Available online at www.sciencedirect.com

SciVerse ScienceDirect

Cognitive Systems Research 19–20 (2012) 62–87

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

The exploration of unknown environments populated with entities by a surprise–curiosity-based agent

Action editor: Ron Sun

L. Macedo *, A. Cardoso

University of Coimbra, Centre for Informatics and Systems of the University of Coimbra, Polo II, 3030 Coimbra, Portugal

Received 27 January 2012; accepted 22 March 2012
Available online 3 May 2012

Abstract

We describe a Belief-Desire-Intention-like architecture for an explorer agent in which the psychological constructs of surprise and curiosity play an important role in decision-making, particularly in the selection of view-points during the process of exploring unknown environments. Taking into account previous studies about the psychological constructs involved in exploratory behaviour, the agent is equipped in advance with the basic desires for maximal information gain (reduce curiosity), and maximal surprise. However, to reflect Berlyne's theory that says that the tendency to explore the environment occurs in the absence of known drives, we considered also the basic desire for minimal hunger as a representative example of those additional basic desires that can restrain exploration. This surprise–curiosity-based exploration strategy was confronted with a "cold" classical exploration strategy in environments populated with entities. The results of this experiment indicate that the classical strategy outperforms slightly the surprise–curiosity-based one with respect to the exploration performance measures of *the time/energy required to explore all the environment completely*, and *the time/energy required to explore all the entities*. However, the classical strategy was outperformed by the surprise–curiosity-based one with respect to the *time/energy required to explore all different entities*, and consequently, with more evidence, with respect to the number of steps (trips between two entities) required to explore all different entities. This is a valuable result for resource-bounded, active learning agents that benefit from choosing the more informative data from which they learn while ignoring time-consuming/expensive, redundant data. This important result is confirmed by the results of the analysis of the agents' behaviour exhibited along the traversing paths in the environment. The experiment also provided results concerning the robustness of the surprise–curiosity-based approach by assessing the influence of surprise and curiosity in several environments of different complexity and with different amplitudes for the visual field of the agent.

Keywords: Exploration of unknown environments; Surprise; Curiosity; Motivational agents; Belief-Desire-Intention agents

1. Introduction

Exploration of unknown environments by artificial agents (usually mobile robots) has been an active field of research. Exploration may be defined as the process of selecting and executing actions so that the maximal knowledge of the environment is acquired (Thrun, 1992a, 1992b, 1993). The result is the acquisition of models of the physical environment. So, exploration of unknown environments involves map-building but it is not confined to this process. Actually, this kind of exploration can be considered as two distinct topics. First, the agent or robot has to interpret the findings of its sensors so as to make accurate deductions about the state of its environment. This is the problem of map-building. The second but not less important aspect of exploration of unknown

* Corresponding author.
E-mail addresses: macedo@dei.uc.pt (L. Macedo), amilcar@dei.uc.pt (A. Cardoso).

1389-0417/$ - see front matter
http://dx.doi.org/10.1016/j.cogsys.2012.04.003

environments is that the agent (or robot in physical environments) has to select its view-points[1] so that the sensory measurements contain new and useful information. This is the problem of exploration itself. It involves guiding an agent in such a way that it covers the environment with its sensors. The accuracy of the map also depends on this choice of view-points during exploration.

Unfortunately, exploring unknown environments requires resources from agents such as time and power. 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). Several techniques have been proposed and tested either in simulated and real, indoor and outdoor environments, using single (Lee, 1996; Lee & Recce, 1994) or multiple agents (e.g., Amat, Mantaras, & Sierra, 1997; Anguelov et al., 2002; Burgard, Moors, & Schneider, 2002; Simmons et al., 2000; Thrun, 1992a, 1992b, 1993, 1995, 2002; Thrun, Burgard, & Fox, 2000; Thrun et al., 2005; Yamauchi, 1998; Yamauchi, Schultz, & Adams, 1999). The exploration domains include planetary exploration (e.g., Mars, Titan or Lunar exploration) (e.g., Bresina, Dorais, Golden, Smith, & Washington, 1999; Simmons et al., 1995; Washington, Bresina, Smith, Anderson, & Smith, 1999), the search for meteorites in Antarctica (e.g., Moorehead, Simmons, Apostolopoulos, & Whittaker, 1999), underwater mapping, volcano exploration, map-building of interiors (e.g., Thrun, 1997, 2002; Thrun et al., 2005), etc. The main advantage of using artificial agents in those domains instead of humans is that most of them are extreme environments making their exploration a dangerous task for human agents. However, there is still much to be done especially in dynamic environments such as those mentioned above.

Exploration strategies have been grouped into two main categories: undirected and directed exploration (Thrun, 1992a). 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.

In spite of these advances in exploration of unknown environments by mobile robots, artificial explorer agents are far from being perfect. In many aspects, their exploratory behaviour is still far from human exploratory behaviour. For instance, the autonomy of agents still needs to be improved, as happens for instance in planetary exploration which is still too human dependent (the plans are determined by a human operator as well as the interesting points to visit) (Bresina et al., 1999; Washington et al., 1999). Moreover, most of those directed strategies rely on the "cold" maximisation of knowledge gain (e.g., Simmons et al., 2000). On contrary, humans possess fully autonomous, successful exploratory skills that are the product of million years of evolution. These skills enable humans to select view-points according to specific psychological constructs that evaluate the environment and inform other mental components to take action appropriately. They perform exploratory behaviour because the activity itself is interesting and spontaneously satisfying. On contrary to robots or softbots, humans avoid dangerous situations because they can experience fear, select the interesting things to visit because they can experience surprise and curiosity or some sort of interest, remember to "recharge batteries" because they can feel hunger, thirst, etc. Therefore, given their success, it makes sense reproducing those human skills in artificial explorer agents. Particularly, an aspect that may be fostered further is the reproduction of the motivation (James, 1890; Maslow, 1987; McDougall, 1908; Weiner, 1980) behind human exploratory behaviour in artificial agents. Of particular interest here is the specific work on basic desires (Havercamp & Reiss, 2003; Reiss, 2000), values (Schwartz, 1992), and self-determination theory (Ryan & Deci, 2000). Even of more particular interest are the works on the motivations for exploratory behaviour, the main examples being the work of Berlyne (1950, 1955, 1960, 1967, 1971), Butler (1953, 1954, 1957, 1958), Freud (1938), James (1890), McDougall (1908), Montegomery (1952, 1953, 1954, 1955), Nunnally and Lemond (1973), and Shand (1914), all of them in some manner connecting exploratory behaviour with psychological constructs such as surprise, curiosity, and novelty.

Specifically, in this article we investigate the role of a surprise–curiosity-based exploration strategy on the performance of agents exploring unknown environments populated with entities. By making use of this surprise–curiosity-based strategy, a resource-bounded explorer agent, that wants to selectively acquire the knowledge of the different entities that populate the environment and that wants to avoid loosing time visiting and studying one type of entity twice or more times, may avoid exploring the whole environment with similar accuracy as if it would do a complete exploration of the environment. Relying on widely accepted philosophical roots (Bratman, Israel, & Pollack, 1988; Dennett, 1987), the model of action adopted for agents is based on the belief-desire theory of action that inspired one of the most well known and studied software agents' architectures: the Belief-Desire-Intention (BDI) architecture (Bratman et al., 1988; Rao & Georgeff, 1991, 1995). Thereby, the actions of the explorer agent, i.e., the selection of the view-points, are the product of cognitive or informational states (beliefs) and motivational states (desires/goals), in this case goals for moving to the target view-points (locations) of the environment (previously unvisited entities, regions of the environment, and places

[1] Locations of the environment from which an agent can acquire information from its environment through its sensors; although all the locations of the environment could be view-points, these are filtered to the locations close to entities as these are the places from which the agent can get all the information about entities.