Reinforcement Sensitivity Theory and Cognitive Architectures

Karl Fua[1,2], **Ian Horswill**[1], **Andrew Ortony**[1,2] and **William Revelle**[1]
karl.fua@gmail.com, {ian,ortony,revelle}@northwestern.edu
[1]Northwestern University, Evanston, Illinois, USA
[2]Computational Cognition for Social Systems, Institute of High Performance Computing,
Agency for Science, Technology and Research, Singapore

Abstract

Many biological models of human motivation and behavior posit a functional division between those subsystems responsible for approach and avoidance behaviors. Gray and McNaughton's (2000) revised Reinforcement Sensitivity Theory (RST) casts this distinction in terms of a Behavioral Activation System (BAS) and a Fight-Flight-Freeze System (FFFS), mediated by a third, conflict resolution system - the Behavioral Inhibition System (BIS). They argued that these are fundamental, functionally distinct systems. The model has been highly influential both in personality psychology, where it provides a biologically-based explanation of traits such as extraversion and neuroticism, and in clinical psychology wherein state disorders such as Major Depressive Disorder and Generalized Anxiety Disorder can be modeled as differences in baseline sensitivities of one or more of the systems. In this paper, we present work in progress on implementing a simplified simulation of RST in a set of embodied virtual characters. We argue that RST provides an interesting and potentially powerful starting point for cognitive architectures for various applications, including interactive entertainment, in which simulation of human-like affect and personality is important.

Keywords: *Reinforcement Sensitivity Theory, Personality, Emotions, Cognitive Architecture, Simulation*

Introduction

Reinforcement Sensitivity Theory (Gray 1970; Gray and McNaughton 2000) was originally developed as a theory of fear and anxiety in rats. One of its goals was to explain individual differences, such as the fact that some rats are relatively confident in large open spaces (where predators are a greater risk) while others are quite timid. RST models the rat behavioral system as containing separate systems for approach (the Behavioral Approach System, or BAS), and avoidance (the Fight-Flight-Freeze System, or FFFS), together with a third system for assessing risk and resolving conflicts between systems (the Behavioral Inhibition System, or BIS). Individual differences can be modeled in terms of differences in the sensitivities of the component systems to rewarding or punishing signals.

However, although initially developed as an animal model, over the years, RST has become increasingly influential in both personality and clinical psychology (e.g. Elliot and Thrash 2002; Depue and Collins 1999; Revelle 2008), the argument being that sensitivity to cues for reward and sensitivity to cues for punishment form two fundamental dimensions of personality, which are often considered to provide the biological basis for Eysenck's Extraversion-Neuroticism dimensions of personality. Extensive research has also been conducted to obtain reliable measures of the sensitivity of each system (BIS and BAS in particular) in humans and to explore how these sensitivities relate to behavior (Torrubia et al. 2001; Carver and White 1994; Ávila and Parcet 2002; Smillie and Jackson 2005). In clinical psychology, RST provides a biological model for the distinctions between Major Depressive Disorder, Generalized Anxiety Disorder, and Panic Disorder, as well as the effects of different pharmacological interventions on them. Panic Disorder and Generalized Anxiety Disorder are thought to correspond to over-activity in the FFFS (panic) system and BIS (anxiety) systems, respectively, whereas Major Depressive Disorder corresponds to under-activity in the approach (BAS) system (Zinbarg and Yoon 2008).

Reinforcement Sensitivity Theory and Artificial Intelligence Systems

RST has a number of features of particular interest to AI. First, the way in which rewards and punishments are handled in RST is very different from the way in which they generally are handled in AI systems. Reinforcement learning in AI systems emphasizes the maximization of rewards, and the outcomes of actions are represented in terms of a single, signed utility value. For example, Q-learning (Watkins 1989) uses observed short-term rewards of actions to estimate the expected (mean) long-term reward, and then chooses actions that maximize expected long-term reward. Multiple reward functions can also be incorporated into reinforcement learning (Whitehead, Karlsson, and Tenenberg 1992), and are used in various AI and/or robotics systems (e.g., Konidaris and Barto 2006; Oudeyer, Kaplan, and Hafner 2007). However, these systems all treat punishment simply as negative reward. In contrast, the animal system as proposed by RST makes an architectural distinction between rewards and punishments; it represents them separately and handles them with distinct behavioral subsystems (the BAS and FFFS, respectively). The BAS is sensitive to all signals of positive reinforcement, including signals of rewards and signals of relief or safety (Wilson, Barrett, and Gray 1989; Smillie and Jackson 2005), while the FFFS is sensitive to all signals of punishment (negative reinforcement). Situations are assessed in terms of two unipolar (unsigned) scales of expected positive outcomes and expected negative outcomes, rather than a single bipolar (signed) scale representing net outcome. In effect, the probability distribution on utility is represented as a two-parameter family rather than a one parameter family. Therefore, unlike reinforcement learning in AI, where the system chooses actions based on the ordering of their utilities without much consideration of their actual values (the highest activation, or most rewarding action, wins), RST takes into account both ordering and absolute values, which means, among other things, that it can distinguish indifference, in which an outcome is not expected to incur significant reward or punishment, from ambivalence, in which both are expected to be possible.

Most computational models of emotions such as (Gratch and Marsella 2005; Canamero 1997; Hudlicka 2004) focus on a separate appraisal process for emotion, whereas RST provides a biological basis for the occurrence of at least some emotions in direct response to external stimuli, by means of separate processes in the architecture. In particular, RST leads to a principled distinction between fear and anxiety – two emotions which emotion theorists often treat as being members of the same family. According to RST, fear and anxiety are subserved by distinct neural substrates, and modulated by distinct drugs. Fear (or panic) is associated with the activation of the avoidance system (the FFFS) when a threat, such as the presence of a predator, is immediate and must be avoided. Based on the concept of defensive distance and defensive avoidance (Blanchard and Blanchard 1990) derived from ethoexperimental studies on rats, activation of the FFFS results in escape behavior, when possible, and freezing or fighting when escape is unavailable. Although the FFFS behaviors are highly reactive, defensive distance can be viewed as one of the many dimensions on which coping potential can be evaluated at a higher, more cognitively driven level. The behaviors resulting from the FFFS can be stimulated through the use of a class of drugs known as panicogenics, and suppressed through the use of another class of drugs, known as panicolytics. Anxiety, by contrast, is correlated with the activation of the Behavioral Inhibition System (BIS) due to conflict between or within the approach (BAS) and avoidance (FFFS) systems. Conflicts result in the inhibition of the BAS and/or FFFS, and the activation of a defensive (cautious) approach behavior related to information gathering and risk assessment. These behaviors are stimulated and inhibited by their own sets of drugs, the anxiogenics and anxiolytics, respectively. The defensive approach behaviors in rats typically involve a heightened sense of the environment, and posture and movement changes. The risk assessment process could also motivate more cognitive processes such as memory retrieval, rumination, and planning. However this relation of risk assessment to cognitive processes has not yet been studied in RST.

Another interesting feature is that even though punishment cues do not always result in fight, flight, or freeze behaviors, the FFFS system nevertheless mediates responses to such cues. Thus a situation, such as realizing that one has forgotten an important deadline, may induce the sympathetic effects on metabolism associated with the fight-or-flight response, such as elevated heart rate and sweating, even though it would be unlikely to produce any inclination toward fighting or fleeing (although, interestingly, it might cause one to freeze up in some sense).

We believe that RST is an interesting theory for AI researchers to examine, partly because it represents a very different functional organization of the cognitive-affective-behavioral nexus from those usually embodied in AI systems. In addition, for applications which specifically need to simulate human affect and personality, it provides one of the better accounts of how at least some aspects of those phenomena might be implemented computationally in humans and other animals. Entertainment systems, such as computer games and interactive drama, are particularly interesting as potential application domains for an RST-based architecture because of the importance in such domains of making characters that have personalities and flaws that an audience will experience as authentic. Here, RST can provide ready-made knobs (parameters) that relate to at least a few widely accepted personality traits and affective states in understandable ways, along with an extensive literature on how different sensitivities of the systems in RST relate to different aspects of human behavior.

That said, it is important to recognize that RST was never intended as a cognitive architecture in the sense generally used in AI and cognitive science. Building a full-blown RST-based architecture for AI purposes will eventually require, among other things, solving the problem of how to integrate the approach/avoid/inhibit organization, which enables the modeling of behaviors such as feeding and predator avoidance, with higher-level deliberative processes. However, we think that a first step is to develop a simulation of only the core RST architecture. We now describe how we have done this in a set of virtual characters.

A Preliminary Implementation of RST

The simulation we have developed, which uses the Twig engine (Horswill 2009) to animate a set of virtual characters, focuses on the main features of RST, namely, the three systems that handle rewards, punishments, and conflicts. In particular, the simulation involves two children (one larger than the other) both of whom are attracted to a ball (reward) and want to play with it. The system architecture is shown in Figure 1.

In our implementation, environmental stimuli take the form of objects (e.g., balls, trees, other children) with which the children interact, as well as the actions the objects perform (e.g., a menacing stare, a growl, or an aggressive approach). These input stimuli can be categorized as cues for reward, non-reward, punishment or non-punishment (relief). For example, a ball is treated as a signal for reward, a tree or

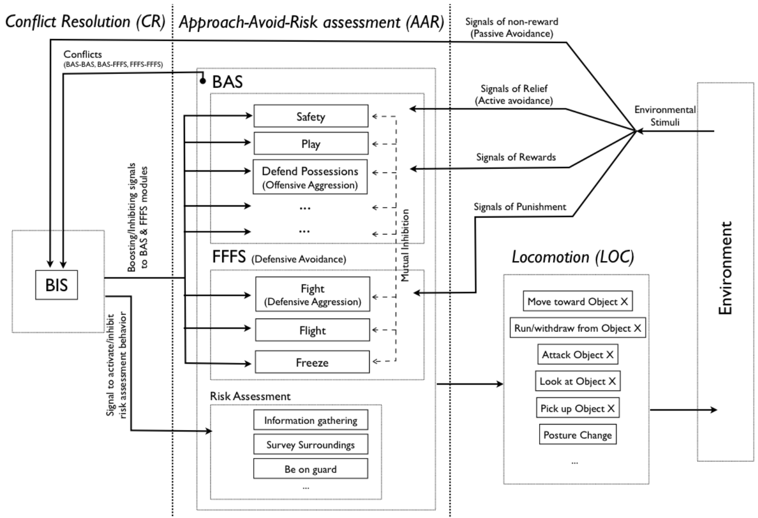

Figure 1: Preliminary architecture based on RST.

a parent as shelter or a signal of relief, and a glare as a signal of punishment.

The Approach-Avoidance-Risk assessment (AAR) part of the architecture contains a set of BAS and FFFS behavior modules that each fire with different intensities in response to different motivations and stimuli. For example, a BAS module responds to the appetitive desire to play with the ball (a cue for reward) when it is present. Another BAS module can respond to the desire for safety and seeks out signs of safety (in this case, for the smaller child, the adult or shelter from a cluster of trees). Each child also has an offensive aggression module which fires when he senses the approach of a rival child who also wants to play with the ball. When this happens the reward derived from playing with the ball is reduced. The FFFS modules take as input signals of punishment, for instance, the presence of or proximity to the bigger, more menacing child. Modules in this part compete with and mutually-inhibit each other, with the highest activating behavior sending the appropriate commands to the locomotion part of the architecture. This part of the architecture also contains the cautious, risk assessment and information-gathering behavior which is implemented as having the smaller child observe the current behavior of the larger child and evaluating how threatening it is, as when, for example, the larger child acts aggressively and moves toward the smaller one.

The Conflict Resolution (CR) part is the BIS module, which responds to conflicting activations and behavioral inclinations generated by modules in AAR. The BIS module responds either by sending an activating signal to the risk assessment behaviors in AAR or by sending signals that strengthen or inhibit the activation levels of the BAS and FFFS behaviors.

Finally, the Locomotion (LOC) part is made up of locomotion controllers that obtain their inputs from behaviors that are activated in the other two parts. The locomotion controllers handle specific actions in response to the behavioral tendencies that are fired in the other two parts. These actions include, for example, the control of movement toward and away from a target, or swinging the fist toward a target. It is through these actions that younger child interacts with the environment (the other child and the ball). This leads to feedback from the environment back to the child, in terms of signals of reward, non-reward, punishment and relief.

Screenshots from the simulation are shown in Figure 2. The first panel (Figure 2(a)) shows the bigger child approaching the ball, while the smaller child watches from a distance because his desire to get to the ball conflicts with his fear of the bigger child. As the simulation progresses and the bigger child does not exhibit any threatening behavior toward the smaller child, the smaller child gradually inches closer to the ball and eventually starts playing with it (Figure 2(b)). Figure 2(c) shows the activation of the flight-behavior in the smaller child when the bigger child exhibits threatening behavior.

More sophisticated adaptation to stimuli, and handling of signals of frustrative non-reward (by BIS) and active avoidance (in BAS) will be implemented in future versions of the simulation. One of the attractive features of this kind of model is the ease with which it allows for interesting individual differences. Differential sensitivities of the modules and behaviors to their respective input signals result not only in routine personality differences, but they can also give rise to socially non-conforming behaviors. For example, an overactive BAS can result in hyperactivity and aggression, and an overactive BIS leads to anxiety disorders.

Conclusions

RST provides a powerful mechanistic account of approach and avoidance behavior while at the same time accounting for important aspects of personality such as anxiety and impulsivity. We believe that implementing RST may be a fruitful alternative for engineering applications that require the simulation of personality and emotions, for instance in en-

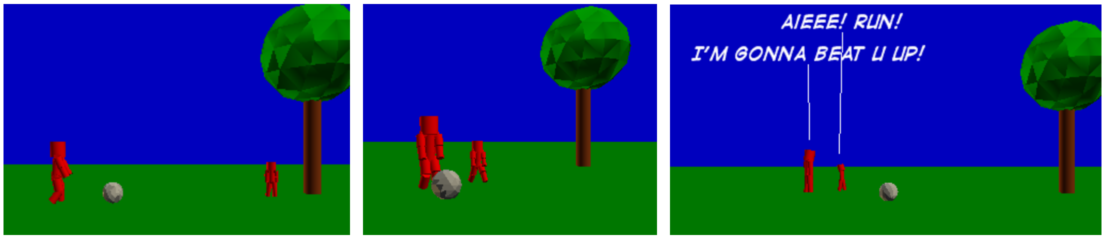

Figure 2: (a) Smaller child watches as bigger child plays with the ball; (b) Smaller child gradually draws closer (cautious risk-assessment approach) as bigger child continues playing with ball and does not exhibit threatening behavior; (c) Offensive aggression (BAS) activates in bigger child, FFFS (Flight) activates in smaller child.

tertainment where the display of certain humanlike qualities is important. We find RST especially compelling because of its established foundation in biological systems and its intuitive accounts of approach and avoidance behaviors. Finally, we wish to emphasize that what we have presented in this paper is only preliminary work, and thus constitutes little more than a pointer to more extensive research we will be conducting in the future.

References

Ávila, C., and Parcet, M. A. 2002. Individual differences in reward sensitivity and attentional focus. *Personality and Individual Differences* 33:979–996.

Blanchard, R. J., and Blanchard, D. C. 1990. An ethoexperimental analysis of defense, fear and anxiety. In McNaughton, N., and Andrews, G., eds., *Anxiety*. Otago University Press. 124–133.

Canamero, D. 1997. Modeling motivations and emotions as a basis for intelligent behavior. In Johnson, L., ed., *Proceedings of the First International Symposium on Autonomous Agents (Agents'97)*. ACM. 148–155.

Carver, C. S., and White, T. L. 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology* 67:319–333.

Depue, R. A., and Collins, P. F. 1999. Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences* 22:491–569.

Elliot, A. J., and Thrash, T. M. 2002. Approach-Avoidance motivation in personality: approach avoidance temperaments and goals. *Journal of Personality and Social Pyschology* 82:804–818.

Gratch, J., and Marsella, S. 2005. Some lessons from emotion psychology for the design of lifelike characters. *Journal of Applied Artificial Intelligence* 19(3-4):215–233.

Gray, J., and McNaughton, N. 2000. *The Neuropsychology of Anxiety*. Oxford: Oxford University Press.

Gray, J. A. 1970. The psychophysiological basis of introversion-extraversion. *Behavior Research and Therapy* 8:249–266.

Horswill, I. 2009. Very fast action selection for parameterized behaviors. *Proceedings of the Fourth International Conference on the Foundations of Digital Games*.

Hudlicka, E. 2004. Two sides of appraisal: Implementing appraisal and its consequences within a cognitive architecture. *Proceedings of the AAAI Spring Symposium, Architectures for Modeling Emotion, Technical Report SS-04-02* 24–31. Menlo Park, CA: AAAI Press.

Konidaris, G. D., and Barto, A. G. 2006. An adaptive robot motivational system. In *S. Nolfi et al., eds., From Animals to Animats 9*. Springer-Verlag. 346–356.

Oudeyer, P.-Y.; Kaplan, F.; and Hafner, V. V. 2007. Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation* 11(2):265–86.

Revelle, W. 2008. The contribution of reinforcement sensitivity theory to personality theory. In Corr, P., ed., *The Reinforcement Sensitivity Theory of Personality*. Cambridge University Press. 508–527.

Smillie, L. D., and Jackson, C. J. 2005. The appetitive motivation scale and other BAS measures in the prediction of approach and active avoidance. *Personality and Individual Differences* 39:981–994.

Torrubia, R.; Ávila, C.; Moltó, J.; and Caseras, X. 2001. The sensitivity to punishment and sensitivity to reward questionnaire (SPSRQ) as a measure of Grays anxiety and impulsivity dimensions. *Personality and Individual Differences* 31:837–862.

Watkins, C. 1989. Learning from delayed rewards. *Ph.D. Thesis*. University of Cambridge, England.

Whitehead, S.; Karlsson, J.; and Tenenberg, J. 1992. Learning multiple goal behavior via task decomposition and dynamic policy merging. In Connell, J. H., and Mahadevan, S., eds., *Robot Learning*. Kluwer Academic Publishers. 45–78.

Wilson, G. D.; Barrett, P. T.; and Gray, J. A. 1989. Human reactions to reward and punishment: A questionnaire examination of Grays personality theory. *British Journal of Psychology* 80:509–515.

Zinbarg, R. E., and Yoon, K. L. 2008. RST and clinical disorders: anxiety and depression. In Corr, P., ed., *The Reinforcement Sensitivity Theory of Personality*. Cambridge University Press. 360–397.