

Exploration of Unknown Environments with Motivational Agents

Luís Macedo^{1,2}, Amílcar Cardoso²

¹*Department of Informatics and Systems Engineering, Engineering Institute, Coimbra Polytechnic Institute, 3030 Coimbra, Portugal*

²*Artificial Intelligence Lab, Centre of Informatics and Systems, University of Coimbra, 3000 Coimbra, Portugal*

lmacedo@isec.pt; {macedo,amilcar}@dei.uc.pt

Abstract

This paper addresses the problem of exploring unknown, dynamic environments with motivational agents. The goal is the acquisition of a model of the environment including models of the entities that populate the environment. We describe the exploration strategy of both single and multiple agents. Each agent performs directed exploration using an action selection method based on the maximization of the intensity of positive feelings and minimization of negative ones. The exploration strategy for multiple agents relies on considering a team leader that integrates the maps and coordinates the actions of the members of the team. We present and discuss the results of an experiment conducted in simulated environments.

1. Introduction

Exploration may be defined as the process of selecting and executing actions so that the maximal knowledge of the environment is acquired [35]. The result is the acquisition of models of the physical environment. Unfortunately, exploring unknown environments requires resources from agents such as time and power. Thus, there is a trade-off between the amount of knowledge acquired and the cost to acquire it. The goal of an explorer is to get the maximum knowledge of the environment at the minimum cost (e.g.: minimum time and/or power). Therefore, strategies that minimize the required amount of these resources and maximize knowledge acquisition have been pursued [35]. Although several exploration strategies have been successfully applied (e.g.: [2, 3, 7, 16, 34-38]), there is still much to be done specially in the exploration of dynamic, three-dimensional environments. These strategies have been grouped into

two main categories: undirected and directed exploration [35]. Strategies belonging to the former group (e.g.: random walk exploration, Boltzman distributed exploration) use no exploration-specific knowledge and ensure exploration by merging randomness into action selection. On contrary, strategies belonging to the latter group rely heavily on exploration specific-knowledge for guiding the learning process. Most of these directed strategies rely on the maximization of knowledge gain (e.g.: [34]). This technique is according with some psychological studies that have shown that novelty and new stimulus incite exploration [5, 20]. Curiosity is the psychological construct that has been closely related with this kind of behavior. Considering that curiosity was innate and that it could also be acquired, Berlyne [5] and McDougall [20] argued that novel stimulus elicits curiosity, which diminishes with continued exposure to the stimulus. Sharing similar ideas with Berlyne and McDougall, Shand [33] defined curiosity as a primary emotion consisting of a simple impulse to know, controlling and sustaining the attention, and evoking the bodily movements that allow one to acquire information about an object. These approaches are closely related to the emotion concept of interest-excitement proposed by the differential emotions theory to account for exploration, adventure, problem solving, creativity and the acquisition of skills and competencies in the absence of known drives [13].

However, as argued by Berlyne [5], in addition to novelty, other variables such as change, surprisingness, complexity, uncertainty, incongruity and conflict also determine this kind of behavior related to exploration and investigation activities. Therefore, curiosity (or novelty) seems to be insufficient to explain all the exploration behavior.

Considered by many authors as a biologically fundamental emotion (e.g.: [13]), surprise may play an

important role in cognitive activities, especially in attention focus and learning ([13, 21, 23]). For instance, Meyer et al.'s [21] experiments provide evidence indicating that surprising-eliciting events initiate a series of mental processes that begin with the appraisal of a cognized event as exceeding some threshold value of unexpectedness or schema discrepancy, and culminate with an analysis and evaluation of that event plus immediate reactions to it and/or schema (beliefs) updating/revision. The model proposed by Ortony and Partridge [23] shares aspects with this one especially in that both defend that unexpected events elicit surprise.

Therefore, taking the ideas above into account, and in order to understand, build or model artificial agents that explore like humans do, it is compulsory that those agents are able to manifest at least surprise and curiosity. However, since it is an activity that involves decision-making, exploration may also be influenced by other motivations. Actually, psychological and neuroscience research over the past decades suggests that motivations play a critical role in decision-making, action and performance, by influencing a variety of cognitive processes (e.g., attention, perception, planning, etc.). On the one hand, recent research in neuroscience [9, 15] supports the importance of emotions on reasoning and decision-making. On the other hand, there are a few theories in psychology relating motivations (including drives and emotions) to action [8, 13]. For instance, in the specific case of emotions, as outlined by [29], within the context of the belief-desire theories of action (the dominant class of theories in today's motivation psychology) there have been proposals such as that emotions are action goals.

Although research in Artificial Intelligence has almost ignored this significant role of motivations on reasoning, several models have been proposed for them recently (e.g.: [4, 10, 17-19, 22, 24-26, 28, 31]).

In this paper we describe an approach for the exploration of unknown, uncertain, and dynamic environments by motivational agents whose decisions are based on the motivations they "feel". The result of the exploration activity is not only the acquisition of maps of the environment where the entities that populate the environment are represented, but also the construction of models of those entities themselves.

The next section presents the architecture of a motivational agent. Subsequently, we describe how motivational agents perform exploration. Then, we present and discuss an experiment carried out to evaluate the role of motivations (surprise, curiosity and hunger) on exploration. Finally, we present conclusions.

2. Architecture of an Agent

The architecture that we adopted for an agent (Figure 1) is based on the belief, desire, and intention (BDI) approach [27]. The deliberative reasoning/decision-making module is in the core of the architecture. It receives internal information (from memory) and environment information (through the sensors) and outputs an action that has been selected for execution. The process of action selection takes into account the states of the environment the agent would like to happen (desires), i.e., it selects an action that leads to those states of the environment the agent prefers. This preference is implicitly represented in a mathematical function that evaluates states of the world in terms of the positive and negative feelings they elicit in the agent. Thus, this function obeys to the Maximum Expected Utility (MEU) principle [30]. In this case, the Utility is positive feelings. The intensities of these feelings (motivations) are computed by the motivations module taking into account both the past experience (the information stored in memory) and the present environment description provided by the sensors. The next subsections describe in more detail the main modules of the architecture.

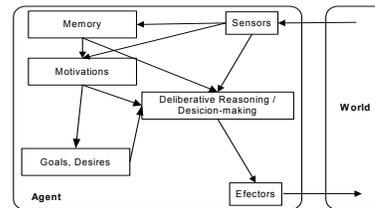


Figure 1. Architecture of an agent.

2.1. Memory

The memory of an agent stores information about the world. This information includes the configuration of the surrounding world such as the position of the entities (objects and other animated agents) that inhabit it, the description of these entities themselves, descriptions of the sequences of actions (plans) executed by those entities and resulting from their interaction, and, in generally, beliefs about the world. This information is stored in several memory components. Thus, there is a metric (grid-based) map [37] to spatially model the surrounding physical environment of the agent. Descriptions of entities (physical structure and function) and plans are stored both in the episodic memory and in the semantic memory [1]. We will now describe in more detail each one of these distinct components.

2.1.1. Metric map. In our approach, a (grid-based) metric map of the world is a three-dimensional grid in which a cell contains the information of the set of entities that may alternatively occupy the cell and the probability of this occupancy. Thus, each cell $\langle x,y \rangle$ of the metric map of an agent i is set to a set of pairs $\phi_{x,y}^i = \{ \langle p_1^i, E_1^i \rangle, \langle p_2^i, E_2^i \rangle, \dots, \langle p_n^i, E_n^i \rangle, \langle p_{n+1}^i, 0 \rangle \}$, where E_j^i is the identifier of the j^{th} entity that may occupy the cell $\langle x,y \rangle$ of the metric map of agent i with probability $p_j^i \in [0,1]$, and such that $\sum_{j=1}^{n+1} p_j^i = 1$. Note that the pair $\langle p_{n+1}^i, 0 \rangle$ is included in order to express the probability of the cell being empty. Cells that are completely unknown, i.e., for which there are not yet no assumptions/expectations about their occupancy, are set with an empty set of pairs $\phi_{x,y}^i = \{ \}$. Note also that each entity may occupy more than a single cell, i.e., there might be several adjacent cells with the same E_j^i .

2.1.2. Memory for entities. The set of descriptions of entities perceived from the environment are stored in the *episodic memory of entities* (Figure 2). Each one of these descriptions is of the form $\langle ID, PS, F \rangle$, where ID is a number that uniquely identifies the entity in the environment, PS is the physical structure, and F is the function of the entity. The sensors may provide incomplete information about an entity (for instance, only part of the physical structure may be seen or the function of the entity may be undetermined). In this case the missing information is filled in by making use of the Bayes's rule [32], i.e., the missing information is estimated taking into account the available information and descriptions of other entities previously perceived and already stored in the *episodic memory of entities*. This means some of the descriptions of entities stored in memory are uncertain or not completely known (e.g.: element 4 of Figure 2).

The physical structure of an entity may be described analogically or propositionally [1]. The analogical representation reflects directly the real physical structure while the propositional representation is a higher level description (using propositions) of that real structure.

The analogical description of the physical structure of an entity comprises a three-dimensional matrix and the coordinates of the centre of mass relatively to the entity and to the environment spaces. Notice that the three-dimensional matrix of the entity is a submatrix of the matrix that represents the metric map.

The propositional description of the physical structure of an entity relies on the representation through semantic features or attributes much like in semantic networks or schemas [1]. Entities are described by a set of attribute-value pairs that can be graph-based represented.

The function is simply a description of the role or category of the entity in the environment. For instance, a house, a car, a tree, etc. Like the description of the physical structure, this may be probabilistic because of the incompleteness of perception. This means, this is a set $F = \{ \langle \text{function}_i, \text{prob}_i \rangle : i=1,2, \dots, n$, where n is the number of possible functions and $P(\text{"function"} = \text{function}_i) = \text{prob}_i \}$.

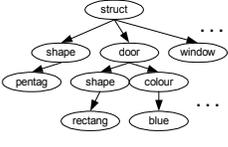
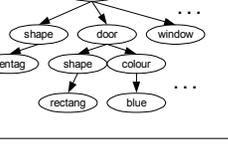
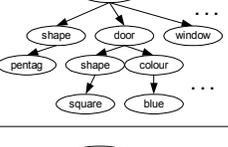
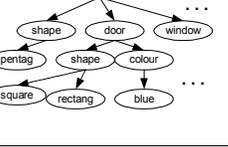
Id	Analogical	Propositional	Function
1	 RG=(2,1,2) AO=(10,10,10) AG=(12,11,12),		<house,1.0>
2	 RG=(2,1,3) AO=(7,3,4) AG=(9,4,7)		<church,1.0>
3	 RG=(1,1,2) AO=(15,17,10) AG=(16,18,12)		<house,1.0>
4	 RG=(2,1,2) AO=(3,12,0) AG=(5,13,2)		<house,.066> <church,.033>

Figure 2. Example of the episodic memory of entities. Although the matrix of the analogical description is of three-dimensional kind, for the sake of simplicity, it is represented here as a two-dimensional matrix corresponding to the upper view of the entity.

2.1.3. Memory for plans. We represent plans as a hierarchy of tasks (a variant of HTNs (e.g., [11]). This structure has the form of a planning tree, i.e., it is a kind of AND/OR tree that expresses all the possible ways to decompose an initial goal task (the root task).

A task t is both conditional and probabilistic (e.g.: [6]). This means each task has a set of conditions $C = \{ c_1, c_2, \dots, c_m \}$ and for each one of these mutually exclusive and exhaustive conditions, c_i , there is a set of alternative effects $\mathcal{E}^i = \{ \langle p_1^i, E_1^i \rangle, \langle p_2^i, E_2^i \rangle, \dots$

$\langle p'_n, E'_n \rangle$, where E'_j is the j^{th} effect triggered with probability $p'_j \in [0,1]$ by condition c_i (i.e., $P(E'_j | c_i) = p'_j$), and such that $\sum_{j=1}^n p'_j = 1$. Each effect

contains information about changes produced in the world by achieving the goal task. Thus, an effect may give information about the amount of power consumed, the new location of the agent, the emotions felt, etc.

2.2. Motivations

This module receives information from the current state of the environment and outputs the intensities of emotions, drives and other motivations. In this paper, this module is confined to the motivations that are related with variables that directly instigate exploration: surprise (elicited by unexpectedness), curiosity (elicited by novelty). In addition, we also consider the influence of the drive “hunger” that reflects the need of a power source.

The agent is almost continuously presented with an *input proposition* [23], which corresponds to some sensorial information of an entity (for instance, “a house with squared windows”). In response to this external stimulus, the surprise and curiosity unit outputs the intensity of these motivations, respectively.

In what concerns to surprise, we have developed a computational model [18] with the collaboration of the psychologists of the University of Bielefeld, Germany [21], and also based on the ideas of Ortony and Partridge [23]. The idea behind this model is that surprise consists of the appraisal of unexpectedness. Actually, there is experimental evidence supporting that the intensity of felt surprise increases monotonically, and is closely correlated with the degree of unexpectedness (see [18] for more details). This means that unexpectedness is the proximate cognitive appraisal cause of the surprise experience. Considering this evidence, we have already proposed [18] that the surprise felt by an agent Agt elicited by an object Obj_k is given by the degree of unexpectedness of Obj_k , considering the set of objects present in the memory of the agent Agt , which is given by the improbability of Obj_k (see [18] for more details):

$$\begin{aligned} SURPRISE(Agt, Obj_k) &= \\ UNEXPECTEDNESS(Obj_k, Agt(Mem)) &= 1 - P(Obj_k) \end{aligned}$$

We define curiosity (following McDougall [20], Berlyne [5] and Shand [33]) as the desire to know or learn an object that arouses interest by being novel, which means that novel objects stimulate actions intended to acquire knowledge about those objects.

Thus, if we accept the above definition, the curiosity induced in an agent Agt by an object Obj_k depends on the novelty or difference of Obj_k relatively to the set of objects present in the memory of Agt :

$$CURIOSITY(Agt, Obj_k) = DIFFERENCE(Obj_k, Agt(Mem))$$

The measure of difference relies heavily on error correcting code theory [12]: the function computes the distance between two objects represented by graphs, counting the minimal number of changes (insertions and deletions of nodes and edges) required to transform one graph into another.

The drive hunger is defined as the need of a source of energy. Given the capacity C of the storage of that source, and L the amount of energy left ($L \leq C$), the hunger elicited in an agent is computed as follows:

$$HUNGER(Agt) = C - L$$

2.3. Deliberative reasoning/decision-making

This module receives information from the internal/external world and outputs an action that has been selected for execution.

Roughly speaking, the agent starts by computing the current world state. This is performed by generating expectations or assumptions for the gaps in the environment information provided by the sensors. Then, new intentions/goals are generated and their Expected Utility (EU) computed. According to this EU, the set of goals of the agent are ranked, and the first one, i.e., the MEU goal is taken and a HTN plan is generated for it and executed.

We will now describe in more detail the steps related with the generation of assumptions/expectations, and generation and ranking of agent’s goals. The generation of plans is performed much like in HTN approaches (see [11]).

2.3.1. Generation of assumptions/expectations. It is very difficult for an agent to get all the information about the surrounding environment mainly because of the incompleteness of their perceptual information. However, taking as evidence the available information it is possible to generate expectations/assumptions for the missing information using the Bayes’ rule [32]:

$$\begin{aligned} P(H_i | E_1, E_2, \dots, E_m) &= \\ &= \frac{P(E_1 | H_i) \times P(E_2 | H_i) \times \dots \times P(E_m | H_i) \times P(H_i)}{\sum_{l=1}^n P(E_1 | H_l) \times P(E_2 | H_l) \times \dots \times P(E_m | H_l) \times P(H_l)} \end{aligned}$$

where E_1, E_2, \dots, E_m are pieces of evidence, i.e., the available information, and $H_i, i=1,2,\dots,n$, are mutually exclusive and collectively exhaustive hypotheses for a

specific piece of the missing information. The set of H_i 's is the exhaustive set of instances assigned to that specific part of the missing information in past cases (of entities, plans, etc.). Each conditional probability $P(E|H)$ is given by the number of times E and H appeared together in past cases divided by the number of times H appeared in those cases. With respect to the entities that populate the environment, in our work, the evidence is the description (propositional) of the physical structure of the entities such as their shape (rectangular, squared, etc.), shape of their constituent parts (in case there are any), colour, etc. The hypotheses could be not only for parts of the descriptions of the physical structure but also for the function or category of the entity. In this case, the result is a probability distribution for the function of the entity (e.g., $P(\text{Function}=\text{house})=0.666$; $P(\text{Function}=\text{church})=0.333$). Based on this distribution, the analogical description of the entity may be now estimated taking into account the analogical descriptions of the entities with those functions. Notice that this resulting analogical description is probabilistic. Thus, considering the episodic memory presented in Figure 2 and the probability distribution for the function of an entity [$P(\text{Function}=\text{house})=.66$, $P(\text{Function}=\text{church})=.33$], the resulting analogical description is similar to that of entity 4 depicted in Figure 2. This is computed as follows. For all function X : (i) take the analogical description of each possible entity with function X and multiply the occupancy value of each cell by $P(\text{Function}=X)$; (ii) superimpose the analogical descriptions obtained in the previous step summing the occupancy values of the superimposed cells.

2.3.2. Generation and ranking of agent's goals/intentions. The motivational system plays an important role in the generation and ranking of goals/intentions. Actually, according to psychologists, motivations are the source of goals in several manners [29]. These motivations may be themselves the goals (e.g., an agent looks for states of the world that elicit certain positive emotions such as happiness or surprise). Therefore, an agent selects actions or sequences of actions that lead to those states of the world. For instance, in exploratory activity, an agent establishes the goal of visiting an object that seems beforehand interesting (novel, surprising) because visiting it will probably make it feel happy by acquiring a considerable amount of information. The algorithm for the generation and ranking of goals/intentions is as follows. First, the set of different goal tasks present in the memory of plans are retrieved and, for each kind, a set of new goals is generated

using the following procedure: given a goal task retrieved from a plan in the memory of plans, the memory and the perception of the agent, similar goals are generated by adapting the past goal to situations of the present state of the world. The adaptation strategies used are mainly substitutions [14]. Thus, for instance, suppose the goal task $visitEntity(e7)$ is present in the memory of the agent. Suppose also that the agent has just perceived three entities present in the environment, $e1$, $e2$ and $e3$. The entity to which $visitEntity$ is applied ($e7$) may be substituted by $e1$, $e2$ or $e3$, resulting three new goals: $visitEntity(e1)$, $visitEntity(e2)$, $visitEntity(e3)$. Then, the EU of each goal task is computed. As said above, a task T is both conditional and probabilistic (e.g.: [6]). Thus, the execution of a goal task under a given condition may be seen according to utility theory as a lottery [30]:

$$Lottery(T) = \left[p^1 \times p_{1,E_1}^1; p^1 \times p_{2,E_2}^1; \dots; p^m \times p_{n_m,E_{n_m}}^m \right]$$

The EU of T may be then computed as follows:

$$EU(T) = \sum_{k,j} p^k \times p_j^k \times EU(E_j^k)$$

The computation of $EU(E_j^k)$ is performed predicting the motivations that could be elicited by achieving/executing the goal task. This means, the motivations felt by the agent when the effect takes place are predicted or estimated (anticipated) based on the information available in the effect about the changes produced in the world. The following function is used to compute $EU(E_j^k)$:

$$\begin{aligned} EU(E_j^k) &= \\ &= \frac{\alpha_1 \times U_{surprise}(E_j^k) + \alpha_2 \times U_{curiosity}(E_j^k) + \alpha_3 \times U_{hunger}(E_j^k)}{\sum_i \alpha_i} = \\ &= \frac{\alpha_1 \times Surprise(E_j^k) + \alpha_2 \times Curiosity(E_j^k) + \alpha_3 \times Hunger(E_j^k)}{\sum_i \alpha_i} \end{aligned}$$

, where, $\alpha_3 = -1$ and α_i ($i \neq 3$) may be defined as follows:

$$\alpha_i = \begin{cases} 1 \Leftarrow C - HUNGER(Agt) - D > 0 \\ 0 \Leftarrow otherwise \end{cases}$$

, where D is the amount of energy necessary to go from the end location of goal task T to the closer place where energy could be recharged. The functions $Surprise(E_j^k)$, $Curiosity(E_j^k)$ and $Hunger(E_j^k)$ are replaced by the functions of curiosity, surprise and hunger defined above and applied for the entities perceived when the effect E_j^k takes place.

This dependence of the parameters α_i ($i \neq 3$) on the hunger of the agent partially models the results of Berlyne's experiments (e.g.: [5]) that have shown that in the absence of (or despite) known drives, humans tend to explore and investigate their environment as well as seek stimulation. Actually, surprise and curiosity are taken into account to compute the EU of a task only when there is enough energy to go from the end location of goal task T to the closest place where an energy source could be found. Otherwise, only hunger is taken into account for the EU of tasks and further ranking. This means that in this situation (when hunger is above a specific threshold), only the goal of recharge has an $EU > 0$. In the other situations (hunger below a specific threshold), hunger plays the role of a negative reward decreasing the utility of a task by the percentage of energy needed after the task is completed. Thus, the more the distance to the end location of a task the more the energy required and the less the utility of that task.

3. Exploration with Motivational Agents

The goal of exploration is twofold: (i) acquisition of maps of the environment – metric maps – to be stored in memory and where the cells occupied by the entities that populate that environment are represented; (ii) construction of models of those entities.

Each agent is continuously performing the deliberative reasoning/decision-making algorithm. Thus, each agent at a given time senses the environment to look for entities and compute the current world state (location, structure and function of those entities) based on the sensorial information and on the generation of expectations for the missing information. Then, a goal of kind *visitEntity* is generated for each unvisited entity within the visual range. In addition, a goal of the kind *visitLoc* is generated for all the frontier cells [38]. Then, these goals are ranked according to their EU, which is computed based on the intensities of motivations predicted as explained above.

Exploration involving a team of agents results from the individually exploration activity of the elements of the team. Although the exploration strategy of each agent almost does not change, the decisions (which entity/location to visit) are centralized in a single agent of the team – the team leader –, which may be itself involved in the exploration. Each agent perceives the environment, generates goals, rank the goals and create plans for these goals as it was exploring the environment alone. The actual plan and map of each agent are then sent to the team leader in order to be

merged and adjusted if necessary, resulting an overall plan and an overall map. In what regards to the plan, these adjustments are made when the leader notices that in the overall plan two or more agents are assigned to perform the same task. In this case, in order to guarantee that a single agent performs that task, this task is deleted from the plans of the other agents, i.e., those other agents are assigned to another goal of their ranking. With respect to the overall map, the maps received from each agent are superimposed and the occupancy probabilities of correspondent cells are appropriately computed. The overall plan and map are then sent back to each agent. This is of primary importance because this way all the agents became aware of the overall knowledge (plans and maps) of the team. Actually, taking into account the overall map built by the team, an agent avoids exploring parts of the environment that were already explored by other agents and, since the map contains the locations of the other agents, it also avoids mistaking the other agents of the team for other entities of the environment.

4. Experiment

We conducted an experiment to test the role of surprise, curiosity and hunger on the efficiency of the exploration of environments populated with entities. Efficiency is measured in terms of the amount of knowledge acquired in a period of time. This amount of knowledge is measured by the number of entities that were visited (i.e. the number of entity models acquired) and by the number of different entities that were acquired (i.e., the number of different entity models acquired) (notice that two distinct entities may be totally equal except the identifier). To do so, we compared the behavior of 7 agents in 5 simulated environments, each one with a different strategy for exploration: Agent 1 performs undirected exploration (random) [35]; Agent 2 performed directed exploration based on curiosity; Agent 3 performed directed exploration based on curiosity and hunger; Agent 4 performed directed exploration based on surprise; Agent 5 performed directed exploration based on surprise and hunger; Agent 6 performed directed exploration based on curiosity, surprise and hunger; and, Agent 7 performed directed exploration based on hunger. In addition, we compared also the exploration efficiency of 3 teams of agents with 2 (team A), 3 (team B) and 4 (team C) agents that perform directed exploration based on surprise, curiosity and hunger. In order to test efficiency, the simulations were conducted with a time limit. The test environments comprised an

average of 25 entities, with an average of 5 identical entities.

The environment simulator was built so that the agents' program could be transferred to a robot with a few improvements. Therefore, a few simplifications were done: a parameter was defined for the visual range of the agents, i.e., objects out of that range are not visible by agents; agents have optic, sonar and infrared simulated sensors; for the sake of simplicity, the visual perception of entities is confined to the shape of the visible part of the structure; the function of an entity is not accessible or can not be inferred from visual information unless the agent is at the same place of the entity; when an entity is perceived, its propositional description is provided directly by the environment simulator to the viewer agent (note that the agent's architecture does not include a module to transform the visual information into propositional information).

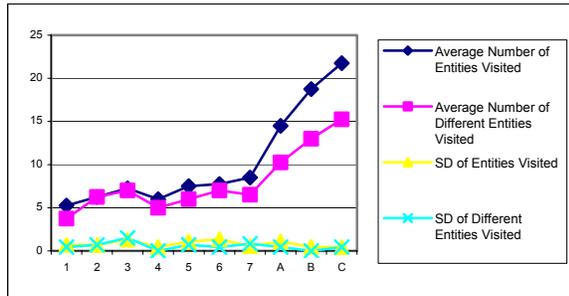


Figure 3. Results of the experiment.

Figure 3 presents the results. As it can be seen, regarding single agent exploration, Agent 3 achieved the best results: it not only visited a large amount of entities but also almost all of them were different. This happened because this agent preferred to visit entities that were both novel and closer (curiosity always directed the agent to novel entities and hunger directed it to the closer ones). Agent 2 only benefited from curiosity. Therefore, it selected always novel entities but their distance was not taken into account and thus it spent most of the time traversing long distances. Consequently, it visited fewer entities. Agent 1 behaved even worse: although the number of entities visited were almost the same, there is a less diversity on them because it randomly selects the entities to visit. Almost similar results were obtained with Agent 4 that took into account solely surprise. Actually, sometimes surprise guided the agent to select entities that were similar to a previous one. However, when hunger was taken into account together with surprise, the number of entities visited increased because of the benefit of selecting the closer ones, although the agent

still visited similar entities. Agent 7 took into account hunger. Therefore, although it visited a relatively large amount of entities, it was not completely successful because similar entities were visited. The relevance of surprise in exploration was better seen in Agent 6 that took into account surprise, curiosity and hunger. Actually, we verified that it is advantageous when the set of entities that could be visited are equally novel and equally distant. In this case the more unexpected one is chosen to be visited.

In what regards to exploration with multiple agents we verified that the more the number of elements of a team the more the number of entities visited. However, although the number of different entity models increases monotonically with the number of elements of the team, this increase is less than that of the number of entity models acquired.

5. Conclusions

We have presented an approach for directed exploration of unknown environments based on surprise, curiosity and hunger. Curiosity might be considered the most relevant of these motivations to exploration. However, surprise and hunger play an important role on exploration, too. Anyway, additional tests are required to evaluate the role of other motivations such as fear on exploratory behavior. Besides, we also concluded that, as in other directed exploration strategies, teams of motivational agents perform better than single motivational agents.

6. Acknowledgements

The PhD of Luís Macedo is financially supported by PRODEP III.

7. References

- [1] A. Aitkenhead and J. Slack, *Issues in cognitive modelling*. London: Lawrence Erlbaum Associates, 1987.
- [2] J. Amat, R. Mántaras, and C. Sierra, "Cooperative autonomous low-cost robots for exploring unknown environments," in *Proceedings of the International Symposium on Experimental Robotics*, 1995.
- [3] D. Anguelov, R. Biswas, D. Koller, B. Limketkai, S. Sanner, and S. Thrun, "Learning hierarchical object maps of non-stationary environments with mobile robots," in *Proceedings of the 17th Annual Conference on Uncertainty in AI*, 2002.
- [4] J. Bates, "The Role of Emotion in Believable Agents," *Communications of the ACM*, vol. 37, pp. 122-125, 1994.

- [5] D. Berlyne, "Novelty and curiosity as determinants of exploratory behavior," *British Journal of Psychology*, vol. 41, pp. 68-80, 1950.
- [6] J. Blythe, "Decision-Theoretic Planning," *AI Magazine, Summer 1999*, 1999.
- [7] W. Burgard, M. Moors, and F. Schneider, "Collaborative exploration of unknown environments with teams of mobile robots," in *Advances in plan-based control of robotic agents*, vol. 4266, *Lectures Notes in Computer Science*, M. Beetz, J. Hertzberg, M. Ghallab, and M. Pollack, Eds. Berlin: Springer Verlag, 2002.
- [8] C. Castelfranchi, R. Conte, M. Miceli, and I. Poggi, "Emotions and goals," in *Perspectives on Cognitive Science*, B. Kokinov, Ed. Sofia: New Bulgarian University, 1996, pp. 131-145.
- [9] A. Damásio, *Descartes'error, Emotion Reason and the Human Brain*. New York: Grosset/Putnam Books, 1994.
- [10] C. Elliott, "The Affective Reasoner: A Process Model of Emotions in a Multi-agent System," Northwestern University, The Institute for the Learning Sciences, Ph.D. Dissertation 32, 1992.
- [11] K. Erol, J. Hendler, and D. Nau, "UMCP: A sound and complete procedure for hierarchical task-network planning," in *Proceedings of the International Conference on AI Planning Systems*, 1994, pp. 249-254.
- [12] R. Hamming, "Error Detecting and Error Correcting Codes," *The Bell System Technical Journal*, vol. 26, pp. 147-60, 1950.
- [13] C. Izard, *The Psychology of Emotions*. NY: Plenum Press, 1991.
- [14] J. Kolodner, *Case-Based Reasoning*. San Mateo, CA: Morgan-Kaufmann, 1993.
- [15] J. LeDoux, *The Emotional Brain*. New York: Simon and Schuster, 1996.
- [16] D. Lee and M. Recce, "Quantitative evaluation of the exploration strategies of a mobile robot," *AAAI-94*, 1994.
- [17] L. Macedo and A. Cardoso, "SC-EUNE - Surprise/Curiosity-based Exploration of Uncertain and Unknown Environments.," in *Proceedings of the AISB'01 Symposium on Emotion, Cognition and Affective Computing*. York, UK: University of York, 2001, pp. 73-81.
- [18] L. Macedo and A. Cardoso, "Modelling Forms of Surprise in an Artificial Agent," in *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, J. Moore and K. Stenning, Eds. Mahwah, NJ: Erlbaum, 2001, pp. 588-593.
- [19] P. Maes, "Artificial Life meets Entertainment: Lifelike Autonomous Agents," *Communications of the ACM. Special Issue on Novel Applications of AI*, 1995.
- [20] W. McDougall, *An introduction to social psychology*. London: Methuen, 1908.
- [21] W. Meyer, R. Reisenzein, and A. Schützwohl, "Towards a process analysis of emotions: The case of surprise," *Motivation and Emotion*, vol. 21, pp. 251-274, 1997.
- [22] E. Oliveira and L. Sarmiento, "Emotional advantage for adaptability and autonomy," in *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*. New York, NY: ACM Press, 2003, pp. 305 - 312.
- [23] A. Ortony and D. Partridge, "Surprisingness and Expectation Failure: What's the Difference?," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*. Los Altos, CA: Morgan Kaufmann, 1987, pp. 106-8.
- [24] A. Ortony, G. Clore, and A. Collins, "The cognitive structure of emotions,". New York: Cambridge University Press, 1988.
- [25] P. Pfeifer, "Artificial Intelligence Models of Emotion," in *Cognitive Perspectives of Emotion and Motivation*, V. Hamilton, G. Bower, and N. Frijda, Eds. Netherlands: Kluwer, 1988, pp. 287-320.
- [26] R. Picard, *Affective Computing*. Cambridge, MA: MIT Press, 1997.
- [27] A. Rao and M. Georgeff, "BDI Agents: From Theory to Practice," in *Proceedings of the First International Conference on Multiagent Systems*, 1995.
- [28] S. Reilly, "Believable Social and Emotional Agents," School of Computer Science, Carnegie Mellon University CMU-CS-96-138, 1996.
- [29] R. Reisenzein, "Emotional Action Generation," in *Processes of the molar regulation of behavior*, W. Battmann and S. Dutke, Eds. Lengerich: Pabst Science Publishers, 1996.
- [30] S. Russel and P. Norvig, *Artificial Intelligence - A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall, 1995.
- [31] J. Schmidhuber, "Curious model-building control systems," in *Proceedings of the International Conference on Neural Networks*, vol. 2: IEEE, 1991, pp. 1458-1463.
- [32] G. Shafer and J. Pearl, "Readings in Uncertain Reasoning,". Palo Alto, CA: Morgan Kaufmann, 1990.
- [33] A. Shand, *The foundations of character*. London: Macmillan, 1914.
- [34] R. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. Younes, "Coordination for Multi-Robot Exploration and Mapping," in *Proceedings of the AAAI-2000*, 2000.
- [35] S. Thrun, "Efficient exploration in reinforcement learning," Carnegie Mellon University, Computer Science Department, Pittsburgh, PA CMU-CS-92-102, 1992.
- [36] S. Thrun, "Exploration in active learning," in *Handbook of Brain Science and Neural Networks*, M. Arbib, Ed., 1995.
- [37] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millenium*, G. Lakemeyer and B. Nebel, Eds. San Mateo, CA: Morgan Kaufmann, 2002.
- [38] B. Yamauchi, "Frontier-based exploration using multiple robots," in *Proceedings of the Second International Conference on Autonomous Agents*. Minneapolis, MN, 1998.