2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology

Join Me with the Weakest Partner, Please

Moshe Mash, Igor Rochlin and David Sarne
Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel

***Abstract*—This paper considers the problem of self-interested agents engaged in costly exploration when individual findings benefit all agents. The purpose of the exploration is to reason about the nature and value of the different opportunities available to the agents whenever such information is a priori unknown. While the problem has been considered for the case where the goal is to maximize the overall expected benefit, the focus of this paper is on settings where the agents are self-interested, i,e, each agent's goal is to maximize its individual expected benefit. The paper presents an equilibrium analysis of the model, considering both mixed and pure equilibria. The analysis is used to demonstrate two somehow non-intuitive properties of the equilibrium cooperative exploration strategies used by the agents and their resulting expected payoffs: (a) when using mixed equilibrium strategies, the agents might lose due to having more potential opportunities available for them in their environment; and (b) if the agents can have additional agents join them in the exploration they might prefer the less competent ones to join the process.**

***Keywords*-Multi-Agent Exploration, Cooperation**

I. Introduction

In many multi-agent settings, agents need to individually engage in exploration of the opportunities available to them [6]. The purpose of the exploration is to reason about the nature and value of the different opportunities whenever such information is a priori unknown [12], [7], [13]. Since such exploration is inherently costly (either involves direct monetary costs or the consumption of some of the agent's resources) the goal of the agents is not necessarily to find the opportunity associated with the maximum value, but rather to maximize the overall benefit, defined as the value of the opportunity eventually picked minus the costs accumulated during the exploration process. As an example, consider the following scenario: Jill wants to buy some magic beans. While Jill knows which stores sell magic beans in the nearest shopping mall, she does not know a priori the price in each store. By visiting the different stores, Jill can find the lowest posted price, however since Jill's time is valuable (e.g., she might as well use the time for other errands) her goal is not necessarily to find the lowest price (i.e., visit all stores) but rather to manage the visits in the different stores such that the sum of the overall exploration cost-equivalent and the price at which the beans are eventually purchased is minimized.

In many cases the results of exploration carried out by one agent can be of benefit to other agents. In our example, Jill can offer her friend Jack, who is also interested in buying magic beans, to join her at the shopping mall and execute the exploration process cooperatively. This way each of them will be individually visiting a different subset of stores, and eventually they will share their findings and buy at the store associated with the minimum price among those visited. While many kinds of cooperative exploration strategies have been suggested in prior work, the assumption they all act on assumption that the goal is to maximize the joint expected benefit or minimize the joint expected cost (i.e., the sum of the value with which the agents end up minus the costs all agents incur along the exploration process). Still, in many settings agents are self-interested and attempt to maximize their individual expected benefit rather than the joint benefit. For example, Jill may find it more beneficial to visit just one store that sells magic beans and then spend her time doing her other shopping. When she meets Jack later, Jill will benefit from the information he obtained through his exploration, however her own spending on exploration will be minimal.

In this paper we supply a cohesive analysis of a model of cooperative exploration with self-interested agents. The analysis derives from the individual expected-benefit-maximizing exploration strategies of the different agents, given the exploration strategies of the others. These facilitate an equilibrium analysis, proving that each agent first determines whether it will engage in exploration at all, and if so it necessarily uses a threshold-based exploration strategy, terminating the exploration only if the value obtained is below (or, depending on the application, above) the threshold. The analysis considers both mixed and pure equilibria, introducing the sets of equations that need to be solved for extracting agents strategies and the conditions that need to be checked for validating the stability of these solutions.

The analysis is used for demonstrating the effect of the different model parameters, in particular the exploration competence of the different agents (i.e., their cost of exploration) and the number of agents that explore cooperatively, over the expected individual and total benefit. One somehow non-intuitive result that is demonstrated is that an increase in the allowed exploration horizon (e.g., the number of potential stores at the shopping mall that sell magic beans) results in a decrease in the individual expected benefit of the agents when mixed equilibrium is used. Another somehow-surprising result that is illustrated in the paper is that if the agents are allowed to add an additional agent to their cooperative exploration, then sometimes the less competent agent (in terms of exploration cost) is the preferable candidate. In the magic beans example, if Jack and Jill have an additional empty seat in their car and need to decide between two of their friends that want to join them (both interested in magic beans as well), then the friend that does not know the structure of the shopping mall might be preferred over the one who knows all the shortcuts and can

978-0-7695-4880-7/12 $26.00 © 2012 IEEE
DOI 10.1109/WI-IAT.2012.155

get to each store faster. These results are in contrast to the case where the agents are fully cooperative and attempt to maximize the overall expected benefit. The fact that the agents are self-interested suggests that many potential solutions, in the form of exploration strategies that are beneficial from the individual and overall expected benefit point of view, become unstable, and the resulting stable solutions are such that the agents find it less beneficial to explore substantially.

In the following section we formally present the model of cooperative exploration with self-interested agents. We present the model analysis in Section III. Section IV illustrates the equilibrium dynamics and resulting performance under different settings. Related work is surveyed in Section V. Finally, we conclude and discuss directions for future research.

II. The Model

We consider a setting where k self-interested individual agents need to engage in exploration of possible available opportunities. Each agent may see a different value in each opportunity, and there is no limit to the number of agents that can benefit from it (e.g., an opportunity can represent a price quote of a seller and any demand can be satisfied). The value of each opportunity is a priori unknown and the agents are only acquainted with the probability distribution function from which the values of the different opportunities available to each agent A_i are derived, denoted $f_i(x)$. In order to obtain the true value of an opportunity (i.e., to "query" the opportunity), an agent needs to consume some of its resources. This is modeled by a cost c_i (agent dependent), expressed in terms of opportunity values. Therefore, the agents are required to conduct an exploration process which takes into consideration the tradeoff between the marginal improvement in the highest value found and the cost incurred along the process.

Since each opportunity is applicable to all agents, the agents have an incentive to explore as a team. This way, each opportunity will be evaluated by only one of the agents and then shared with all others. This form of exploration necessarily dominates individual explorations, since any strategy used by the different agents when exploring individually can be adopted by the team, resulting in the same cost of exploration however with the best value found shared by all agents. The benefit of an agent at the end of the process is therefore the best value obtained by the group minus the costs accumulated individually along the agent's exploration process. The model assumes that the agents are constrained by an exploration horizon of n time periods and that in each exploration period each agent can query at most one opportunity (while incurring the appropriate cost), i.e., the agents can cooperatively query kn opportunities at most along their cooperative exploration.

While there are many communication schemes that can be considered for the model, in this paper we assume no communication between the agents along their cooperative exploration. Once all agents have individually decided to terminate their exploration, the values found by each agent are shared with all of the other agents and can be used by any of them. Since the agents are self-interested, they cannot be forced to explore all possible n opportunities (or any desired subset), and their extent of exploration is the one that maximizes their individual expected benefit given the exploration strategies used by the other agents. Finally, the overall number of opportunities available to the agents is assumed to be greater than the maximum number of opportunities the agents can potentially query, kn, and the agents are assumed to be capable of exploring the different opportunities with no overlap (i.e., no two agents will query the same opportunity).

III. Analysis

We first present the expected-benefit-maximizing exploration strategy of a single agent with no cooperative exploration. The single agent's strategy is then augmented to the case of cooperative exploration, taking into consideration the distribution of the best value resulting from the exploration of the other agents. This leads to equilibrium analysis, in which a stable set of strategies are found, from which none of the agents has an incentive to deviate individually. Finally, we also develop the optimal cooperative exploration strategy, to be used later for comparing the loss, in terms of aggregate expected benefit, due to the fact that the agents are self-interested rather than cooperative.

Since there is no communication between the agents and the best values obtained are revealed at the end of the individual explorations, the agents, if choosing to take part in the exploration, will all initiate their individual explorations at time $t = 1$. The benefit of having all agents start at the first period is that this way each agent will be less constrained by the exploration horizon n. Since opportunity values are derived from a common distribution function, any a priori order according to which opportunities are explored is likely to produce the same result. An agent's exploration strategy is therefore a mapping $S(v) \rightarrow \{resume, terminate\}$, where v is the best value found so far, $resume$ suggests the exploration of another opportunity and $terminate$ suggests terminating the exploration process and returning the value v. If the agent relies solely on the best value it achieves throughout its exploration, then the model can be mapped to the legacy sequential exploration model found in literature and the optimal (expected-benefit-maximizing) exploration strategy is reservation-value (threshold) based [19]. According to this strategy, each agent A_i calculates a reservation-value r_i, and resumes its exploration as long as the best value found so far is below r_i. The expected-benefit-maximizing r_i value is extracted from:

$$c_i = \int_{y=r_i}^{\infty} (y - r_i) f_i(y) dy \quad (1)$$

Intuitively, r_i is the value where the exploring agent A_i is precisely indifferent: the expected marginal benefit from obtaining the value of the opportunity exactly equals the cost of obtaining that additional value. It is notable that the value of r_i does not depend on the number of remaining oppertunities which values are still unknown.

Now consider the case where the value with which the agent ends up is the maximum among any of the opportunities evaluated by any of the agents. Theorem 1 suggests that in this case, the agents' exploration strategy is still reservation-value (threshold) based and that the reservation value used is stationary, i.e., the agent continues evaluating opportunities until obtaining an opportunity which value is above some reservation value r.

Theorem 1: The expected-benefit-maximizing exploration strategy for agent A_i is to set a reservation value r_i, where r_i is the solution to:

$$c_i=\int_{y=r_i}^{\infty} f_i(y)\int_{x=-\infty}^{\infty}(max(y,x)-max(r_i,x))\bar{f}_i(x)dxdy \quad (2)$$

where $\bar{f}_i(x)$ is the probability distribution function of the maximal value obtained along the exploration of the other agents.

The agent should always choose to obtain the value of another opportunity if the highest value obtained so far is below r_i and terminate the exploration once the maximum value obtained so far is greater than r_i. The individual exploration also terminates once all opportunities have been explored.

Proof: The agents' strategies are necessarily reservation-value based because, in the absence of any other new information along the exploration process, each agent can base its strategy only on the best value obtained so far (hence a reservation-value based strategy). We now prove that the reservation value used by each agent remains stationary along its exploration and calculated according to Equation 2. We use r_i^j to denote the reservation value to be used in the j-th time period ($j \leq n$). We first prove that $r_i^j \leq r_i^{j-1}$ for any $1 < j \leq n$. This proof is simple — if an agent finds it beneficial to explore an additional opportunity when $n-j$ opportunities remaining, then it is necessarily beneficial for the agent to explore an additional opportunity if $n-j+1$ opportunities remain. Now consider the expected benefit of the agent upon reaching the j-th opportunity (i.e., at time j) if the best value it has found so far is x. If $x \leq r_i^{j+1}$ then the agent necessarily needs to explore the i-th opportunity (since $r_i^j \leq r_i^{j-1}$). If $x > r_i^{j+1}$, then the agent can either terminate its exploration, in which case its expected benefit is $\int_{z=-\infty}^{\infty} max(x,z)\bar{f}_i(z)dz$, or explore the $j-th$ opportunity. If exploring the $j-th$ opportunity then it is guaranteed that the agent will terminate exploration right after, as $x > r_i^{j+1}$.[1] Therefore, the expected benefit from the additional exploration is given by: $-c_i+\int_{y=-\infty}^{x} f_i(y)\int_{z=-\infty}^{\infty} max(x,z)\bar{f}_i(z)dzdy+\int_{y=x}^{\infty} f_i(y)\int_{z=-\infty}^{\infty} max(y,z)\bar{f}_i(z)dzdy$. The first term relates to the case where the value obtained in the j-th exploration is above x, whereas the second term relates to the case where it is below x (however, given that the agent already has value $x > r_i^{j+1}$ at hand, the exploration terminated without exploring the $j+1$ opportunity). The agent should therefore explore the j-th opportunity if:

$$\int_{z=-\infty}^{\infty} max(x,z)\bar{f}_i(z)dz \geq -c_i + \int_{y=-\infty}^{x} f_i(y)\int_{z=-\infty}^{\infty} max(x,z)\bar{f}_i(z)dzdy + \int_{y=x}^{\infty} f_i(y)\int_{z=-\infty}^{\infty} max(y,z)\bar{f}_i(z)dzdy \quad (3)$$

[1]And similarly, in the case $j=n$ the exploration will terminate right after, in the absence of further exploration opportunities.

The reservation value r_i^j is, by definition, the value x for which the expected benefit from resuming the exploration process equals the best value obtained so far throughout the exploration. Therefore, r_i^j is the value x for which Equation 3 becomes an equality. The derivative of both sides of Equation 3 reveals that the right-hand-side increases at a lower rate compared to the left-hand-side term, therefore there is a single reservation value that satisfies the equality. Substituting $x = r_i^j$ in Equation 3 in its equality form and applying some standard mathematical manipulations obtains Equation 2. ■

Theorem 1 specifies the best response strategy of any of the agents, given the strategies used by the other agents. Since the agents are self-interested, they may find it beneficial to avoid exploration at all entirely benefit from the findings of others. Any given solution, therefore, needs to be checked for stability, and the equilibrium is a set of strategies from which none of the agents will want to deviate individually. In order to extend the equilibrium analysis, we allow also mixed-equilibrium solutions. A mixed equilibrium in this case is based on a set $\{(p_1, r_1), ..., (p_k, r_k)\}$ where p_i is the probability that agent A_i will engage in exploration, and r_i is the reservation value the agents will use if exploring. It is notable that if the agent finds it beneficial to engage in exploration, then it will necessarily use the expected-benefit-maximizing strategy according to Theorem 1 (i.e., will never randomize along its exploration).

In order to formulate $\bar{f}_i(x)$, we use the probability that the maximum value obtained along the exploration process of an agent A_i that uses a reservation value r_i is lesser than x, denoted $F_i^{return}(x)$, calculated according to:

$$F_i^{return}(x)=\begin{cases} F_i(x)^n & x<r_i \\ \frac{1-F_i(r_i)^n}{1-F_i(r_i)}(F_i(x)-F_i(r_i))+F_i(r_i)^n & x\geq r_i \end{cases} \quad (4)$$

The case where $x < r_i$ requires that all n explored opportunities result in a value below the reservation value r_i. In the case $x \geq r_i$, if the exploration was terminated at the j-th exploration then the maximum value obtained throughout the exploration will be smaller than x only if the value obtained at the exploration is between x and r_i (as otherwise the exploration resumes) and all the former $j-1$ opportunities queried returned a value smaller than r_i (as otherwise the j-th opportunity is not reached). This results in the geometric series $\prod_{j=1}^{k}(F(x)-F(r_i))F(r_i)^j$ and the final result requires the addition of the case where all k opportunities turn out to be associated with a value smaller than r_i, i.e., $F_i(r_i)^n$.

Similarly, the function $f_i^{return}(x)$, which is the probability the maximum obtained throughout agent A_i's exploration, is calculated as the derivative of $F_i^{return}(x)$:

$$f_i^{return}(x) = \frac{d(F_i^{return}(x))}{dx} \quad (5)$$

Using $F_i^{return}(x)$ we can now formulate the probability that the maximum found by all of the agents except for A_i is smaller or than equal to x, denoted $\bar{F}_i(x)$:

$$\bar{F}_i(x) = \prod_{j=1 \wedge j\neq i}^{k} (p_j F_j^{return}(x) + (1-p_j)) \quad (6)$$

The calculation is based on having the best value recieved by any of the agents found in the exploration below x. The probability that agent j's best value found throughout its exploration is below x is $F_j^{return}(x)$ if the agent actually engaged in exploration (i.e., with p_j chance) and 1 otherwise (i.e. with $(1-p_j)$ chance). Consequently, the probability function $\bar{f}_i(x)$ is the derivative of $\bar{F}_i(x)$:

$$\bar{f}_i(x) = \frac{d(\bar{F}_i(x))}{dx} \tag{7}$$

These enable calculating the expected benefit of agent A_i when the other agents are using the set of strategies $\{(p_1, r_1), ..., (p_k, r_k)\}$. If the agent chooses to engage in exploration then its expected benefit, denoted $EB_i(explore)$, is given by:

$$EB_i(explore) = -c_i \frac{1 - F_i(r_i)^n}{1 - F_i(r_i)} + \int_{y=-\infty}^{\infty} f_i^{return}(y) \int_{x=-\infty}^{\infty} max(y, x) \bar{f}_i(x) dx dy \tag{8}$$

and when the agent opts not to explore at all, its expected benefit, denoted $EB_i(\neg explore)$, is:

$$EB_i(\neg explore) = \int_{x=-\infty}^{\infty} x \bar{f}_i(x) dx \tag{9}$$

Equation 8 includes the cost associated throughout the exploration of A_i, $-c_i \frac{1-F_i(r_i)^n}{1-F_i(r_i)}$, as this becomes a Bernoulli sampling process with a success probability of $F(r_i)$. The second term is the expected maximum value between the best value found by the agent itself (i.e., associated with a distribution $f_i^{return}(y)$) and the best value returned by the other agents (associated with a distribution $\bar{f}_i(x)$). The term on the right-hand-side of Equation 9 is simply the expected value of the maximum value returned by the other agents.

At this point, we have everything that is needed in order to formulate the equilibrium stability conditions. A set of strategies $\{(p_1, r_1), ..., (p_k, r_k)\}$ will be in equilibrium only if the following conditions are held: (a) for every agent A_i for which $p_i = 0$, $EB_i(explore) \leq EB_i(\neg explore)$; (b) for every agent A_i for which $p_i = 1$, $EB_i(explore) \geq EB_i(\neg explore)$; and (c) for every agent A_i for which $0 < p_i < 1$, $EB_i(explore) = EB_i(\neg explore)$. Therefore, in order to find the equilibrium, one needs to check the stability of 3^k possible solutions of the type $\{(p_1, r_1), ..., (p_k, r_k)\}$ differing in the value each p_i obtains ($p_i = 0$, $p_i = 1$ and $0 < p_i < 1$). For every combination, the reservation values and the value p_i of the agents that do explore ($0 < p_i < 1$) should be calculated by solving a set of equations of type (2) and $EB_i(explore) = EB_i(\neg explore)$, according to (8) and (9). Once the appropriate reservation values and probabilities are extracted, the stability conditions need to be validated.

We note that there is no guarantee that an equilibrium will actually exist (either pure or mixed). Also, there is no guarantee that if one exists there will be no other equilibria (i.e., a multi-equilibria is possible). In the latter case, if there is one equilibrium that dominates the others in terms of the individual expected benefit each and every agent obtains then it will likely be the one used. Otherwise, there is no way of deciding which of the equilibria is the one to be used, and we leave this question beyond the scope of the current paper.

Finally, we develop the cooperative overall expected-benefit-maximizing strategy (denoted "fully cooperative") that is used in the following section for comparative purposes. In the cooperative setting, all agents aim to maximize the sum of their expected benefits. Naturally, the exploration strategies that maximize the latter are different from those used for the self-interested case. Furthermore, while the overall expected benefit increases when all agents are exploring cooperatively, there is often an incentive for some of the agents to deviate from the cooperative strategy in order to improve their individual expected benefit.

When the agents are cooperative and attempt to maximize overall expected benefit, the only change that needs to be made in Equation 3 is the multiplication of $max(x, z)$ and $max(y, z)$ in k, as any improvement in the best value found improves the benefit of all k agents. Therefore, the overall expected-benefit-maximizing reservation value r_i to be used by A_i is the solution to:

$$c_i = k \int_{y=r_i}^{\infty} f_i(y) \int_{x=-\infty}^{\infty} max(y, x) - max(r_i, x) \bar{f}_i(x) dx dy \tag{10}$$

The overall expected-benefit-maximizing strategy is thus the set of reservation values $(r_1, ..., r_k)$ that are the solution to the set of k equations obtained from substituting $i = 1, ..., k$ in Equation 10. Since the optimal solution in this case may be to have only some of the agents engage in exploration, one needs to iterate over all of the possible combinations of having $k' < k$ agents engage in exploration and solve for $(r_1, ..., r'_k)$, choosing the combination for which the overall expected benefit is maximal.

Obviously the overall expected benefit obtained when the agents cooperatively attempt to optimize that measure is greater than in the case where each self-interested agent attempts to maximize its own expected benefit.

IV. Equilibrium Dynamics

In this section we demonstrate the resulting equilibrium exploration strategies and the individual and overall expected benefit under different settings. In particular we consider the effect of differences in the exploration competence of the different agents (i.e., their cost of exploration c_i), the number of agents and the exploration horizon n over individual and overall benefit.

Figure 1 depicts the expected benefit of the agents as a function of the exploration horizon n. The setting used for this example considers three homogeneous agents ($k = 3$), each associated with an exploration cost of 0.1 ($c_1 = c_2 = c_3 = 0.1$) and a uniform distribution of values ($f_i(x) = 1$ $\forall 0 \leq x \leq 1$, and $f_i(x) = 0$ otherwise, $i = 1, 2, 3$). In this case, the equilibrium calculation, based on the analysis given in the former section, reveals three different equilibria. The first two are pure equilibria: (a) when $n = 1$, two of the agents explores and one does not; (b) when $n > 1$, one of the agents explore, and the two others do not. The third equilibrium is a

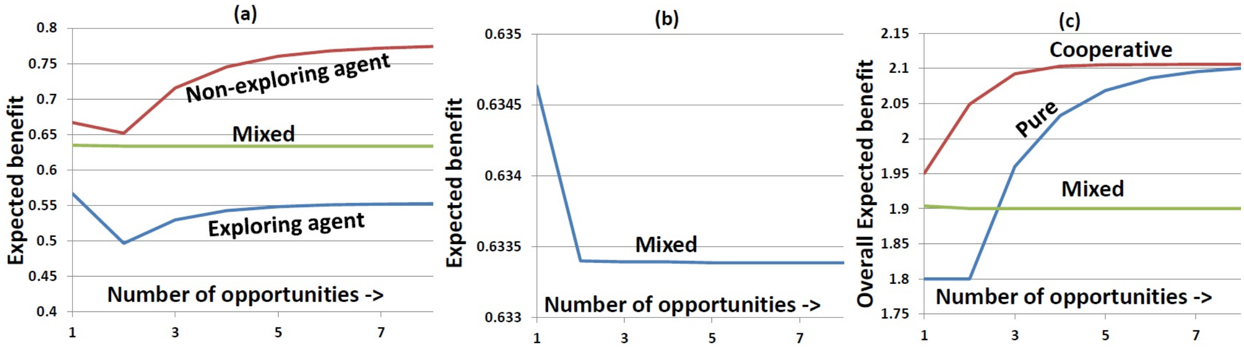

Figure 1. The agents' expected benefit as a function of the exploration horizon n, for the setting: $k = 3$, $c_1 = c_2 = c_3 = 0.1$, and a common uniform distribution. The two possible equilibria are a mixed one and a pure one (where for $n = 1$ only two agents engage in exploration and for $n > 1$ only one agent engages in exploration). (a) Individual expected benefit. (b) A finer-grained representation for the mixed equilibrium. (c) Overall expected benefit with the equilibria and when the agents are fully cooperative.

mixed one, where $p_1 = p_2 = p_3$. The expected benefit of the agents with each of the equlibria is depicted in Figure 1(a). The upper curve corresponds to the expected value of the agent(s) that is not doing any exploration in the pure equilibrium (one agent in the case where $n = 1$ and two agents otherwise). The bottom curve corresponds to the expected value of the agent(s) that engages in exploration (two agents in the case where $n = 1$ and one agent otherwise). Generally, the increase in the number of opportunities that the agents can explore has a positive effect over the performance of the agents in their individual explorations. This explains the increase in the expected benefit, as a function of n when $n > 1$, in both curves. The decrease in both curves when transitioning from $n = 1$ to $n = 2$, though, is explained by the change in the equilibrium structure — in this case the negative effect of having only one agent explore in equilibrium outweighs the positive effect of the increase in the number of opportunities the agents can explore. The middle curve corresponds to the expected value when using the mixed equilibrium (hence equal to all agents).

While the issue of which equilibrium is likely to hold is beyond the scope of the current paper, we note that neither of the equlibria (the pure and the mixed) dominates the other. In this example the pure equilibrium does not result in a better individual expected benefit to all agents (the agent(s) that does not engage in exploration ends up with an expected benefit inferior to the one received with the pure equilibrium). The advantage of the mixed equilibrium in this context is that it guarantees that all agents end up with the same expected benefit, and since the agents are homogeneous, self-interested and fully rational, it seems like a natural selection.

Figure 1(b) enlarges the middle curve of (a), i.e., depicts the agents' individual expected benefit when using the mixed equilibrium (using a more fine-grained scale). Unlike with the pure equilibrium (for $n > 1$), the individual benefit obtained with the mixed equilibrium actually decreases as the number of opportunities increases. This non-intuitive behavior results from the decrease in the value of p used in equilibrium. The agents can potentially explore to a greater extent, and if exploring they actually do so, however the decrease in the chance that they will engage in exploration (as a result of the stability considerations) results in a total decrease of the individual expected benefit. Figure 1(c) depicts the overall expected benefit of the three agents when using the different equilibria and when acting according to the cooperative exploration strategy (which is inherently unstable). As observed from the figure, neither of the equilibria generally dominates the other, and as expected both are dominated by the cooperative (though non-stable) strategy. The overall expected benefit with the pure equilibrium converges to the case of cooperative exploration. This is explained by the fact that as the cooperative exploration becomes less constrained by the number of opportunities available, the optimal strategy is to have just one of the agents engage in exploration (as sequential exploration dominates parallel exploration whenever there is no restriction over the number of exploration steps [21]). The difference in the overall expected benefit in this case, compared to the pure equilibrium in which only one agent engages in exploration, derives from the fact that in the latter case the agent attempts to maximize its own expected benefit rather than the overall of the three.

Figure 2 presents an analysis similar to the one given in Figure 1, however having the number of agents, k, as the influencing parameter. In this case the exploration cost is 0.1 for all agents, the exploration horizon is $n = 3$ and the distribution of values is uniform for all agents as in Figure 1. Here again there is a single pure equilibrium (in which only one of the agents engages in exploration) and a mixed equilibrium. As observed from the figure, the individual expected benefit of the agents when the mixed equilibrium solution is used increases as the number of agents increases. This suggests that the improvement achieved by the addition of agents (in terms of the number of opportunities that can now be explored in parallel and overall) is greater than the decrease in the exploration extent resulting from the reliance of each of the agents on the others.

Finally, Figure 3 depicts the expected individual benefit of two homogeneous agents (associated with a cost of exploration $c_1 = c_2 = 0.2$) when adding a third agent whose exploration cost, c, is given by the horizontal axis. The exploration horizon

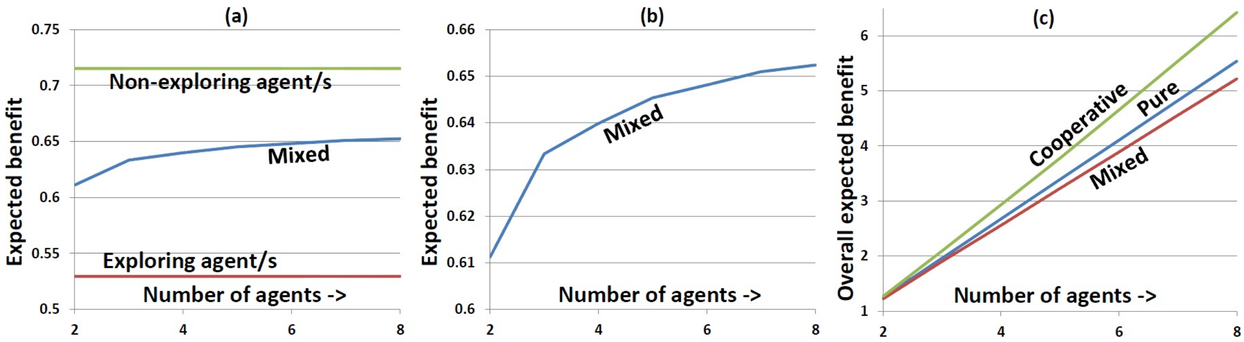

Figure 2. The agents' expected benefit as a function of the number of agents, k, for the setting: $n = 3$, $c_1 = c_2 = c_3 = 0.1$, and a common uniform distribution. The two possible equilibria are a mixed one and a pure one (where only one agent engages in exploration). (a) Individual expected benefit. (b) A finer-grained representation for the mixed equilibrium. (c) Overall expected benefit with the equilibria and when the agents are fully cooperative.

is limited to a single opportunity, i.e., $n = 1$, and the distribution of values is uniform for all agents as in Figure 1. In this case there is a mixed equilibrium if $0.1 < c < 0.5$. Two additional pure equilibrium also hold for some c values, both consisting of only one out of the three agents engaging in exploration. In the first, depicted in (a), the exploring agent is one of the two homogeneous agents. This equilibrium holds only for c values in the interval $(0.17, 0.5)$. It is notable that in this case the expected benefit of the different agents does not depend on the cost c (which is relevant only to the third agent) since the agent engaged in exploration is one of the two homogeneous agents. The expected benefit of the two homogeneous agents and the third (joining) agent, when the mixed equilibrium is used, are distinguished by the markings "Mixed (M)" and "Mixed (T)", respectively. In the second pure equilibrium, depicted in (b), the exploring agent is the third agent (the one who joins the two homogeneous agents). This equilibrium holds for any $c < 0.5$ and hence is the only equilibrium (and necessarily the one used) when $c < 0.1$. With this equilibrium the cost of the exploring agent, which is the third agent, decreases as its cost of exploration c increases (since neither of the other two agents explore). As with the former examples, the two pure equilibria in this case suggest an expected benefit smaller than with the mixed equilibrium for the agent engaged in exploration and vice-versa for the agent that does not explore.

A finer-grained representation for the individual expected benefits obtained by the two homogeneous agents, when using the mixed equilibrium, is given in Figure 3(c). As observed from the figure, as the cost of exploration of the third partner increases, the expected benefit of the two homogeneous agents increases. However, starting from a certain c value, any further increase in the exploration cost of the joining agent results in a decrease in the two agents' expected benefit. This result suggests that the agents sometimes should prefer to join the less competent agent (i.e., the one associated with a smaller exploration cost compared to other potential candidates) to their cooperative exploration. The non-intuitiveness of this result derives from the fact that exploration costs are traditionally considered "inefficiencies" as they represent the lack of transparency in which of the environment the agents are operating [30], [3]. In the presence of exploration costs a rational player will not aim to find the best option, but rather settle for "good enough", beyond which the marginal cost of exploring exceeds the marginal benefit of continuing the exploration. Thus, exploration costs promote sub-optimal results (or so it would seem). As such, the traditional wisdom is that when designing a MAS environment, exploration costs should be reduced to a minimum, and indeed (as also reflected in the figure) the overall expected benefit with the fully cooperative case (i.e., when the agents obey the overall expected-benefit-maximizing strategy) increases as the exploration cost c decreases. However, when the agents are self-interested, such generally beneficial solutions often do not hold and eventually the resulting equilibrium suggests an inferior expected benefit to the agents when joining an agent with a relatively small exploration cost.

Figure 3(d) presents the overall expected benefit with the three equilibria and when the agents are fully cooperative. From this figure we learn that the choice of equilibrium that maximizes the overall expected benefit depends on the cost of exploration of the third agent. In our case the overall expected-benefit-maximizing equilibrium is the pure equilibrium in which the third agent is doing all the exploration for $c < 0.1$, the mixed equilibrium when $0.1 \leq c \leq 0.25$ and the other pure equilibrium otherwise.

V. Related Work

The model analyzed in this paper is based on two important concepts that are extensively researched in multi-agent systems literature. The first is cooperation between agents and the second is costly exploration. Cooperation evolves in multi-agent environments whenever agents can benefit from cooperating and coordinating their actions [15]. Consequently, group-based cooperative behavior has been suggested in various domains [29], [31]. The advantages encapsulated in teamwork and cooperative behaviors are the main driving force of many coalition formation models in the area of cooperative game theory and MAS [16] theory. Yet, the majority of cooperation and coalition formation MAS-related research tends to focus on the way coalitions are formed and consequently concerns issues such as the optimal division of agents into disjoint exhaustive coalitions [26], [31], division of coalition payoffs

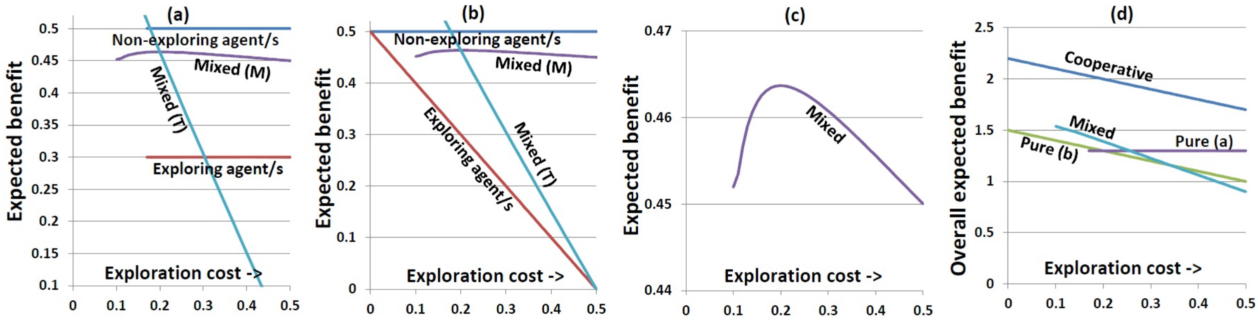

Figure 3. The agents' expected benefit as a function of a potential partner's exploration cost, c. The setting consists of two agents characterized by $c_1 = c_2 = 0.2$ and a third agent characterized by c. The distribution of opportunities values is a uniform distribution, common to all three agents. There are three possible equiliria in this case, two pure ones and one mixed. (a) Individual expected benefit in the case where the mixed equilibrium is used and in the case of using a pure equilibrium where the only agent engaged in exploration is one of the two homogeneous agents. (b) Individual expected benefit in the case where the mixed equilibrium is used and in the case of using a pure equilibrium where the only agent engaged in exploration is the third agent (which exploration cost differs from the others). (c) A finer-grained representation of the expected benefit of the two homogeneous agents with the mixed equilibrium. (d) Overall expected benefit with the three equilibria and when the agents are fully cooperative.

[31] and enforcement methods for interaction protocols [20]. In this paper, however, the focus is on how the cooperation is carried out, once formed.

Exploration is important in MAS, in particular when there is no central source that can supply full immediate reliable information on the environment and the state of the other agents that can be found. The introduction of exploration costs into MAS models leads to a more realistic description of these environments. This is because agents are typically required to invest/consume some of their resources in order to obtain information concerning opportunities available in their environment [3], [13].

Optimal exploration (often referred as "search") strategies for settings where individuals need to search for an applicable opportunity while incurring a search cost have been widely studied, prompting several literature reviews [17], [19]. These models have developed to a point where their total contribution is referred to as "search theory". Over the years, many search model variants have been considered, differing in the decision horizon (finite versus infinite) [17], the presence of the recall option [19], the distribution of values and the extent to which findings remain valid along the process [14]. Nevertheless, search theory literature investigates mainly the extraction of an optimal stopping rule for the individual exploring agent, and most often does not consider the system-wise analysis. Few studies have attempted to extend the exploration to a multi-agent (or a multi-goal) exploration model, e.g., attempting to purchase several commodities while facing imperfect information concerning prices or operating several robots in order to evaluate opportunities in different locations [12], [27], [11], [4], [10]. However, these works consider cooperative agents that attempt, through their coordinated exploration, to maximize the overall (rather than individual) benefit. One exception to the above is our recent work that considers a coordinated exploration model in which the agents are self-interested [25]. Nevertheless, in that work the agents are constrained by the results of the exploration carried out by the other agents rather than benefiting from it. Consequently, the nature of the equilibrium set of exploration strategies used is substantially different.

An equilibrium analysis of models where self-interested agents are engaged in costly exploration can be found in "two-sided" search literature [2], [5], [28]. The standard two-sided distributed search model postulates an environment populated by fully rational self-interested agents, searching for appropriate partners to form mutually acceptable pairwise partnerships. Nevertheless, in this model the focus is, once again, on the way the agents decide with whom to partner rather than how to act cooperatively.

Finally, we note that the non-intuitive finding according to which the agents should sometimes prefer joining an agent associated with a greater exploration cost to their group (rather an agent with a smaller exploration cost) follows, in spirit, previous results in other settings. In particular, ones in which it has been shown that so-called "inefficiencies" can increase market performance, under certain circumstances. For example, Masters [18] shows that an increase in the minimum wage, which is often considered by economics as inefficiency, can have positive employment effects. In transportation economics (e.g. congestion games), equilibrium is frequently not the overall optimum. In such cases, it has been shown that taxation can change the equilibrium to a more desirable one [23], [22], [9]. Similarly, taxes can facilitate more desirable equilibria in Boolean games [8]. Here we show that a somehow similar phenomena also happens in the context of costly exploration, though the model and analysis are, of course, totally different from the above mentioned.

VI. Discussion and Conclusions

Unlike prior works that consider extensions of traditional single-agent sequential exploration models to the multi-agent case, the analysis given in this paper considers the agents to be self-interested rather than fully cooperative. As such, the analysis is essentially equilibrium-based — instead of extracting a strategy that maximizes the overall expected benefit, the goal is to find a stable solution while each of the agents attempts to maximize its individual expected benefit, given the exploration strategy used by the others. Taking the agents to be self-interested makes the model more applicable (than the fully cooperative one) whenever the agents represent different individuals with different goals or that cannot be forced to obey some external solution which is "socially beneficial". Naturally, when it comes to implementation, many complementary aspects of multi-agent collaboration that are beyond the scope of the current paper need to be considered, such as ontologies and semantics [1], [24]. This paper is focused in the way the cooperative exploration itself is executed.

The equilibrium analysis considers both pure and mixed equilibria. As demonstrated in the former section, typically none of the equilibria dominates the other, and the determination of which will hold in the case of multi-equilibria is unresolved (as in game-theory in general). The mixed equilibrium solution might seem appealing, as it does not imply that some of the agents will engage in exploration whereas others will not. Furthermore, in the case of rather homogeneous agents it provides a rather balanced distribution of expected benefits.

Two interesting phenomena illustrated using the analysis of the mixed equilibrium are the decrease in the expected benefit of the agents when the exploration horizon increases and the preference of additional "weak" rather than "strong" agents as far as cooperative exploration is concerned. Both are explained by the stability requirement: while better solutions can be extracted whenever the number of opportunities available to the agents increases and when the more competent agents join the collective exploration effort, these solutions cannot hold, as some of the agents have an incentive to individually deviate from them.

One possible extension of the model suggested for future study is a cooperative exploration model with self-interested agents that can communicate along the process. For instance, in our example, Jack and Jill can communicate using their mobile phones and share findings as they individually visit the different stores. The analysis in this case can rely, in part, on the principles given in this paper. In the new model, however, instead of assigning an exploration probability to any individual agents' exploration as a whole, each agent will need to set its exploration probability per round, given the best value received so far by any of the agents and the time remaining for the exploration. The equilibrium strategies can thus be calculated in this case using dynamic programming.

Acknowledgments

This work was partially supported by ISF/BSF grants 1401/09 and 2008-404.

References

[1] M. Afsharchi, B. Far, and J. Denzinger, "Ontology-guided learning to improve communication between groups of agents," *ACM*, pp. 923–930, 2006.

[2] A. Atakan, "Assortative matching with explicit search costs," *Econometrica*, vol. 74, no. 3, pp. 667–680, 2006.

[3] Y. Bakos, "Reducing buyer search costs: Implications for electronic marketplaces," *Mgmt. Science*, vol. 42, pp. 1676–92, 1997.

[4] K. Burdett and D. A. Malueg, "The theory of search for several goods," *JET*, vol. 24, pp. 362–376, 1981.

[5] K. Burdett and R. Wright, "Two-sided search with nontransferable utility," *RED*, vol. 1, pp. 220–245, 1998.

[6] M. Chhabra, S. Das, and D. Sarne, "Expert-mediated search," in *Proc. AAMAS*, 2011, pp. 415–422.

[7] S. Choi and J. Liu, "Optimal time-constrained trading strategies for autonomous agents," in *Proc. MAMA*, 2000.

[8] U. Endriss, S. Kraus, J. Lang, and M. Wooldridge, "Designing incentives for boolean games," in *Proc. AAMAS*, 2011.

[9] D. Fotakis, G. Karakostas, and S. G. Kolliopoulos, "On the existence of optimal taxes for network congestion games with heterogeneous users," in *Proc. SAGT*, 2010, pp. 162–173.

[10] A. Frieder, R. Lin, and S. Kraus, "Agent-human coordination with communication costs under uncertainty," in *Proc. AAAI*, 2012.

[11] J. Gatti, "Multi-commodity consumer search," *JET*, vol. 86, no. 2, pp. 219–244, 1999.

[12] N. Hazon, Y. Aumann, and S. Kraus, "Collaborative multi agent physical search with probabilistic knowledge," in *Proc. IJCAI*, 2009, pp. 167–174.

[13] J. Kephart and A. Greenwald, "Shopbot economics," *JAAMAS*, vol. 5, no. 3, pp. 255–287, 2002.

[14] M. Landsberger and D. Peled, "Duration of offers, price structure, and the gain from search," *Journal of Economic Theory*, vol. 16, no. 1, pp. 17–37, 1977.

[15] K. Lermann and O. Shehory, "Coalition formation for large-scale electronic markets," in *Proc. ICMAS*, 2000, pp. 167–174.

[16] C. Li, S. Chawla, U. Rajan, and K. Sycara, "Mechanisms for coalition formation and cost sharing in an electronic marketplace," in *Proc. EC03*, 2003, pp. 68–77.

[17] S. Lippman and J. McCall, "The economics of job search: A survey," *Economic Inquiry*, vol. 14, pp. 155–189, 1976.

[18] A. M. Masters, "Wage posting in two-sided search and the minimum wage," *International Economic Review*, vol. 40, no. 4, pp. 809–826, 1999.

[19] J. McMillan and M. Rothschild, "Search," in *Proc. Handbook of Game Theory with Economic Applications*, vol. 2, 1994, pp. 905–927.

[20] P. Michiardi and R. Molva, "Analysis of coalition formation and cooperation strategies in mobile ad hoc networks," *Ad Hoc Networks*, vol. 3, pp. 193–219, 2005.

[21] P. Morgan and R. Manning, "Optimal search," *Econometrica*, vol. 53(4), pp. 923–944, 1985.

[22] M. Penn, M. Polukarov, and M. Tennenholtz, "Random order congestion games," *Math. Oper. Res.*, vol. 34, no. 3, pp. 706–725, 2009.

[23] ——, "Taxed congestion games with failures," *Ann. Math. Artif. Intell.*, vol. 56, no. 2, pp. 133–151, 2009.

[24] Z. W. Ras and A. Dardzinska, "Solving failing queries through cooperation and collaboration," *World Wide Web*, pp. 173–186, 2006.

[25] I. Rochlin, D. Sarne, and G. Zussman, "Sequential multilateral search for a common goal," in *Proc. IAT*, 2011, pp. 349–356.

[26] T. Sandholm, K. Larson, M. Andersson, O. Shehory, and F. Tohme, "Coalition structure generation with worst case guarantees," *AIJ*, vol. 111, pp. 209–238, 1999.

[27] D. Sarne, E. Manisterski, and S. Kraus, "Multi-goal economic search using dynamic search structures," *JAAMAS*, vol. 21, pp. 204–236, 2010.

[28] R. Shimer and L. Smith, "Assortative matching and search," *Econometrica*, vol. 68(2), pp. 343–370, 2000.

[29] N. Tsvetovat, K. Sycara, Y. Chen, and J. Ying, "Customer coalitions in electronic markets," in *Proc. AMEC*, 2000, pp. 121–138.

[30] V. Vaishnavi and W. Kuechler, "An approach to reducing e-commerce coordination costs in evolving markets using a semantic routing protocol," *J. Electron. Commerce Res.*, vol. 4, no. 3, pp. 113–127, 2003.

[31] J. Yamamoto and K. Sycara, "A stable and efficient buyer coalition formation scheme for e-marketplaces," in *Proc. Agents'01*, 2001, pp. 576–583.