Web Intelligence and Agent Systems: An International Journal 0 (2013) 1–0
IOS Press

Evaluating the Applicability of Peer-Designed Agents for Mechanism Evaluation

Avshalom Elmalech
Computer Science Department
Bar-Ilan University
Ramat-Gan, Israel
elmalea@cs.biu.ac.il

David Sarne
Computer Science Department
Bar-Ilan University
Ramat-Gan, Israel
sarned@cs.biu.ac.il

Abstract.

In this paper we empirically investigate the feasibility of using peer-designed agents (PDAs) instead of people for the purpose of mechanism evaluation. This approach has been increasingly advocated in agent research in recent years, mainly due to its many benefits in terms of time and cost. Our experiments compare the behavior of 31 PDAs and 150 people in a legacy eCommerce-based price-exploration setting, using different price-setting mechanisms and diverse performance measures. The results show a varying level of similarity between the aggregate behavior obtained when using people and when using PDAs. In some settings similar results were obtained, in others the use of PDAs rather than people yielded substantial differences. This suggests that the ability to generalize results from one successful implementation of PDA-based systems to another, regarding the use of PDAs as a substitute for people in system evaluations, is quite limited. The decision to prefer PDAs for mechanism evaluation is therefore setting dependent and the applicability of the approach must be re-evaluated when switching to a new setting or using a different measure. Furthermore, we show that even in settings where the aggregate behavior is found to be similar, the individual strategies used by agents in each group highly vary. Finally, we report the results of an extensive comparative analysis of the level of optimality reflected in people's and PDAs' individual decisions in our decision making setting. The results show that the decisions of both groups are far from optimal, however the use of PDAs results in strategies that are more than twice as close to the optimal ones.

Keywords: Peer-Designed Agents; System Evaluation; Bounded Rationality; Decisions' Optimality; Simulation.

1. Introduction

Recent research increasingly relies on Peer-Designed Agents (PDAs), i.e., computer agents developed by people, as an efficient means for replacing people in evaluating systems [44]. Examples of such implementations can be found in various domains, such as negotiation [31], security systems [30] and parking allocation [12]. Similar to agent technologies in general, the use of PDAs for system evaluation encompasses many advantages, such as allowing the simulation of high-level types of information-oriented behaviors, interaction with other individuals, influencing one another to make separate decisions, and simulation of large-scale systems, due to the relatively small cost of cloning agents [64]. On top of these capabilities, PDAs may directly address one of the key challenges in using agents to replace people for system evaluation which is having the ability to accurately capture people's behavior in the simulated domain. People are known to be associated with diverse behaviors, which makes it difficult to capture behavior patterns in a monolithic model [33]. The use of PDAs in comparison to other agent design methods may offer a better representation of the rich set of behaviors used by people.

Indeed, if PDAs could capture people's behavior adequately, a PDA-based system is likely to perform similar to the case where the system is populated with people. Nevertheless, empirical investigation of the level of similarity observed between PDAs and people is not conclusive; some work suggests a relatively strong correlation between the behaviors of the two [12], while in other work the PDAs are reported to act differently than people to some extent [18,44]. Yet, even in cases where individual differences between the behavior of PDAs and people were reported, the performance of mechanisms applied directly to people and PDAs were similar [31,30]. Consequently, the conclusion of these works is that PDAs can alleviate the evaluation process of mechanisms (replacing people) and facilitate their design.

Unlike prior work, the underlying hypothesis of the research reported in this paper is that the extent of similarity between the average behavior observed when using people and PDAs is mechanism and measure dependent. The reported experimental results demonstrate how the use of the same set of PDAs leads to various different conclusions regarding the applicability of the "PDAs as a substitute for people" approach, simply by changing either the measures used or the mechanism evaluated. While complete similarity is reported for some combinations, in others the difference is substantial.

In order to test our hypothesis, a set of experiments with people and PDAs was carried out. The experimental design we used was the classical exploration-versus-exploitation setting of price-search in eCommerce [62,32,35,43]. In this setting, buyers need to decide which sellers to query for the prices of a specific item, where the price-querying incurs some cost and prices are a priori uncertain. The performance of people and PDAs in this domain is measured by three different parameters (measures): the exploration extent, the agent's overall expense and the seller's revenue. The evaluation of performance recurs with three different price setting mechanisms, using 150 human subjects and 31 PDAs.

The paper also presents the results of two complementary experiments that are of great value to the research of PDAs and the relationship between the behavior they exhibit compared to people's behavior in similar decision-making situations. The first experiment is used for comparing the level of optimality measured in PDAs' and people's decisions for the same decision-making situations. The results show that both populations perform far from optimal, yet PDAs decisions are, on average, twice as close to the optimal decisions than people's decisions in the domain used. Furthermore, a drill-down analysis of the results reveals that the most apparent difference between the two in terms of decision optimality is that people find it more difficult than PDAs to recognize which opportunity to explore next, whenever the exploration process should resume according to the optimal strategy. Nonetheless, they are not that far behind agents in their ability to recognize the benefit in terminating and resuming the exploration. The second experiment is to validate the fitness of PDA-programmers as legitimate representatives of the general population.

The importance of the results is in showing the great caution that should be taken when attempting to generalize specific results obtained in this line of research. Namely, one cannot deduce from the success of PDAs in reflecting an average behavior similar to the one obtained with people in a specific domain for other domains. Furthermore, even if the same domain is used, there is no guarantee that the same level of similarity will hold when the evaluated mechanism is changed. Even if only one mechanism is evaluated, and only the performance measure is changed, there is no guarantee that similar results will be obtained. Instead, whenever considering the use of PDAs as a substitute for people in system evaluation, one needs to test the applicability of this approach for the specific mechanism and measures of interest, using a pilot study or other means. The comparative evaluation of the optimality level exhibited in people's and PDAs' decisions provides several additional insights from this important point of view, relating to irrationality and sub-optimality in these two populations in costly-exploration domains.

We note that beyond the main contribution to the research of PDAs as substitutes for people in decision-making studies, the results reported in the paper suggest several secondary contributions to the research of search theory and pricing. For example, it suggests that results obtained in prior research (focused on simplistic exploration settings, indicating that people's search efficiency is close to the theoretical optimal strategy [37]) do not hold when considering more realistic exploration opportunities. The results also suggest that in this important domain agents cannot be assumed to search optimally, despite their substantial computational capabilities. This is important since much of the theoretical work dealing with agent-aided search-based ecommerce (especially in the area of pricing) assumes that the agents are likely to exhibit fully rational behavior. An additional implication of the results is that having the ability to distinguish between human and agent consumers can substantially improve the effectiveness of price setting methods in costly-exploration based markets.

In the following section we review related work. Then, in Section 3 we introduce the exploration problem that is used as a framework for evaluation. Section 4 details the evaluation methodology, including the different mechanisms and measures, the experimental infrastructure and the problem sets used. The results and their analysis are given in Section 5. The complementary experiments to compare the level of optimality exhibited in people's and PDAs' decision making in the costly-exploration domain, and their analysis and implications, are given in Section 6. Finally, we conclude with a discussion and directions for future work.

2. Related Work

The use of PDAs in literature is quite extensive, mainly due to their twofold advantage: First, being autonomous agents, PDAs offer many benefits in terms of simulation design and maintenance. Second, since their strategy is determined by a large set of individuals, PDAs can potentially enable representation of a rich set of realistic real-life strategies. Therefore the review given in this section touches on both topics, including the use of autonomous agents in simulation in general, the challenge of and different approaches used for capturing people's behavior and the use of PDAs for various purposes.

Recognizing autonomous agents' capabilities to interact among themselves and scale, many researchers have suggested the use of them in simulation [64,7,51,38].[1] The increasing interest in agents for simulating complex adaptive systems has led to several multi-agent software platforms, such as SWARM [36], JADE [3], Robocup-Rescue [55] and Repast Symphony [41]. While these systems supplied the infrastructure for facilitating the agent-based simulation development, the task of reliably constructing the individual agents' behaviors remained beyond their scope. For this reason, simulation designers have used various methods for setting human-like behavior in the agents they developed. This include among others, statistical-data based modeling [56], pre-specification of agents' roles using a set of parameters according to which the agents act [34], using pre-defined events and reactions [40], defining a number of basic behaviors from which all of the agents' more complex behaviors are constructed [50,58] or using a combination of rules and finite state machines to control an agent's behavior using a layered approach [59]. The main difficulty of these methods is usually in situations different from those used to collect the real data from which the strategies were derived.

The success of simulation designers constructing human-like behavior in agents is controversial. For example the transaction price path of agent traders designed using bounded rational theories was found to bear some resemblance to the transaction price path of human subjects in a double auction environment [17]. In recent years, however, much evidence has been reported on the discrepancies between the behavior of such agents and people's behavior. For example it was shown that the resemblance between people and agents reported for the double auction environment does not hold once the value of one of the market parameters slightly changes [60,9]. Moreover, the simulation designer and even domain experts are quite limited in the number of different individual behaviors they can generate.

In some respects, the idea of using people to program agents as a means for achieving a reliable set of strategies for specific decision situations relies on the "strategy method" paradigm from behavioral economics [49]. In the strategy method people state their action for every possible situation that may arise in their interaction. The question of whether or not the strategy method can ensure real-life behaviors has been extensively researched in the last decade. Several authors reported no substantial differences between the choices made using the strategy method and

[1] The first use of agents in this context is reported in [52].

those made by people in experiments, e.g., in the sequential dictator game [11], in measuring people's willingness to pay for insurance [5,6] and even in sequential exploration problems (a bit more simplistic than the one we use) [54]. Others, however, reported significantly different results with the strategy method compared to ones obtained by people playing the same game, e.g., with trust game variations [10], trust-punishment games [8] and social good allocation games [39]. One of the explanations given for the difference between the two methods is the tendency of people to be influenced by emotions, where in strategy method people are in a "cold" state and are less emotionally arouse [10]. The main difference between the strategy method and PDAs technology is that in the first participants need to enumerate their choices for each possible state of the game, whereas with PDAs the requirement is to express a cohesive formulation of their strategy [12,44,14]. This entails various implications related to the time it takes to capture one's strategy (an advantage for the PDAs in cases where the possible system's state is large and an advantage for the strategy method when the game is extremely simple, e.g., in the ultimatum game), the ability to understand one's strategy (an advantage for PDAs, as their code can be analyzed afterwards) and the ability to use the strategy when the setting slightly changes (impossible with the strategy method).

The PDAs technology has been widely used in recent years. For example, in Kasbah [13] PDAs that buy and sell were used for evaluating an electronic marketplace. In Colored Trails [18], PDAs were used for reasoning about players' personalities in uncertain environments. Other works, e.g., [31,30,12,44] used PDAs for evaluating specific mechanisms in various domains such as evaluating security algorithms and evaluating automated negotiators.

As mentioned above, the main motivation for having people program agents' behaviors is the possibility that the resulting strategy will correspond to their own. This, however, is not straightforward. Evidence of discrepancies between actual and reported human behavior is a prevalent theme in research originating in other various domains, in particular in metacognition research [19]. Examples of such discrepancies include over-reporting of political participation [4] and contrasting results between self-reported and performance-based levels of physical limitations [27]. Indeed, part of the PDA-based literature simply does not consider the PDAs-people similarity question as an issue or attempt to make any claims regarding this aspect (e.g., in TAC [63,48,25]). Yet, much of the PDA literature tends to assume that people can successfully (to some extent) capture their real-life strategy in a given domain when programming an agent [14,44]. Even in cases where some discrepancy between PDAs and people's behavior is reported, the average performance is reported to be similar, suggesting that PDAs can replace people in mechanism evaluation [31,30]. Common to all the above work is that they draw their people-PDAs similarity-related conclusions based on the specific implementation they use. The current work attempts to present a more targeted evaluation, which compares the same populations while varying both the mechanism and the measure used.

3. The Exploration Model

As the underlying framework for the research, we consider the canonical sequential exploration problem described by Weitzman [62] to which a broad class of search problems can be mapped. In this problem, a searcher is given a number of possible available opportunities $S = \{s_1, ..., s_n\}$ (e.g., to buy a product) of which she can choose only one. The value v_i of each opportunity s_i (e.g., expense, reward, utility) is a priori unknown to the searcher. Only its probability distribution function, denoted $f_i(v)$, is known to the searcher. The true value of opportunity s_i can be obtained ("disclosed"), but only by paying a fee, denoted c_i, which is possibly different for each opportunity. Once the searcher decides to terminate her search (or once she has disclosed the value of all opportunities) she chooses the one with the minimum or maximum value (depending on whether values represent costs or benefits) among the opportunities whose values were obtained.

The exploration problem as formulated above is generic and applies to a variety of real-world situations. For example, when looking for a used car, the information given in ads reflects a noisy signal, and the true value of cars can be revealed only through a test drive or a costly inspection. Similar exploration problem characteristics can be found in job-search [16], multi-robot systems [20] and assignment of jobs to servers in computer systems [45]. The literature on costly-exploration-based settings reveals the use of various assumptions regarding the nature of the exploration executed by searchers (human or agent). These range from no exploration at all [29,23], through obtaining the value of a pre-defined fixed number of opportunities [28], to optimal exploration [24,57].

The optimal (overall-expense minimizing) strategy for the above exploration problem is given in "search theory" literature [62].[2] According to this solution the entire structure of the optimal policy can be reduced to a simple statement about reservation prices (thresholds); the searcher should set a threshold, denoted r_i, for each opportunity (e.g., a seller) s_i, calculated according to:

$$c_i = \int_{x=-\infty}^{r_i} (r_i - x) f_i(x) dx \tag{1}$$

Intuitively, r_i is the value where the searcher is precisely indifferent: the expected marginal benefit from obtaining the value of the opportunity exactly equals the cost of obtaining that additional value. The searcher should always choose to obtain the value of the opportunity associated with the minimum threshold and terminate the exploration once the minimum value obtained so far is less than the minimum threshold of any of the remaining opportunities.

While this exploration setting is common, the nature of its optimal strategy is non-intuitive as thoroughly discussed in [62]. For example, an important property of the above solution is that the threshold calculated for each opportunity does not depend on the number and properties of the other opportunities, but rather merely on the distribution of the value of the specific opportunity and the cost of evaluating it (i.e., considers the myopic gain of revealing the value of that opportunity and terminating, rather than the full gain of exploring the opportunity and continuing in an optimal manner). Also the reservation value is completely insensitive to the probability distribution values at the higher end of the tail (as the integral limits are between $-\infty$ and r_i), meaning that any rearrangement of the probability mass located above the reservation value leaves the reservation value unaltered (however it does change the expected value). Finally, when other things are equal, it is optimal to first sample from distributions which are more spread out or riskier. These low-probability high-payoff situations should be prime candidates for early investigation even though they have a smaller chance of ending up as the source ultimately yielding the best reward when the exploration ends. Based on the non-intuitiveness of the solution characteristics, people's exploration strategies when facing the problem are likely to be different from the optimal strategy and diverse.

The specific sequential exploration problem that is used as a framework for testing our hypothesis considers a buyer searcher operating in an electronic marketplace populated by N sellers. The sellers are assumed to be active in other various markets (i.e., multichannel retailers [22]) and frequently change their prices through various dynamic pricing techniques [26,29]. Their pricing is therefore taken to be external, unaffected by the buyer's exploration behaviors in this market. Any seller s_i is therefore assumed to be associated with a different distribution function $f_i(v)$, from which its price is drawn at any given time, if queried for its price.[3] The distributions $f_i(v)$ $(i = 1, ..., N)$ are assumed to be known to the buyer (or can be learned using past experience, Bayesian update, etc.). The buyer's price-querying process itself is assumed to be costly, as it requires the consumption of some resources (either CPU time, communication bandwidth, etc., in the agents' world or time, parking fees, transportation costs, etc., in the physical world) [29]. This cost is assumed to be seller-dependent — the cost of querying the price of seller s_i is denoted c_i. Buyers are assumed to always prefer the minimum quote among those received from the sellers they query (i.e., sensitive to price [61]).

Given the above problem description, a buyer needs to set its expense-minimizing exploration strategy based on the set of sellers, their distribution of prices and querying costs $(f_1(v), ..., f_N(v)\ , c_1, ..., c_N)$. The buyer's strategy thus determines which seller to query next (if at all) based on the best (lowest) price obtained so far and the set of sellers that have not been queried yet.

Reliably capturing the exploration strategies of buyers in such settings is important for evaluating various market design mechanisms, among which we focus on price setting mechanisms. In particular, we consider the problem of a single seller that operates only in this market (i.e., a non-multichannel retailer whose pricing does not affect the prices set by the other sellers). A key challenge for a seller whenever considering new price-setting mechanisms (i.e., ones that set a price as a function of the costs of query and distribution of prices of the other sellers) or when having to choose among several existing ones is to reliably evaluate the expected revenue from each such

[2] In this paper we use the expense-minimizing version of the problem, as our experimental infrastructure is based on the price-search problem.

[3] This assumption, which is commonly used in exploration models [62], has evidence in a large body of empirical research in the form of the persistence of price dispersion both in traditional and online retail markets [15,2].

mechanism. Evaluating the mechanisms by actually applying them is risky, as it may be associated with substantial losses. The use of PDAs in this case may seem quite appealing if it can be guaranteed that the resulting system performance (i.e., in terms of average expected revenue) will be similar to the one obtained in the case of evaluating the mechanisms with human buyers.

Overall, the use of the exploration problem described above has many advantages in the context of this research. First, it considers a real-life setting with which most people are familiar (or experienced with, and thus are likely to have a well-established strategy for it). Second, sellers in such domains are required for frequent evaluations of different price setting methods due to the dynamic nature of eCommerce which makes it ideal for PDA-based evaluation, if indeed PDAs exhibit similar performance (to people). This also implies that various price setting mechanisms can be considered for our experiments. Finally, the domain offers several possible measures for capturing overall exhibited exploration behavior (i.e., overall performance), according to which the similarity between PDAs and people can be evaluated.

4. Evaluation Methodology

The goals set for the evaluation were twofold: First, we wanted to compare the system's performance when the exploration is carried out by PDAs and people using different price-setting mechanisms and diverse measures. Second, we wanted to supply a drill-down analysis of the exploration strategies used by people and PDAs, to possibly explain the differences observed in system performance. Along with these two goals, we also evaluated the performance of an agent that uses the optimal (i.e., expected-expense-minimizing) exploration strategy. The performance of this agent, and in particular its individual decisions in similar decision situations experienced by people and PDAs were used as a benchmark for evaluating the extent to which people's and agents' strategies were suboptimal, and how significant the differences between the two were.

The evaluation was based on three different measures and three different price setting mechanisms. Each measure-mechanism combination was used both with PDAs, people searchers, and the optimal agent. In the following paragraphs we describe these measures and mechanisms, the experimental infrastructure that was developed and the sets of exploration problems that were designed for the experiments.

4.1. Pricing Mechanisms

The goal of having different pricing mechanisms in our experimental design was to enable the production of different sets of exploration settings with a slight, though consistent, variation between sets. With each different pricing method, buyers faced the same sets of problems, however the price of one of the sellers was determined according to different logic.

We emphasize that for the purpose of testing the research hypothesis any set of mechanisms that act consistently (i.e., that always return the same price if the same problem instance is given again) is appropriate. Still, in an effort to improve the realism of our experiments, we attempted to come up with "reasonable" pricing mechanisms, of the kind that sellers who would potentially adopt the PDA-based methodology might choose to evaluate. The three different pricing methods that were designed for our experiments are: Theoretic-Optimal Pricing (denoted "M1"), Mean-Fonders Pricing (denoted "M2") and Cost-Probability Tradeoff Pricing (denoted "M3"). All three are designed to output a deterministic price q (or alternatively, a degenerated distribution with all of its mass around q). Also, all three methods assume that no prior information about the buyer is available and rely solely on the distribution of the other sellers' prices and exploration costs. Since the model assumes that the buyer must query the seller prior to making the purchase, then the buyer's expense directly associated with seller i when purchasing the product from that seller at price q is $c_i + q$.

4.1.1. Theoretic-Optimal Pricing (M1)

This price setting method assumes that the buyer is fully rational and uses the optimal (threshold-based) exploration strategy described in the previous section. According to this exploration strategy, the buyer will continue to query sellers as long as the best price she has obtained so far is greater than the reservation value of any of the remaining non-queried sellers. Therefore, if "our" seller sets a price q, the buyer will certainly not buy the product at that price as long as other sellers associated with a reservation value (calculated according to Equation 1) lower than $q + c_i$ exist. Only if all the sellers of the latter type are queried and reveal prices greater than $q + c_i$ then the buyer will buy from "our" seller at a price q, with the necessary expense of $q + c_i$. (Sellers associated with a reservation value greater than $q + c_i$ do not need to be considered as they will not be queried according to the optimal strategy rule.) The probability that a seller j associated with a price distribution $f_j(x)$ reveals a price higher than $q + c_i$ is $(1 - F_j(q + c_i))$, where $F_j(q)$ is the appropriate cumulative distribution function. Therefore the expected benefit from setting a price q is given by $q \prod_{r_j < q+c_i} (1 - F_j(q + c_i))$. The deterministic price q that maximizes seller s_i's expected revenue can therefore be extracted from:

$$argmax_q \big(q \prod_{r_j < q+c_i} (1 - F_j(q + c_i)) \big) \tag{2}$$

4.1.2. Mean-Fonders Pricing (M2)

The Mean-Fonders Pricing is based on experimental evidence relating to the tendency of people to overemphasize mean values in problem solving, by reasoning about one feature of a distribution rather than all of the distribution's features [37]. A buyer whose exploration is driven by mean values is likely to follow a modification of the above threshold-based optimal exploration rule, whereby the thresholds used are merely based on means. Specifically, the threshold that will be assigned to each seller s_i in this case is $r'_i = \mu_i + c_i$, where μ_i is the expectancy of $f_i(q)$. Intuitively, this would mean greedy querying, i.e., always preferring to query the next seller for which the sum of her expected requested price and the expense of querying her is the lowest (among those for which that sum is lower than the current best price known to the buyer). The sub-optimality of this strategy derives from the fact that by taking the mean value the higher prices of the distribution "override" the best price obtained so far, instead of avoiding them and weighing in the calculation only those cases where the best price improves based on the new price obtained.

A buyer following the mean-based exploration strategy will thus buy the product from "our" seller that uses a deterministic price q and is associated with an exploration cost c_i, only if former queries made to all sellers associated with thresholds lower than $q + c_i$ have yielded prices greater than $q + c_i$. The seller using the Mean-Fonders Pricing heuristic will therefore set a price q^* according to:

$$q^* = argmax_q \big(q \prod_{r'_j \leq q+c_i} (1 - F_j(q + c_i)) \big) \tag{3}$$

It is notable that the above rule assumes that the searcher fully relies on mean values. Seemingly, PDAs are likely to exhibit this kind of exploration behavior more frequently than people, as their computational capabilities better enable means calculation in comparison to people. Nevertheless, the extent to which PDAs' strategies are based on means is inconclusive, and while it is possible that many of them take into consideration mean values there is no guarantee that PDAs fully follow the above strategy. In fact, the attempt to utilize this property as a general rule for improving PDAs' exploration was reported to be unsuccessful [45]. Still, the use of means in PDAs' search, even to a limited extent, makes it a good candidate for basing our second pricing heuristic on it.

4.1.3. Cost-Probability Tradeoff Pricing (M3)

The Cost-Probability Tradeoff Pricing heuristic also assumes that the buyer follows a threshold-like decision rule; however, thresholds are dynamically set according to the best value obtained during any stage of the exploration. Specifically, the buyer is assumed to calculate a threshold for each seller s_i, denoted r''_i, that averages two terms. The first is the probability that the price that will be obtained from that seller will be below the best price known, and the second is the relative magnitude of the cost of querying that seller compared to the maximum possible querying cost:

$$r''_i(x) = \frac{P(q_i < x)}{2} + \frac{MaxCost - c_i}{2 \cdot MaxCost} \tag{4}$$

where x is the minimal price obtained so far along the exploration, $MaxCost$ is the maximum querying cost of any of the sellers in the problem setting. At any stage of its exploration, the buyer will query the seller associated with the maximum threshold as long as it is greater than 0.5 which is, in fact, the threshold assigned to any of the sellers that have already been queried.[4]

The intuition in considering this exploration strategy relies on former evidence of people's use of thresholds (even in the form of rules of thumb) in their search behavior [46,37], and at the same time experimental evidence indicating that people's exploration behavior does not seem to be related to risk aversion [53] but rather to loss aversion [47]. The heuristic thus attempts to balance between the probability of finding a better price (better than the minimum found so far) and the cost incurred as part of such exploration.

Following the solution concept of the other two methods, the benefit-maximizing price q^* when buyers use the cost-probability tradeoff approach is given by:

$$q^* = argmax_q(q \prod_{r_i''(q+c) \geq 0.5} (1 - F_i(q+c))) \tag{5}$$

4.2. Measures

The performance measure of interest when evaluating an exploration setting of the type used in this paper depends on the evaluator's goal. For example, if a seller's best interests are of concern, the seller's expected revenue should be the measure of interest. If a buyer's welfare is of concern then the buyer's expected overall expense should be considered. We used three different performance measures of the individual and system performance. These included the *exploration extent*, measuring the number of sellers a buyer explored throughout its exploration process; the buyer's *overall expense*, measuring the minimum value obtained plus the exploration costs accumulated along the process; and the *seller's revenue*, measuring the payment received by the seller using the evaluated pricing mechanism. If PDAs indeed exhibit behaviors similar to those of people in this domain, then all three measures should reflect a consistent pattern.

4.3. Experimental Infrastructure

Two experimental infrastructures were developed, simulating a price-search environment. The first was designed to experiment with PDAs and a theoretic-optimal buyer and the second for experimenting with people.

4.3.1. Evaluating PDAs

The PDAs' evaluation infrastructure enables our system to instantiate agents with the appropriate exploration problem input, receive their exploration choices and supply them with values based on their selections (according to the distribution of values of the different options available). The agents, upon receiving the problem input, which includes the costs of querying and the distribution of values of the different opportunities available to them, have to decide at each stage of the process whether to terminate the exploration and end up with the lowest value revealed so far, or to resume exploration. In the case of continuing exploration, they also need to inform the system who to query next.

To facilitate the evaluation of the pricing mechanisms with the theoretic-optimal buyer, an additional agent was developed, applying the optimal exploration principles (i.e., using a set of thresholds according to Equation 1). This agent implemented the same API as the PDAs and thus could be used with the same experimental infrastructure.

Our experimentation used agents designed by computer science students in a core Operating Systems course. As part of a regular course assignment in which an agent needed to be programmed, the students also had to program the agent's strategy layer according to the above guidelines. A portion of the assignment's grade was correlated with the agent's performance, i.e., the overall querying costs plus the value it ended up with. As part of their assignment, students provided documentation that described the algorithm used for managing the exploration. An external proxy program was used to facilitate communication with the different stores. The main functionality of the proxy was to randomly draw a store's price based on its distribution, if queried, and to calculate the overall querying costs and the price paid. The above procedure complies with the common practice in PDA-based research [31,30,12,14].

[4]Since the probability for a seller that has already been queried of having a price lower than the best price found is zero (cancelling the first term of (4)) and there is no need to spend any further resources in its exploration (hence the second term of (4) becomes 0.5).

Overall, we used 31 PDAs that the students developed, each tested with all of the problems from the set of problems described below with each of the three pricing methods.[5]

4.3.2. Evaluating People

The second evaluation infrastructure developed was a JavaScript web-based application, emulating an exploration problem with 8 stores, each associated with a different distribution of prices and a cost for obtaining the true price (represented as a "parking cost" for parking next to that store). Figure 1 depicts a screenshot of the system. In this example, the price of two stores is already known. Querying a store is done by clicking the "Check" button below it, in which case the true price of the store becomes known and the parking cost of that store is added to the accumulated cost. The game terminates when clicking the "Buy" button (available only for stores whose prices are known), upon which a short summary of the overall expense is presented to the player (divided into the accumulated exploration cost and the price paid for the product itself).

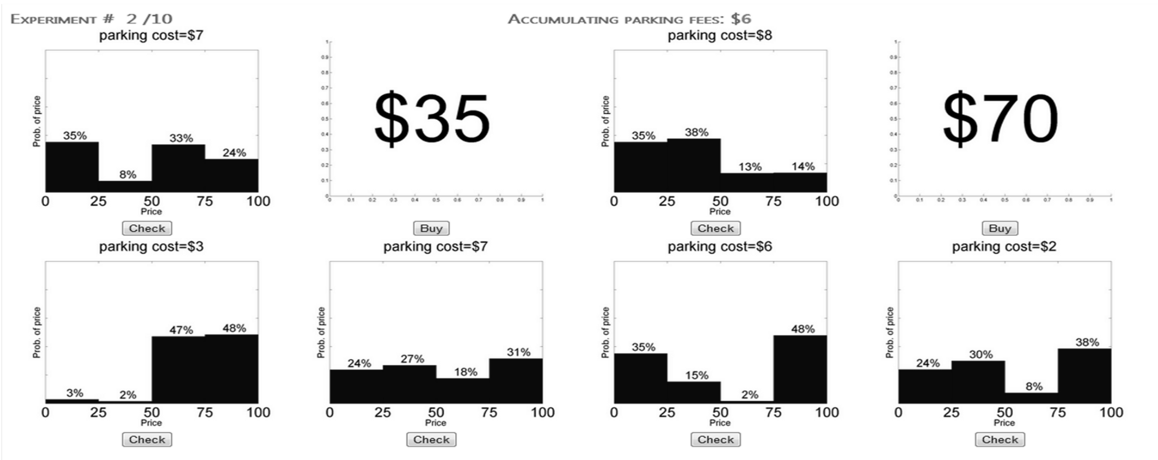


Fig. 1. A screenshot of the system designed for experimenting with people.

Subjects were recruited using Amazon's Mechanical Turk service [1].[6] When logged in, subjects received a short textual description of the experiment, emphasizing the exploration aspects of the process and the way costs are accumulated, followed by a short video clip. Then, a series of practice games were played in order to make sure that the subject understood the experiment. Participants had to play at least three practice games; however, they could continue practicing until they felt ready to start the experiment. Once the subjects finished practicing, the system randomly drew 10 problems from its problem repository of 100 problems (see details below) for them to play sequentially. To prevent the carryover effect, a "between subjects" design was used, allowing each participant to participate only in one experiment using the same pricing method in all of the 10 problems presented to her. To motivate players to exhibit efficient exploration behavior, and possibly also extend their practice section, they were told that players among the 40% with the minimal average overall expense would receive double the payment for their participation in the experiment. As a means of precaution, the time it took each participant to make each selection and the overall time of each game played was logged. Two participants with unusually low times were removed from the database and new participants were recruited to replace them. Overall, 150 people participated in the experiments, 50 for each pricing method (mechanism).

[5] The PDAs can be re-used across the different mechanisms as they were originally designed for producing an exploration strategy given a general multi-seller uncertain offerings. Since the mechanisms merely deal with price setting, the PDAs were perfectly suitable for all of them.

[6] For a comparison between AMT and other recruitment methods see [42].

4.3.3. The Choice of Using a General Population

We note that PDA programmers and the Amazon Turk participants are not from the same population. The former are all computer science students. The latter are from a more general population. This choice of participants was intentional to comply with the general idea which motivates PDA literature, whereby the PDAs (whose development requires some knowledge in programming) can replace people (from the general population) in system evaluation. Therefore, our group of computer science students fairly represents the typical population of PDA programmers, whereas the group from Mechanical Turk corresponds to a general group of individuals one would likely recruit for experimental evaluation of a tested system.

To broaden the understanding of the effects that the use of these two different populations have on the similarity results (between people and PDAs) a complementary experiment was designed, aimed to repeat the system evaluation with one of the three mechanisms (M1), this time with computer science students rather than people recruited through Amazon Turk. None of the students recruited for this experiment had participated in the PDA programming task (to avoid any carryover effect). Overall, 50 computer science students participated in this experiment, whereby each was given 10 problems (similar to the experiment with Amazon Turk). By comparing the performance of computer science students and people from the general population, we were able to reason about whether the source of the difference between PDAs and people in our experiments (whenever such a difference was observed) was the result of the PDA designers belonging to a different population than the Amazon Turk or if it could be attributed to an inherent difference between PDAs' and people's behaviors.

4.4. Problem Sets

The set of problems consisted of 100 randomly generated problems. The distributions of prices these problems used were formed as multi-rectangular distribution functions. In a multi-rectangular distribution function, the interval is divided into subintervals $(x_0, x_1), (x_1, x_2), .., (x_{n-1}, x_n)$, and the probability distribution is given by $f(x) = \frac{p_i}{x_i - x_{i-1}}$ for $x_{i-1} < x < x_i$ and $f(x) = 0$ otherwise ($\sum_{i=1}^{n} P_i = 1$). The benefit of using a multi-rectangular distribution function is its simplicity and modularity, in the sense that any distribution function can be modeled through it with a relatively small number of rectangles. Furthermore, the multi-rectangular distribution function is easier for people to grasp, as they can be given the explicit probability captured by each rectangle (equally distributed over the interval), as illustrated in Figure 1, rather than having them struggle with a distribution function which values are difficult to interpret.

Each problem in the set contained 8 stores, where one of them was the store for which the evaluated pricing techniques apply. The distribution of prices of the other seven stores was randomly generated under the constraint of having exactly four rectangles, each defined over an equal-size interval. This, again, in an effort to facilitate people's understanding of the price-distribution and their ability to reason about the problem. The overall interval of prices was $(0 - 100)$, as illustrated in Figure 1. The exploration cost for each store (including the store to which the pricing techniques applied) was randomly picked from the interval $(1 - 10)$. Three variants of each problem were generated, one according to each of the pricing mechanisms described above. The variants of each given problem thus differed only in the price set for the store for which the pricing applies.

5. Results

We briefly review the set of PDAs received and then describe the performance of PDAs compared to people as reflected in the experiments according to the different measures and evaluated mechanisms. Then we present a drill down analysis of the performance obtained when using people and PDAs problem-wise and individual-wise. Statistical significance was tested, whenever applicable, using a t-test (assuming unequal variance).

5.1. PDA strategy

When using PDAs instead of experimenting with people the analysis of individual strategies becomes relatively easy, as one only needs to review the PDAs' code in order to reveal their strategy. The strategies used by the PDAs in our experiments reveal several characteristics with which agent designs varied, including: relying on variance (16%) and the expected value (96%) of each store, weighing the parking costs (querying costs) as an affecting factor (93%), a preliminary selection of stores for querying (54%), the inclusion of the cost incurred so far (i.e., "sunk cost") in the decision-making process (12%) and the use of the probability of finding a store with a lower price than the minimum found so far (6%). It is notable that most of these characteristics do not affect (or directly relate to) the optimal strategy (see Section 3). In particular, many of the PDAs (67%) use the mean price of a store as a parameter that directly influences the exploration strategy, to some extent, even though the optimal strategy is not affected directly by means.

We note that no similar analysis is offered for people's strategies in our experiments, as it is impossible to understand the strategies people used based on their actions only. Furthermore, even if attempt is made to extract the strategy based on people's description of the strategy they used, there is no guarantee that these were indeed the strategies used (see related work section for evidence on the discrepancies between actual and reported human behavior). The analysis of the difference between the two populations is thus based entirely on their performance in the experiments (and in the next section, based on the individual decisions they made).

5.2. *Performance Comparison*

Figures 2-4 depict the average overall buyer's expense (payment for the product plus accumulated costs along the exploration), average exploration extent and seller's revenue, according to the type of buyer used (theoretic-optimal, person (Mechanical-Turk participant) or PDA) for each of the different pricing mechanisms (M1, M2 and M3). As can be observed from the figures, the magnitude of the difference in performance between any two types of buyers substantially varies. The following table summarizes the similarity in terms of the average behavior between PDAs and people for each measure-mechanism combination:[7]

	M1	M2	M3
Overall expense	+ (0.34)	- (<0.001)	+ (0.065)
Exploration extent	- (<0.001)	- (<0.001)	- (<0.001)
Seller's revenue	- (<0.001)	- (<0.001)	+ (0.019)

As can be observed from the table, the only mechanism for which a consistent result was obtained (cross-measures) was $M2$ (whereby there was a substantial difference between the overall behavior of PDAs and people). Similarly, the only measure for which a consistent result was obtained (cross-mechanisms) was the exploration extent (wherein once again there was a substantial difference between the overall behavior of PDAs and people). For all other mechanisms and measures, there was no consistency in the determination of whether or not the performance of people is similar to the performance of PDAs across measures and across the tested mechanisms. For example, with PDAs a similar average exploration expense and seller's revenue were obtained when the seller set her prices according to M3; however, substantial differences were observed in the performance with the two groups in both measures when M2 was used. Similarly, there was a substantial difference in the seller's revenue when playing against PDAs and when playing against people, if the pricing mechanisms M1 and M2 were used; however, no significant differences were noted if pricing mechanism M3 was used.

When comparing PDAs and people to the theoretic optimal agent, the performance of the latter was found to be substantially different ($p < 0.001$) from those of PDAs and people in all measure-mechanism combinations. We will discuss this result and its alignment with prior work in the next section, where we analyze the optimality aspect in more detail.

[7] Similarity is determined whenever there is no statistically significant difference ($p < 0.001$) and is marked as "+". The number in brackets is the appropriate $p - value$.

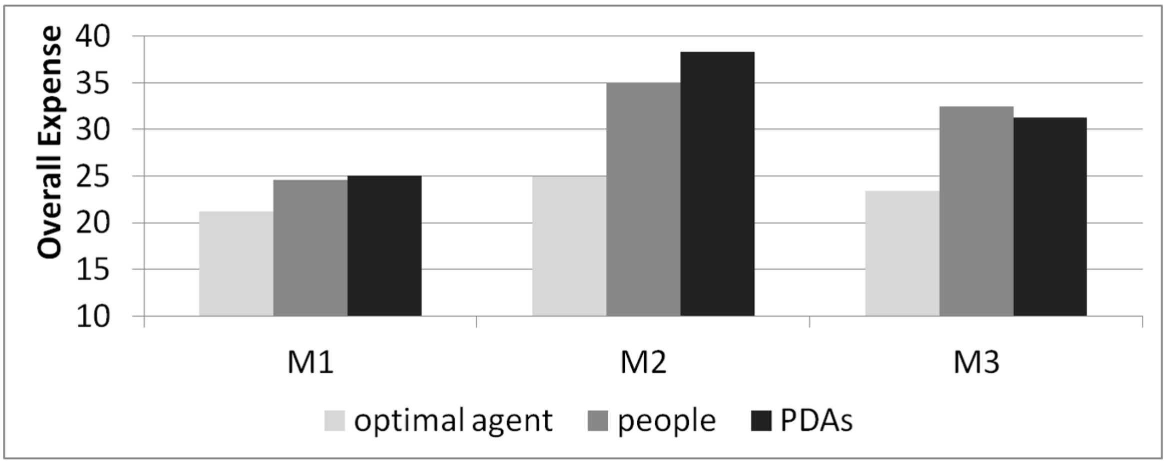


Fig. 2. Average buyer’s overall expense by the three pricing methods.

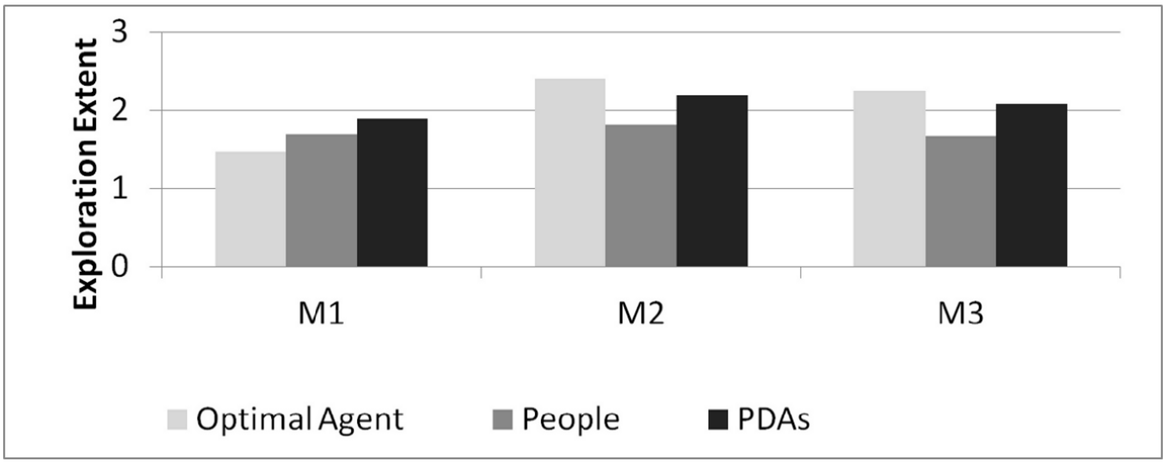


Fig. 3. Average buyer’s exploration extent by the three pricing methods.

The substantial differences obtained for some of the measure-mechanism combinations when using PDAs and when using people suggest that individuals from the two populations use different exploration strategies. To support this latter claim we present Figure 5, which illustrates a drill-down comparison of the buyer’s average expense according to their type and the individual problem used, for each of the price-setting mechanisms. Notice that under each mechanism (different pricing method) the buyer confronted a different problem, as the price of the seller whose price was set by the specific pricing method was different. For illustration purposes, the problems in each of the graphs are sorted according to the average performance achieved by human buyers, in ascending order. Each data point relating to people is the average of the overall expense achieved by human subjects who encountered the specific problem variant to which it relates (on the horizontal axis). The data points relating to PDAs depict the average performance of the 31 PDAs when encountering the same problem, and those relating to the theoretic-optimal agent depict the expected performance of the latter when given the problem. As can be observed in the figure, the average expense per-problem incurred by the PDAs as buyers was substantially different from the one incurred by people as buyers. Indeed, for some of the problems, the performance of PDAs was relatively close to that of human buyers (especially with M3); however, different results were obtained for the majority of the cases. As expected, the results of the theoretic-optimal agent were substantially better than the other two for the majority of the exploration problems tested. In a small portion of problems, people and/or PDAs managed to perform slightly better than the theoretic-optimal agent. This can be attributed to the number of participants that played each problem

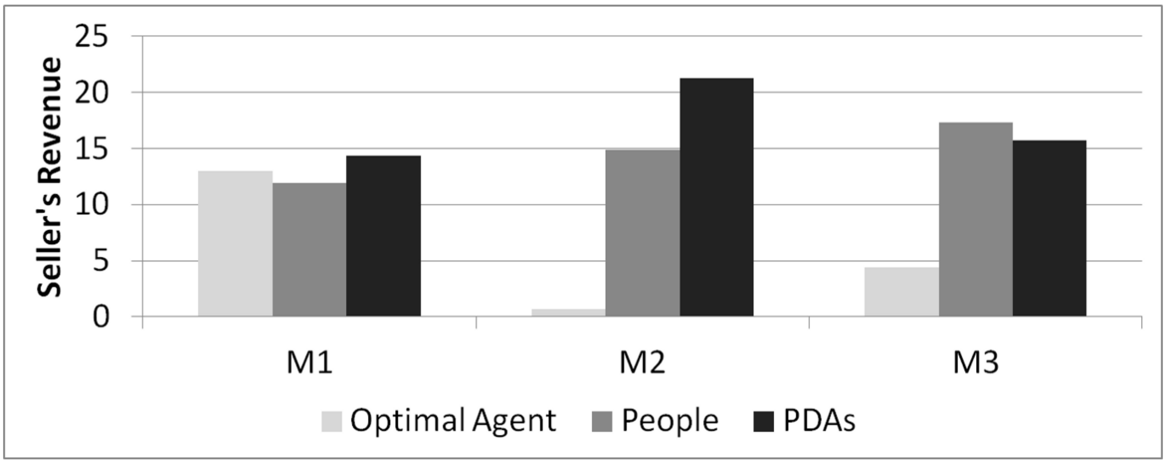


Fig. 4. Average seller's revenue by the three pricing methods.

and the fact that for a single play of a problem instance it is possible that a player using a non-optimal strategy will end up with a lower total expense compared to a player using the optimal strategy. If the number of people playing each problem is substantially large, the theoretic-optimal agent's average performance will inevitably be better as it minimizes the expected overall expense. Similar patterns were obtained in the drill-down analysis according to problem type for the exploration extent and the seller's revenue measures.

5.3. Performance Comparison

Finally, we report the results of the experiment aimed at testing the difference between computer science students and the general Amazon Turk population in costly-exploration-based environments. The following table details the average performance of the two populations according to the three different measures:

	Amazon Turk participants	Computer science students
Overall expense	24.6	23.8
Exploration extent	1.69	1.62
Seller's revenue	11.9	11.8

A two-sided ANOVA with no replications reveals no difference between the two populations for $\alpha = 0.05$ in all three measures. Furthermore, the comparison of the results of the computer science students and the PDAs revealed results similar to those given in the summary table above for the $M1$ mechanism (for exploration extent and seller's revenue measures $p < 0.001$ and for the overall expense measure $p > 0.001$).

6. Measuring the Level of Optimality in People and PDAs' Decisions

In this section we utilize our experimental infrastructure to present a thorough complementary investigation of the level of optimality reflected in people's and PDAs' decisions. The goal of this comparative evaluation is primarily to reason about irrationality and sub-optimality within these two populations, which might explain some of the differences observed between people's and PDAs' performance in this important domain.

As reported in the previous section, the performance of the optimal agent was found to be substantially different from that of PDAs and people in all measure-mechanism combinations. This result gains importance in light of the large body of research from past decades, attempting to experimentally investigate people's exploration patterns. Based on the results reported in this literature, the determination of whether or not people follow the theoretic-optimal strategy remains inconclusive. While some papers find that in general human searchers exhibit behavior that is consistent with that predicted in theory [46], others give evidence to differences between the behaviors exhibited

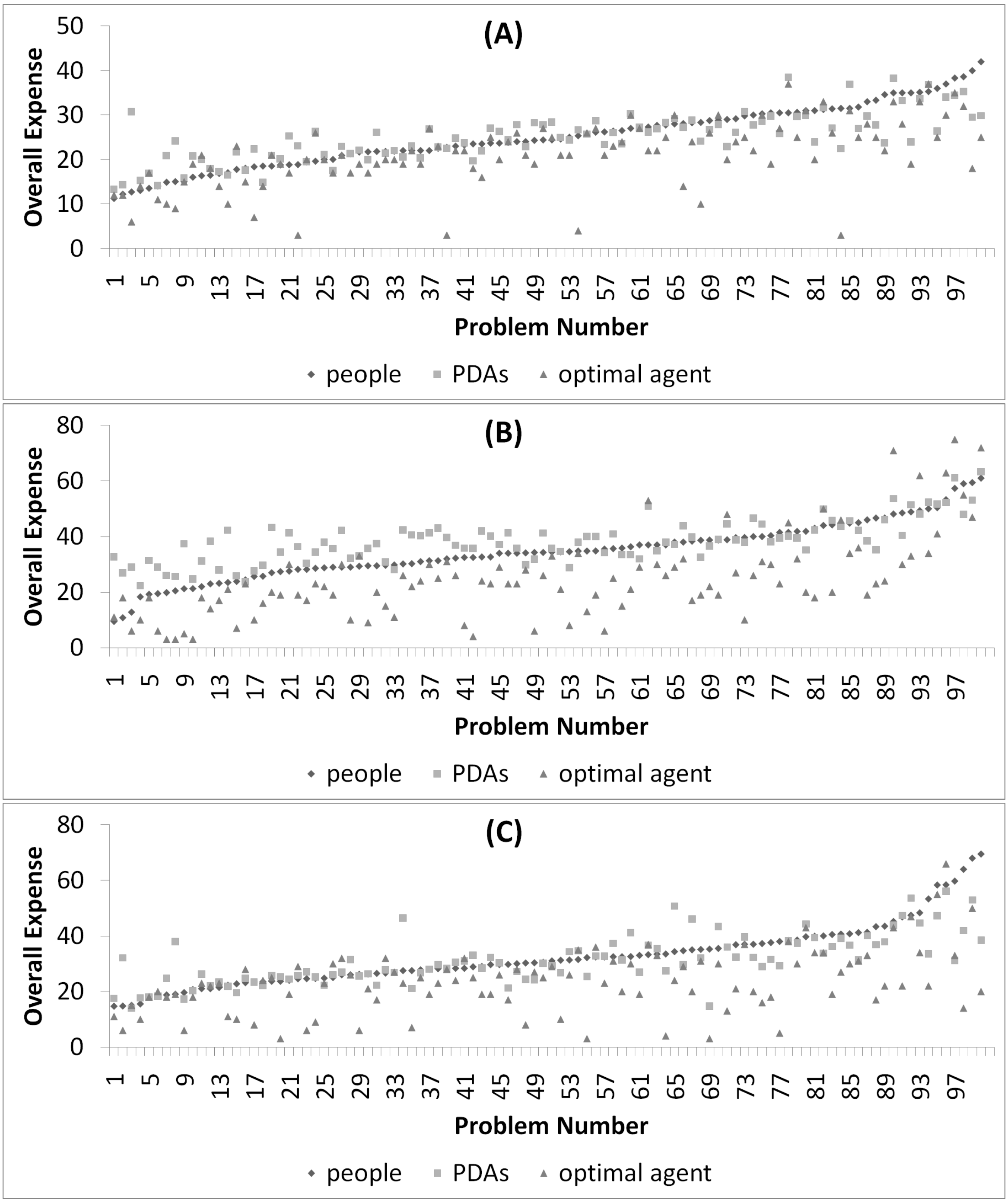


Fig. 5. The average expense of buyers according to the 3 pricing methods: (a) M1; (b) M2; and (c) M3.

by people and the theoretic-optimal strategy in exploration settings [21,37]. The major difference reported relates to the tendency of people to terminate their exploration before the theoretic-optimal strategy does [53,47]. Regarding the efficiency of people's exploration, there is consent between researchers that, regardless of whether or not people actually use the theoretic-optimal exploration strategy, there is no substantial difference between the performance of people in costly-exploration and the optimum [21,37,53]. Nevertheless, these works use a somewhat simplistic problem formulation. In particular, the underlying assumption in these works is that the values of the opportunities available to the searcher are drawn from a common distribution and that the cost of obtaining the true value of an opportunity is similar. While these tractable assumptions simplify the analysis and the experimental design, they over-constrain the problem. Taking the product-search problem domain as an example, in real-life different sellers are often characterized with different distributions of prices and costs of exploration. For example, in discount chains and outlets, prices are generally lower; however, the cost of revealing them is usually substantial (e.g., cost of time and transportation to visit them). Even in online markets, the cost of obtaining price information (mostly in terms of time spent) varies between sellers as it highly depends on the web-site design, its supported search capabilities, the product classifications used, ease of navigation, etc. When using the simplified exploration assumptions, there is no importance to the order by which opportunities are queried, and the exploration rule becomes simply a stopping rule. This enables people to produce simple strategies, based on rules of thumb of the type usually used in real life. Naturally, as the strategy space becomes larger, as in the underlying model used in this paper, the difference between the strategies people use or program into PDAs becomes more visible.

The analysis given in this section goes beyond the comparison of the aggregated measures for reasoning about the optimality in PDAs and people's decisions. Instead, it compares their decisions with those of the optimal agent in the same decision situations. For this purpose we define a decision situation as a game snapshot that captures the set of values already revealed to the decision maker and the set of opportunities which values are unknown (i.e., only their distribution of values and cost of revealing the true value are known). While the optimal strategy is not affected by the set of values obtained but rather relies solely on the best value obtained so far (and the reservation values of the remaining opportunities), people and PDAs may base their strategies either partially or fully on the set of values obtained along the exploration; hence the inclusion of the full set in the decision situation description is mandatory for our purposes.

In order to determine the level of optimality of one's decisions, we use the average similarity between the decisions made by that individual and the optimal ones in the same decision situations. Therefore, for each decision situation faced by a person or a PDA we extract the optimal decision (terminate the exploration or continue to a specific opportunity which has an unknown value) and compare it with the PDA's/person's decision. The optimality similarity measure for a decision situation o, denoted $d_{sim(o)}$, is defined as binary variable which receives 1 if both the decision maker and the optimal agent picked the same alternative (i.e., both terminated or both chose to explore the same opportunity) and otherwise 0. In order to determine the similarity between a PDA/person and the optimal agent, an average over the values of $d_{sim(o)}$ over the set of decision situations O was used: $\frac{\sum_{o \in O} d_{sim(o)}}{|O|}$. Therefore the average optimality of a decision maker equals the percentage of decision situations in which the decision made was identical to the one made by the optimal agent.

Since the experiments so far were mechanism-oriented, and the evaluation of optimality we report in this section is general, we generated a new set of 100 problem instances according to the same guidelines reported in Section 4.4, except that this time the values of all opportunities were a-priori unknown. These problems were used with 22 new people recruited through Amazon Mechanical Turk (each given 20 random problems from the set). A larger set of 5000 problems (generated in a similar manner) was used for the PDAs. This experimental design choice was made to improve statistical significance; since now decisions were compared to the optimal agent's decisions, to test optimality, we were no longer limited to the set of problems used by people and we could benefit from the ability of PDAs to solve any number of exploration problems. All 31 PDAs were used in these experiments. The optimality similarity of PDAs and people in each decision situation in any of the problems encountered was recorded and averaged according to the above guidelines. Overall there were $484,827$ decision situations confronted by the different PDAs and $1,596$ faced by people.

Figures 6 and 7 depict the average optimality level in people's and PDAs' decisions for each of the 31 PDAs and 22 people who participated in the experiment. As can be observed from the figures, both PDAs and peoples' decisions were far from optimal. Still, PDAs had a greater percentage of optimal decisions than people. The average percentage of decision situations where the optimal decision was made, per agent, was 41.2% (equivalent to 40.7% of all decisions made by PDAs), whereas the average, per-person was 22.3% (equivalent to 22% of all decisions made by people). Furthermore, from the figures we can conclude that there were no exceptionally better or exceptionally worse PDAs or people in terms of the exhibited optimality in their decisions. The best PDA, in terms of decision optimality, achieved an average similarity of 70.4% and the worst achieved 20%.[8] More than 77% of the PDAs had an average of $25\% - 50\%$. Similarly, the best person, in terms of decision optimality, achieved an average similarity of 34.5% and the worst achieved 13.6%. More than 73% of the human subjects had an average of $13\% - 25\%$.

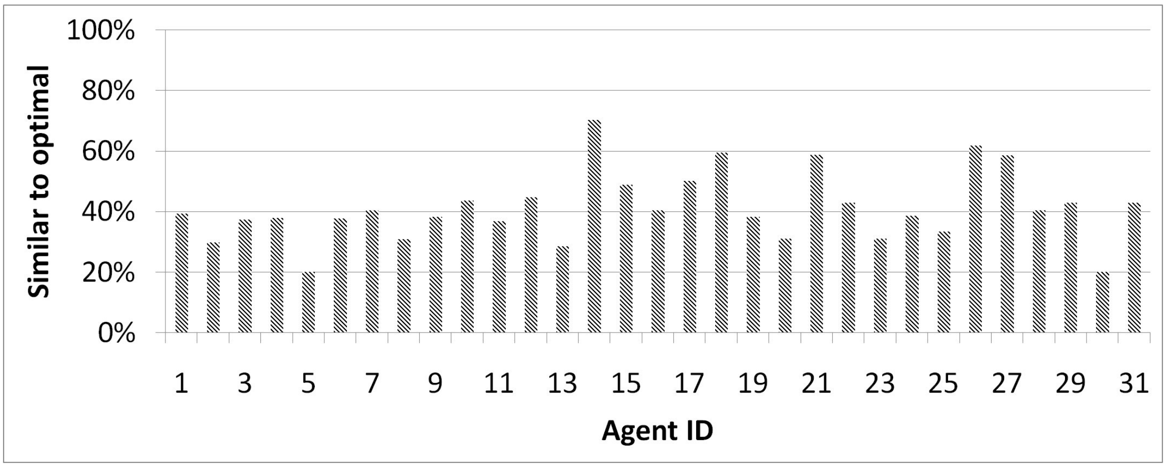


Fig. 6. Average similarity between PDAs and the optimal agent.

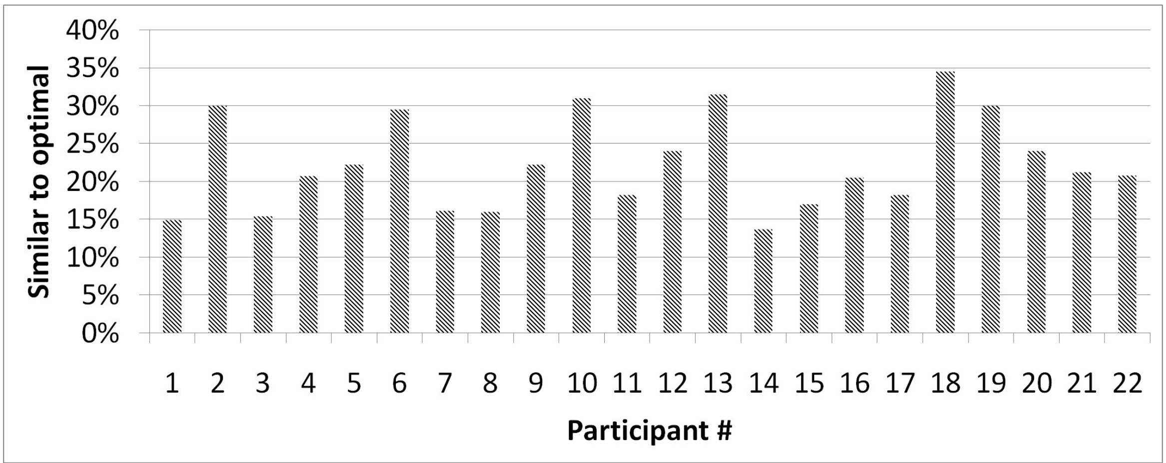
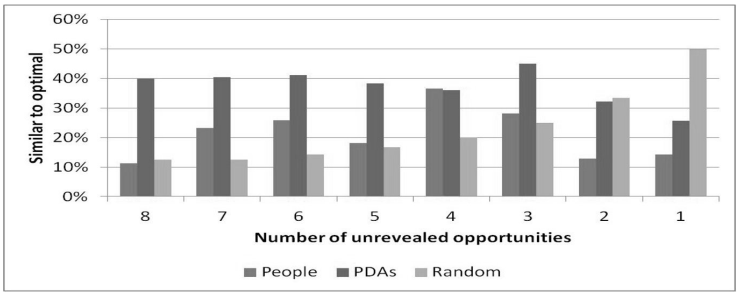


Fig. 7. Average similarity between people and the optimal agent.

Figure 8 depicts a break-down of the optimality similarity measure according to the number of opportunities for which the value has not been revealed yet to the decision maker (the horizontal axis), comparing PDAs, people and a random decision rule. The result for the random decision rule is simply the inverse of the number of opportunities for which the value has not been revealed yet plus one (which is the option of terminating the exploration). Therefore, the greater the number of "unrevealed" opportunities the lower the optimality of the random decision rule. The table to the right of the graph details the number of relevant decision situations the data applies to (for each number of opportunities available), where the number in parenthesis is the number of the unique participants to whom the latter value applies. The last column in the table gives the chi-test p-value for the similarity between people and PDAs in this case. Several important observations can be made based on the figure. First, it is striking that when facing

[8]Indeed the difference between the "best" and "worst" PDAs is substantial, however none of them was "exceptionally better" than the rest of the population, as there were other agents with close performance.

eight unknown opportunities, PDAs perform four times better than people (40% compared to 10%) in terms of decision optimality. The uniqueness of the eight choices case is that no termination is allowed (as no value has been revealed yet) hence the decision maker is forced to pick any of the eight opportunities. When facing seven unknown opportunities, the decision maker once again has eight choices (pick any of the seven opportunities or terminate the exploration); however in the latter case people's optimality level rises to 23.3%. These last two results indicate that people are very good at identifying the right time to terminate exploration, however very bad at picking the right opportunity to explore. This conclusion is supported by Figures 9 and 10 that depict a breakdown of the similarity results given in Figure 8 to the case where the optimal strategy is to terminate exploration (Figure 9) and when it is optimal to resume exploration (Figure 10). The case of having eight unrevealed opportunities is irrelevant in these figures, as terminating the exploration at this point is not applicable. As can be observed from the figures, people's ability to correctly choose between terminating and resuming the exploration is quite impressive (more than 60% in all situations where the optimal strategy is to resume exploration (Figure 9) and in situations where the number of unrevealed opportunities is relatively large (6 and 7) and the optimal strategy is to terminate exploration (Figure 10)). As the number of unrevealed opportunities decreases, people's correct identification of the benefit in termination drops (Figures 9) and the percent of correct identification of the benefit in resuming exploration increases (Figure 10). This probably accounts for an increased tendency of those that got that far in their exploration to further resume exploration.[9] The difficulty of people is thus to choose the appropriate opportunity to explore, when resuming exploration, and this indeed accounts for their poor performance in terms of similarity to optimal decisions. This is illustrated in Figure 11 which depicts the similarity measure based only on those decision situations in which both the decision maker and the optimal agent chose to resume exploration. In this figure, we observe a remarkably low similarity between people's decisions and the optimal strategy in these situations, suggesting that even when a person manages to identify the benefit in further exploration, she is most likely to fail in identifying the appropriate opportunity that needs to be explored.[10]



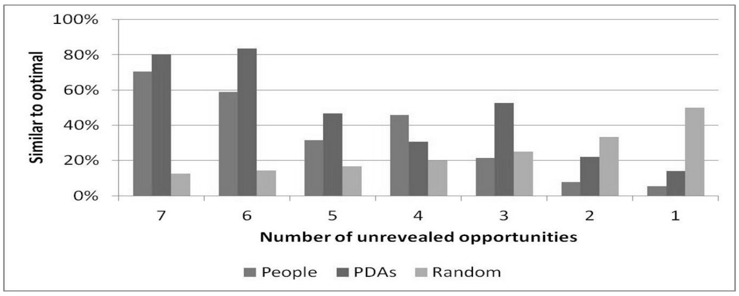
#	People(unique)	PDAs(unique)	chi-test
8	460(22)	155000(31)	<0.001
7	451(22)	155000(31)	<0.001
6	248(20)	82054(30)	<0.001
5	149(19)	28718(17)	<0.001
4	104(15)	21366(16)	0.92
3	57(10)	17444(14)	0.01
2	47(6)	9892(12)	<0.001
1	42(5)	8189(12)	0.09

Fig. 8. Level of similarity to optimal strategy, as a function of the number of opportunities for which the value has not been revealed yet, for people, PDAs and a random decision rule.

As far as PDAs are concerned, the optimality level exhibited in their decision remained steady (give and take), and significantly greater than people's, as the number of unrevealed opportunities available in the decision situation changes (see Figure 8). From Figures 9 we can observe an impressive ability of PDAs to correctly identify situations where it is optimal to terminate the exploration (more than 80%) when having 6 or 7 opportunities with unknown values and then a gradual decrease in the optimality similarity measure. This decrease is explained by the low number of unique PDAs that account for the results with less than 6 available opportunities (12-17 of the 31), providing strong evidence for the existence of a sub-group with a tendency to over-explore, which substantially influences the aggregated result. In terms of correctly identifying situations where exploration should be resumed (Figure 10), PDAs reflect impressive capabilities with no general dominance over people (depending on the number of available opportunities, the PDAs can either be significantly better than, significantly worse than or insignificantly different

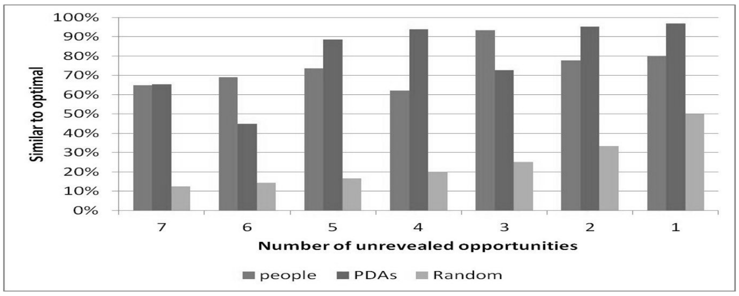
[9]This latter conclusion is supported by the sharp drop in the number of unique participants that account for these results, suggesting that indeed there is a small group of participants with a tendency towards "over exploration".

[10]The increased similarity achieved for people in situations with 1-3 unrevealed opportunities is insignificant due to the substantially low number of decision situations (and in particular the number of unique participants) on which it is based.



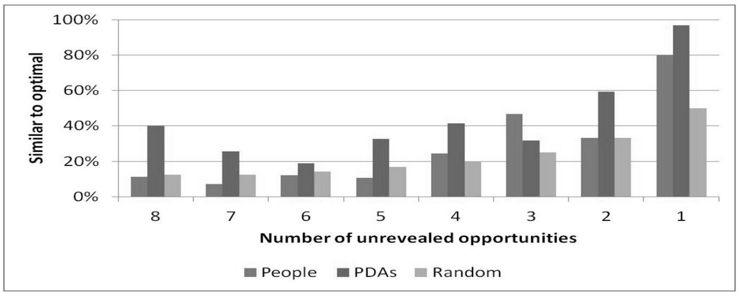
#	People(unique)	PDAs(unique)	chi-test
7	115(21)	42441(31)	<0.001
6	73(18)	28282(30)	<0.001
5	54(15)	11553(17)	0.02
4	59(14)	10677(16)	0.01
3	42(8)	11056(14)	<0.001
2	38(6)	7183(12)	0.03
1	37(5)	7042(12)	0.13

Fig. 9. Level of similarity to optimal strategy, as a function of the number of opportunities for which the value has not been revealed yet, for people, PDAs and a random decision rule, in decision situations where the optimal decision is to terminate exploration.



#	People(unique)	PDAs(unique)	chi-test
7	336(22)	112559(31)	0.82
6	175(20)	53772(30)	<0.001
5	95(19)	17165(17)	<0.001
4	45(13)	10689(16)	<0.001
3	15(7)	6388(14)	0.07
2	9(4)	2709(11)	0.01
1	5(3)	1147(11)	0.03

Fig. 10. The percentage of resume exploration decisions, as a function of the number of opportunities for which the value has not been revealed yet, for people, PDAs and a random decision rule, in decision situations where the optimal decision is to resume exploration.



#	People(unique)	PDAs(unique)	chi-test
8	460(22)	155000(31)	<0.001
7	336(22)	112559(31)	<0.001
6	175(20)	53772(30)	0.01
5	95(19)	17165(17)	<0.001
4	45(13)	10689(16)	0.02
3	15(7)	6388(14)	0.21
2	9(4)	2709(11)	0.11
1	5(3)	1147(11)	0.03

Fig. 11. Level of similarity to optimal strategy, as a function of the number of opportunities for which the value has not been revealed yet, for people, PDAs and a random decision rule, in decision situations where both the optimal decision and the decision-maker's choice were to resume exploration (i.e., measuring the ability to choose the appropriate option).

from people). Finally, the ability of PDAs to optimally choose the opportunity to explore in cases where further exploration is indeed the optimal choice (Figure 11) sharply decreases (till the case of six unrevealed opportunities) and then increases (except for the case of three opportunities). The decrease is very surprising as one would expect that a decrease in the number of choices available to the decision maker would enable a more optimal decision. This is indeed the case when there are less than six unexplored opportunities. A possible explanation for the decrease in decision optimality with six and more opportunities is that when all opportunities are available, the opportunities that are substantially better than the others are easily recognized, whereas once they are removed (explored) it is more difficult to identify the more beneficial ones among those remaining. This effect decreases as the decision maker advances in its exploration process and the differences between the benefit of choosing any of the remaining opportunities become less apparent.

The above detailed analysis, complemented by Figure 12, can explain the aggregated results given in Figure 8. Figure 12 depicts the division of the total set of decision situations into situations where the optimal strategy is to terminate exploration and where it is to resume exploration. Naturally, as the number of remaining unrevealed opportunities decreases, it is more likely that the optimal decision is to terminate exploration, as observed in the figure. The fact that the same pattern (in terms of the division between optimal terminate and resume decisions) is observed for both sets of problems (the 100 problems that were used with people and the 5000 problems used with PDAs) is not surprising as both sets were generated using similar parameters and hence represent the same population of problems. In general Figure 8 demonstrates a relatively stable level of optimality in PDAs' decisions, despite the decrease in the number of available unrevealed opportunities. The explanation for this phenomenon is as follows. For 8 opportunities, the optimality rate is the same as depicted in Figure 11, since terminating the exploration at this stage is not feasible, thus the only question is whether or not the PDA manages to pick the right alternative among the eight choices available. Then, with 6 and 7 alternatives, the PDAs' optimality when resuming exploration drops (Figure 9). Nonetheless, this is compensated by their relatively impressive ability to terminate, whenever this choice is optimal (Figure 9). As the number of available opportunities decreases below 6, the PDAs' ability to correctly identify the benefit in terminating the exploration drops to a lower level; however this is compensated both by better performance in cases where the exploration should be resumed (Figure 11) and an increase in the dominance of the latter situations within the overall number of decision situations (Figure 12(a)). A similar explanation can be provided for people. For 8 opportunities, the optimality rate is the same as depicted in Figure 11 as explained above. The general overall increase in people's optimality level as demonstrated in Figure 8 is explained by their relative strength in identifying optimal termination situations (Figure 9) and an increase in the weight of decisions of the latter type (as the number of unrevealed alternatives decreases) in the overall decision situations (Figure 12(b)).

Overall, we can conclude from the analysis given in this section that both PDAs and people's choices are far from optimal, and that the difference between the optimality level reflected in these two groups' decisions is significant (except for a very small number of exceptions, as shown by the $p - values$ given next to each figure). The difference between people's and PDAs' ability to correctly identify situations in which exploration should be resumed is relatively small. Nevertheless, PDAs are better than people in identifying situations where it is optimal to terminate exploration and much better than people in identifying the appropriate opportunity to explore when further exploration is favorable. Surprisingly, for some of the cases, the level of optimality reflected in people's and PDAs' decisions is lower than the one achieved when randomly picking one of the alternatives available (see Figure 9, in the case of a small number of options).

7. Discussion and Conclusions

The results reported in this paper suggest that indeed the determination of whether or not PDAs can be used as a substitute for people for mechanism evaluation is mechanism and measure dependent. These findings are very different from claims made in prior work on the usefulness of the PDA-based approach and results reported concerning the similarity of people and PDAs in specific domains. Our results demonstrate the risks in generalizing, based on the "average behavior" observed by applying a specific mechanism and using a specific set of measures for comparison. Thus, the use of PDAs should be carefully handled, and the similarity between the behavior of people and the PDAs used should be verified for every new mechanism that needs to be evaluated using the exact same measures of interest. This verification process can be either based on a pilot study or other means aiming to compare the accuracy of capturing one's strategy through the PDA.

Specifically, in our case, had the seller tested the three methods with PDAs, the preferred pricing method would have been Mean-Fonders Pricing (M2); whereas, if the buyers are people then the preferred pricing method would have been Cost-Probability Tradeoff (M3). Moreover, even if the seller had decided to run a pilot test with one of the methods to evaluate the usefulness of using PDAs as a substitute for people to solve its price-setting problem, she might have reached the wrong conclusions. For example, if the Cost-Probability Tradeoff (M3) pricing mechanism was chosen for the pilot study and the expected revenue measure was used, the seller would have assumed that PDAs could be used also for the evaluation of the other two methods and the wrong choice would have been chosen. Even if the seller is given data on the extent of exploration and the average overall expense of PDAs in comparison to people (which are the two main measures that capture exploration), it should not assume that the use of PDAs with the same mechanisms would yield an expected revenue similar to the one obtained with human buyers.

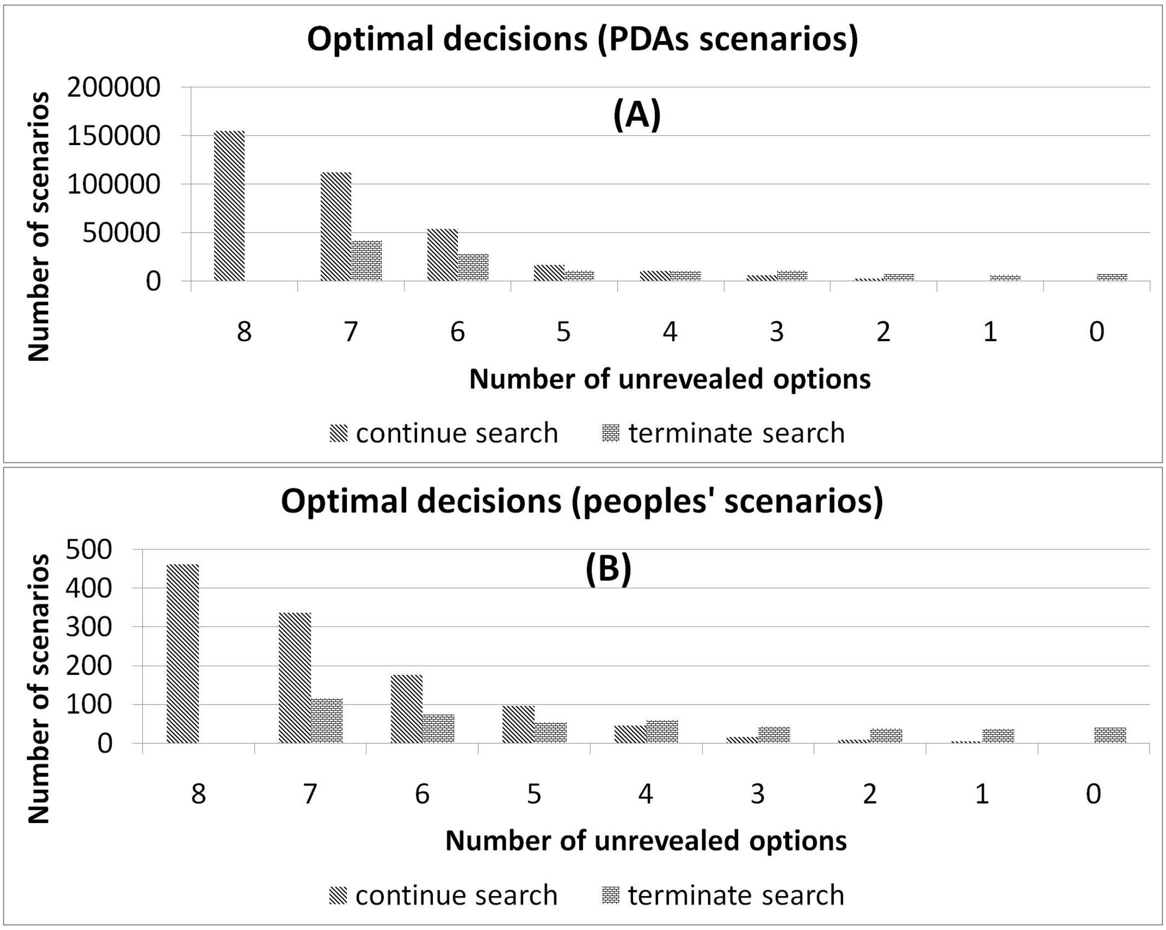


Fig. 12. Number of terminate-exploration and continue-exploration decisions of an optimal decision-maker: (a) over PDA's problem set; (b) over people's problem set.

Another interesting finding is that from the seller's point of view the dominant pricing method among the three methods evaluated, as reflected in Figure 4, is buyer-type-dependent (the Cost-Probability Tradeoff (M3) method is preferred with human buyers, the Mean-Fonders Pricing (M2) is preferred with PDAs and the Theoretic-Optimal Pricing (M1) is preferred with theoretic-optimal buyers). While this matter is not the essence of this paper, this result is important due to the nature of eCommerce. Since sellers are likely to encounter a mixture of human and agent buyers, if the buyer type can be identified (e.g., using Captchas) a finer grained selection of the price setting mechanism can be used. We emphasize that the paper does not attempt to argue the applicability of the pricing method, nor their optimality or validness. These were developed using common sense and used only for the purpose of testing the (PDA-related) research hypothesis. As discussed earlier in the paper, any set of mechanisms that act consistently would have been appropriate for our purposes.

The results of the additional experiment that compares the performance of computer science students themselves and people in general in exploration-based environments suggest two additional contributions. First, they negate a possible difference between the two populations (people in general and computer science students) as a possible explanation for the difference observed between PDAs and people in the problem domain used. This strengthens the assumption that in many cases there are substantial differences between PDAs and people, even if the PDAs programmers are of the same population in which observations are made. Second, it strengthens the validness of using Mechanical Turk as part of our evaluation methodology.

The results concerning the level of optimality observed in people's and PDAs' decisions in this important domain of costly exploration reveals several additional insights. None of the PDAs reached a substantial optimality level in its decisions (though compared to people, the PDAs' level of optimality is quite impressive). This low similarity level, suggests that even when PDAs are programmed using a substantially different strategy than the one used by people, they are far from optimal and their designer fails to make use of their relative computational and storage strengths to achieve a near-optimal behavior. This implies that even in real-life settings where agents are key players (e.g., in future and in some current eCommerce markets) solutions that assume theoretic-optimal exploration strategy of agents (e.g., those that are commonly found in agent-based pricing theory literature [24]), are likely to fail. In such settings, it is therefore important to base the evaluation of any developed mechanism aiming to enhance market performance on a large set of diverse agents, developed by different individuals and groups. In this sense, the work on TAC [63] and the development of trading agents within this framework is highly favorable. As for people's level of optimality, indeed prior work has investigated this aspect, though, as discussed in the previous section, the analysis suggested there relies on a comparison of the average performance of people to the one achieved using the optimal strategy according to different measures. The analysis of this aspect in this paper relies on a comparison at the decision situation level, hence it directly touches on the optimality issue. The results reveal a substantial low level of optimality in people's decisions in this domain.

Finally, we note that despite the inability to generalize the applicability of PDAs for mechanism evaluation, this technology is extremely useful in settings where they are found suitable as substitutes for people. The use of PDAs in such cases can save substantial resources and facilitate the evaluation of numerous potential configurations, in a relatively short time, without having to recruit people over and over again for expensive experiments. In particular, this technology is useful for simulating and researching large-scale systems due to the relatively low cost of cloning agents. Therefore, an important direction for future research is the development of methodologies for facilitating an effective and efficient ad hoc evaluation of PDAs' suitability to substitute people in simulating a given domain. Similarly, it would be interesting to propose insights on the class of mechanisms for which PDAs could be "safely" used, and those for which humans are more adequate, as well as means for facilitating the design of PDAs that abstract as much as possible from the details of the specific systems.

Acknowledgments

Preliminary results of this research appeared in the proceedings of the 2012 IEEE/WIC/ACM International Conference on Intelligent Agent Technology.

References

[1] AMT. *http://www.mturk.com/*, 2010.
[2] M. Baye, J. Morgan, P. Scholten, and D. Jansen. Persistent price dispersion in online markets. *The New Economy and Beyond: Past, Present and Future (DW Jansen, ed.). New York: Edward Elgar Publishing*, pages 122–143, 2006.
[3] F. Bellifemine, A. Poggi, and G. Rimassa. Jade: a fipa2000 compliant agent development environment. In *Proceedings of the fifth international conference on Autonomous agents*, pages 216–217. ACM, 2001.
[4] M. Bertrand and S. Mullainathan. Do people mean what they say? implications for subjective survey data. *American Economic Review*, 91(2):67–72, 2001.
[5] A. Bosch and J. Silvestre. Does risk aversion or attraction depend on income? an experiment. *Economics Letters*, 65(3):265–273, 1999.
[6] A. Bosch and J. Silvestre. Risk aversion and embedding bias. *Materia (s)*, 1:01–2006, 2006.
[7] T. Bosse, C. Gerritsen, M. Hoogendoorn, S. Jaffry, and J. Treur. Agent-based vs. population-based simulation of displacement of crime: A comparative study. *Web Intelligence and Agent Systems*, 9(2):147–160, 2011.
[8] J. Brandts and G. Charness. Truth or consequences: An experiment. *Management Science*, 49(1):116–130, 2003.
[9] P. Brewer, M. Huang, B. Nelson, and C. Plott. On the behavioral foundations of the law of supply and demand: Human convergence and robot randomness. *Experimental economics*, 5(3):179–208, 2002.
[10] M. Casari and T. Cason. The strategy method lowers measured trustworthy behavior. *Economics Letters*, 103(3):157–159, 2009.
[11] T. Cason and V. Mui. Social influence in the sequential dictator game. *Journal of Mathematical Psychology*, 42(2):248–265, 1998.
[12] M. Chalamish, D. Sarne, and R. Lin. Enhancing parking simulations using peer-designed agents. *IEEE Transactions on Intelligent Transportation Systems*, 14(1):1–7, 2012.

[13] A. Chavez and P. Maes. Kasbah: An agent marketplace for buying and selling goods. In *Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology*, pages 8–12, 1996.

[14] K. Cheng, I. Zuckerman, D. Nau, and J. Golbeck. The life game: Cognitive strategies for repeated stochastic games. In *Proceedings of the third international conference on social computing (socialcom)*, pages 95–102. IEEE, 2011.

[15] K. Clay, R. Krishnan, E. Wolff, and D. Fernandes. Retail strategies on the web: Price and non–price competition in the online book industry. *The Journal of Industrial Economics*, 50(3):351–367, 2002.

[16] S. Gal, M. Landsberger, and B. Levykson. A compound strategy for search in the labor market. *International Economic Review*, 22(3):597–608, 1981.

[17] D. Gode and S. Sunder. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of political economy*, 101:119–137, 1993.

[18] B. Grosz, S. Kraus, S. Talman, B. Stossel, and M. Havlin. The influence of social dependencies on decision-making: Initial investigations with a new game. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 782–789. IEEE Computer Society, 2004.

[19] C. Harries, J. Evans, and I. Dennis. Measuring doctors' self-insight into their treatment decisions. *Applied Cognitive Psychology*, 14:455–477, 2000.

[20] N. Hazon, Y Aumann, and S. Kraus. Collaborative multi agent physical search with probabilistic knowledge. In *Proceedings of the 21st international joint conference on Artifical intelligence (IJCAI'09)*, pages 167–174, 2009.

[21] J. Hey. Still searching. *Journal of Economic Behavior and Organization*, 8(1):137 – 144, 1987.

[22] A. Hodrick and R. Xing. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.

[23] M. Janssen and J. Moraga-Gonzalez. Strategic pricing, consumer search and the number of firms. *Review of Economic Studies*, 71(4):1089–1118, 2004.

[24] M. Janssen, J. Moraga-Gonzalez, and M. Wildenbeest. Truly costly sequential search and oligopolistic pricing. *International Journal of Industrial Organization*, 23(5-6):451–466, 2005.

[25] P. Jordan, C. Kiekintveld, and M. Wellman. Empirical game-theoretic analysis of the tac supply chain game. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 193:1–193:8. ACM, 2007.

[26] J. Jumadinova and P. Dasgupta. Firefly-inspired synchronization for improved dynamic pricing in online markets. In *Proceedings of Self-Adaptive and Self-Organizing Systems (SASO' 08)*, pages 403–412. IEEE Computer Society, 2008.

[27] G. Kempen, M. Van Heuvelen, R. Van den Brink, A. Kooijman, M. Klein, P. Houx, and J. Ormel. Factors affecting contrasting results between self-reported and performance-based levels of physical limitations. *Age and Ageing*, 25(6):458–464, 1996.

[28] J. Kephart and A. Greenwald. Shopbot economics. *Game theory and decision theory in agent-based systems*, 5(3):119–158, 2002.

[29] J. Kephart, J. Hanson, and A. Greenwald. Dynamic pricing by software agents. *Computer Networks*, 32(6):731 – 752, 2000.

[30] R. Lin, S. Kraus, N. Agmon, S. Barrett, and P. Stone. Comparing agents success against people in security domains. In *Proceedings of the 25th Association for the Advancement of Artificial Intelligence (AAAI-2011)*, 2011.

[31] R. Lin, S. Kraus, Y. Oshrat, and Y. Gal. Facilitating the evaluation of automated negotiators using peer designed agents. In *Proceedings of the 24th Association for the Advancement of Artificial Intelligence (AAAI-2010)*, pages 817–822, 2010.

[32] S. Lippman and S. Ross. Variability is beneficial in marked stopping problems. *Economic Theory*, 35(2):333–342, 2008.

[33] P. Maes. *Designing Autonomous Agents*, pages 105–122. MIT Press, Cambridge, MA, 1990.

[34] D. Massaguer, V. Balasubramanian, S. Mehrotra, and N. Venkatasubramanian. Multi-agent simulation of disaster response. In *Proceedings of First International Workshop on Agent Technology for Disaster Management*, pages 124–130, 2006.

[35] J. McMillan and M. Rothschild. Search. *Proceedings of Handbook of Game Theory with Economic Applications*, 2:905–927, 1994.

[36] N. Minar, R. Burkhart, C. Langton, and M. Askenazi. The swarm simulation system: A toolkit for building multi-agent simulations. Working Papers 96-06-042, Santa Fe Institute, 1996.

[37] P. Moon and A. Martin. Better heuristics for economic search, experimental and simulation evidence. *Behavioral Decision Making*, 3(3):175–193, 1990.

[38] Hans-Joachim Mosler. Modelling social behavior with a socio psychological simulation approach. *Web Intelligence and Agent Systems*, 2(3):185–200, 2004.

[39] R. Murphy, A. Rapoport, and J. Parco. Credible signaling in real-time trust dilemmas. *University of Arizona, mimeo*, 2007.

[40] S. Musse and D. Thalmann. Hierarchical model for real time simulation of virtual human crowds. *IEEE Transactions on Visualization and Computer Graphics*, 7(2):152–164, 2001.

[41] M. North, E. Tatara, N. Collier, and J. Ozik. Visual agent-based model development with repast simphony. In *Proceedings of the Agent Conference on Complex Interaction and Social Emergence*, pages 173–192, 2007.

[42] G. Paolacci, J. Chandler, and P. Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419, 2010.

[43] I. Rochlin, D. Sarne, and G. Zussman. Sequential multi-agent exploration for a common goal. *To Appear in Web Intelligence and Agent Systems*, 2013.

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

[45] D. Sarne, A. Elmalech, B. Grosz, and M. Geva. Less is more: Restructuring decisions to improve agent search. In *Proceedings of the Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pages 431–438, 2011.

[46] A. Schotter and Y. Braunstein. Economic search: an experimental study. *Economic Inquiry*, 19(1):1–25, 1981.

[47] D. Schunk and J. Winter. The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment. *Economic*

Behavior Organization, 71(2):347 – 360, 2009.
[48] L. Schvartzman and M. Wellman. Learning improved entertainment trading strategies for the tac travel game. *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*, 59:195–210, 2010.
[49] R. Selten, M. Mitzkewitz, and G. Uhlich. Duopoly strategies programmed by experienced players. *Econometrica*, 65(3):517–555, 1997.
[50] W. Shao and D. Terzopoulos. Autonomous pedestrians. *Graphical Models*, 69(5–6):246–274, 2007.
[51] I. Sharpanskykh and J. Treur. Abstraction relations between internal and behavioural agent models for collective decision making. *Web Intelligence and Agent Systems*, 10(4):465–484, 2012.
[52] V. Smith. An experimental study of competitive market behavior. *Journal of Political Economy*, 70:111–137, 1962.
[53] J. Sonnemans. Strategies of search. *Journal of Economic Behavior and Organization*, 35(3):309 – 332, 1998.
[54] J. Sonnemans. Decisions and strategies in a sequential search experiment. *Journal of Economic Psychology*, 21(1):91–102, 2000.
[55] S. Tadokoro, H. Kitano, T. Takahashi, I. Noda, H. Matsubara, A. Shinjoh, T. Koto, I. Takeuchi, H. Takahashi, F. Matsuno, M. Hatayama, J. Nobe, and S. Shimada. The robocup-rescue project: A robotic approach to the disaster mitigation problem. *IEEE International Conference on Robotics and Automation*, 4:4089–4094, 2000.
[56] T. Takahashi, S. Tadokoro, M. Ohta, and N. Ito. Agent based approach in disaster rescue simulation - from test-bed of multiagent system to practical application. In *Proceedings of RoboCup 2001: Robot Soccer World Cup V*, pages 102–111, 2002.
[57] Z. Tang, M. Smith, and A. Montgomery. The impact of shopbot use on prices and price dispersion: Evidence from online book retailing. *International Journal of Industrial Organization*, 28(6):579–590, November 2010.
[58] D. Terzopoulos, X. Tu, and R. Grzeszczuk. Artificial fishes: Autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life*, 1(4):327–351, 1994.
[59] B. Ulicny and D. Thalmann. Towards interactive real-time crowd behavior simulation. *Computer Graphics Forum*, 21(4):767–775, 2002.
[60] M. Van Boening and N. Wilcox. Avoidable cost: Ride a double auction roller coaster. *The American Economic Review*, 86:461–477, 1996.
[61] Y. Wan and G. Peng. What's next for shopbots? *IEEE Computer*, 43:20–26, 2010.
[62] M. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–54, 1979.
[63] M. Wellman, P. Wurman, K. O'Malley, R. Bangera, S. Lin, D. Reeves, and W. Walsh. Designing the market game for a trading agent competition. *IEEE Internet Computing*, 5(2):43–51, 2001.
[64] Y. Zhang, K. Biggers, L. He, S. Reddy, D. Sepulvado, J. Yen, and T. Ioerger. A distributed intelligent agent architecture for simulating aggregate-level behavior and interactions on the battlefield. In *Proceedings of the 5th World Multi-Conference on Systemics, Cybernetics, and Informatics (SCI-01), Orlando, Florida*, pages 58–63, 2001.