\}$ **do**
6: $\qquad\quad \sigma \leftarrow$ SenseAuctionState()
7: $\qquad\quad w \leftarrow$ PickReport($\epsilon, \sigma, Q, \hat{Q}, W$)
8: $\qquad\quad$ Report(w)
9: $\qquad\quad p_d \leftarrow$ ObserveCost()
10: $\qquad\quad (d, s, j'_d) \leftarrow$ SenseJourneyState()
11: $\qquad\quad$ **if** $d = n$ **then**
12: $\qquad\qquad a_d \leftarrow \text{argmax}_{a \subseteq j'_d \mid \gamma(a) \leq s} \sum_{j \in a} v_j$
13: $\qquad\quad$ **else**
14: $\qquad\qquad a_d \leftarrow \text{argmax}_{a \subseteq j'_d \mid \gamma(a) \leq s} \left[\left(\sum_{j \in a} v_j \right) + \text{argmax}_{w \in W} Q((d+1, s-\gamma(a)), w) \right]$

15: $\qquad\quad v \leftarrow \sum_{j \in a_d} v_j$
16: $\qquad\quad$ **if** $(\sigma, w) \notin \hat{Q}$ **then**
17: $\qquad\qquad Q(\sigma, w) \leftarrow v - p_d$
18: $\qquad\qquad \hat{Q} \leftarrow \hat{Q} \cup \{(\sigma, w)\}$
19: $\qquad\quad$ **else**
20: $\qquad\qquad Q(\sigma, w) \leftarrow Q(\sigma, w) + \alpha \cdot (v - p_d - Q(\sigma, w))$
21: $\qquad\quad$ **if** Random(0,1) $< \epsilon$ **then**
22: $\qquad\qquad a_d \in \{a \subseteq j'_d \mid \gamma(a) \leq s\}$ $\qquad\qquad\qquad$ ▷ Pick random journey set
23: $\qquad\quad$ CompleteJourneys(a_d)

Algorithm 3 Report Selection Function for Fully-Expressive

1: **procedure** PickReport($\epsilon, \sigma, Q, \hat{Q}, W$)
2: $\quad q \leftarrow \{w \mid (\sigma, w) \in \hat{Q}\}$
3: $\quad$ **if** Random(0,1) $< \epsilon \vee |q| = 0$ **then**
4: $\qquad$ **if** $|q| = 0$ **then**
5: $\qquad\quad w \leftarrow [0, 0, \ldots, 0]$ $\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ $s_{\max}$ elements
6: $\qquad$ **else**
7: $\qquad\quad w \in q$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ Random report
8: $\qquad$ **return** ModifyRandomElement(w)
9: $\quad$ **else**
10: $\qquad$ **return** $\text{argmax}_w Q(\sigma, w)$

Algorithm 4 Report Selection Function for FINITE and SMV

1: **procedure** PICKREPORT($\epsilon, \sigma, Q, \hat{Q}, W$)
2:    **if** RANDOM(0,1) $< \epsilon$ **then**
3:       **return** $w \in W$ ▷ Random report
4:    **else**
5:       **return** $\text{argmax}_w Q(\sigma, w)$

Appendix C. Mechanical Turk Instructions

This appendix shows the instructions given to all participants at the beginning of the second experiment. The instructions shown in the first experiment are almost identical. The game is available at http://www.bid2charge.com/jair, where both sets of instructions can be viewed and where it is possible to play the game as experienced by the participants.

Specifically, all figures here are shown as a single scrollable webpage. Figures 19—21 are shown to all participants, with one of Figures 22, 23 or 24 shown next, depending on the participant's treatment (for FINITE, SMV or Fully-Expressive, respectively). Figures 25 and 26 are then shown to all participants again.

Figure 19: Instructions shown to participants (part 1) with initial overview of the game.
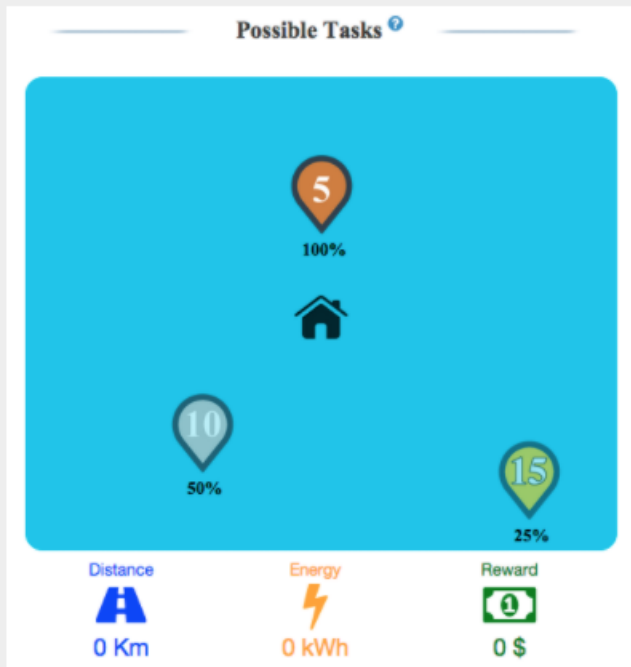
Figure 20: Instructions (part 2) with overview of planning stage.
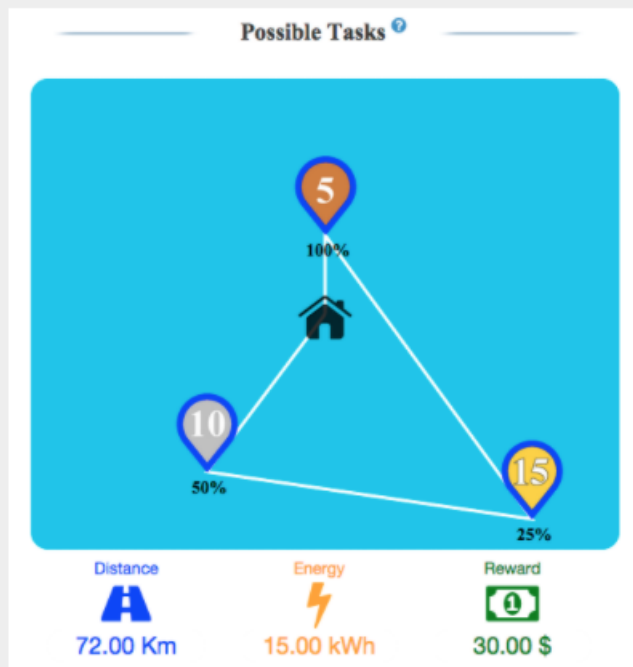
Figure 21: Instructions (part 3) continuing the planning stage.

Figure 22: Instructions (part 4) shown to participants with the FINITE interface.

Figure 23: Instructions (part 4) shown to participants with the SMV interface.

Figure 24: Instructions (part 4) shown to participants with the Fully-Expressive interface.

Figure 25: Instructions (part 5) covering task selection and participant consent.

Quiz

So you can check your understanding of the rules, please answer the following questions:

What is the goal of the game?

○ Complete as many tasks as possible.

○ Achieve as high a profit as possible.

○ Preserve your battery as much as possible.

○ Use your battery as much as possible.

○ Win as much electricity in the auctions as possible.

Given the image on the right, what is the maximum reward you can achieve with a battery charge of 12kWh?

○ $5

○ $10

○ $15

○ $30

○ I would need to select a different combination of tasks to find out.

The bids you place in the auction represent:

○ The exact amount I will pay if I win.

○ The minimum amount I will pay if I win.

○ The maximum amount I will pay if I win.

[Submit Answers]

Figure 26: Instructions (part 6) showing final quiz.

Appendix D. Mechanical Turk Questionnaire

This appendix shows the questionnaire shown to all participants at the end of the second experiment. As for the instructions shown initially, they are represented as a single scrollable page. The first set of questions (Figure 27) collects demographic information, while the second set collects data about the user experience (Figure 28).

Figure 27: Final questionnaire (part 1) to collect demographic information.

To what extent do you agree with the following statements?

I would be happy to use this type of auction system for charging a real EV.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I found the auction system unnecessarily complex.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I thought the auction system was easy to use.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I was successful in accomplishing what I was asked to do.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I had to work hard to achieve my level of performance.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I felt irritated, stressed or annoyed during the game.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I needed to learn a lot of things before I could get going with this game.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

I felt very confident in playing the game.

○ Strongly Disagree ○ Slightly Disagree ○ Neutral ○ Slightly Agree ○ Strongly Agree

Figure 28: Final questionnaire (part 2) to collect data about user experience.

References

Ahn, C., Li, C.-T., & Peng, H. (2011). Optimal decentralized charging control algorithm for electrified vehicles connected to smart grid. *Journal of Power Sources, 196*(23), 10369 – 10379.

Ash, T. (2011). *Landing Page Optimization: The Definitive Guide to Testing and Tuning for Conversions.* John Wiley & Sons.

Ausubel, L. M. (2004). An efficient ascending-bid auction for multiple objects. *The American Economic Review, 94*(5), 1452–1475.

Baarslag, T., & Gerding, E. H. (2015). Optimal incremental preference elicitation during negotiation..

Bartram, L. (2015). Design challenges and opportunities for eco-feedback in the home. *Computer Graphics and Applications, 35*(4), 5262.

Benisch, M., Sadeh, N., & Sandholm, T. (2008a). A Theory of Expressiveness in Mechanisms. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence.* AAAI.

Benisch, M., Sadeh, N., & Sandholm, T. (2008b). The cost of inexpressiveness in advertisement auctions. In *Proceedings of ACM EC Workshop on Advertisement Auctions.*

Bergemann, D., Shen, J., Xu, Y., & Yeh, E. (2012). Multi-dimensional mechanism design with limited information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pp. 162–178, New York, NY, USA. ACM.

Blumrosen, L., & Feldman, M. (2006). Implementation with a bounded action space. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, pp. 62–71, New York, NY, USA. ACM.

Blumrosen, L., & Feldman, M. (2013). Mechanism design with a restricted action space. *Games and Economic Behavior, 82*, 424–443.

Brooke, J. (1996). Sus-a quick and dirty usability scale.. *Usability evaluation in industry, 189*(194), 4–7.

Clement-Nyns, K., Haesen, E., & Driesen, J. (2010). The impact of charging plug-in hybrid electric vehicles on a residential distribution grid. *IEEE Transactions on Power Systems, 25*(1), 371–380.

Coskun, A., Zimmerman, J., & Erbug, C. (2015). Promoting sustainability through behavior change: A review. *Design Studies, 41*(B), 183–204.

Cramton, P. (2010). Simultaneous ascending auctions. In Cochran, J. J., Cox, L. A., Keskinocak, P., Kharoufeh, J. P., & Smith, J. C. (Eds.), *Wiley Encyclopedia of Operations Research and Management Science.* John Wiley and Sons, Inc.

Davis, M., Alexander, M., & Duvall, M. (2013). Total cost of ownership model for current plug-in electric vehicles. Tech. rep., Electric Power Research Institute.

Dütting, P., Fischer, F., & Parkes, D. C. (2011). Simplicity-expressiveness tradeoffs in mechanism design. In *Proceedings of the 12th ACM Conference on Electronic Commerce*, EC '11, pp. 341–350, New York, NY, USA. ACM.

Dütting, P., Henzinger, M., & Starnberger, M. (2013). Valuation compressions in vcg-based combinatorial auctions. In Chen, Y., & Immorlica, N. (Eds.), *Web and Internet Economics: 9th International Conference, WINE 2013, Cambridge, MA, USA, December 11-14, 2013, Proceedings*, pp. 146–159. Springer Berlin Heidelberg, Berlin, Heidelberg.

Elmalech, A., Sarne, D., & Grosz, B. J. (2015a). Problem restructuring for better decision making in recurring decision situations. *Autonomous Agents and Multi-Agent Systems*, *29*(1), 1–39.

Elmalech, A., Sarne, D., Rosenfeld, A., & Erez, E. S. (2015b). When suboptimal rules. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pp. 1313–1319.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*, 32–64.

Flach, J. (1995). Situation awareness: Proceed with caution. *Human Factors*, *37*, 149–157.

Flath, C. M., Ilg, J. P., Gottwalt, S., Schmeck, H., & Weinhardt, C. (2014). Improving electric vehicle charging coordination through area pricing. *Transportation Science*, *48*(4), 619–634.

Gajos, K. Z., Weld, D. S., & Wobbrock, J. O. (2010). Automatically generating personalized user interfaces with supple. *Artificial Intelligence*, *174*(12), 910 – 950.

Hart, S. G., & Stavenland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, chap. 7, pp. 139–183.

Hartline, J. D., & Roughgarden, T. (2009). Simple versus optimal mechanisms. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC'09)*, pp. 225–234. ACM.

Hayakawa, K., Gerding, E., Stein, S., & Shiga, T. (2015). Online mechanisms for charging electric vehicles in settings with varying marginal electricity costs. In *24th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2610–2616.

Hodes, T. D., Katz, R. H., Servan-Schreiber, E., & Rowe, L. (1997). Composable ad-hoc mobile services for universal interaction. In *Proceedings of the 3rd Annual ACM/IEEE International Conference on Mobile Computing and Networking*, MobiCom '97, pp. 1–12, New York, NY, USA. ACM.

Howard, R. A. (1960). *Dynamic Programming and Markov Processes*. MIT Press.

Ipakchi, A., & Albuyeh, F. (2009). Grid of the future. *Power and Energy Magazine, IEEE*, *7*(2), 52–62.

Jung, M. F., Sirkin, D., Gür, T. M., & Steinert, M. (2015). Displayed uncertainty improves driving experience and behavior: The case of range anxiety in an electric car. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 2201–2210, New York, NY, USA. ACM.

Kahneman, D. (2000). A psychological point of view: Violations of rational rules as a diagnostic of mental processes (commentary on stanovich and west). *Behavioral and Brain Sciences*, *23*, 681–683.

Kluckner, P. M., Weiss, A., Schrammel, J., & Tscheligi, M. (2013). *Exploring Persuasion in the Home: Results of a Long-Term Study on Energy Consumption Behavior*, pp. 150–165. Springer International Publishing.

Larson, K., & Sandholm, T. (2005). Mechanism design and deliberative agents. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and MultiAgent Systems*, pp. 650–656.

Lund, H. (2007). Renewable energy strategies for sustainable development. *Energy, 32*(6), 912 – 919. Third Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems.

Mason, W., & Suri, S. (2012). Conducting behavioral research on amazon's mechanical turk. *Behavior Research Methods, 44*(1), 1–23.

Mets, K., D'hulst, R., & Develder, C. (2012). Comparison of intelligent charging algorithms for electric vehicles to reduce peak load and demand variability in a distribution grid. *Journal of Communications and Networks, 14*(6), 672–681.

Milgrom, P. (2010). Simplified mechanisms with an application to sponsored-search auctions. *Games and Economic Behavior, 70*(1), 62 – 70. Special Issue In Honor of Ehud Kalai.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature, 431*(7010), 760–767.

Nichols, J., Myers, B. A., Higgins, M., Hughes, J., Harris, T. K., Rosenfeld, R., & Pignol, M. (2002). Generating remote control interfaces for complex appliances. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology*, UIST '02, pp. 161–170, New York, NY, USA. ACM.

Nisan, N., Roughgarden, T., Tardos, E., & Vazirani, V. V. (2007). *Algorithmic game theory*, Vol. 1. Cambridge University Press Cambridge.

Onal, E., Schaffer, J., O'Donovan, J., Marusich, L., Yu, M. S., Gonzalez, C., & Höllerer, T. (2014). Decision-making in abstract trust games: A user interface perspective. In *2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pp. 21–27.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making, 5*(5), 411–419.

Petrou, K., Quiros-Tortos, J., & Ochoa, L. F. (2015). Controlling electric vehicle charging points for congestion management of uk lv networks. In *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power Energy Society*, pp. 1–5.

Pierce, J., & Paulos, E. (2012). Beyond energy monitors: Interaction, energy, and emerging energy systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 665–674.

Ramchurn, S. D., Vytelingum, P., Rogers, A., & Jennings, N. R. (2011). Agent-based homeostatic control for green energy in the smart grid. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(4), 35.

Ramchurn, S. D., Vytelingum, P., Rogers, A., & Jennings, N. R. (2012). Putting the 'smarts' into the smart grid: A grand challenge for artificial intelligence. *Commun. ACM*, *55*(4), 86–97.

Robu, V., Gerding, E. H., Stein, S., Parkes, D. C., Rogers, A., & Jennings, N. R. (2013). An online mechanism for multi-unit demand and its application to plug-in hybrid electric vehicle charging. *Journal of Artificial Intelligence Research*, *48*, 175–230.

Rosenfeld, A., & Kraus, S. (2012). Modeling agents based on aspiration adaptation theory. *Autonomous Agents and Multi-Agent Systems*, *24*(2), 221–254.

Royal Academy of Engineering (2010). *Electric Vehicles: Charged with potential*. Royal Academy of Engineering.

Samadi, P., Mohsenian-Rad, H., Schober, R., & Wong, V. W. S. (2012). Advanced demand side management for the future smart grid using mechanism design. *IEEE Transactions on Smart Grid*, *3*(3), 1170–1180.

Sandholm, T., & Boutilier, C. P. (2006). Preference elicitation in combinatorial auctions. In Cramton, P., Shoham, Y., & Steinberg, R. (Eds.), *Combinatorial Auctions*, pp. 233–263. MIT Press.

Seuken, S., Charles, D., Chickering, M., & Puri, S. (2010a). Market design & analysis for a p2p backup systems. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pp. 97–108, New York, NY, USA. ACM.

Seuken, S., Jain, K., Tan, D. S., & Czerwinski, M. (2010b). Hidden markets: UI design for a P2P backup application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 315–324, New York, NY, USA. ACM.

Seuken, S., Parkes, D. C., Horvitz, E., Jain, K., Czerwinski, M., & Tan, D. (2012). Market user interface design. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pp. 898–915, New York, NY, USA. ACM.

Simon, H. A. (1972). Theories of bounded rationality. In McGuire, C. B., & Radner, R. (Eds.), *Decision and organization: A volume in honor of Jacob Marschak*, pp. 161–176. North-Holland.

Stein, S., Gerding, E. H., Nedea, A., Rosenfeld, A., & Jennings, N. R. (2016). Bid2charge: Market user interface design for electric vehicle charging. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 882–890. International Foundation for Autonomous Agents and Multiagent Systems.

Stern, N. (2006). *Stern Review on the Economics of Climate Change*. HM Treasury.

Strömberg, H., Andersson, P., Almgren, S., Ericsson, J., Karlsson, M., & Nåbo, A. (2011). Driver interfaces for electric vehicles. In *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '11, pp. 177–184, New York, NY, USA. ACM.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press Cambridge.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*(3-4), 279–292.