# Arguments for a Computational Model for Forms of Selective Attention based on Cognitive and Affective Feelings

Luis Macedo
University of Coimbra
Centre for Informatics and Systems of the University of Coimbra
Coimbra, Portugal
Email: macedo@dei.uc.pt

***Abstract*—We present arguments for a computational model of selective attention that relies on the assumption that uncertain, surprising and motive congruent/incongruent information demands attention from an intelligent agent. This computational model that we propose has been integrated into the architecture of a Belief-Desire-Intention artificial agent so that this can autonomously select relevant, interesting information of the (external or internal) environment while ignoring other less relevant information. The advantage is that the agent can communicate only that interesting, selective information to its processing resources (focus of the senses, decision-making, etc.) or to its human owner's processing resources so that these resources can be allocated more effectively. We provide both theoretical and empirical arguments for that computational model.**

## I. INTRODUCTION

The advent of information technology is a primary reason for the abundance of information with which humans are inundated, due to its ability to produce more information more quickly and to disseminate this information to a wider audience than ever before. Surprisingly, a lot of recent studies confirmed what Toffler [1] predicted a few decades ago: the overabundance of information instead of being beneficial is a huge problem having many negative implications not only in personal life but also in organizations, business, and in general in the world economy. Research proves that the brain simply does not deal very well with this multitasking process: there is a waste of time as the brain switches from one task to another and back again [2]. This explains why decision quality and the rate of performing tasks degrades with increases in the amount of information being considered.

A fundamental strategy for dealing with this problem of information overload [3] should include making devices that incorporate themselves selective attention agents in order to decrease the amount of information considered in their own reasoning/decision-making processes or decrease the amount of information provided by them to humans, preventing these from a number of interruptions.

But how to model selective attention in artificial agents? Although selective attention has been thoroughly researched over the last 100 years in psychology and more recently in neuroscience (e.g., [4], [5]), at present there is no general theory of selective attention. Instead there are specific theories for specific tasks such as orienting, visual search, filtering, multiple action monitoring (dual task), and multiple object tracking.

In spite of this, a number of models of selective attention has been proposed in Cognitive Science (e.g., [6], [7]). Particularly related with these models is the issue of measuring the value of information. A considerable amount of literature has been published on these measures, especially from the fields of active learning and experimental design. Most of those measures rely on assessing the utility or the informativeness of information (e.g., [8], [9], [10], [11]). However, little attention has been given to the surprising and motive congruence value of information, giving the beliefs and desires of an agent.

Opposed to other approaches (e.g., [12], [13], [14], [15]) relying on low-level, raw information, Macedo, Reisenzein and Cardoso (e.g., [16], [17]), and Lorini and Castelfranchi [18] proposed, independently, computational models of surprise that are based on the mechanism that compares newly acquired beliefs to preexisting beliefs. Both models of artificial surprise were influenced by psychological theories of surprise (e.g., [19]), and both seek to capture essential aspects of human surprise (see for a comparison [20]). In agreement with most theories of human surprise, both models of artificial surprise conceptualize surprise as a fundamentally expectation- or belief-based cognitive phenomenon, that is, as a reaction to the disconfirmation of expectations or, more generally, beliefs. Furthermore, in both models, beliefs are understood as propositional attitudes (e.g., [21]), and a quantitative belief concept (subjective probability) is used. Both artificial surprise models draw a distinction between two main kinds of expectations or beliefs whose disconfirmation causes surprise (see also [22]): active versus passive expectations. Although Macedo and Cardoso initially used the same surprise intensity function, according to which the intensity of surprise about an event is proportional to its unexpectedness, Macedo, Reisenzein and Cardoso subsequently opted for a "contrast model" of surprise intensity. This model assumes that the intensity of surprise about an event reflects its probability difference to the contextually most expected event (see also [23]).

In this paper we describe an artificial selective attention mechanism that may be used by artificial agents so that only cognitively and affectively, interesting/relevant information is selected and forwarded to reasoning/decision-making units. Our approach relies on the psychological and neuroscience studies about selective attention which defend that variables such as unexpectedness, unpredictability, surprise, uncertainty, and motive congruence demand attention (e.g., [24], [4], [22]).

The next section describes the computational model of selective attention, focusing on how the multidimensional value of information is computed. Section 3 presents both theoretical and empirical arguments for that computational model of selective attention. Finally, in Section 4 we present conclusions.

## II. A COMPUTATIONAL MODEL FOR FORMS OF SELECTIVE ATTENTION

Selective attention may be defined as the cognitive process of selective allocation of processing resources (focus of the senses, etc.) on relevant, important or interesting information of the (external or internal) environment while ignoring other less relevant information. The issue is how to measure the value of information. What makes something interesting?

We developed an architecture for a personalized, artificial selective attention agent (see Figure 1). We assume: (i) this agent interacts with the external world receiving from it information through the senses and outputs actions through its effectors; (ii) the world is described by a large amount of statistical experiments; (iii) the agent is a BDI agent [25], exhibiting a prediction model (model for generating expectations, i.e., beliefs about the environment), a desire strength prediction model (a model for generating desire strengths for all the outcomes of the statistical experiments of the world that are known given the desires of the agent – profile of the agent which include basic desires), as well as the intentions (these define the profile of the agent); (iv) the agent contains other resources for the purpose of reasoning and decision-making.

The first of the modules of the architecture (module 1 in Figure 1) is concerned with getting the input information. The second is the computation of the current world state. This is performed by generating expectations or assumptions (module 2), based on the knowledge stored in memory, for the gaps of the environment information provided by the sensors (module 1). We assume that each piece of information resulting from this process, before it is processed by other cognitive skills, goes through several sub-selective attention devices, each one evaluating information according to a certain dimension such as surprise (module 4), uncertainty (module 5), and motive-congruence/incongruence – happiness (module 6). For this task the selective attention mechanism takes into account some knowledge container (memory — preexisting information (module 7)), and the intentions and desires (motives — module 8). There is a decision-making module (module 9) that takes into account the values computed by those sub-selective attention modules and decides if a piece of information is relevant/interesting or not. Then, this module of decision-making selects the more relevant pieces of information so that other resources (reasoning, decision-making, displaying, communication resources, etc.) (module 10) can be allocated to deal with them.

The process of making the right decision depends heavily on a good model of the environment that surrounds agents. This is also true for deciding in which information should the agent focus. Unfortunately, the real world is not crystal clear to agents. Agents almost never have access to the whole environment, mainly because of the incompleteness and incorrectness of their perceptual and understanding components. In fact, it is too much work to obtain all the information from a complex and dynamic world, and it is quite likely that the accessible information suffers distortions. Nevertheless, since the success of agents depends heavily on the completeness of the information of the state of the world, they have to pursue alternatives to construct good models of the world even (and especially) when this is uncertain. According to psychologists, cognitive scientists, and ethologists [26], [27], humans and, in general, animals attempt to overcome this limitation through the generation of assumptions or expectations to fill in gaps in the present or future observational information. When the missing information, either of the present state of the world or of the future states of the world, becomes known to the agent, there may be an inconsistency or conflict between it and the assumptions or expectations that the agent has. As defended by Reisenzein [28], Gardenfors [29], Ortony and Partridge [22], etc., the result of this inconsistency gives rise to surprise which in our model of selective attention and according to previous studies plays a central role in selective attention. It also gives rise to the process of updating beliefs, called belief revision (e.g., [30]).

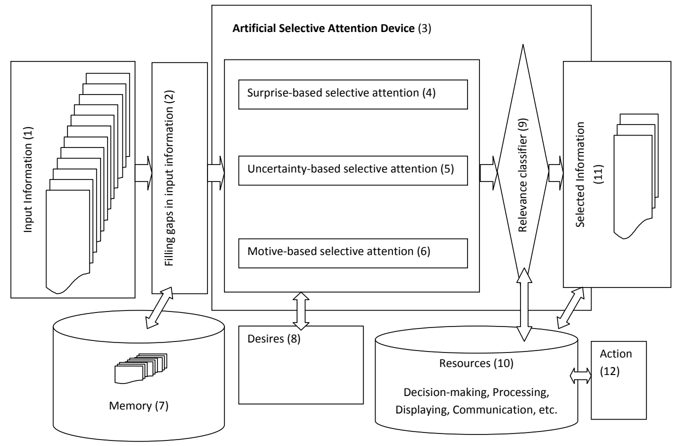


Fig. 1. Architecture of an artificial selective attention agent.

The representation of the memory contents (beliefs) relies on semantic features or attributes much like in semantic networks [31] or schemas [32]. Each attribute, $attr_i$, viewed by us as a statistical experiment, is described by a probabilistic distribution, i.e., a set $A_i = \{< value_j, prob_j, desireStrength_j >: j = 1, 2, \ldots, n\}$, where $n$ is the number of possible values of the attribute, $P(attr_i = value_j) = prob_j$, and $desireStrength_j$ is the desirability of $attr_i = value_j$ (for a related work see [33]).

While the belief strengths are inferred from data using a frequentist approach and updated as new information is acquired, the desirability of the outcomes can be previously set up or learned based on the intentions and contexts of the agent on which it depends, suffering changes whenever the agent is committed with a new intention and/or in a new context. For modelling this dynamics, we make use a desire strength prediction model, i.e., a model for generating desire strengths for all the outcomes of the statistical experiments of the world that are know given the desires of the agent, the intentions, as well as the context of the user (for more details see [34]). As seen before, the desire strength is associated with each attribute together with the belief strength.

Much like the motivation system of Clarion [35], the module of desires encompasses explicit (goals) and implicit motives (basic desires). Following the pluralist view of motivation [36], [37], [38], [39], the sub-module of basic desires (basic motivations/motives) contains a set of basic desires that drive the behaviour of the agent by guiding the agent to reduce or to maximize a particular feeling [40]. Among the basic desires we can find surprise and curiosity.

The module of feelings receives information about a state of the environment and outputs the intensities of feelings. Following Clore [41], we include in this module affective, cognitive, and bodily feelings. The latter two categories are merged to form the category of non affective feelings. This means that this module is much broader than a module of emotion that could be considered. Feelings are of primary relevance to influence the behavior of an agent, because computing their intensity the agent measures the degree to which the desires are fulfilled. In this paper, we highlight the feelings of surprise and pleasantness/unplesantness.

Although the architecture of the computational model of selective attention includes all those above-mentioned sub-selective attention modules, we reserve some room in the architecture of the model for other sub-selective attention components, such as coping potential, complexity.

The next sub-sections describe each one of the dimensions for evaluating information, namely surprise, uncertainty, and motive congruence/incongruence. While the dimensions of surprise and uncertainty are related to the value of information to the belief store of the agent, the dimension of motive congruence/incongruence is related to the value of information to the goals/desires of the agent (these dimensions are related to the concepts of cognitive and affective feelings of [41] and belief-belief and belief-desire comparators of [33]).

### A. *Surprise Value of Information*

We adopted the computational model of surprise of [16], [17] which is formally defined in Definition 1 (for related models see [20]). Macedo, Cardoso and Reisenzein computational model of surprise suggests that the intensity of surprise about an event $E_g$, from a set of mutually exclusive events $E_1, E_2, \ldots, E_m$, is a nonlinear function of the difference, or contrast, between its probability and the probability of the highest expected event $E_h$ in the set of mutually exclusive events $E_1, E_2, \ldots, E_m$.

*Definition 1:* Let $(\Omega, A, P)$ be a probability space where $\Omega$ is the sample space (i.e., the set of possible outcomes of the experiment), $A = A_1, A_2, .., A_n$ is a $\sigma$-field of subsets of $\Omega$ (also called the event space, i.e., all the possible events), and $P$ is a probability measure which assigns a real number $P(F)$ to every member $F$ of the $\sigma$-field $A$. Let $E = \{E_1, E_2, \ldots, E_m\}$, $E_i \in A$, be a set of mutually exclusive events in that probability space with probabilities $P(E_i) >= 0$, such that $\sum_{i=1}^{m} P(E_i) = 1$. Let $E_h$ be the highest expected event from $E$. The intensity of surprise about an event $E_g$ from $E$ is given by:

$$S(E_g) = \log(1 + P(E_h) - P(E_g)) \tag{1}$$

The probability difference between $P(E_h)$ and $P(E_g)$ can be interpreted as the amount by which the probability of $E_g$ would have to be increased for $E_g$ to become unsurprising.

*Proposition 1:* In each set of mutually exclusive events, there is always at least one event whose occurrence is unsurprising, namely, $E_h$.

### B. *Uncertainty-based Value of Information*

Information is a decrease in uncertainty which, according to information theory, is measured by entropy [42]. When new information is acquired its amount may be measured by the difference between the prior uncertainty and the posterior uncertainty.

*Definition 2:* Let $(\Omega, A, P_{prior})$ be a probability space where $\Omega$ is the sample space (i.e., the set of possible outcomes of the experiment), $A = A_1, A_2, .., A_m$ is a $\sigma$-field of subsets of $\Omega$ (also called the event space, i.e., all the possible events), and $P_{prior}$ is a probability measure which assigns a real number $P_{prior}(F)$ to every member $F$ of the $\sigma$-field $A$. Let $E = \{E_1, E_2, \ldots, E_m\}$, $E_i \in A$, be a set of mutually exclusive events in that probability space with probabilities $P_{prior}(E_i) >= 0$, such that $\sum_{i=1}^{m} P_{prior}(E_i) = 1$. Let $P_{post}$ be the posterior probability measure, after some data is acquired, which assigns a real number $P_{post}(F)$ to every member $F$ of the $\sigma$-field $A$ such that it assigns $P_{post}(E_i) >= 0$ with $\sum_{i=1}^{m} P_{post}(E_i) = 1$. According to information theory, the information gain of an agent after some data is acquired, $IG(E)$, is given by the decrease in uncertainty:

$$\begin{aligned} IG(E) &= H_{prior}(E) - H_{post}(E) \\ &= -\sum_{i=1}^{m} P_{prior}(E_i) \times \log(P_{prior}(E_i)) - \\ &\quad (-\sum_{i=1}^{m} P_{post}(E_i) \times \log(P_{post}(E_i))) \end{aligned} \tag{2}$$

$H_{post} = 0$ if and only if all the $P_{post}(E_i)$ but one are zero, this one having the value unity. Thus only when we are certain of the outcome does $H_{post}$ vanish, otherwise it is positive.

$IG$ is not normalized. In order to normalize it we must divide it by $\log(m)$ since it can be proved that $IG \leq \log(m)$:

$$IG(E) = \frac{H_{prior}(E) - H_{post}(E)}{\log(m)} \tag{3}$$

### C. *Motive Congruence/Incongruence-based Value of Information*

While the measure of surprise takes into account beliefs that can be confirmed or not, the pleasantness function that we describe in this subsection takes as input desires that, contrary to beliefs, can be satisfied or frustrated. Following the belief-desire theory of emotion [33], we assume that an agent feels happiness if it desires a state of affairs (a proposition) and firmly beliefs that that state of affairs obtains. The intensity of happiness about an event is a monotonically increasing function of the degree of desire of that event as formally defined in Definition 4.

*Definition 3:* Let $(\Omega, A)$ be a measurable space where $\Omega$ is the sample space (i.e., the set of possible outcomes of the experiment) and $A = A_1, A_2, .., A_m$ a $\sigma$-field of subsets of $\Omega$ (also called the event space, i.e., all the possible events). We define the measure of desirability of an event on $(\Omega, A)$ as $D : A \rightarrow [-1, 1]$, i.e., as a signed measure which assigns a real number $-1 \leq D(F) \leq 1$ to every member $F$ of the $\sigma$-field $A$ based on the profile of the agent, so that the following properties are satisfied:

- $D(\emptyset) = 0$
- if $A_1, A_2, \ldots$ is a collection of disjoint members of $A$, in that $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

$$D(\bigcup_{i=0}^{\infty} A_i) = \sum_{i=0}^{\infty} D(A_i) \quad (4)$$

The triple $(\Omega, A, D)$ is called the desirability space.

*Definition 4:* Let $(\Omega, A, P)$ and $(\Omega, A, D)$ be the probability and the desirability spaces described, respectively, in Definition 1 and Definition 3. Let $E = \{E_1, E_2, \ldots, E_m\}$, $E_i \in A$, be a set of mutually exclusive events in that probability space with probabilities $P(E_i) >= 0$, $\sum_{i=1}^{m} P(E_i) = 1$. If $P(E_g) = 1$, the intensity of happiness, i.e., motive congruence, about an event $E_g$ from $E$ is given by:

$$MC(E_g) = D(E_g) \quad (5)$$

### D. *The Principle of Selective Attention*

Having defined the motive, the uncertainty-based, and surprise-based selective attention modules, we are now in a position to formulate, in a restricted sense (without the inclusion of other information measures such as complexity), the principle that a resource-bounded rational agent should follow in order to avoid an overabundance of information and interruptions in the absence of a model for decision-making. Note that if this model is known, the problem is reduced to the classical computation of the value of information that has been extensively studied (e.g., [8], [31]).

*Definition 5:* A resource-bounded rational agent should focus its attention only on the relevant and interesting information, i.e., on information that is congruent or incongruent to its motives/desires, and that is cognitively relevant because it is surprising or because it decreases uncertainty.

We may define real numbers $\alpha$, $\beta$, and $\gamma$ as levels above which the absolute values of motive congruency, surprise, and information gain (decrease of uncertainty), respectively, should be so that the information can be considered valuable or interesting. These are what we called the triggering levels of alert of the selective attention mechanism. Note that, making one of those parameters null is equivalent to removing the contribution of the corresponding component from the selective attention mechanism (for a different approach see Martinho and Paiva's attention grabbing mechanism [7] which main feature is not relying on tuned parameters but on expectation and prediction error).

## III. Theoretical and Empirical Arguments

The value of a theory is a function of the problems it resolves and the phenomena it explains. The following arguments for the computational model described in this paper permits to assess better its value.

### A. *Theoretical Arguments*

In cognitive science, attentional focus is linked with expectation generation and failure, i.e., with surprise [22]. Therefore, it is reasonable to consider that any model of selective attention should rely on a cognitive model of surprise. This justifies the presence of such a model in the selective attention model proposed in this paper.

However, surprise is not enough. Happiness/pleasantness, which according to cognitive theories of emotion and specifically to belief-desire theories of emotion [33] is directly related to congruence between new information and the human agent's motives/desires, may also play a fundamental role on attention [43], [44], [45]. This ideas are matched by our computational model of selective attention by incorporating a measure of the expected satiation of the desires, i.e., the expected reward or utility of the information for a specific human agent, based on her/him particular intentions and desires at hand.

Moreover, it is generally agreed that surprise and curiosity/interest play an essential role in selective attention [46], [4], [19], [22], [14], [47]. In fact situations that include novelty (different, unfamiliar), incongruity, unpredictability, surprise, uncertainty, change, challenging and complexity (hard to process, challenging, mysterious) (e.g., [46], [24], [48], [49], [50], [43], [44], [45]) certainly demands greater attention than a stimulus distinguished by none of these properties. More precisely, these properties are also those assigned to situations that cause curiosity [46], [19]. However, among those variables, only novelty and uncertainty are taken into account in our selective attention model. In fact, while novelty and uncertainty are captured by the uncertainty-based and surprise-based measures of information, complexity and coping potential are not yet considered, although there is room for them in the present model. This is therefore part of our future work.

Besides being consistent with theoretical work, the computational model of selective attention might be used to explain the basis of other cognitive abilities of the agent in that it decides in which information those other cognitive abilities should focus. In fact, the selective attention model selects the more relevant pieces of information so that other resources (reasoning, decision-making, displaying, communication resources, etc.) (module 10) can be allocated to deal with them.

The selective attention model also explains other phenomena in various domains. To illustrate, consider the traffic information domain in which we are applying our model.

Our Intelligent Travel Information System receives information about the traffic conditions and sends it to the mobile devices of the travelers. All that collected information is stored in the knowledge base/memory of the system. There is a personal selective attention agent for each registered traveler. Each one of these personal agents has information about the expectations of its owner based on their travel history.

Let us illustrate how the value of information is computed by the selective attention mechanism. Suppose that a traveller's navigation system provided the pre-route path containing a road $A$ for an agent (a driver) based on its profile (e.g., preference for shortest routes). Suppose the agent has the following expectations for the traffic conditions of road $A$, for a certain period/time of the day for a certain day of the week: 60% of probability of "good traffic conditions" (event $E_1$), 30% of probability of "moderate traffic conditions" (event $E_2$), and 10% of probability of "bad traffic conditions" (event $E_3$). Suppose the desire strengths of these events are 1, -0.5, and -1, respectively. Given that the agent plans to go trough that route, suppose its module for generating/managing desires assigns a null desire strength for the other routes as it does not care about the traffic conditions of the other roads that are not part of its planned route. What is the relevance of becoming aware that the current traffic conditions of road A are good (event $E_1$)? Considering solely the motive-based component, the outcomes (events $E_1$, $E_2$, and $E_3$) elicits happiness (motive congruence) with intensity 1, -0.5 and -1, respectively. $E_1$ is congruent/consistent with the goals of the agent, while $E_2$ and $E_3$ are incongruent with the goals of the agent.

According to Equation 1, the surprise value of $E_1$, $E_2$, and $E_3$ are, respectively, 0, 0.38, and 0.58. Illustrating for the case of $E_3$:

$$\begin{aligned} Surprise(E_3) &= \log(1 + P(E_1) - P(E_3)) \\ &= \log(1 + 0.6 - 0.1) = 0.58 \end{aligned} \tag{6}$$

According to Equation 3, the normalized information gain value of $E_1$, $E_2$, or $E_3$ is:

$$\begin{aligned} IG(E) &= \frac{H_{prior}(E) - H_{post}(E)}{\log(m)} = \frac{H_{prior}(E) - 0}{\log(3)} \\ &= \frac{-\sum_{i=1}^{3} P_{prior}(E_i) \times \log(P_{prior}(E_i))}{\log(3)} \\ &= 0.82 \end{aligned} \tag{7}$$

Assume the Principle of Selective Attention described above, with parameters $\alpha = 0.3$, $\beta = 0.5$, and $\gamma = 0.6$. Are all these events interesting? Considering the motive-based component all those events are interesting. However, from the perspective of the surprise-based selective attention component, the answer is "no" to the question related with the events $E_1$ and $E_2$ in that their surprise values, 0 and 0.38, respectively, are below $\beta$. With respect to $E_3$ the answer is "yes" given that its surprise value is 0.58. Taking the uncertainty-based component into account, the answer is "yes" for all the events because their occurrence gives a normalized information gain of 0.82 which is above $\gamma$.

By enabling the possibility of considering different values for the parameters $\alpha$, $\beta$, and $\gamma$, the model permits the existence of different selective attention models. Thus, if the parameters are learned for a specific person (e.g., by using machine learning techniques), it is possible to model the subjectivity of selective attention models of different persons.

### B. *Empirical Arguments*

We applied the selective attention model to the domain of traveling information. As mentioned before, we are developing an Intelligent Travel Information System in which there is a personal selective attention agent for each registered traveler. Each one of these personal agents has information about the expectations of its owner based on their travel history. There is also a master agent whose main function is gathering travel information from various sources (e.g., Points of Interest (POI) databases, real-time level of traffic, Google API) and make it available to the personal agents.

We did an exploratory study in order to compare the value of information computed by the selective attention agent in the three dimensions described above and the value of interest assigned by humans to traffic information and POI recommendation. While the value of interest rated by humans is of subjective nature, the value of interest computed by the artificial selective attention agents is based rigorously on expectations computed from statistical data collected from previous travelling situations. The artificial agent used Equations 1, 3, and 5 to compute the cognitive and affective value of information about traffic events. The results provide evidence of the value of the model (for more details see [34]).

## IV. Conclusion

We presented a computational model for selective attention based on cognitive and affective feelings. We found theoretical and empirical evidence indicating that the mechanism models to a certain extent human selective attention. The integration of such mechanism in artificial agents opens a wide range of applications. In the traveling information domain that we mentioned as in others, the selective attention mechanism may contribute for decreasing the amount of unnecessary information while maintaining acceptable the performance of the owner (a human).

## Acknowledgment

Work funded by Fundação para a Ciência e Tecnologia — Project PTDC/EIA-EIA/108675/2008, and by FEDER through Programa Operacional Factores de Competitividade do QREN — COMPETE:FCOMP-01-0124-FEDER-010146.

## References

[1] A. Toffler, *Future Shock*. Bantam Book, 1970.

[2] T. Klingberg, *The overflowing brain: information overload and the limits of working memory*. Oxford University Press, 2008.

[3] N. O'Connell, "Interruption overload," *Strategic Direction*, vol. 24, no. 10, pp. 3–5, 2008.

[4] D. Kahneman, *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[5] R. D. Wright and L. M. Ward, *Orienting of Attention*. Oxford, UK: Oxford University Press, 2008.

[6] E. Horvitz, A. Jacobs, and D. Hovel, "Attention-sensitive alerting," in *Proceedings of the Fifteenth Conference on Uncertainty and Artificial Intelligence*. Morgan Kaufmann, 1999, pp. 305–313.

[7] C. Martinho, "Using anticipation to create believable behaviour," in *Proceedings of the Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*. AAAI Press, 2006, pp. 175–180.

[8] E. Horvitz and M. Barry, "Display of information for time-critical decision-making," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 1995, pp. 296–305.

[9] D. MacKay, "Information-based objective functions for active data selection," *Neural Computation*, vol. 4, no. 4, pp. 590–604, 1992.

[10] D. Lindley, "On a measure of information provided by an experiment," *The Annals of Mathematical Statistics*, pp. 986–1005, 1955.

[11] B. Settles, "Curious machines: Active learning with structured instances," Ph.D. dissertation, University of Wisconsin, Madison, USA, 2008.

[12] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," *Advances in Neural Information Processing Systems (NIPS 2005)*, vol. 19, pp. 1–8, 2006.

[13] M. Peters, "Towards artificial forms of intelligence, creativity, and surprise," in *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Madison, Wisconsin, USA: Erlbaum, 1998, pp. 836–841.

[14] J. Schmidhuber, "Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts," *Connection Science*, vol. 18, pp. 173–187, 2006.

[15] P. Oudeyer, F. Kaplan, and V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 2, pp. 265–286, 2007.

[16] 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. Edinburgh, Scotland, UK: Erlbaum, 2001, pp. 588–593.

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

[18] E. Lorini and C. Castelfranchi, "The cognitive structure of surprise: looking for basic principles," *Topoi: An International Review of Philosophy*, vol. 26, no. 1, pp. 133–149, 2007.

[19] W. U. 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.

[20] L. Macedo, A. Cardoso, R. Reisenzein, E. Lorini, and C. Castelfranchi, "Artificial surprise," in *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*, J. Vallverdú and D. Casacuberta, Eds. Hershey: USA: IGI Global, 2009, pp. 267–291.

[21] J. Searle, *Intentionality*. Cambridge: Cambridge University Press, 1983.

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

[23] K. H. Teigen and G. B. Keren, "Surprises: Low probabilities or high contrasts?" *Cognition*, vol. 87, pp. 55–71, 2003.

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

[25] A. Rao and M. Georgeff, "BDI agents: from theory to practice," in *Proceedings of the First International Conference on Multiagent Systems*. San Francisco, CA, USA: MIT Press, 1995, pp. 312–319.

[26] C. Kline, "Observation-based expectation generation and response for behavior-based artificial creatures," MSc Thesis, Massachusetts Institute of Technology, 1999.

[27] J. Piaget, *The origins of intelligence in children*. New York: International Universities Press, 1952.

[28] R. Reisenzein, "The subjective experience of surprise," in *The message within: The role of subjective experience in social cognition and behavior*, H. Bless and J. Forgas, Eds. Philadelphia, PA: Psychology Press, 2000.

[29] P. Grdenfors, "The role of expectations in reasoning," in *Knowledge Representation and Reasoning Under Uncertainty*, M. Masuch and L. Polos, Eds. Berlin: Springer-Verlag, 1994, pp. 1–16.

[30] ——, "Belief revision: An introduction," in *Belief Revision*, P. Grdenfors, Ed. Cambridge, UK: Cambridge University Press, 1992, pp. 1–20.

[31] S. Russell and P. Norvig, *Artificial intelligence - a modern approach*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 2010.

[32] D. Rumelhardt and A. Ortony, "The representation of knowledge in memory," in *Schooling and the Acquisition of Knowledge*, R. Anderson, R. Spiro, and W. Montague, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977, pp. 99–135.

[33] R. Reisenzein, "Emotions as metarepresentational states of mind: Naturalizing the belief-desire theory of emotion," *Cognitive Systems Research*, vol. 9, 2008.

[34] H. Costa, B. Furtado, D. Pires, L. Macedo, and A. Cardoso, "Context and Intention-Awareness in POIs Recommender Systems," in *Proceedings of the $6^{th}$ ACM Conf. on Recommender Systems (RecSys'12), $4^{th}$ Workshop on Context-Aware Recommender Systems (CARS'12)*. ACM, 2012.

[35] R. Sun, "The motivational and metacognitive control in clarion," in *Modeling Integrated Cognitive Systems*, W. Gray, Ed. Oxford University Press, 2007.

[36] S. M. Havercamp and S. Reiss, "A comprehensive assessment of human strivings: testretest reliability and validity of the Reiss Profile," *Journal of Personality Assessment*, vol. 81, p. 123 132, 2003.

[37] W. McDougall, *An introduction to social psychology*. London: Methuen, 1908.

[38] S. H. Schwartz, "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries." in *Advances in experimental social psychology*, M. P. Zanna, Ed. NY: Academic Press, 1992, vol. 25, pp. 1–65.

[39] B. Weiner, *Human motivation*. New York: Holt, Rinehart & Winston, 1980.

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

[41] G. Clore, "Cognitive phenomenology: Feelings and the construction of judgment," in *The Construction of Social Judgments*, L. Martin and A. Tesser, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, pp. 133–163.

[42] C. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.

[43] P. J. Silvia, "Interest - the curious emotion," *Current Directions in Psychological Science*, vol. 17, pp. 57–60, 2008.

[44] ——, *Exploring the psychology of interest*. New York: Oxford University Press, 2006.

[45] ——, "What is interesting? exploring the appraisal structure of interest," *Emotion*, vol. 5, pp. 89–102, 2005.

[46] D. Berlyne, "Arousal and reinforcement," in *Nebraska Symposium on Motivation*, D. Levine, Ed. Lincoln, Nebraska: University of Nebraska Press, 1967, vol. 15, pp. 1–110.

[47] J. Schmidhuber, "Curious model-building control systems," in *Proceedings of the International Conference on Neural Networks*. Singapore: IEEE, 1991, vol. 2, pp. 1458–1463.

[48] D. Berlyne, "The arousal and satiation of perceptual curiosity in the rat," *Journal of Comparative and Physiological Psychology*, vol. 48, pp. 238–246, 1955.

[49] ——, "Novelty and curiosity as determinants of exploratory behavior," *British Journal of Psychology*, vol. 41, no. 1, pp. 68–80, 1950.

[50] D. E. Berlyne, *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts, 1971.