The Practical Advantage of Surprise-based Agents

(Extended Abstract)

Luis Macedo
Centre for Informatics and Systems of the University of Coimbra
Polo II, Pinhal de Marrocos
3030 Coimbra, Portugal
macedo@dei.uc.pt

ABSTRACT

We argue that taking surprise into account in the artificial agents's reasoning may have advantageous implications in various situations. Relying on theoretical and empirical evidence, our arguments are supported by the application of surprise-based agents to three different domains, namely exploration of unknown environments, divergent production and evaluation of creative products, and selective attention to travel information.

Categories and Subject Descriptors

I.2.m [**Artificial Intelligence**]: Miscellaneous

General Terms

Experimentation

Keywords

Surprise, Emotion, Intrinsic Motivation, BDI Architecture, Selective Attention, Exploration of Unknown Environments, Creativity, Intelligent Transportation Systems, Anticipation

1. INTRODUCTION

Surprise is considered a mental state that results from unexpected events [10, 9]. In addition to the the general informational and motivational functions of surprise in humans, it has been recognized that surprise plays an important role specifically in cognitive activities, especially in attention focusing [10], learning [1] and creativity [2]. Therefore, it seems reasonable to allow artificial agents to take advantage of surprise. In this paper, we show evidence for this advantage by presenting illustrative practical applications of surprise-based agents to three different domains, namely exploration of unknown environments, divergent production and evaluation of creative products, and selective attention to travel information. The computational model of surprise integrated into the agents is that of Macedo, Cardoso and Reisenzein (e.g., [8]), which, likewise Lorini and Castelfranchi's model (for their comparison see [7]), is influenced by psychological theories of surprise (e.g., [9, 10]), and seek to capture essential aspects of human surprise.

Cite as: The Practical Advantage of Surprise-based Agents (Extended Abstract), Luis Macedo, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1401-1402

2. PRACTICAL APPLICATIONS OF SURPRISE-BASED AGENTS

The architecture that we adopted for an agent (Figure 1) is based on the BDI approach. Of key relevance for the agent's behavior are the basic desires, feelings, goals and plans with which it is equipped.

Following the pluralist view of motivation, the module of basic desires (basic motivations/motives) contains a set of basic desires that drive the agent to reduce or to maximize a particular feeling [6]. The intensity of feelings is, therefore, important to compute the degree of satisfaction of the basic desires.

In this paper we focus on agents that exhibit the basic desire of surprise that directs the agent to feel surprise, i.e., to satisfy that basic desire the agent selects moving to states of the world in which it expects feeling surprise (e.g., in the case of exploration, the agent desires to visit previously unvisited entities, regions of the environment and places where it can feel surprise; in the case of creative production, the agent desires building surprising products; in the case of selective attention to travel information, the agent desires attending to surprising travel information).

The computational model of surprise developed by [8] suggests that the relation between subjective probability and the intensity of surprise about an event E_g from a set of mutually exclusive events $E_1, E_2, \ldots, E_m$ is described by:

$$Surprise(E_g) = \log(1 + P(E_h) - P(E_g))$$

Taking the ideas of [1] into account, and in order to understand, build or model artificial agents that explore like humans do, it might be advantageous for those agents taking into account not only novelty, entropy, but also surprise. In order to prove this we developed an agent equipped with the basic desires for minimal hunger, maximal information gain (reduce curiosity), and maximal surprise that drive the agent to reduce the feeling of hunger, to reduce the feeling of curiosity (by maximizing information gain) and to maximize the feeling of surprise. The agent explores an artificial environment consisting of buildings located at specific positions, that differ in their structural properties (concerning e.g., the shape of the roof, the door, and the windows) and their functions (e.g., home, hotel, church) by selecting those objects that are expected to elicit more surprise, less hunger and that provides maximal information. More details about this application may be found in [3]. The experimental results show evidence for a significant influence of the factor of surprise on agent's exploratory performance.

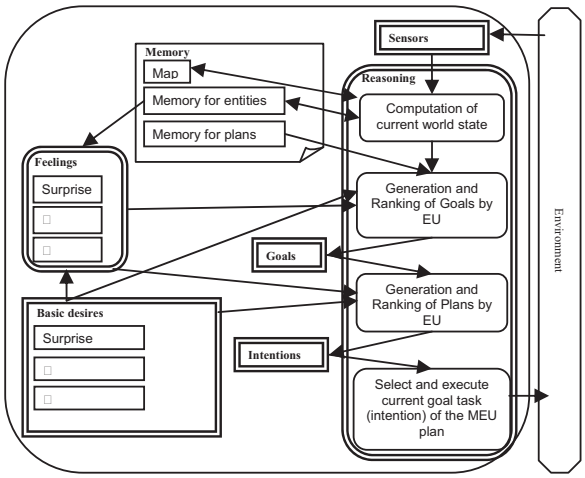

Figure 1: Architecture of a surprise-based agent.

The nature of the link between surprise and creativity is very strong. Boden [2] argued that there is a distinction between mere novelty and creativity: creative products are not only novel but also unpredictable, unexpected and therefore surprising. Creative products amaze us, shock us and delight us mainly because they are unexpected or unpredictable. This critical role of surprise is present in the multi-agent system that we developed and which is composed by surprise-based agents. There are two categories of agents: those that produce creative products and those that evaluate the products that populate the environment. More details about the role played by surprise both in the production of creative products and in the evaluation of creative products may be found, for instance, in [5].

It is generally agreed that surprise and curiosity/interest (that may be elicited by surprise) play an essential role in natural selective attention [1, 9, 10]. Inspired on these natural selective attention studies, we developed an artificial selective attention model [4]. This model is of critical importance in a number of applications. Take the example of complex environments such as urban spaces, in which mobile devices can help humans to perform better by providing them with useful information about other agents, buildings, weather conditions, transportation systems, etc. Furthermore, this information can be shared among different devices to improve the efficiency of urban information systems. However, if the information provided to users becomes too high, instead of being beneficial, it can become a problem. With the expected increase in the number of information devices and the countless kinds of information that will be provided, this problem of information overload is bound to become worse in the future. This is even more problematic because most of the time this information is provided in a way that affects especially the high level natural selective attention, which is involved in strategic cognitive choices such as the preference or shift of a task or activity over another. This means that humans might have to interrupt whatever they are doing to deal with the information provided by those devices. This phenomena is sometimes referred as "Interruption overload" and it is especially problematic (or dangerous) if the human agent is performing critical tasks like driving a car. Our approach to solve this problem involves the integration of the artificial selective attention model on technological devices so that these devices can autonomously select and provide only the relevant information for the human agent, preventing this human agent from a superabundance of information and unnecessary interruptions.

In conclusion, we showed evidence for the advantage of surprise for the reasoning process of agents by referring to empirical and theoretical work in the exploration of unknown environments, creative production and evaluation, and selective attention to travel information. Other potential applications can be explored that could provide additional evidence in favor of our claims that surprise is not an optional extra but a functional necessity by playing an important role in vital mechanisms of resource-bounded agents such as anticipation, selective attention and learning.

3. REFERENCES

[1] D. Berlyne. *Conflict, arousal and curiosity*. McGraw-Hill, New York, 1960.

[2] M. Boden. Creativity and unpredictability. *Stanford Humanities Review*, 4(2):123–139, 1995.

[3] L. Macedo. *The Exploration of Unknown Environments by Affective Agents*. Phd Thesis, University of Coimbra, 2007.

[4] L. Macedo. A computational model of selective attention. Technical Report 2010/01, University of Coimbra, 2010.

[5] L. Macedo and A. Cardoso. Assessing creativity: the importance of unexpected novelty. In *Proceedings of the ECAI'02 Workshop on Creative Systems: Aproaches to Creativity in AI and Cognitive Science,*, pages 31–37. University Claude Bernard - Lyon, Lyon, France, 2002.

[6] L. Macedo and A. Cardoso. Exploration of unknown environments with motivational agents. In N. Jennings and M. Tambe, editors, *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 328 – 335. IEEE Computer Society, New York, 2004.

[7] L. Macedo, A. Cardoso, R. Reisenzein, E. Lorini, and C. Castelfranchi. Artificial surprise. In J. Vallverdú and D. Casacuberta, editors, *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*. IGI Global, Hershey: USA, 2009.

[8] L. Macedo, R. Reisenzein, and A. Cardoso. Modeling forms of surprise in artificial agents: empirical and theoretical study of surprise functions. In K. Forbus, D. Gentner, and T. Regier, editors, *Proceedings of the 26th Annual Conference of the Cognitive Science Society*, pages 873–878. Lawrence Erlbaum Associates, Inc., Chicago, Illinois, USA, 2004.

[9] U. Meyer, R. Reisenzein, and A. Schützwohl. Towards a process analysis of emotions: The case of surprise. *Motivation and Emotion*, 21:251–274, 1997.

[10] A. Ortony and D. Partridge. Surprisingness and expectation failure: what's the difference? In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 106–8. Morgan Kaufmann, Milan, Italy, 1987.