

DISSERTATION

Modeling Emotional Effects on Decision-Making by Agents in Game-Based Simulations

Marc Spraragen

Submitted to the University of Southern California
Computer Science Department as Partial Fulfillment of
Degree Requirements for Doctor of Philosophy

This dissertation is dedicated to my parents, who never gave up on me.

Acknowledgements

First and foremost, I would like to acknowledge my dissertation committee, without whom this work would not exist. Michael Arbib, Jim Blythe, Azad Madni ("outside member" in name, but a true mentor), and my committee chair and steadfast advisor Michael Zyda. For rigorous and incisive critique, for in-depth aid with content and form, for making sure I thought and wrote like a scientist, and for patience and visionary guidance, I thank you all.

I would also like to acknowledge people who helped me greatly with my research both before and during the dissertation. Fellow "Games PhD" students Peter Landwehr (CMU), Jerry Lin, Balki Ranganathan, and Powen Yao; research scientists Kathleen Carley, Yolanda Gil, Jon Gratch, Jihie Kim, Mike Obal, Paul Rosenbloom, Ning Wang, and Patrick Winston (MIT); and three outside USC on whom my work relied: Pauline Felder, Kate Harlan, and Janell Rothenberg.

Last but not least, I would like to thank my friends and family for all of their wonderful support and understanding.

Table of Contents

Chapter 1. Introduction.....	6
1.1. Motivation for Research.....	6
1.2. Problem Statement.....	10
1.3. Research Hypothesis.....	12
Chapter 2. Background and Literature Review.....	12
2.1. Cognitive Modeling.....	12
2.2. Appraisal and Dimensional Emotional Theory.....	14
2.3. Emotional Effects on Decision-Making.....	16
2.4. Previous Work: Review of Relevant Computational Models.....	26
2.5. State of the Art in Game AI Emotion Research.....	32
Chapter 3. Approach: Model and Implementation.....	33
3.1. Essential Components.....	34
3.1.1. Theory of Emotion Supported by Cognitive Science.....	35
3.1.2. Memory Network Evaluated for Emotional Content.....	35
3.1.3. Humanlike Cognitive Process Model Sensitive to Emotional State.....	37
3.1.4. Range of Known, Quantified Effects of Emotional State on Decision-Making.....	45
3.2. Emergent Effects Leading to Emotional State-Dependent Decision Making.....	47
Chapter 4. Key Experiments and Results.....	49
4.1. Background.....	49
4.2. AX Continuous Performance Task Experiment.....	50
4.2.1. Memory Contents and Game Elements.....	50
4.2.2. Task Procedure.....	51
4.2.3. Training and Testing.....	52
4.2.4. Additional System Constants and Constraints.....	52
4.2.5. Results.....	53
4.2.6. Correlation of Results to Human Data.....	54
4.3. Tower of London Experiment.....	55
4.3.1. Memory Contents.....	56
4.3.2. Decision Making Procedure.....	57
4.3.3. Experimental runs.....	58
4.3.4. Results.....	58
4.3.5. Results' Correlation to Human Data.....	60
4.4. Discussion of Results from Calibration Experiments.....	63
4.5. Nuclear Power Plant Experiment.....	64

4.5.1. Overview.....	64
4.5.2. Domain Model for Nuclear Power Plant	65
4.5.3. Procedures and Experimental Methods.....	66
4.5.4. Results and Analysis.....	70
Chapter 5. Discussion.....	72
5.1. Objective and Expected Outcome.....	72
5.2. Scope	73
5.2.1. Theoretical Scope.....	74
5.2.2. Experimental Scope.....	74
5.2.3. Modeling and Architectural Scope	74
5.3. Research Findings.....	75
5.3.1. Summary of Methods.....	75
5.3.2. Experimental Results Summary.....	76
5.3.3. Discussion of Findings	77
5.4. Implications and Conclusions.....	77
5.4.1. Key Findings and Research Contribution.....	78
5.4.2. Current Experiment Redesign.....	81
5.4.3. Enhancement: Emotional Feedback and Emotional State Change	84
5.4.4. Additional Cognitive Modeling Experiments.....	85
5.4.5. Towards a Nuclear Power Plant Turing Test.....	87
5.4.6. Agent Enhancements	90
5.4.7. Model Augmentation	91
5.4.8. Applications of Future Work.....	94
Appendix A: Further Background and Discussion.....	96
A.1. Human Planning, Problem-Solving, and Decision-Making	96
A.1.1. Reactive and Deliberative Planning.....	98
A.1.2. Case-based Planning.....	99
A.2. Additional Modeling Background	100
A.2.1. Activation and Association Maintenance Modeling.....	100
A.2.2. Unified Model of Process Cooperation.....	101
A.3. Further Discussion of Emotional Theory	101
A.4. Further Effects to be Modeled.....	107
Appendix B: Implementation.....	110
B.1. Code specification Including Memory for Experimental Domains	110
B.2. Sample Output from Nuclear Power Plant Decision-Maker.....	110
References	114

Index of Tables

<i>Table 1. Existing human experiments of emotional state's effects on decision-making....</i>	<i>19</i>
<i>Table 2. Representative computational models of emotional effects on decision.....</i>	<i>27</i>
<i>Table 3. High-level model of deliberative subsystem processes</i>	<i>41</i>
<i>Table 4. Effects modeled by Emo</i>	<i>48</i>
<i>Table 5. Summary of further experiment modeling.....</i>	<i>85</i>
<i>Table 6. Augmented emotional model.....</i>	<i>93</i>
<i>Table 7. Map of appraisal dimensions to predicted emotion generation.....</i>	<i>104</i>
<i>Table 8. Further emotional effects for future work.....</i>	<i>109</i>

Index of Figures

<i>Figure 1. High-level conceptual diagram of Emo agent framework.....</i>	<i>34</i>
<i>Figure 2. System errors in AY case averaged by emotional state.....</i>	<i>53</i>
<i>Figure 3. Correlation between agent and human AY errors.....</i>	<i>54</i>
<i>Figure 4a. Tower of London extra moves averaged by emotional state level.....</i>	<i>59</i>
<i>Figure 4b. Tower of London plan step rate per move averaged by emotional state level</i>	<i>60</i>
<i>Figure 5a. Correlation between agent and human Tower of London extra moves.....</i>	<i>62</i>
<i>Figure 5b. Correlation between agent and human Tower of London plan step rate</i>	<i>63</i>
<i>Figure 6. Nuclear power plant pipe rupture scenario outcomes</i>	<i>71</i>
<i>Figure 7. Nuclear power plant reactor readout panel.....</i>	<i>87</i>
<i>Figure 8. Nuclear power plant control panel</i>	<i>88</i>
<i>Figure 9. Nuclear power plant help manual screen</i>	<i>89</i>

Chapter 1. Introduction

1.1. Motivation for Research

Intelligent agents are playing an increasingly important role in game-based simulations[1][2][3][4][5]. Agent behavior has continued to improve from the purely robotic to more humanlike models based on ACT-R and other cognitive architectures[1]. An important aspect of humanlike behavior is changing emotional states [6][7][8]. Human cognitive science, neuroscience, and psychology have extensively studied phenomena relating to the effects of emotion on cognition for centuries. Examples include Plato's concept of emotion as a distraction from rationality, the Yerkes-Dodson Law illustrating enhancement of problem-solving given moderate emotional arousal [9], the discovery of emotion positively impacting decision-making in the Iowa gambling task [10], and beyond. In the case of Phineas Gage, a physical lack of emotional processing capacity correlated with poor judgment [11]. Cognitive scientists have studied emotional effects on human planning, decision, and recall, among other processes, providing a reference point for computationally modeling such effects. Relatively little work has been done on creating computational architectures that incorporate emotional effects with an integrated set of cognitive processes [12], and even less on modeling effects of emotion on the human capacity for decision-making [13].

Within the comparatively younger field of Artificial Intelligence, the subfield of computational cognitive modeling has often focused on deliberative processes such as decision-making, problem-solving or planning, in order to better understand human intelligence. These processes are often represented as independent of emotions. For instance, the ACT-R and Soar systems modeled human cognition for decades before they were used to run emotional simulation experiments.

The study and evaluation of videogame AI characters with humanlike simulated emotion is part of a more general effort to model more realistic software agents. Computational exploration of the interactions between emotion and other cognitive processes is important for developing architectures for general intelligence, and for the fields of human social and behavioral modeling, game AI, and human-computer interaction. For example, in a diplomacy training simulation, a human player could realistically anger an AI agent non-player character (NPC) through cultural insensitivity, causing the NPC to make decisions hostile to the player, even those detrimental to the NPC [14]. However, particularly in game AI, computational modeling of emotional effects on decision-making has yet to catch up to the range and depth of human studies' results.

There are several distinguishing features, both broad and specific, of suitable intersections between game AI and computational emotional modeling. Generally, if one goal of including intelligent non-player characters in a videogame is to create a

psychological proximity with the player [15], then realistic emotional expression for NPCs would help solidify that bond and lend more engaging qualities to the game [16]. More specifically, emotions can play a part in coloring or shaping NPCs' interactions with the player, the game world, the storyline, and each other. An internal emotional system gives NPCs a more specific context for their actions and choices, and also impacts the quality of their cognitive processes. There is a need for a fast, lightweight, portable, and flexible model that can fulfill these objectives realistically.

In serious games, the stakes of agent realism are raised from engagement for its own sake to engagement for the sake of a greater purpose such as learning or training[2][3][4]. One example of a serious game in which realistic agent emotions would be very useful is a therapy training game. The therapist in training would learn how to temper his or her own dialogue and actions in the real-time presence of emotionally normal and abnormal agents. Another motivating example for simulation is a nuclear power plant operator overseeing an emergency plant shutdown. Such procedures require strong recall ability and precise decisions [17]. How might the operator's emotional state affect his decisions, and how might emotion help or hinder his planning process under tighter time constraints [18]? An operator in a positive emotional state might be more optimistic than statistically realistic estimates suggest, and would not feel the need to look into alternative plans; if unforeseen negative results get in the way during execution, the plan might not succeed. Also, in urgent situations, the dependence on emotional cues for decision-making

becomes stronger, leading to less “rational” choices [19]. To cite a specific case, irrational operator action was a major factor of the Three Mile Island nuclear plant accident in 1979 [20]. One potential real-world application of my model is to simulate the maintenance and shutdown of a similar nuclear plant, and to compare its performance against that of human subjects planning out and executing the same sets of (simulated) tasks during common crises such as a pipe rupture.

Very few AI agents have been designed from the ground up to model a wide set of effects of emotion on deliberative human processes such as decision-making. In terms of structure, one of the closest matches to the research presented in this dissertation is H-CogAff [21], which has a cognitive architecture incorporating the potential for pervasive emotional effects, as well as a “deliberative layer” of cognition that includes decision. However, H-CogAff has not been implemented, and does not clearly define a working memory structure that can facilitate the interaction between emotion and deliberation, and also does not delineate a range of emotional effects. Similarly, systems such as EM [22] and Tabasco [23] model a planning function without dependent cognitive processes that might be affected by emotion. Other systems like certain ACT-R extensions [24][25][26], though built on an integrated cognitive architecture, were designed to model one emotional effect on recall or decision, and do not use a more general model of emotion and cognition to produce a wider set of results across effects or problem domains. Moreover, ACT-R, Soar, and other major cognitive architectures are not particularly

portable for embedding in various game agents that simply need supplementary emotional state-based emotional subsystems for realistic decision-making. Emotional state is disambiguated here as a high-level or consolidated variable based on the emotional contents of working memory [27], leaving constant other factors such as temperamental predispositions and physical comfort. Emotional state is also a temporal structure, relatively long-lasting as compared to the vagaries of event-specific emotional episodes [13] but shorter-term with respect to affective temperament. With that definition, emotional state is a state readily measured in human experiments by self-report combined with physiological signal analysis. Several human experiments use emotional state to provide straightforward quantitative correlation with emotional effects on decision-making. Emotional state is therefore a useful variable for validation of claims of humanlike realism in a computational model of such effects. However, effects of emotional state on decision-making have not been sufficiently addressed in agent architectures used in game-based simulations. Fortunately, such effects are well studied in cognitive science. That body of work provides a sound basis for creating computational models of emotionally sensitive agents.

1.2. Problem Statement

The research problem addressed in this dissertation is the development of computational agent models that reflect the influence of human emotional state on decision-making in

game-based simulations. Consider again an agent operating a simulated nuclear power plant. The operator agent needs to respond to various anomalous situations that arise during power plant operation. For example, the operator agent, observing a sudden drop in cooling water pressure, would need to make several correct assumptions and decisions in short order. Those decisions require attention focus, recall ability, and precise choices, all of which are human processes susceptible to the effects of emotional state. For instance, an operator in a positive emotional state is more likely to be optimistic and underestimate the likelihood of a critical cause (e.g., pipe rupture) for the loss of water pressure, whereas an operator in a negative emotional state is more likely to suspect such a cause and act accordingly.

My contribution is to enrich intelligent agent behavior in games through incorporating emotional state-based effects on decision-making. My agent architecture model incorporates four main integrated components: an associative semantic network model of memory operated on by modified spreading activation functions as per ACT-R, a model of human deliberative processes based on Newell's Unified Cognitive Architecture and Kahneman's two-system cognitive theory, a hybrid appraisal / dimensional model of human emotional, and a small set of emotional state-dependent mechanisms affecting the agent directly. Together, these components are used to model a set of emotion-based effects on decision-making behavior. The implemented agent and its underlying model

are also designed to contribute to the general understanding of how emotion affects cognition and how to simulate and validate these effects computationally.

1.3. Research Hypothesis

An emotional agent architecture based on a precise combination of principles from cognitive science and computational modeling exhibits realistic behaviors in complex decision-making tasks performed across simulation domains. Such an agent computationally models documented and quantifiable emotional state-dependent emotional effects with a high degree of fidelity to human data, enabling realistic humanlike decision-making. Evaluation is by t-testscorrelating existing human experiment data with results of computationally modeling these experiments.

Chapter 2. Background and Literature Review

2.1. Cognitive Modeling

Unified theories of cognition [28] are designed to explain and model all known aspects of human cognition in a single system. According to the “Newell Test” [29], cognitive aspects to be unified include flexible, dynamic and adaptive behavior, as well as natural language processing and others. Study of unified cognitive theories gave rise to several

well-known architectures such as Soar, ACT-R, and CLARION. In such systems, emotion can influence cognitive processes by means of providing goal-based cues and biases.

The computational modeling of human decision-making as part of cognition is cast as an interdependent set of deliberative cognitive processes [30][31], according to some studies. Kahneman and others refer to these deliberative processes as belonging to “Deliberative subsystem” of cognition, as distinct from the reactive processes of “Associative subsystem” [18][32]. Associative subsystem incorporates emotional heuristics [33] and is described as fast, intuitive, associative, and parallel. By contrast, Deliberative subsystem incorporates slower, rule-based, serial cognitive processes. The two systems interact and interrupt one another extensively. Some models build in a “System 3” of reflective (metacognition) processes [34][35]; the model presented in this dissertation relegates meta-management and associated processes to Deliberative subsystem.

The concept of an associative memory network underlying the interaction between emotion and cognition was pioneered by Bower [36] to better define the relationship between emotional state and memory. Anderson’s ACT and ACT-R teams developed a similar theory to formalize activation strength (i.e., importance and relevance) of linked concept “nodes” in a memory network. The activation strength formula in ACT-R represents a node’s activation strength as its “base” activation (i.e., how recently and

frequently the node has been activated) plus the node's strength of association with adjacent nodes. The formula does not take into account emotional impact of a node to generate activation strength. A later expansion of ACT-R modeling the Iowa Gambling Task [25] uses a modified equation that includes emotional weight as risk probability, particularly on the link between a node representing a deck of cards and the previous positive or negative outcomes of choosing a card from that deck.

Bower [36] represents the human memory as a semantic network of associations. In Bower's theory, a node represents a semantic concept (or an aggregated chunk of concepts and links). Nodes are connected by directed links which themselves have semantic specifications, for instance a "causal link" from node A to node B, denoting that A causes B. Other semantic link types might include "negates" or "enhances." The semantic link concept enables straightforward processing of emotional value between connected nodes. For instance, if A was a very displeasing event and B is found to negate A, B would thereafter be regarded as more pleasing than before. Links can also provide referents for directed emotion: if situation B is frustrating, and person A is found to have caused B, the frustration may develop into anger at A.

2.2. Appraisal and Dimensional Emotional Theory

Dimensional theories of emotion generation are similar to appraisal theories in that both map emotion-evoking events to emotional states. Appraisal theories relate discrete appraisal elements to discrete emotional states, while dimensional theories posit a non-relational “core affect” (or emotional state) emotional state tracked as a single uniquely determined point along a number of continuous, orthogonal dimensions [37]. The 2-dimensional circumflex theory [38] places various emotional states around the axes of Pleasure (a.k.a. Valence) and Arousal. Pleasure represents the positive or negative reaction to an entity or situation, and Arousal represents the intensity or importance of same. The Pleasure, Arousal and Dominance (PAD) dimensional theory is named for its three dimensions of Pleasure, Arousal, and Dominance. Dominance is defined as the degree to which a person feels powerful or in control of their situation, analogous to coping potential in appraisal theory [39]. PAD is analogous to a 3-dimensional expansion of the circumflex. For instance, in PAD theory both anger and anxiety arise from similar low-Pleasantness and high-Arousal events. However, anger and anxiety are on opposite sides of the Dominance dimension: an anxious person feels less in control of their situation than does an angry person. Like Challenge and Threat theory, but unlike the concept of primary vs. secondary appraisals, PAD posits both Pleasure and assessment of coping potential (Dominance) at the same level. This paradigm allows an agent to quickly evaluate primal (e.g., fight vs. flight, anger vs. fear) circumstances [40] on a single node without resorting to semantic traversal of the network to make a decision; for

instance, to choose an action that will raise both Pleasure and Dominance, especially if either is critically low.

2.3. Emotional Effects on Decision-Making

Human emotion can cause critical interrupt signals to cognitive processes [41]. An emotional signal would be responsible for focusing cognitive attention onto an emotionally compelling stimulus [42]. Emotional intensity grants a heightened priority to relevant concepts attended to during an emotional episode. In a similar vein, the cue utilization theory [43] states that under higher levels of emotional intensity (as with similar stressors like task urgency or difficulty [44]), cognitive cues not central to the arousing agent or situation would be increasingly ignored. This “tunneling” effect could lead to overlooking subtle but important details, or leave the subject open to misdirection and other forms of deliberate manipulation [7].

Emotion may refocus cognition away from a task at hand, causing distraction, even though the emotional episode is incidental / irrelevant to the task. The recorded theories of incidental emotion as distraction date back to the ancients, who likened emotion to two horses (aka associative subsystem) drawing “the chariot of the soul,” with reason or rationality (deliberative subsystem) as the charioteer. Too much independent activity by the horses, especially the black one that represented negative emotions, could disrupt the

movement of the chariot as it proceeded towards enlightenment. Emotional as an alarm signal also primes Deliberative subsystem cognitive processes to cope with an emotionally charged stimulus. In planning, the signal may cause a change of plan, or a change of goal, or a reappraisal of the stimulus and associated concepts.

Distraction of attention from a concept related to a task at hand may also play an adaptive role in decision-making. Some human cognitive science studies have shown that the nature of the distraction is the key. If the distractor is also task-related, then there can be associative creativity in the solving [45].

Many experiments use an agent's overall emotional state to measure emotional effects on human cognition, as it provides a straightforward correlation between gestalt physiological measurements and self-reported prevailing emotional state [36][46]. Negative emotional state (displeasure) can lead to narrow-minded but careful decision making; positive emotional state (pleasure) can lead to broad decisions that attempt to achieve multiple goals with less attention to detail [47] and more heuristic processing [48].

One study showed that for a subject in a positive emotional state, plan goals seem less costly, and positive outcomes of choices seem more likely [49][50]. Conversely, people in a negative emotional state pessimistically overestimate cost and planning time, allowing

them in some cases (when the plan is not abandoned) to plan more optimally than people in neutral or positive emotional states. To explain the overall positive emotional state phenomenon, Mayer suggests that people in positive emotional states devote less time to planning because they expect favorable outcomes, and so do not look much further than the first plausible course of action [48]. Another experiment demonstrates that either positive emotional state or negative emotional state (as opposed to neutral emotional state) impairs a subject's planning by reducing planning steps (i.e., search breadth and depth) with respect to plan execution actions. Both non-neutral emotional states also cause an underestimation of steps needed to achieve a goal, leading to fewer optimal solutions generated [46].

Bower's work on emotional state and memory [51][52] shows that concepts associated with a positive or negative aspect are more easily recalled when learned during a congruent emotional state (i.e., sustained emotional state). The emotional state-congruent recall effect is hypothesized to be due, ultimately, to congruence building emotional arousal, although matching emotional state is not necessarily the cause of the recall congruence.

A recalled concept may be strongly arousing enough that it wrests attention from the task at hand. This ties in to the previous discussion of distraction vs. creative association, as emotional state-related distractors may or may not also be task-related.

Table 1. Existing human experiments of emotional state's effects on decision-making

Study	Emotional state Induction Method	Task	Result vs. Neutral Emotional state
[53] Spies, 1996	Autosuggestive self-reference + music	Working memory span	Positive emotional state: impairment. Negative emotional state: impairment
[54] Hesse, 1996	Autosuggestive self-reference + music	Word vs. non-word recognition	Positive emotional state: no effect Negative emotional state: facilitation
[55] Oaksford, 1996	Film clips	Tower of London	Positive emotional state: impairment
[56] Gasper, 2002	Autobiographical memory	Switching to novel word-finding strategy	Positive emotional state: no effect Negative emotional state: facilitation after switch instruction
[46] Phillips, 2002a	Film clips and music	Tower of London (young subjects 19-37)	Positive emotional state: no effect Negative emotional state: no effect

[46] Phillips, 2002a	Film clips and music	Tower of London (older subjects 53-80)	Positive emotional state: impairment Negative emotional state: impairment
[57] Phillips, 2002b	Story	Initial letter fluency (syntax association)	Positive emotional state: no effect
[57] Phillips, 2002b	Story	Uses of objects fluency (semantic association)	Positive emotional state: facilitation
[57] Phillips, 2002b	Autobiographical memory	Switching syntax and semantics processing fluency; Stroop color test; switching Stroop color and word identification	Positive emotional state: impairment
[58] Dreisbach, 2004	Viewing pictures	Switching to novel stimulus	Positive emotional state: facilitation Negative emotional state: no effect
[58] Dreisbach, 2004	Viewing pictures	Switching to inhibited stimulus	Positive emotional state: impairment Negative emotional state: no effect
[59] Dreisbach, 2006	Viewing pictures	AX Continuous Performance Task: maintain and	Positive emotional state: impairment

		execute goal	Negative emotional state: no effect
[59] Dreisbach, 2006	Viewing pictures	AX Continuous Performance Task: switch goals	Positive emotional state: facilitation Negative emotional state: no effect

As shown in Table 1, survey of experimental evidence [60] on the facilitation or impairment of emotional state on various decision-making tasks presents mixed results, dependent on several extra factors such as the nature of the task and the population being tested. The survey's main findings are summarized in Table 1, along with supplemental studies [54][56][59]. All emotional states were estimated by subjects' self-reported rating scales. The combination of these results supports the claim that context (environment, task at hand, memory contents) influences how emotional state affects decision-making[56].

Similar to results from the Tower of London experiments [46], another study [53] showed that emotional state impaired attention function when memorizing words or sets of words. Their results also showed that working memory capacity was adversely affected: fewer words in a list were retained short-term in both positive and negative emotional states.

The Tower of London emotional state experiments [55] were similar to those of Phillips et al. in that both showed positive emotional state as detrimental to a focused planning task. However, Oaksford's study did not show a significant difference between neutral and negative emotional state, though the tendency was towards impairment. The differences in the experiments were the emotional state induction specifics: Oaksford used only film clips, while Phillips used music as well as film clips. Also, there was no separation of age groups in Oaksford's experiment.

The 1996 study by Hesse and Spies [54] was an early demonstration of negative emotional state facilitating decision-making. Negative-emotional state participants were found to persevere on a very cut-and-dried priming stimulus (e.g., PERUSE, a synonym of the target word READ), which helped them differentiate more readily whether the target was a word or a scrambled non-word, e.g. REDA, in such cases. Positive-emotional state participants, on the other hand, did better with free-association priming stimuli (e.g., WRITE and READ), as expected.

Gasper did several experiments in 2002 [56] that led to a generally similar conclusion regarding how negative emotional state can facilitate decision-making. Gasper showed that negative-emotional state subjects were quicker to switch to a new specific strategy that broke a set pattern of word letters in a string (e.g., every other letter as in RXEXAXD before the set-breaking item, and every third letter RXXEXXAXXD after).

Positive-emotional state subjects, being less bound to a single strategy due to increased free-association flexibility on the previous data, did not take the set-breaking item to mean the beginning of a single new pattern set, and kept trying their initial strategy or strategies in addition to the correct new one.

The battery of experiments (aside from the Tower of London experiment, discussed later) done by Phillips et al. in 2002 [57] studied the effect of positive emotional state on several color vs. word identification tasks. Particularly, they used the Stroop test, wherein the name of a color is written in incompatibly colored ink. The results showed that positive emotional state impaired forced strategy switching, i.e., alternately having to read the word vs. say its color. The researchers' rationale was that attention focus would need to be strong in order to maintain a new strategy in order to make the correct decisions in shorter time. In the same paper, Phillips et al. conducted several syntax vs. semantics fluency experiments. Positive emotional state impaired syntax fluency (e.g., "name as many words as you can that start with R"), but facilitated semantic fluency (e.g., "name as many uses for a newspaper as you can"). The result echoes the finding of Hesse and Spies; initial letter syntax fluency requires a strong feature-matching capacity, while semantic fluency calls for more free association.

Dreisbach's and Goschke's experiments [58] also demonstrated the perseveration vs. distractibility effect of emotional state, using two letter set switching tasks: learned

irrelevance and perseveration. In an example of the learned irrelevance experiment, the gray letter on a card with gray and black letters is initially the target stimulus, but new white letters become the target and black letters become distractors. In a similar case of the perseveration experiment, the gray letter on a card with gray and white letters is initially the target stimulus, but then gray letters become the distractors and new black letters become the target stimulus. Results for perseveration (wherein the novel stimulus is a distractor) were impaired by positive emotional state, while results for learned irrelevance (wherein the novel stimulus is the target) were facilitated. The hypothesis of the researchers was that positive emotional state increases a bias towards novel stimuli.

Phillips' Tower of London results [46] are unique in the survey in that they show negative emotional state impairing planning capability vis-a-vis neutral emotional state results, particularly for the older group of test subjects. The Dreisbach AX Continuous Performance Task study [59] is a prime example of how positive emotional state can facilitate decision-making ability with task-related distractors. These experiments, and my modeling thereof, are described in more detail in the Key Experiments and Results chapter, and an analysis aimed towards modeling all of the survey's studies is in the Discussion chapter.

One analysis that encompasses several of the above survey's results is that positive emotional state leads to increased distractibility [61]. When distractors are irrelevant to a

cognitive task at hand, increased distractibility impairs decision speed and accuracy [59]. However, when distractors are task-relevant, increased distractibility leads instead to cognitive flexibility and creative solutions[62].

Aside from the effects illustrated by the survey, the Affect Heuristic [7] is another emotional state-dependent and time-dependent effect of emotion on decision-making: “gut feelings” during emotionally arousing moments can be a useful heuristic to making a decision in a timely manner, bypassing deliberative evaluation or elaboration. Illustrations of the heuristic in action might include a quarterback making a split-second choice to throw to a particular receiver: “the situation just felt right.” Similarly, “the situation just felt wrong” to a military commander making the time-sensitive call to shoot an object out of the sky even though it was possibly not an enemy missile but a friendly fighter plane. This heuristic would be useful for game agents when time-sensitivity is a factor in decisions.

Limiting the time to judge a situation’s risk and benefit induces a sense of urgency and also increases human reliance on the affect heuristic. Under time pressure, the inverse relationship between expected risks and benefits (higher risk implying lower benefits and vice versa) becomes more pronounced in decision-making [19].

2.4. Previous Work: Review of Relevant Computational Models

The impact of the human psychological and cognitive studies outlined above has contributed to an increase in computational modeling of emotion and cognition. Some affect-antecedent AI systems, for example, model plan-based appraisal: how thinking of a plan changes emotional state [8]. Affect-consequent systems[63]—computational models of the effects of emotion on cognition— may be categorized along several criteria. One clear demarcation is between “behavior-consequent” and “cognitive-consequent” models, although many systems include both of these functions.

A behavior-consequent model maps an agent’s emotional state / emotional state to embodied physical actions or other direct outward or social expression, for instance smiling when happy or turning on a light if afraid of the dark. Behavioral-consequent models are often used to synthesize human-like emotional or social behavior in embodied robots like Kismet [64] or in virtual agents such as Greta [65].

As this dissertation deals with the effects of emotional state on decision-making, the following section focuses on cognitive-consequent models: those which produce emotional effects on internal cognition, modeled after cognitive phenomena observed in human subjects. Cognitive change may be manifested as behavior change, but not

necessarily. Prominent cognitive-consequent models are summarized in Table 2, with several elaborated on in the following paragraphs.

Table 2. Representative computational models of emotional effects on decision

System (Creator)	Emotional Cognitive Effects Modeled	Computational Processes For Modeling Effects
[66] ACRES / WILL (Moffat & Frijda)	Attention focus shift (coping strategy). Planning, Decision: Goal shift (coping strategy)	Competition for slots in priority queue, based partially on intensity and urgency of emotion
[67] ActAffAct / BehBehBeh (Rank)	Choice of Relational Action Tendency affected (coping strategy)	Processes related to coping with emotional episode activated, assigned priority and resources
[24] ACT-R extension (Cochran)	Recall ability of any memory decays over time if memory has low arousal, grows with high arousal	Arousal parameter added to base activation formula for ACT-R memory chunks
[25] ACT-R extension (Fum & Stocco)	Memories associated with risk have higher recall strength for emotion-enabled agents	Emotional strength added to relevance parameter in ACT-R memory activation formula
[26] ACT-R extension (Belavkin)	Positive and negative pleasure aid decision processes, up to a certain level of arousal	Goal relevance + noise parameters added to ACT-R production rule selection equation
[68] (no name) (Ahn & Picard)	Different emotions affect risk-aversion when choosing actions	Experienced-utility and expected-utility function parameters change based on emotion
[69] ALEC (Gadanhó)	Rules learned based on past emotional experience	Goal conduciveness parameter added to CLARION rule selection equation, based on Q-Learning with emotion as heuristic
[22] EM (Reilly)	Planning action attributes change based on pleasure	Goal conduciveness parameter in action selection process
[37] EMA (Gratch & Marsella)	Attention focus shift (coping). Plan change, BDI change, action tendencies change (coping strategies)	Soar-based state propositions gain monitoring annotations (attention), see below for

		BDI-based coping mechanisms
[70]Émile (Gratch)	Plan choices, plan selection criteria	(Discussed) Focus planning algorithm on intense appraisal areas, increase/decrease importance of negative plan interaction
[71] EM-ONE (Singh)	Modification of plans, desires, beliefs.	Top-down modification of decision functions (“Critics”)
[72] FearNot! / (Fatima Dias, Paiva, & Aylett)	Attention focus shift. Plan and goal choices change (coping strategy)	Processing resource reallocation; modification of outcome values and probabilities based on emotional state
[73] FLAME (El-Nasr)	Plan choice and constraints change. Reinforcement learning biased	Emotional state-congruence weighting factor added to Q-Learning-based action selection equation
[74] (no name) (Gmytrasiewicz & Lisetti)	Emotional state-congruent choice evaluation (negative emotional state makes negative outcomes seem worse and also more likely)	State utility shifts, state evocation probability shifts, allotted planning time changes
[21] H-CogAff (Sloman)	Attention focus shift (alarms). Emotional state-based decision biases	Bottom-up override and bias signals between normally top-down levels of processes
[75] (Malfaz & Salichs)	Reinforcement learning biases	Emotional state weighting added to Q-Learning-based action evaluation
[76] MAMID (Hudlicka)	Attention focus shift. Decision biases and cognitive ability based on current emotional state	System-wide working memory capacity and inference/recall speed parameters altered, read/write emotional parameters on mental constructs
[77] (no name) (Meyer)	Plan/agenda changes (e.g., fear causes cautious planning)	Addition of emotional state parameters to KARO-based logical formulas
[78] Soar-Emote (Marinier)	Attention focus shift. Planning goal shift / abandonment. Reinforcement learning and recall biases.	PEACTIDM-based modules modified to handle incoming emotional data; System-wide emotional state also used by modules
[23]Tabasco	Planning action choice biases	RAP (Reactive Action Plan)

(Staller & Petta)		chosen based on goal relevance
[40] WASABI / MAX (Becker-Asano)	Plan and action utility choice/evaluation process biased by optimism or pessimism	BDI planning processes affected by specific emotional states triggered by PAD-based emotional settings

In current systems, attention and focus shift is frequently modeled. Systems that address attention/focus shift include MAMID [76] and H-CogAff [21]. One of the sequential modules of MAMID is devoted to cognitive attention focus, which selects a subset of incoming data for further processing. H-CogAff, similarly, has an oversight mechanism for sensing pattern-driven “alarms” from all levels of its cognitive processing: reactive, deliberative, and reflective—i.e., Systems 1, 2, and also 3 [32]. This mechanism in both models redirects cognition to process the stimulus that invoked the alarm.

Effects on decision-making are often cast as constraints on goal and action choices, though there are other types of effects as well. Planning effects represent a form of coping in EMA [37] and Émile [70], among other systems. Emotion can affect Émile’s planning algorithm so that, for example, it focuses on the more intensely emotional plan steps. In EMA, appraisal and coping are interdependent in a closed loop and the strategy for building a plan to cope with a particular emotional stimulus is subject to change following the next round of appraisal. Meyer’s system [77] takes a different approach: emotions cause global effects on search control during planning; for instance, a sad agent

is more likely to look for alternative plans or goals, whereas a fearful agent will be cautious and perform more checks on its environment during planning and execution.

Choice and decision biases are modeled by Becker-Asano's WASABI [40], Rank's BehBehBeh [67], Gadanho's ALEC [69], and the system designed by Gmytrasiewicz and Lisetti [74]. In WASABI, the agent's overall emotional state (used throughout this paper synonymously with emotional state) constrains the set of possible next actions and goals. BehBehBeh and other models of Frijda's theory such as ACRES/WILL [66], use the concept of Relational Action Tendencies (RATs) in a similar manner to constrain decisions; RATs are formed as a direct result of appraisals and narrow the set of next action choices. ALEC uses a fast emotional system that operates asynchronously with its cognitive system to model the Somatic Marker Hypothesis during decisions. Gmytrasiewicz and Lisetti's system also incorporates action tendencies affected by emotional state state changes, formally modeled using decision theory.

Emotional effects on human learning are typically memory-based, and may be used to reinforce recall and decision biases. Memories are evaluated as, or become associated with, particular emotional experiences; cognitive effects follow from these evaluations and associations. Recent work at MIT [68] modeling decision under emotional influence also demonstrates leverage of an agent's previous emotional experience for predictive purposes, using prospect theory. The result is fast, subjective reinforcement learning, and

decision biases result from previous experience. FLAME [73] uses a fuzzy logic method for similar purposes, conditioning an agent by mapping emotional states to remembered events.

Emotion as a recall heuristic has been handled in different ways by the systems that have modeled it. ACT-R, with its well-tested model of associative memory, has been a natural starting point for these systems. Fum and Stocco's ACT-R extension [25], for example, takes advantage of ACT-R's associative memory to reproduce the Iowa Gambling Task's results (though with a skeptical view towards the Somatic Marker Hypothesis, as the researchers do not assume that the saved heuristic references are physiological as opposed to conceptual). MAMID also models emotional effects on cognitive recall and inference, particularly changes to the speed and capacity of those processes based on emotional appraisal.

In summary, though their designs are widely varied, existing computational models of emotional effects on decision-making address one or more of the following basic components:

1. Theory of emotion supported by cognitive science[39]
2. Memory network evaluated for emotional content[7][19]
3. Humanlike cognitive process model sensitive to emotional state[18]

4. Range of known, quantified effects of emotional state on behavior[36][50][61][79][79a]

However, no individual model contains all four components. A further gap in computational agent modeling research relates to point 4. None of the systems designed to date has developed a model to underpin an open-ended set of realistic emotional effects on the decision-making behavior of humanlike agents across simulation domains. My work addresses these gaps by developing a computational agent architecture that is informed by cognitive science and based on computational modeling methods. Specifically, the research performed in this dissertation hypothesizes that an emotional agent architecture that correctly combines the aforementioned four principles from cognitive science and existing computational modeling can exhibit realistic decision-making behaviors in tasks required in complex system operation (simulation).

2.5. State of the Art in Game AI Emotion Research

In addition to advances made in the above work, AI made specifically for games is also starting to embrace emotional behavior and decision-making, particularly in serious games. There are several commercial emotional game agent models or “emotional game middleware”: for instance, EKI-One / MASA Life [80] and XAITment [81], which use emotional state to inform discrete action choices in finite state machines, behavior trees,

and path-finding algorithms. Several experimental games also feature emotional agent models, as in Rosalind Picard's and Hyungjil Ahn's affective gaming research at the MIT Media Lab [68], and in Michael Mateas' narrative intelligence game *Façade* [16]. The Virtual Humans group at USC/ICT, including Gratch and Marsella's serious game work as well as that of Profs. Rizzo, Rosenbloom and others, incorporates cognitive systems such as Soar and EMA and also brings the emotional aspects of players into games via biofeedback [37][82]. The annual Artificial Intelligence In Digital Entertainment (AIIDE) conference typically showcases new models of emotional game agents, such as those found in the multiplayer game *Pataphysic Institute* [83]. One gap in game AI research is that none of the games or middleware to date has developed a cross-domain model to underpin an open-ended set of realistic emotional effects on the decision-making behavior of humanlike agents. My agent model is meant to address that gap.

Chapter 3. Approach: Model and Implementation

The new agent framework, dubbed "Emo," is an integrated architecture (Figure 1) designed to model emotional state effects on the process of decision-making in humanlike game agents.

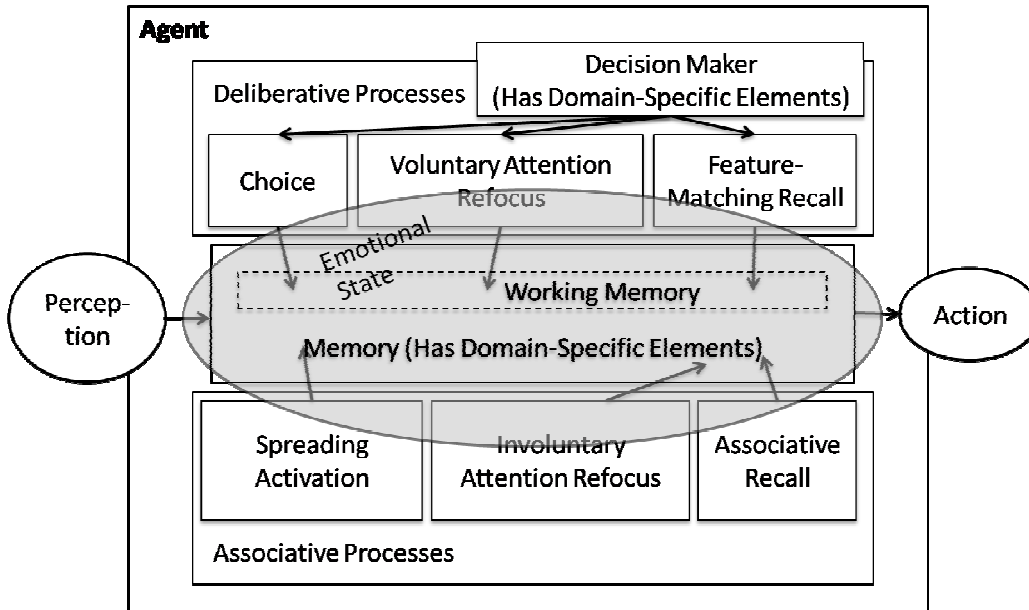


Figure 1. High-level conceptual diagram of Emo agent framework

The approach embodied in this figure is in line with findings that emotion influences deliberative processes [51][84]. Emotion is fundamental to the model and is implemented in Emo. Emotion is pervasive in its memory structure, and in the algorithms underlying its processing modules.

3.1. Essential Components

The framework combines the four essential components distilled from the cognitive science and computational modeling literature. These components are detailed in the following subsections.

3.1.1. Theory of Emotion Supported by Cognitive Science

Emo's emotional state and emotional memory are based on a hybrid dimensional and appraisal theory of emotion. Using dimensional theory, agent emotional state is modeled as a global value ranging from -1.0 (negative) through 0.0 (neutral) to 1.0 (positive). Based on appraisal theory, each node in memory also has a similar emotional value. The justification for this modeling choice is flexibility. Many emotional decision-making effects, for instance emotional state-congruent recall, are based on the resonance between the emotional content of memory and the overall emotional state.

3.1.2. Memory Network Evaluated for Emotional Content

The central data structure in Emo is the associative-semantic network that forms its long-term memory [36] featuring conceptual nodes, e.g. “physical object” or “plan step.” Each node contains an emotional rating. Emo’s cognitive processes are influenced by the emotional ratings of the memory network’s individual nodes, as well as by overall emotional state, which is modeled as an aggregate of the individual nodes’ ratings. In this sense the system has no single appraisal module; each cognitive process is affected in its own way by the emotional cues it encounters.

Links between two nodes include exactly one undirected associative strength link, and any number of directed semantic links [85] such as “causality.” The association links allow for humanlike memory association during cognitive processes. In my model, emotional cues and their consequent cognitive effects can be evoked by the associations and choices made during the process of decision-making, as opposed to emotional changes resulting only from the execution (successful or not) of a plan or partial plan [37].

The model's working memory network is a dynamic, high-activation virtual subset of the larger long-term memory network. Working memory also functions as a central repository acted upon by the interdependent cognitive process modules of the deliberative subsystem. Deliberative process modules like choice and feature-matching recall partially depend on the top-level (self-invoking) deliberative subsystem decision-maker module to focus attention on a particular node. The deliberative subsystem may be interrupted if the associative subsystem involuntarily refocuses attention on the node in working memory with the highest activation rating. Some processes are subject to both associative and deliberative subsystem operation: for instance, involuntary recall of highly arousing memories vs. voluntary recall of a task-related concept.

The network was partially derived from Bower's work on emotion and cognition, with two important distinctions. First, Bower's network centralizes emotions in specially

designated nodes such as “Joy” or “Sadness,” to which other conceptual nodes link to signify emotional content. My agent model, by contrast, carries the components for emotional content on each node. This was done to enable fast processing (e.g., no network traversal needed) of basic emotional reactions, such as fear when encountering a shark. Decentralized emotional encoding also allows quick, parallel processing of emotional association between nodes. The second distinction from Bower’s work is the labeling of links with semantic information. This provides the basis for referential and inferential processing of emotional information (e.g., if B is unpleasant, and A causes B, then A is unpleasant), traversable by associative and deliberative subsystem processes.

As mentioned above, working memory is a subset of long-term memory with one main distinction: the nodes in working memory are accessible by all deliberative subsystem processes, but other long-term memory nodes can only be accessed by associative subsystem processes. The two networks are handled separately, as they are in human cognitive science, because there could be translation incompleteness or error both forward and backward between long-term and working memory [86].

3.1.3. Humanlike Cognitive Process Model Sensitive to Emotional State

Aside from the previously mentioned cognitive processes acting on the memory network, the system’s memory is made “realistically dynamic” by means of

associations maintenance and spreading activation. These processes occur within the network continuously (i.e., on an ongoing basis). Association maintenance strengthens or weakens the associative links between nodes, depending on how often and how recently the nodes appear together in working memory. Spreading activation maintains the activation strength of linked nodes based on the ACT-R model: if node A has high activation strength (which is partially based on emotional ratings), associated node B will also gain in activation strength to a degree depending on the strength of the associative link between A and B [36][87]. The nature of the spreading activation may also depend on the semantic link between A and B—if action A negates unpleasant event B, A’s emotional rating might increase. Spreading activation also includes algorithms for determining decay of activation over time in nodes that are not in working memory. For instance, as new stimuli are brought to focus, older concepts and their emotional information diminish out of working memory, simulating loss of activation strength over time [87a].

The associative subsystem process of spreading activation is a weighted version of the ACT-R model of memory, similar to but distinct from the ACT-R extension used to model the Iowa Gambling Task [25]. The base activation strength of node i depends on overall emotional state, as well as on the number of recent appearances of i in working memory:

$$B_i = n * m_i \left(\ln \left(\sum_{(k)} t_k^{-d} \right) \right),$$

where t_k is the elapsed time since the k th appearance of node i in working memory; d represents a constant decay factor; m_i is a bias giving higher activation strengths to nodes whose pleasure ratings match that of overall emotional state, and n is a weighting constant incorporating a normalizing factor. n is partially based on whether i is currently the single node in the system's attention focus.

B_i is then merged with the spread activation strengths of directly associated nodes (depending in part on strength of association) into the node's overall activation strength A_i :

$$A_i = B_i + \sum_{(j)} (n_j * S_{ij} * A_j),$$

where B_i is the base activation of node i ; A_j is the activation strength of node j ; S_{ij} is the association strength of the link between nodes i and j ; n is a spread factor partially based on whether node j is in working memory, or attention, or neither.

A_i is used to determine the likelihood that node i will involuntarily draw attention away from current deliberative processes (Associative subsystem's involuntary attention refocus, an "alarm" interrupt of the deliberative subsystem) [21][41]:

$P(\text{Attention})_i = n(A_i - A_j)$, where A_j is the activation strength of currently attended node j ; and n is a weighting factor n representing the system's overall susceptibility to involuntary attention shift (a variable based on emotional state).

Attention provides an additional increase of a newly attended node's activation strength. Competition for cognitive resources (e.g., attention, working memory) is based on the winner-take-all behavior of some neural networks [88].

A_i also determines the ease with which node i will be involuntarily recalled into working memory (associative subsystem influence on deliberative subsystem via working memory [88a]):

$P(\text{Recall})_i = nA_i$, where n is a weighting factor based on the system's overall susceptibility to involuntary recall, which can also change with emotional state. If working memory is already at capacity (the system specifies a node capacity limit L , "7, plus or minus 2" [89]), then the node with the least arousal is moved out of working memory to make room for the new node. Capacity itself can potentially be subject to change, depending on the experiment being modeled.

Association management, also modeled on ACT-R, proceeds in parallel with spreading activation through working and long-term memory. Each node separately and continuously updates its own link strengths to adjacent nodes:

$$S_{ij(new)} = n * S_{ij(current)} * T_{ij},$$

where n is a weighting factor based on whether the link's source and destination nodes i and j are in working memory or not, and T_{ij} is the number of times that i and j have both appeared in (or out of) working memory together.

Whenever a node's activation strength is incremented, the node is checked for involuntary recall into working memory and potentially into attention focus. Decrementing the node's activation strength invokes similar checks to see if the node remains in attention or working memory.

Table 3. High-level model of deliberative subsystem processes

Module	Function	Computational Approach
Decision-maker	Solve task at hand (varies)	Varies (e.g., search algorithm, goal switching)
Choice	Choose which goal to pursue, and which action or strategy to use to achieve that goal	Probabilistic utility function (cost * probability of success).
Feature-matching recall	Bring actions, goals, and other concepts from long-term memory to working memory	Primes the associative subsystem's spreading activation to bring task-associated concepts into working memory and potentially refocus attention involuntarily
Voluntary attention refocus	Bring actions, goals, and other concepts into attention focus for	Place task-related node into attention focus

	use by other Deliberative subsystem processes	
--	--	--

The model's deliberativesubsystem processing modules[28][90] are described in Table 3. The set of modules may expand in the future, and individual modules may scale to become more complex, so I have built the system using the following implementation principles: 1) parallel, flexible, modular processing, 2) parsimony of mechanisms, and 3) plausible mapping from human emotional-cognitive theory and experimental results.

The cognitive structure of process modules asynchronously accessing working memory is based on a variant of architectures used to model systems or societies of multiple agents [91][92]. My design allows modules at any granularity to be scoped as single processes, which in turn allows an easy separation of parallel processes, such as association maintenance, into multiple potentially competing threads.

Moreover, the decision-maker module is designed to be loosely coupled with Emo's subsystems, allowing cross-domain flexibility. For instance, in the three different experiments outlined in Chapter 4, I was able to use three different decision-maker algorithms: a simple decision tree, an AStar search algorithm, and a mental modeling process. The remaining subsystems of Emo all remained exactly the same throughout the experiments.

Deliberative subsystem processes revolve around the high-level decision-maker (which, as mentioned above, varies according to the problem at hand), with function calls made from the decision-maker to choice, feature-matching recall, and voluntary attention refocus. Currently, the deliberative subsystem's functionality is modeled as a continuously running single process. This reflects the serial nature of the deliberative subsystem (vs. the associative subsystem's parallel processing). Without associative subsystem or emotional memory cues involved, the deliberative subsystem would simply select a node from working memory and run a decision-making operation based on that node.

Choice, incorporating judgment (evaluation) [93], uses a simple goal-based utility function (likelihood * cost) [94] as a rationale to choose among alternative options. The function may be modified by emotional state and by the emotional attributes of the nodes in question.

In feature-matching recall, a node in working memory whose attributes most closely match the values needed by the decision-maker is brought into attention focus. The more comprehensive the search criteria, the more precise the matching recall will be. The attention refocusing may also trigger associative subsystem-based associative recall of high activation strength nodes into working memory as well, as per Minsky's K-Lines theory [95]; particularly useful if a completely matching node is not immediately found

in working memory. Conversely, if more than one node from working memory matches the feature set closely enough, the choice process weighs in to select the best candidate.

The reasoning behind the division of labor for recall in my model (as a partially targeted and partially involuntary gleaning from long-term memory into working memory) is that this emulates the human paradigm of task focus and distraction. The recall process can continue over several passes of the associative subsystem until the correct (or close enough by feature set) node is found. This approach ties in with the concept of emotional state-congruent recall: if current emotional state is part of the search criteria, concepts learned in a remembered matching emotional state will be more likely to be recalled [36].

Voluntary attention refocus is directed by the top-level decision-making procedure onto a node to be deliberated upon by the deliberative subsystem. This function can be hijacked by the associative subsystem involuntarily, as described above, if an unattended node attains enough activation strength.

Perceptions of the world state and action execution are also parts of the model, dependent on the nature of Emo and its environment. Executing an action is done such that changes made to the state of the world are then detectable at the next perception step thereafter. The decision-making procedure then checks to see which of its goals have been successfully resolved vis-à-vis the new state of the world.

3.1.4. Range of Known, Quantified Effects of Emotional State on Decision-Making

The emotional state-based effects on decision-making in my model are based on three mechanisms. Singly these mechanisms are simple, but in combination they produce complex and realistic effects. The currently modeled set of effects depends mainly on the activation strength and pleasure attribute of a given node, and on emotional state-based global system values (distractibility, and others). Description of each direct mechanism follows.

1. The key variables of emotional state-dependent activation strength are presented below:

Trigger: continuous update by associative subsystem

Input: Emotional state m , Node n ,

Process (in spreading activation):

$n.\text{base_activation_strength} += a * m * n.\text{emotional_rating};$

//where a is a normalizing constant

Output: Node n

In summary, the base activation strength of a given node at any update cycle is augmented by the similarity of its emotional rating to that of the current emotional state[25].

2. The key variables of emotion-dependent success probabilities are presented below:

Trigger: Choice called

Input: Node_List candidate_nodes, emotional state m

Process (in choice):

For all candidate_nodes.Node n:

$$\text{probability_success}(n) += a * m * n.\text{emotional_rating}$$

Output: Node_List candidate_nodes

Conceived probability of success of an action is directly proportional to how closely the current emotional state pleasure matches the pleasure of the action Node. The emotional state-augmented probability of success is then used by the utility-based Choice function (cost * probability of success) to produce a utility value for the action. Tighter time urgency may increase the effect of emotional state-congruence (i.e., Affect Heuristic) on the success probabilities, when time is a factor [18][19][93][94].

3. The key variables of emotional state-dependent attention flexibility are presented below:

Trigger: Perception or recall raises a node's activation strength above the node currently in attention focus

Input: Working_memory wm, emotional state m

Process (in attention):

float flexibility = a * m //where a is a normalizing constant

if (flexibility > random())

Node attended_node =

argmax<wm.Nodes>(Node.activation_strength,

flexibility)

Output: Node attended_node

This third mechanism is behind involuntary attention refocus (to a node whose activation strength was highly augmented by the associative process of spreading activation). The flexibility of Emo (i.e., the chance that such a refocus will be successful) is directly proportional to the emotional state pleasure rating [21][37][46].

3.2. Emergent Effects Leading to Emotional State-Dependent Decision Making

To assure parsimony and flexibility, the set of emotion-dependent mechanisms was reduced to a combination of the mechanisms described in the previous section and the emergent effects that result from them. This was done because increasing

the number of mechanisms could mean more unforeseen interactions and corrective hacks. The following table shows how the deliberative processes, the mechanisms defined above, and the simpler effects in the table itself contribute to the emergence of further effects.

Table 4. Effects modeled by Emo

Emergent Effect of Emotional State on Decision-Making	Trigger	Summary Description	Contributing Processes and Mechanisms
[36] Emotional state-Congruent Recall	Emotional state at certain level	Nodes with emotional rating aligned with emotional state more likely to be recalled	Emotional State-Congruent Activation, Recall
[50] Optimistic / Pessimistic Choice	Emotional state at certain level, more than one node is candidate for choice	Node with emotional rating aligned with emotional state more likely to be chosen	Utility-Based Choice, Emotional State-Congruent Success Probabilities
[61][79] Emotional state-based Distraction OR Cognitive Flexibility	High-activation node appears in working memory, other than the node selected by deliberative subsystem.	New node(s) attended to by deliberative subsystem processes and checked via feature-matching for task relevance. The more relevant the newly attended node, the more this process resembles creativity than distraction	Emotional State-Congruent Recall, Emotional State-Dependent Attention Flexibility
[All of the above] Emotional state-Dependent Decision-Making	Deliberative task step taken	Task-dependent effects	All of the above, as the deliberative subsystem attempts to decide on an action to take

The entries in Table 4 are well-cited cognitive effects from studies of normal human emotions that affect decision-making, and that I replicated experimentally to underscore the realism of the model.

Chapter 4. Key Experiments and Results

4.1. Background

Experiments that can link emotional state to effects on cognition and that lend themselves to computational modeling are precious few. This is the challenge addressed in this research. Specifically, I developed an agent-based modeling framework to address this challenge. My model is sufficiently robust to support experimentation that would show the effects of emotional state changes on decision-making. A related challenge was to find appropriate human data as evidence for quantitative correlation. The strength of the framework lies in its ability to handle a variety of effects discussed in the cognitive science literature. The model experimentally demonstrates very different and even apparently contradictory results. In different experiments, neutral, positive, or negative emotional state proved optimal, depending on the decision-making task for the given experiment.

4.2. AX Continuous Performance Task Experiment

To ensure that Emowas calibrated to human data, I modeled Dreisbach's 2006 experiment [59]. The problem domain, the “AX Continuous Performance Task,” comprised a series of letter pairs presented one letter at a time. The letter A would appear first in 80% of all pairs, followed by the letter X in ~86% of the “A” cases. In 10% of all pairs, the X would appear after a different priming letter besides A. The subject was to press the left arrow key on a keyboard when the X appeared as the second letter in any pair, and the right arrow key if any other letter appeared as the second letter of a pair.

For the system's results to match the human experiment, positive emotional state would facilitate the “AY” cases (in which the A was followed by any letter other than X), but have little effect on, or even impair, performance in other cases. Overall, the AX experiment was meant to illustrate that greater flexibility (in positive emotional state) may lead to unexpected proficiency, depending on the task-relevance of the new node in focus.

4.2.1. Memory Contents and Game Elements

The memory contained one node for each single letter involved in the experiment: A, B, D, E, F, G, M, P, S, U, X, and Z. The ordered set of 200 “probe-target” letter pairs, read

by the perception module in random order, fit the distribution used in the human experiment: 70% AX, 10% AY (where Y is any of the above letters other than A or X), 10% BX (where B is any letter other than A or X), and 10% BY (where B and Y are not the same letter, and neither A nor X).

4.2.2. Task Procedure

The decision-making process for the AX experiment involved the influence of emotional state on simple choice. The choice was based on the expectation of the second (“probe”) letter in a “cue-probe” pair. After each “cue” letter was read, attention focused on the memory node representing that letter, and the decision-maker made its prediction as to the “probe” letter, at which point the actual “probe” letter was pulled into competition for attention. The associative system of Emo then performed its operations. If, at the next perception step, the perceived letter was different from the predicted letter, the two corresponding nodes would most likely compete for attention. The node that gained attention focus caused the correct or incorrect button to be pressed. The emotional state-based effect on this process was that a more positive emotional state allowed the attention focus to be switched more readily from the predicted node to the perceived node, which facilitated correct responses in the less predictable “AY” cases.

4.2.3. Training and Testing

As was done in the human AX experiment, the system was trained on the first 100 pairs of letters with the distribution as described above, in order to learn prediction strategies. The system was then tested on 100 additional letter pairs with the same distribution, for the same 9 constant emotional state settings as (-1.0 to 1.0 with 0.25 increments). The results recorded during testing were the numbers for correct and incorrect guesses for each of the four cases: AX, AY, BX, and BY.

4.2.4. Additional System Constants and Constraints

There were several important system-wide constant values that were set during training and remained the same for all experiments, for the sake of agent consistency across domains. Most importantly, an attention threshold value was set. This value refers to the difference in activation strengths between a potentially “interrupting” node and the node currently attended to by the deliberative subsystem. If the difference was greater than the threshold in favor of an interrupting distractor node, attention would involuntarily shift to that node.

Several bias values affected nodes in memory, mainly operating on activation strengths. There was a base activation bias for a node in working memory, and an additional base

activation bias for the node in attention. Spreading activation was similarly affected: there was a spreading activation bias from a node in working memory, and an additional spreading activation bias from a node in attention. There was also a decay factor lessening the spreading activation between any two nodes not in working memory.

4.2.5. Results

The results from the experiment are summarized in Figure 2:

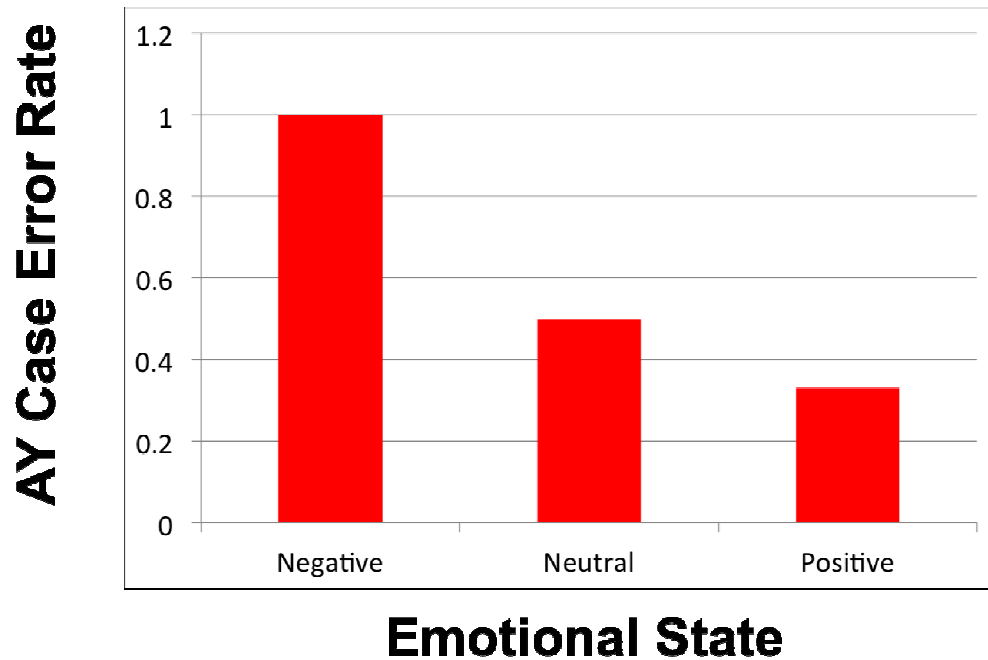


Figure 2. System errors in AY case averaged by emotional state

4.2.6. Correlation of Results to Human Data

Some caveats apply to the AX experimental results. First, in the human results emotional state was not quantified beyond positive, negative, and neutral. Second, the longer the time between cue and probe in the human experiment, the more exaggerated the effects. This time differential was not taken into consideration; the system correlated more to the 250ms lapse than the 1250ms lapse.

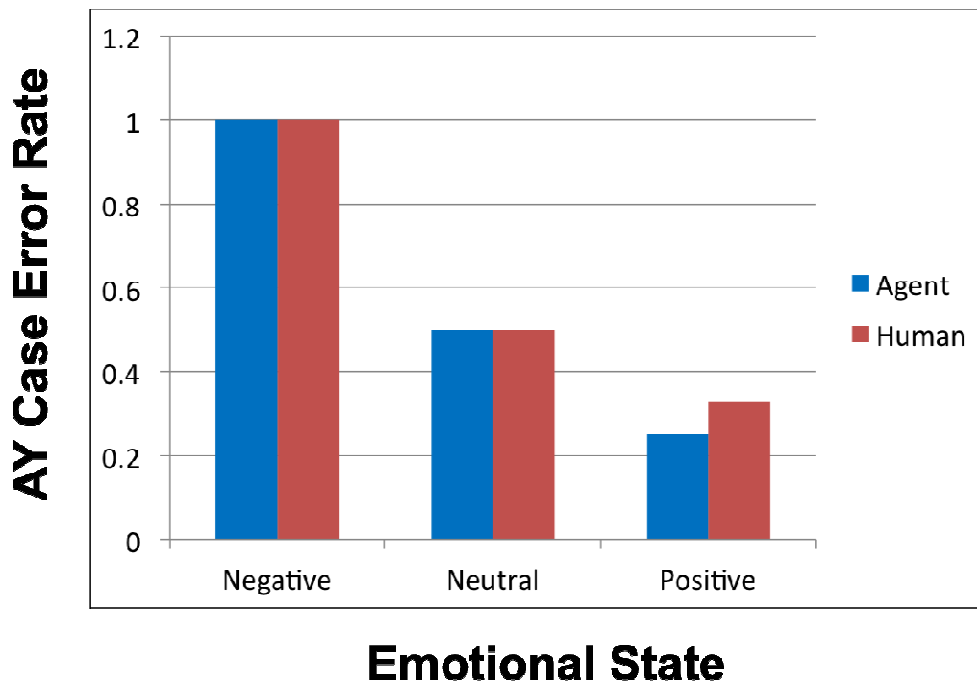


Figure 3. Correlation between agent and human AY errors

The result shown in Figure 3, with Emo generally matching the trend of the human experimental data, is that positive emotional state in combination with task-relevant distractors facilitates predictive decision-making when goals are abruptly switched during testing, but impairs the same when the original goal needs to be maintained. Again, without access to the full set of human data the numerical t-test can only show a surface correlation; that said, for the average emotional state-based values the t-test probability of the correlation hypothesis being correct was 0.987113.

4.3. Tower of London Experiment

To further calibrate Emo, I modeled Phillips et al's 2002 human experiment using the "Tower of London" (ToL) test of planning ability [46]. The experiment was designed to show that involuntary recall of emotionally charged nodes could distract Emo more often in either a positive or negative emotional state (as opposed to neutral emotional state). To summarize, the Tower of London problem consists of three posts on which are stacked several disks (in the human experiment's case, five disks) of equal size but different colors. The subject is given a starting configuration of the disks on the three posts, and a desired end configuration. The player's actions consist of moving one disk at a time from the top of one stack to the top of another. Unlike the similar Tower of Hanoi game, any disk can stack atop any other.

The expected outcome was that results obtained by the system would align with results from the human experiment. The human results showed that strong positive or negative emotional state leads to less planning time and to more trial-and-error moves made during decision-making. Moreover, the experiment was meant to show that one effect of distraction on the decision-making process was that fewer options were presented during choice, leading to decreased proficiency by an agent in a non-neutral emotional state.

4.3.1. Memory Contents

The setup for the experiment first involved loading a network of concept nodes and links into Emo's long-term memory. The concept nodes consisted of the seven colors of the rainbow along with seven similarly colored disks, plus the general concepts “Color” and “Disk,” and seven “associated” color-specific distractornodes, each one strongly (0.9) linked to a single color (see Appendix B for code and memory details for all experiments).

The ordered set of five disk colors was the same for each ToL game: 1=Red Disk 2=Orange Disk 3=Yellow Disk 4=Green Disk 5=Blue Disk. The gameplay rules and the start and end positions of Tower of London games were also given to the system, though these were not shared in working memory for this experiment.

4.3.2. Decision Making Procedure

Emo's decision-maker for ToL was a modified A* search algorithm to consider and decide on the next move to be made. After each deliberative step considered by Emo, attention was refocused on the particular disk to be moved. The associative subsystem then performed operations over the entire contents of memory (working and long-term): spreading activation, link association maintenance, and finally attention refocus check. These processes could cause the contents of working memory to change, as well as cause a shift in attention focus. In a case where a node other than the disk to be moved became the focus of attention, Emo was interrupted from its planning and would make a move based on the currently considered set of possible moves. Otherwise, Emo would consider all six possible moves, pick the best of the set, and then begin planning again.

Adding a wrinkle to these procedures was emotional state-congruent recall. Emotional state-congruent recall means that the more that emotional state varied from a neutral pleasure rating, the more nodes with similar emotional state ratings would be activated during spreading activation. As the only highly pleasant / unpleasant concepts in memory were not disks, but instead task-irrelevant distractor concept nodes like "Blood" and "Sky," strong emotional state in either direction was more likely to produce attention focus shift away from the disk being considered for a move.

4.3.3. Experimental runs

Emo was tested on 450 ToL games ranging from 4 to 11 minimum moves each. For comparison, the psychology researchers used three games with 5, 7, and 9 minimum moves for each of their 96 human subjects in positive, negative, or neutral emotional state (per subject self-report). The values tracked during the experimental runs were the number of moves made per game, the number of planning steps per move, the number of interrupts to the planning process and the number of times attention was wrested back to the correct disk for planning.

4.3.4. Results

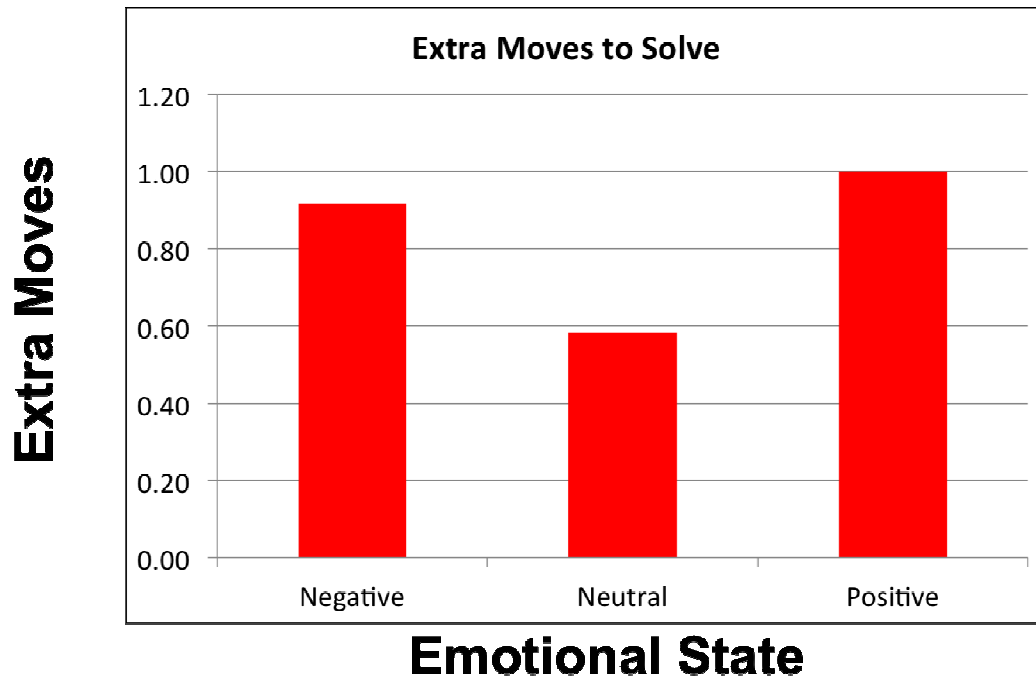


Figure 4a. Emo's Tower of London extra moves averaged by emotional state level

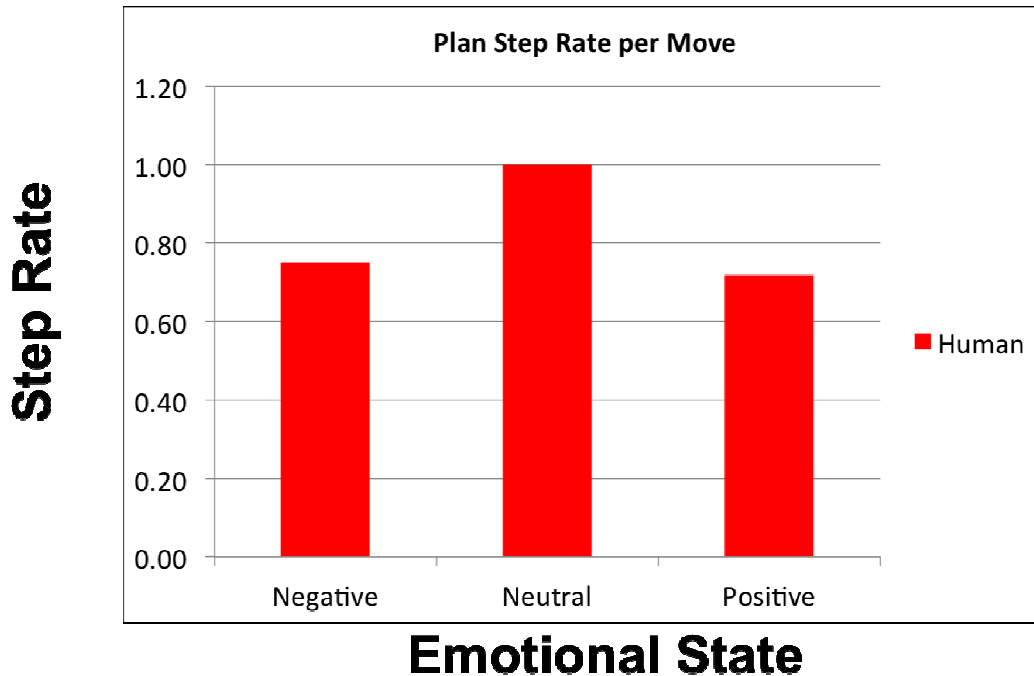


Figure 4b. Emo's Tower of London planstep rate per move averaged by emotional state level

4.3.5. Results' Correlation to Human Data

In analyzing the results and their correlation to the psychology researchers' data, three major comparison caveats must be noted. First, in the human results, only the emotional state self-reported immediately before the set of games was used for reference in the correlation. Second, planning time in the human data refers to the number of seconds taken before the first move, not the number of planning steps taken between moves. I correlated this pre-move measurement to the system's planning steps per move because

Emo was not given two separate phases for planning and moving, and Emo's only task between moves was the planning of the next move. Third, only the older age group's results (n=48, ages 53-80) from the human experiment were used in the correlation, as the emotional state difference effects were more prominent in that group.

For the data correlation, as shown in Figures 5a and 5b below, I used Emo's emotional state ratings of positive = 0.75, neutral = 0.0, and negative = -0.75. This corresponds to the human self-reports of induced emotional state ranging from 5 to 15 on a 20 point scale, as well as representing the average emotional states from the first, second, and third groups of system results. Without access to the full set of human data, it is difficult to produce a deep t-test correlation, but for the number of moves averages, the probability of the correlation hypothesis being correct was 0.809519. The steps vs. time in seconds correlation was not similarly not suitable for a full t-test without full human data, but for the averagecase, the probability of the correlation hypothesis being correct was 0.989123.

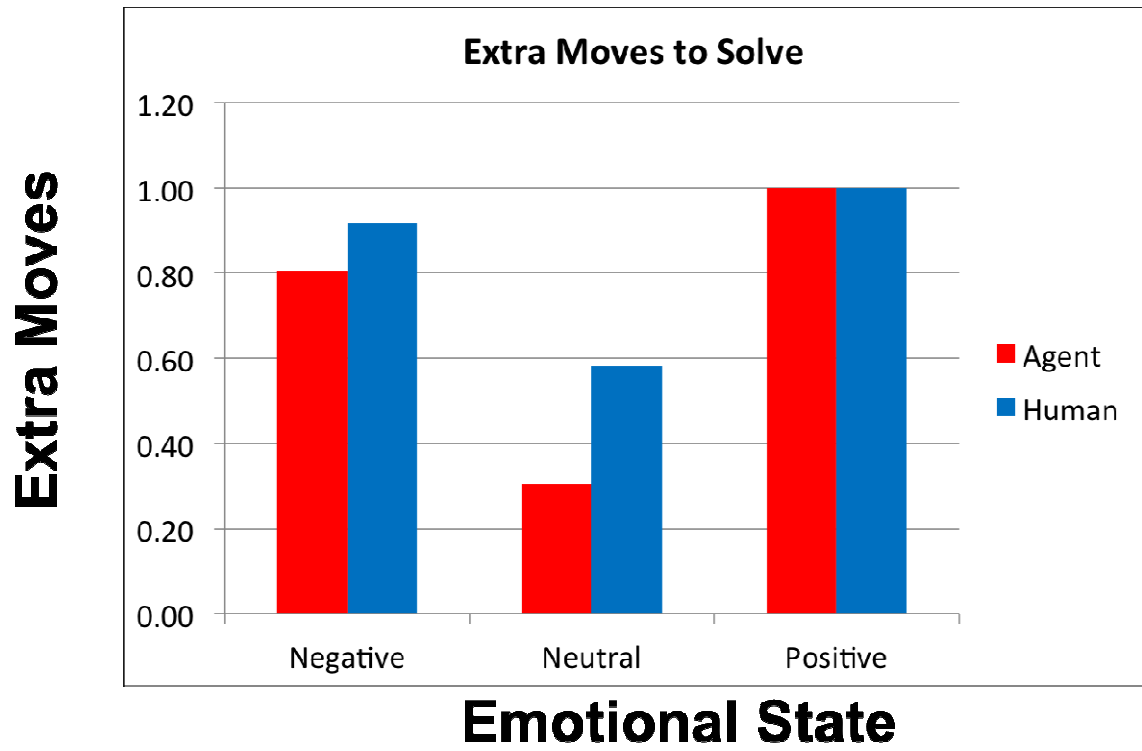


Figure5a. Correlation between agent and human Tower of London extra moves

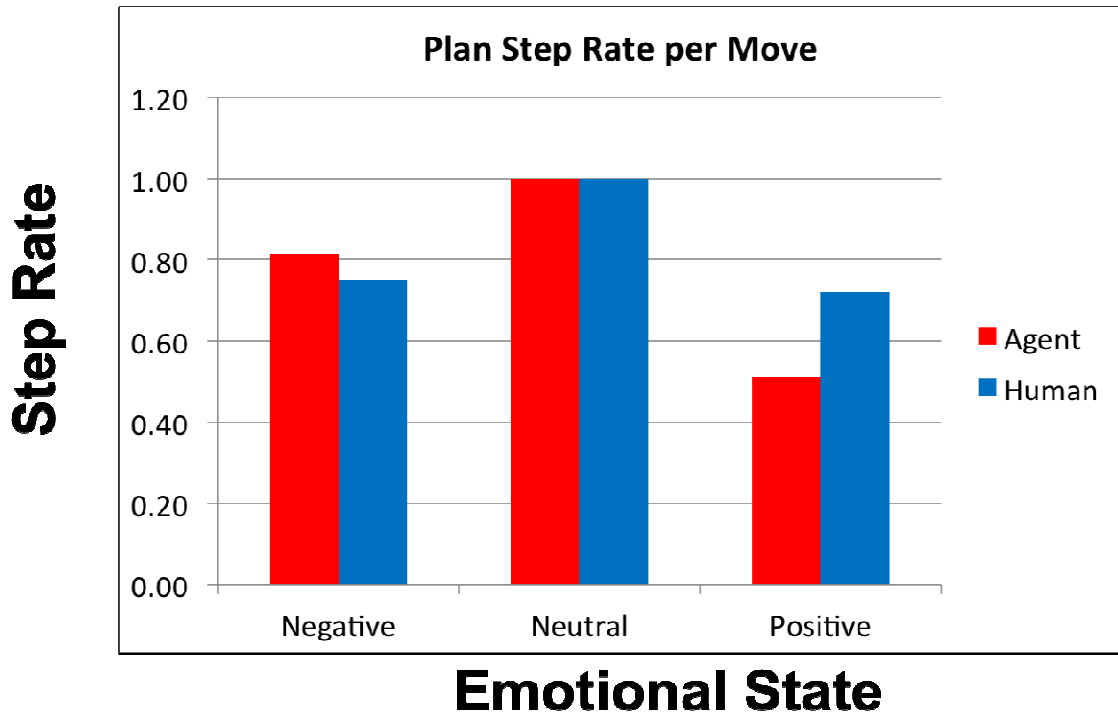


Figure 5b. Correlation between agent and human Tower of London plan step rate

4.4. Discussion of Results from Calibration Experiments

Within Emo, the perception module came into play during the AX experiment, as the perceived probe and cue letters were unexpected though predictable, which was the basis of the test. Perceiving a letter, similar to the deliberative subsystem's recalling of a letter, simply put the perceived letter into competition for attention focus. If the perceived “cue” letter was A, the predicted “probe” letter was likely X. If the perceived “probe”

letter was then not X, depending on the emotional state-based attention threshold, the correct letter may have been attended to (correct result) or not (incorrect result).

Results show the emergence of emotional state-dependent decision-making across these two experimental domains. The trend correlations to human data, even without changing the system's parameters, indicates the capability of the system to model human emotional state-based effects in a more important and sophisticated, yet less well-documented decision-making domain, such as operation of a nuclear power plant.

4.5. Nuclear Power Plant Experiment

4.5.1. Overview

The Nuclear Power Plant experiment was designed to demonstrate that Emo was capable of producing a range of emotional, and particularly emotional state-dependent, effects on making decisions during gameplay for problems with real-world impact and implications. There has been work in the human-computer interaction field towards developing [96] and evaluating [17] intuitive and comprehensive interfaces for plant operators to smoothly manage maintenance, prevention, and emergency procedures. These studies are in turn based on the procedural manuals from the IAEA and other official sources [97]. Other work has attempted to isolate the primary human factors that have led to power

plant accidents such as 1979's Three Mile Island disaster [17][20] but without the aim of computationally modeling the cognitive issues involved. Though panels have investigated Fukushima owner company Tepco's training and operation practices [98], the operator action data from the 2011 Japanese reactor meltdown has not been analyzed in any academic paper to my knowledge.

Human error encompasses several issues in addition to emotional state (e.g., vigilance, cognitive overload). Keeping Emo's perception ability and working memory size constant controlled these issues. Emotional state was the sole independent variable in the experiment.

4.5.2. Domain Model for Nuclear Power Plant

The problem domain as it appears in Emo's perception and memory are sets of potential action nodes and (partial) world state proposition nodes; some of the latter may be chosen as goals by the decision-maker. Along with Emo's memory nodes' standard emotional ratings, each action node, e.g., "turn off pump," includes estimated cost and success likelihood for utility function purposes. Each state node such as "pump on" also includes a cost associated with that state. The multiple nodes comprising the current world state are instantly perceivable by Emo, although checking actions for particular values may be required.

The Nuclear Power Plant scenario demonstrates the effects of emotional state on decision-making in a crisis simulation with implications for critical real-world operations.

4.5.3. Procedures and Experimental Methods

During the scenario enactment, Emo aims to achieve and maintain a perceived world state in which all goals are satisfied. It is important to note that even if there are no high-priority crisis-averting goal(s) at present, there are always recurring maintenance-level goals. To simulate this, some goal state nodes are set back to “needs checking” status. The user can inject faults in the scenario such as low water pressure, which initially Emo is unaware of. The fault is caused by a pipe rupture that Emo is also unaware of. Pursuant to checking the water pressure, Emo became aware of the pressure problem for the first time, and could begin to decide if a pipe rupture was present.

The scenario was run across Emo, its decision-maker, and a model of the nuclear plant itself.

1. Emo uses its deliberative subsystem's voluntary attention refocus, feature-matching recall and choice processes to select and assign weights to candidate actions. Emo's associative subsystem also affects these procedures as elaborated below

2. The decision-maker, based on a hypothesis deduction model written in Prolog, maintains a set of beliefs and goals based on the state of the reactor model, and can construct and revise hypothetical scenarios based on its beliefs in order to decide upon the best action (from a emotional state-dependent, weighted candidate set of actions provided by the main system) to satisfy its current goal.
3. The nuclear power plant reactor model, which has the actual state and the controls to alter that state. See Appendices for code specification.

At each step, Emo and subordinate models update themselves. The decision-making module within Emo initiates the cross-model activity, and the following event sequence occurs:

1. The decision-maker checks the state of the reactor model, and updates its goals.
2. The decision-maker calls Emo's attention to voluntarily refocus on the pertinent aspects of the reactor model state along with the proximate goal. This refocus triggers a search for a suitable action by Emo.
3. Emo runs its feature-matching recall in an attempt to recall into working memory a list of candidate actions to provide to the decision-maker. The feature-matching criteria are that a goal nuclear power plant model state is a known outcome of the needed action, and that the current nuclear power plant model state is a known precondition of the needed action. During this process, Emo's deliberative subsystem is more or less susceptible to distraction (for better or worse) by its

associative subsystem, depending on emotional state. If Emo refocuses attention on a node or nodes other than a list of actions, then the decision-maker loses its train of thought and goes back to the beginning of step 2 above. If not distracted by irrelevant nodes, Emo runs its choice process to assign weights to the candidate actions. The weights are based on the emotional ratings and success probabilities of the actions' outcomes. The ratings and probabilities themselves are adjusted up or down based on emotional state. Emo then returns the weighted action list to the decision-maker.

4. The decision-maker generates hypothetical plans incorporating each candidate action and its weight. The action that most likely leads to the most positive outcome fitting the most likely hypothetical scenario is chosen.
5. The decision-maker passes the chosen action to the reactor model subsystem, which updates the actual state to match the probabilistically determined outcome of the action.

The system as described above was run ~20,000 times on the same testing scenario: 100 times at emotional state values ranging from -1.0 to 1.0, with an increment of 0.01. The memory contents of Emo had uniform starting activation strengths of 0.5, and starting emotional ratings of 0, with one exception: the “pipe rupture” node had an emotional rating of -0.5, to demonstrate a case involving beneficial effects of negative emotion (which case only appears anecdotally in the cognitive science literature). An ablated

version of Emo was also run the same number of times. The distinguishing features of the ablated agent were: a constant neutral emotional state rating of 0.0, and immunity to any emotional-state-based effects.

The testing scenario was a nuclear power plant operator agent detecting low water pressure, modeled after a known potential power plant crisis involving operator decision and judgment[96][99]. During a routine system check, water pressure begins dropping rapidly, therefore increasing coolant temperature and threatening meltdown. Emo needs to realize first that the water pressure is low, and then perhaps judge that a pipe rupture may have occurred, based only on the numerical readouts. If a rupture is hypothesized, Emo must isolate and repair the ruptured pipe by using the bypass valve and emergency sealant spray. However, Emo may instead hypothesize that there is a non-pipe rupture related emergency that can only be stopped by shutting down the reactor. Inaction, or the inability to choose, will cause meltdown.

The pipe rupture fix, if there is a pipe rupture, is the less costly option, but a rupture is less probable than circumstances requiring a full shutdown. Moreover, fixing a pipe rupture that does not exist would still require a shutdown, and thus would cost the sum of pipe fix + shutdown. Therefore, a strictly utility function-based “by the book” operator would resort to the shutdown unless the pipe rupture somehow seemed to be more likely.

Emo was also subject to effects of distraction, which could lead to inaction and, potentially, meltdown of the plant. On the other hand, distraction by a task-relevant concept, (water pressure or pipe rupture, in this example) could lead to a better solution than might be found by focusing on the wrong aspect of the plant.

The nuclear power plant experimental hypothesis has four sub-hypotheses: 1) An agent in a negative emotional state is more likely to hypothesize that there is a pipe rupture, and act accordingly to avert the crisis. 2) An agent in a positive emotional state is more likely to hypothesize there is nothing wrong with the plant but will perform a shutdown rather than have the plant melt down due to inaction. 3) An agent in either an extremely positive or negative emotional state is more subject to distraction than an agent in a more neutral emotional state. 4) The emotionless version of Emo will always exhibit the same set of outcomes.

4.5.4. Results and Analysis

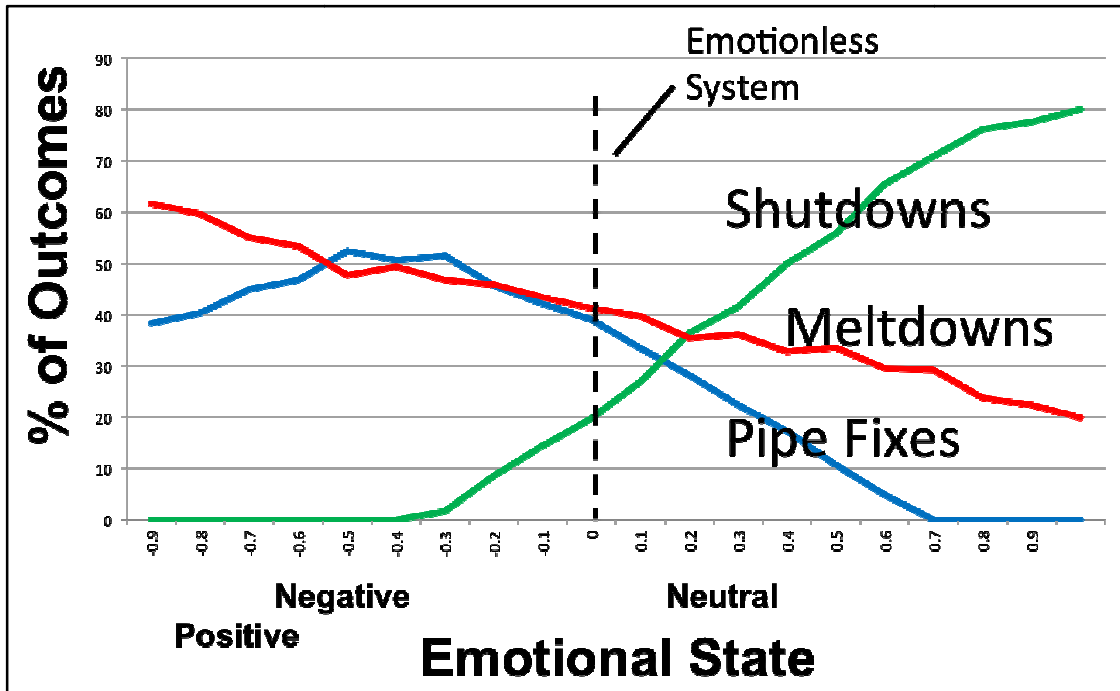


Figure 6. Nuclear power plant pipe rupture scenario outcomes

As shown in Figure 6, the sub-hypotheses are mostly borne out by the results. 1) The emotional agent is able to detect and resolve the pipe rupture more readily in a somewhat negative emotional state. Any negative emotional state stronger than -0.5 causes more meltdowns due to over-focusing on irrelevant details and subsequent inaction. 2) As Emo's emotional state increases to neutral and positive, shutdowns begin to occur more frequently, eventually outpacing pipe fixes. This effect is partially due to the negative pipe rupture concept having less activation strength when Emo is in a positive emotional state. 3) Shutdowns outstrip meltdowns (which represent pernicious distraction and inaction) for agents in more positive emotional states as well. This is because an agent in

a positive emotional state will readily choose a non-meltdown hypothesis and take appropriate action. However, in positive emotional states the pipe rupture hypothesis rarely increases in likelihood to warrant Emo to initiate a pipe fix; therefore, there are invariably more shutdowns. 4) The emotionless agent's performance was always the same, as it never had the effects of varying attention flexibility or various hypothesis likelihoods.

Collectively, the results imply that emotional state is an important factor in a decision-making scenario modeled on a real-world critical problem space.

Chapter 5. Discussion

5.1. Objective and Expected Outcome

I will begin by recapitulating the objectives and expected outcomes. Intelligent agents in games tend to exhibit behaviors that do not reflect humanlike qualities. In particular, they do not exhibit human emotional state effects on decision-making behavior in games. Even when emotional behavior is expressed in a few game agent architectures, such behavior is not informed by an underlying theory of emotion, nor is it quantitatively validated using human emotional behavior data. The objective of this dissertation was to present a new emotional agent architecture with both theoretical and experimental

underpinnings, and that manifests a range of emotional state effects on behavior, especially real-time decision-making behavior. The approach is informed by the appraisal and dimensional theories of emotion, which together ensure that emotionally appraised concepts in memory correspond with the emotional state of Emo, and that such resonance produces multiple realistic effects on Emo's decision-making behavior. The approach was validated in a series of experiments, of increasing sophistication in terms of both scenario and methods employed. The two calibration experiment designs were suitable analogs leading up to the nuclear experiment. The AX experiment modeled simple decision, and the Tower of London modeled planning; both of these were part of the decision-making process used by the decision-maker in the nuclear power plant experiment. In addition, both calibration experiments modeled Emo's distraction, emotional state-congruent recall and attention refocus, which were all also part of the effects on Emo in the nuclear power plant experiment. The results of the calibration experiments were correlated against human data from similar cognitive science experiments. The expected outcome of these results was that lightweight intelligent agents can exhibit realistic humanlike behavior in arbitrarily complex real-time games, across various domains.

5.2. Scope

5.2.1. Theoretical Scope

There are many competing theories of human emotion and decision-making behavior, and of the interplay between the two. In choosing which theories to explore in my model, I partially worked backwards from the effects that informed the experimental design. This led to using a hybrid theory in which emotional state has direct effects on cognitive processes, as well as a resonance with the emotional content of memory. A future challenge could be to work with different theories to produce similarly humanlike results.

5.2.2. Experimental Scope

Limiting the evaluation plan for the model is the relative lack of human emotional state vs. decision-making studies that quantitatively display results for all emotional state cases over a wide range of effects. The theoretical data is rich, but can only go so far in validating a computational model.

5.2.3. Modeling and Architectural Scope

Computational modeling of even a small subset of an organic system as complex as the human brain, without sacrificing critical elements of fidelity as regards the research hypotheses, is very difficult. The fidelity at the behavioral level of modeling needed to

be carefully calibrated, which is what I did. Deeper levels of fidelity were beyond the scope of this work; PC hardware is vastly removed from the physical nature of the human brain, and the brain's neural signals have physical and chemical nature very different from digital data transfer.

At a level of slightly greater abstraction, modeling the emergence of the conceptual level from low-level neural activity would also require an entirely separate dissertation. To study the relationships among emotional effects on cognition, I chose the phenomenological level for correlation with human experimental data, and to a lesser degree the conceptual level of data structures and process modeling based on cognitive theory and experimentation. I focused on the effects of emotional state on decision-making as, especially for game AI, these effects are relatively observable, quantifiable, and meaningful.

5.3. Research Findings

5.3.1. Summary of Methods

The experiments employed in this dissertation used a combination of computational modeling and statistical hypothesis testing. Results were evaluated by t-tests, correlating data from existing human experiments against data from agent models reflecting human

emotional state. My validation experiment, involving a nuclear power plant operator agent, exploited results from two calibration experiments.

To calibrate the system against human data, I modeled Dreisbach's 2006 experiment on emotional state and distractibility featuring the "AX Continuous Performance Task," and Phillips et al's 2002 human experiment on emotional state-dependent planning using the "Tower of London" test. The model can also support several other experiments from cognitive science that show direct effects of emotional state on decision-making; the two modeled experiments present the most quantifiable results over a range of effects.

The nuclear power plant experiment was used to further validate the calibration results on a far more complex problem space with real-world impact and implications.

5.3.2. Experimental Results Summary

In the AX experiment, although Emo was primed to expect X after A, in cases where X did not appear Emo pressed the correct button more often when in a positive emotional state. The t-test correlation between agent and human results was ~90%. In the Tower of London experiment, performance was measured by the number of extra moves needed by Emo to solve the problem. The t-test correlation between agent and human results was ~90% for this experiment as well. For the nuclear power plant experiment, the emotional

agent was able to successfully hypothesize and troubleshoot a pipe rupture, based on a low water pressure reading scenario, more often and more readily in a somewhat negative emotional state; these results were not attainable by the emotionless version of Emo.

5.3.3. Discussion of Findings

Collectively, the results confirm that emotional state is an important factor in decision-making scenarios in game domains, scalable to model on a real-world critical problem space. The videogame industry and human development initiatives have become increasingly related in recent years, with more call for effective, engaging, and realistic serious games for education, professional training, and therapy. Specifically, social and behavioral models have been sought after to validate and enrich the experience of these games for users, developers, and researchers alike [100].

Computational modeling of emotion is essential for realistic humanlike AI decision-making in games and simulations. My research provides such a model implemented in a new system suitable for a game AI. The system realistically and flexibly models emotional state-based effects on deliberative and associative processes, validated by its capability to model several human emotional state-based experiments.

5.4. Implications and Conclusions

5.4.1. Key Findings and Research Contribution

My dissertation is concerned with developing computational models of human decision-making that are sensitive to emotional state. Specifically, the model developed for this dissertation incorporates emotional state effects on human decision-making behavior. My contribution is to enrich intelligent agent behavior in games. My agent architecture computationally models documented and quantifiable emotional state-dependent effects with a high degree of fidelity to human data, enabling realistic humanlike decision-making. T-tests between existing human experiment data and system model results of AX and Tower of London experiments show ~80-100% correlation fidelity.

Also, the combination of direct emotional state effects produces emergent behaviors that model human decision-making. The approach is applicable to multiple game AI domains. Implemented experiments for AX, Tower of London, and Nuclear Power Plant operation illustrate different aspects of AI game agent decision-making, and show how emotional state effects combine to produce positive and negative emergent decision-making behaviors across different game domains.

The experimental findings justify the approach that I have taken, particularly in terms of the four components from cognitive science and computational modeling that I used as guidelines for designing Emo's architecture.

The first component is a theory of emotion supported by cognitive science. Emo uses a hybrid dimensional and appraisal theory of emotion, as supported by cognitive science. Working from dimensional theory, agent emotional state (core affect) is a global value. Based on appraisal theory, each node in memory also has a similar emotional value. The justification for this modeling choice is flexibility. All of the behavioral and cognitive effects in this work, for instance the emotional state-dependent recall and attention refocus that informed distraction in the Tower of London experiment, are based on the interaction between the emotional content of memory and the overall emotional state.

Another necessary component was the memory network that was evaluated for emotional content. The central data structure in Emo architecture was the associative-semantic node network that represents long-term memory. An arbitrarily complex network allows scalability of the experimental domain from AX to Tower of London to a nuclear power plant. The links allowed the network to model a humanlike dual system of associative (unconscious) and searchable / deliberative (conscious) memory.

A humanlike cognitive process model sensitive to emotional state acts upon Emo's memory network. In particular, Emo uses an associative/deliberative dual process model as outlined by Stanovich and West [101]. Together, the dual subsystems grant humanlike realism to Emo's process model. Unconscious and conscious work together or in conflict, and all processes are suffused by and susceptible to multiple emotional effects. The interplay between the two systems was most directly shown in the nuclear power plant experiment, as Associative subsystem-based optimism and pessimism played a direct role in the Deliberative subsystem deliberative decision process.

The model uses only three direct emotional state-dependent effects whose combinations lead to other, emergent behaviors. Increasing the number of hard-coded parameters would likely mean more unforeseen interactions and corrective hacks. The first direct effect is emotional state-dependent activation strength adjustment. Memory nodes with emotional rating closest to Emo's emotional state gain proportional activation strength, and vice versa. This leads to increased recall and attention focused on such nodes. The second effect is emotional state-dependent attention flexibility; the chance of involuntary attention refocus to a new stimulus is directly proportional to Emo's emotional state. The third direct effect is emotional state-dependent success probabilities. Emo's predicted probability of an outcome (of an action or an event) is directly proportional to how closely the current emotional state matches the emotional rating of the node associated with that outcome. The predicted probability is then used by the choice utility function.

As seen in the nuclear power plant experiment, one behavior indirectly dependent on emotional state is optimism / pessimism (e.g., in a negative state, a negative event like pipe rupture seems more likely, and is also more likely to be recalled). Another is distraction; a deliberative task can be interrupted by attention refocus to a node with higher activation strength, which led to meltdowns in the nuclear scenario. Similar to distraction is creative solution finding, where the distractor is actually task-relevant, as when in more positive emotional states shutdowns and pipe fixes began to outpace meltdowns.

To present a richer decision-making process integrated more tightly with emotional state, the implemented experiments can be revised with several enhancements, which give expected results based on theory, but are not quantifiably comparable against human data.

5.4.2. Current Experiment Redesign

First, the AX experiment's decision maker could be generalized to recognize bigrams in addition to single letters as concepts in memory. These bigrams may be emotionally

charged two-letter words [102], and resultantly the anticipation of a bigram based on the first card in the two-card sequence may depend on emotional state. The mechanism used to track these bigrams would be a two-dimensional interaction matrix: the first letter on one axis, and the second letter on the other. This method could potentially scale to trigrams and beyond, and also to interactions between bigrams as they appear consecutively in the overall sequence of an experimental run.

The Tower of London could have a tighter binding of the AStar-based decision-maker to emotional state, meaning that the planning actions and potential positions, as well as the disks themselves, are concepts in memory subject to emotional effects. The mechanism would be similar to the enhancement to the AX decision-maker, in that a multidimensional interaction matrix will cover all possible chunks of disk position (e.g., recognition of single stack status vs. multiple stack), and actions taken that affect those chunks. Even sequences of multiple actions could be tagged with emotional ratings, though that would be beyond the scope of the initial revision.

The nuclear power plant experiment could be upgraded to include multiple and sequential states and actions in memory, maintained in an interaction matrix. The decision-maker would have several alternative hypotheses to choose from, to determine the cause of low water pressure.

Given Emo's flexibility to use many different high-level decision-makers, Emo's nuclear power plant operation decision-maker could then be augmented to use a multi-level troubleshooting hypothesis deduction model. At the ground truth level, a pipe has ruptured in the power plant. At the evidentiary level, the only relevant readout on the console is that the water pressure is very low. On the level of cause discernment, competing hypotheses can account for this evidence:

1. The console is incorrectly displaying low water pressure and the plant is actually functioning normally.
2. There is an undetectable loss of primary pump power.
3. There is a pipe rupture somewhere along the water route.

An operator weighing the likelihood of each hypothesis against the others may resolve the competition outright if one of the hypotheses seems far more plausible. On the other hand, none of the hypotheses may be likely enough to warrant a particular course of action (or inaction). In such a case, Emo would try a test action that could lend evidence to support or deny one of the hypotheses. For instance, if starting the emergency pump does not raise the water pressure, this removes credence from the hypothesis that the primary pump is deficient. If enough credence is gained, a hypothesis can be confirmed and further action taken.

A completely rational operator would follow the above troubleshooting procedure and balance the credence of each hypothesis by evidence alone. However, Emo's emotional state would resonate with the emotional rating of each concept encountered during troubleshooting, causing distortion of rational behavior. For example, the pipe rupture hypothesis (being the most negatively charged of the three) would seem more likely in a negative emotional state, perhaps to the point where no contradictory evidence, no matter how strong, could dissuade Emo. Conversely, in a positive emotional state the attention flexibility of Emo might make its belief waver among several hypotheses, given even marginally supportive evidence for any of them.

This enhancement would allow Emo to choose courses of action more as a human troubleshooter would do, stepwise, and based on evidence as opposed to relying on hardcoded probabilities.

5.4.3. Enhancement: Emotional Feedback and Emotional State Change

The revised experiments can provide an opportunity to generate changes to the emotional state[37]. In the AX experiment, these changes would be based on the appearance of an emotionally charged letter and/or bigram. The new Tower of London experiment would include alteration of emotional state based on repeated moves and positions. Emotional state may also change based on Emo getting closer to or further from the solution based

on heuristic assessment. In the nuclear power plant experiment, the emotional state would change depending on the criticality of readings as well as the success or failure of actions taken.

5.4.4. Additional Cognitive Modeling Experiments

Several adjustments to Emo's processes could be made in order to model further cognitive science experiments involving decision-making and emotional state. The additional experiments are listed below, along with the adjustments needed to model the experiments in Emo.

Table 5. Summary of further experiment modeling

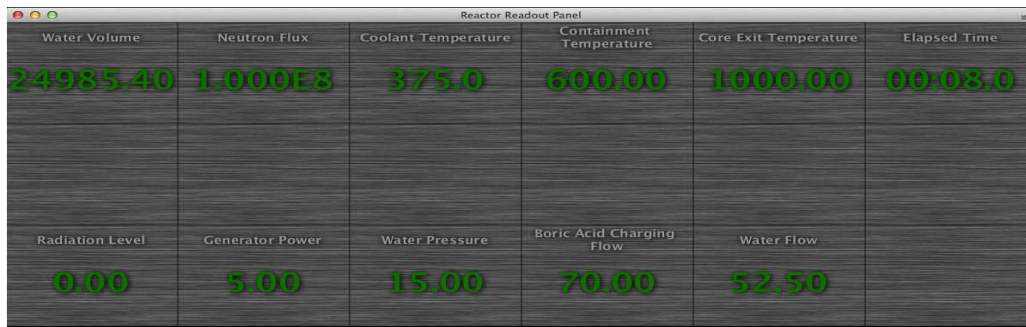
Study	Task	Result	Implementation Notes
[53] Spies, 1996	Working memory span	Positive emotional state: impairment. Negative emotional state: impairment	Emotional state-based working memory size adjustment
[54] Hesse, 1996	Word vs. non-word recognition	Neutral emotional state: no effect Negative emotional state: facilitation	Feature-matching faster but stricter in negative emotional state
[55] Oaksford, 1996	Tower of London	Positive emotional state: impairment Negative emotional state: no effect (tends towards impairment)	None needed

[56] Gasper, 2002	Switching to novel word-finding strategy	Negative emotional state: facilitation after switch instruction	Feature-matching strategy switches more readily in positive emotional state
[46] Phillips, 2002a	Tower of London (young subjects 19-37)	Positive emotional state: no effect Negative emotional state: no effect	Lower distraction threshold for all emotional states
[57] Phillips, 2002b	Initial letter fluency (syntax association)	Positive emotional state: no effect over neutral emotional state	Feature-matching faster but stricter in negative emotional state
[57] Phillips, 2002b	Uses of objects fluency (semantic association)	Positive emotional state: facilitation over neutral emotional state	Feature-matching faster but stricter in neutral or negative emotional state
[57] Phillips, 2002b	Switching syntax and semantics processing fluency; Stroop color test; switching Stroop color and word identification	Positive emotional state: impairment vs. neutral emotional state	Feature-matching strategy switches more readily in positive emotional state
[58] Dreisbach, 2004	Switching to novel stimulus	Positive emotional state: facilitation Negative emotional state: no effect	Feature-matching bias toward novel stimuli in positive emotional state
[58] Dreisbach, 2004	Switching to inhibited stimulus	Positive emotional state: impairment Negative emotional state: no effect	Feature-matching bias toward novel stimuli in positive emotional state

Generally, the agent framework that I have created is capable of accommodating all of the above experiments, though the modeling of each decision-maker and the corresponding problem domain and memory contents would need customization.

5.4.5. Towards a Nuclear Power Plant Turing Test

At the present, the framework is in place to run a Turing Test-inspired experiment for further evaluation of the humanlike qualities displayed by the system. In the pilot study, two groups of human subjects would be involved: “Testers” and “Observers.” The Testers would use a software-based procedural help manual and a GUI-based console that I have implemented (shown in Figures 7-9 below), to simulate an interface for maintenance and emergency operations in a nuclear power plant control room. All operations that a Tester initiates would have a cost associated with them (potential irradiation), and some operations would have a limited chance of success.



The image shows a screenshot of a software interface titled "Reactor Readout Panel". It features a dark grey background with a grid of six columns and two rows of data. The data is displayed in a green monospace font. The first row contains: Water Volume (24985.40), Neutron Flux (1.000E8), Coolant Temperature (375.0), Containment Temperature (600.00), Core Exit Temperature (1000.00), and Elapsed Time (00:08.0). The second row contains: Radiation Level (0.00), Generator Power (5.00), Water Pressure (15.00), Boric Acid Charging Flow (70.00), and Water Flow (52.50).

Water Volume	Neutron Flux	Coolant Temperature	Containment Temperature	Core Exit Temperature	Elapsed Time
24985.40	1.000E8	375.0	600.00	1000.00	00:08.0
Radiation Level	Generator Power	Water Pressure	Boric Acid Charging Flow	Water Flow	
0.00	5.00	15.00	70.00	52.50	

Figure 7. Nuclear power plant reactor readout panel

Secondary Control Panel						
Auxiliary Cooling Rods	Backup Generator	Bypass Valve	Containment Chamber Absorption Gas	Containment Isolation Valve	Containment Spray Flow	Control Rods
Undeployed ?	Offline ?	Closed ?	Undeployed ?	Open ?	Offline ?	Undeployed ?
Coolant Makeup Flow	Emergency Boric Acid Charging Flow	Emergency Boric Acid Dilutant	Emergency Boric Acid Relief Valve	Emergency Condensate Chamber Reflux	Emergency Containment Pressurizer	Emergency Core Partitions
Offline ?	Offline ?	Undeployed ?	Closed ?	Offline ?	Offline ?	Undeployed ?
Emergency Decrease Core Pressurizer	Emergency Feedwater Pump	Emergency Fuel	Emergency Fuel Damper	Emergency Hydrogen Absorption Rods	Emergency Hydrogen Flow	Emergency Increase Boric Acid Pressure
Offline ?	Offline ?	Undeployed ?	Offline ?	Undeployed ?	Offline ?	Offline ?
Emergency Increase Core Pressurizer	Emergency Neutron Absorption Material	Emergency Sealant Spray	Emergency Steam Release	Emergency Throughput Valve	Emergency Water Pipe Rerouting	Emergency Water Release
Offline ?	Undeployed ?	Offline ?	Offline ?	Closed ?	Offline ?	Offline ?
Extra Fuel	Filtration Fan	Generator Auxiliary Power	Generator Electricity Feedback Loop	Generator Emergency Power	Heat Residual Exchanger	Insulator Dump
Undeployed ?	Offline ?	Offline ?	Offline ?	Offline ?	Closed ?	Undeployed ?
Reactor Coolant Pump	SCRAM Button	Secondary Feedwater Pump	Secondary Feedwater Valve		Water Pump	Relief Valve
Offline ?	Undeployed ?	Offline ?	Closed ?		Offline ?	Closed ?

Figure 8. Nuclear power plant control panel

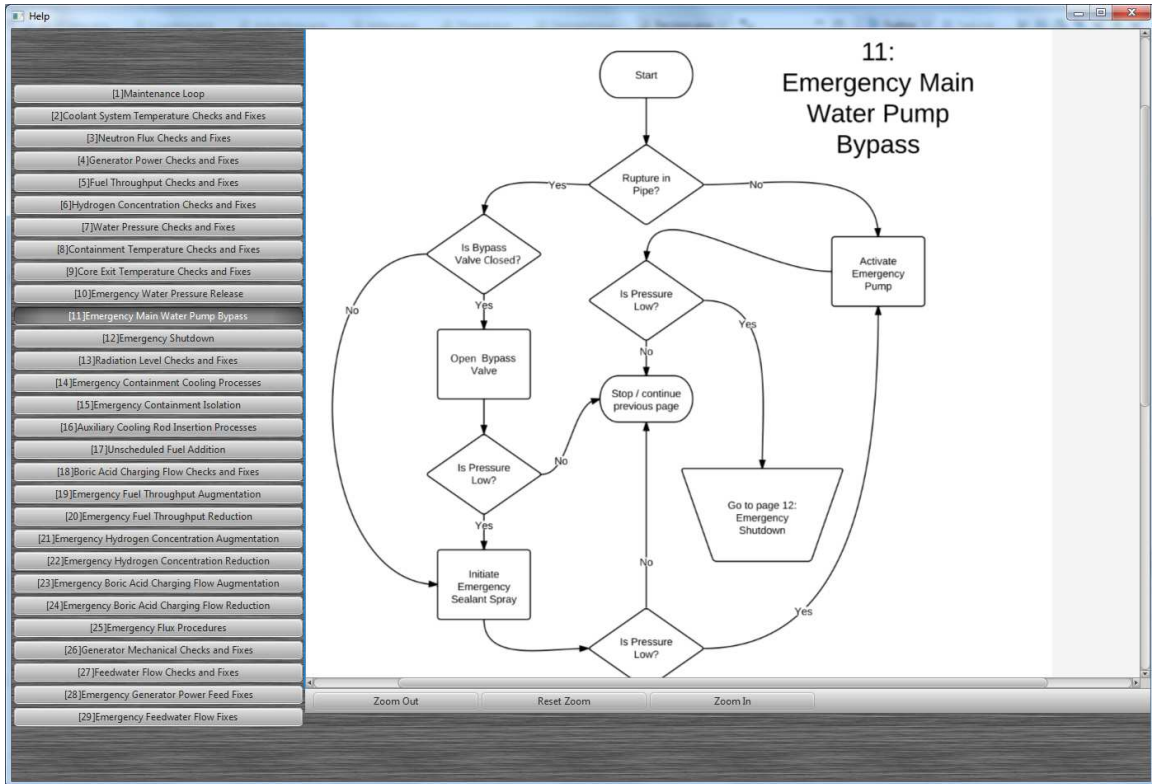


Figure 9. Nuclear power plant help manual screen

The experiment scenario for the system would cover three phases. Beforehand, the Testers would be asked some background information, including current emotional state, and then the console interface and help flowchart system would be described to them. A human cognitive science experimenter would carefully vet the test scenario and the experimental setup. After the experiment, the Testers would be given a survey of their emotional states and emotional responses during the tests.

The two versions of the system—the full emotional version and the ablated “emotionless” version—would be confronted with the same scenario as the Testers. The ablated system would be modified, as in the previous experiments, to ignore emotional cues and to remain in a constant neutral emotional state, minimizing emotional effects on decision-making.

The hypothesis behind the experiment is that the emotional version of the system is more likely to be mistaken for human than the emotionless version, based on the judgment of human Observers.

Observers would see text readout reports and visual playback of console operations from the Testers, mixed randomly with similar session results from the two system versions. System playback would be calibrated to match the speed of a human Tester, in order to control that variable for Observers. The Observers’ goal will be to determine which reports and playback represent human Testers, and which were generated by the system.

5.4.6. Agent Enhancements

The previous chapters outline Emo’s combination of dimensional and appraisal theory in memory and processes. The state of the art in agent architectures using both dimensional theory and appraisal theory is represented by systems such as EMA [37] and WASABI

[40], though these systems were designed for different purposes than was my agent model: emotional plan evaluation and embodiment / expression, respectively. Based on study of these systems and on my own work, a hybrid model of emotional theory is quite expressive in terms of generating effects on decision-making: the resonance between agent emotional state and emotional memory contents, which informs many effects, is derived from the combination.

5.4.7. Model Augmentation

Learning would become a primary element in upcoming revisions of the model. My current model uses preset, static emotional memory, whereas human emotional state-congruent recall and choice are dynamically primed by learning under particular emotional states. Perception-based experiential learning, as well as inference, would make the system's memory truly dynamic by creating and updating nodes and links. The learning module would be reinforcement-based, from outcomes of perception, cognition and action, enabling case-based planning and episodic memory [28][29][103]. Learning could be invoked voluntarily or involuntarily to write to long-term memory. The inference module would likely be built as a rule-based production system. Both new modules would be subject to emotional state effects as are the currently implemented deliberative subsystem processes.

Sophisticated time awareness would also be a key aspect of future versions of Emo. A rich set of past experiences and future expectations could inform the modeling of several more emotional effects on decision-making. Additionally, the time-based Affect Heuristic effect would be enabled, which would cause Emo to rely more heavily on emotional criteria for decision making given a tighter time schedule to select an action.

The current belief in cognitive science is that arousal is as important as emotional state for determining and modulating effects on decision-making[104]. High-arousal states lead to narrowed attention, while low-arousal states lead to broadened focus (e.g., more than one concept node in attention focus at a time). The Yerkes-Dodson law [9] illustrates that moderate arousal can aid recall via motivated learning, especially when the arousal is integral to the learning task at hand. However, higher levels of arousal act as a cognitive load [105] and thereby become too distracting to allow focused attention to a task, and recall suffers accordingly. Other studies show a more nuanced tradeoff of positive and negative effects on learning and recall due to arousal. According to a 1963 experiment [106], learning under high arousal facilitates long-term recall of the learned subject, but short-term recall is impaired. The opposite effect occurred in the subjects' learning under low arousal. The arousal factor and its effects would be tightly integrated in future iterations of the model.

There are also a number of discrete, labeled emotions that share high or low emotional state and arousal ratings, e.g. fear and anger. Even with the inclusion of dominance as a factor in the affective model, emotional states like foreboding or directed anger are not easy to model without using the semantic links between nodes over an associative memory network to specify future events or other referents. Future work would augment my agent's hybrid appraisal dimensional / emotional model[107][108], using Table 6 below as a starting point:

Table 6. Augmented emotional model

Emotional State	Arousal	Dominance	Time	Referent	Provisional Emotion Label
+	open	+	Future	Y	Anticipation
+	open	-	Future	open	Hope
-	+	+	Future	open	Courage
-	+	-	Future	Y	Foreboding
-	-	open	Future	open	Neutrality
+	open	+	Future	N	Confidence
+	open	-	Future	N	Optimism
-	open	+	Future	N	Fatalism
-	open	-	Future	N	Pessimism
+	open	+	Present	Y	Appreciation

+	open	-	Present	Y	Awe
-	+	+	Present	Y	Anger
-	+	-	Present	Y	Fear
-	-	open	Present	Y	Disinterest
+	+	+	Present	open	Joy
+	open	-	Present	open	Thrill
+	-	+	Present	open	Satisfaction
-	+	+	Present	N	Frustration
-	+	-	Present	N	Anxiety
-	-	open	Present	N	Boredom
+	open	-	Past	Y	Nostalgia
-	open	open	Past	Y	Guilt
-	open	-	Past	open	Regret
+	open	open	Past	Y	Relief

5.4.8. Applications of Future Work

Generally, I intend to integrate Emo with other ongoing research such as emotion generation processes. This would begin with the creation of a feedback loop between

cognition and emotion. For instance, success or failure of an action may alter emotional state, and this state may in turn affect the formation of future decisions.

The model may also be applied to emotion-enabled serious game engines and virtual training environments, providing more realistic and valuable interactions with humanlike characters than are currently available. Emo may also be used as a module in other cognitive architectures.

The main prospective application for the nuclear plant simulator is behavior modeling, leading to training. Numerous power plant operators, at various levels of experience and in various emotional states, would use the system to run several crisis-averting scenarios in their professional training. The operators' data would be recorded and incorporated into parameters that could calibrate Emo towards a predictive model of standard operator behavior. Emo would later perform many runs of the operators' test scenarios in various emotional state settings; analysis of Emo's results compared with individual operators' results would suggest which operators could benefit from further training under emotionally stressful conditions.

Besides nuclear power plant operation, other real-world domains have similar requirements for training: complex human-computer interaction combined with emotional stress. Air traffic control trainees could use the model very similarly to nuclear

operators. The aspects of Emo that would need to change significantly would be the world state (modeling air traffic interactions and interface), the memory contents (airplanes and weather patterns), and decision-making rules. The main system would remain the same as the one used in the nuclear experiment.

One intriguing challenge for the Emo model is that it could be used in conjunction with human emotion detection to enable an affect-based level of interaction between human users and AI. For example, a training application for therapists who work with Post-Traumatic Stress Disorder (PTSD) patients could incorporate the model into a system that tracks when and how an AI-based patient is emotionally triggered, then advises a therapist how to manage trigger-based crisis situations. Concurrently, the therapist's own affective signals would be monitored by the system. Those signals could engender calming or excitatory effects on a PTSD patient, and when appropriate the system would provide suggestions to the therapist for coping with emotional stress in themselves as well as their patients.

Appendix A: Further Background and Discussion

A.1. Human Planning, Problem-Solving, and Decision-Making

The fields of human cognitive science and artificial intelligence define the basic process of planning similarly [109][110], as creating and executing procedures for achieving particular goals or desired outcomes. The basic creation of plans has also been described as a search through a space (or set of related spaces) of connected problem states [111], or as future-oriented problem-solving [112]. Beyond problem solving, planning can involve predicting outcomes, scheduling, testing solutions, plan revision, evaluation after the fact, and learning from previous outcomes [113]. Planning domains for both humans and computers can include well-defined problems such as chess as well as ill-defined domains like grocery shopping [114].

Human planning uses a limited working memory (a structure wherein plans can be retained temporarily when they are being formed or transformed or executed), including a “problem space,” and larger long-term memory which includes goal and action ‘libraries’ [115], possibly augmented by reference to externally stored information [116]. Planning involves complex processes such as decision (i.e., how/when to solve a problem) and simple processes such as recall [90][109]. Attention, time (urgency) and other heuristics guide decision, search, recall, and other planning-related processes. Case-based reasoning is also in effect during planning: this amounts to expression of recalled plan preferences [117]. Further, expression of planning is determined by the influences of prejudices and preferences. [117][118][119].

Other computational characteristics of human planning involve incremental interleaving of planning (refinement) and execution phases, allowing cognition to opportunistically react to environmental changes [118][120]. Human planning is also partial-order and bidirectionally sequential, e.g., backward-chaining from the goal state and concurrently forward-chaining from start state, or top-down and bottom-up processing if a hierarchical task network or similar is to be modeled. Global/abstract changes are made to higher-level goals, and local/concrete details can also be changed at the lower levels.

Three specific computational models of human planning appear in the cognitive architectures of Soar [121], ACT-R [122], and CLARION [123]. Soar models planning as a specific case of cognitive activity, i.e. successive refinement of state structures by operators within multiple problem spaces. A significant aspect of ACT-R is its modeling of memory as a network of nodes bound by weighted associative links, which can inform search, recall and decision during the planning process. CLARION's modular cognitive processing architecture contains subsystems of various processing speeds that can conflict with and interrupt each other. Particularly pertinent to my work is CLARION's motivational subsystem (Associative subsystem) having the capacity to interrupt the deliberations of the action-centered subsystem (Deliberative subsystem).

A.1.1. Reactive and Deliberative Planning

The human capacity for planning developed in a partially unpredictable environment. One form of planning suited for that type of environment is an integration of reactive and deliberative planning [124]. Deliberative planning manages the long-term goals and formulation of plans to achieve those goals, while reactive planning does periodic maintenance on the plan such that if changes to the world state have obviated a plan step or made a goal inaccessible, the plan can be repaired or overhauled. Though sophisticated, this “one-shot” type of planning has shortcomings, as the planner does not automatically learn from experience.

A.1.2. Case-based Planning

Planning with an experience-based learning module enables “case-based planning” in which previous uses of a plan, plan step, or goal can be evaluated from past experience [103]. Integrating the above reactive/deliberative planner with a case-based system is more in line with the human use of memory and experience in both reaction and deliberation, and particularly allows emotionally charged experiences to influence these processes. Sufficient training of a case-based planner allows the creation or modification of rules to follow given a familiar problem space. In recent versions of Soar [125], combined episodic / procedural / semantic memory structures, supported by reinforcement learning, have been used to enable case-based reasoning.

A.2. Additional Modeling Background

A.2.1. Activation and Association Maintenance Modeling

The standard ACT-R [122] equation for the activation strength A_i of node i ,

$$A_i = B_i + \sum_{(j)} W_j S_{ij},$$

where B_i is the “base activation” of node i , W_j is a weighting factor of associated node j , and S_{ij} represents the strength of association between i and j ; modified in my model to incorporate the emotional impact of associated nodes.

In ACT-R, the value of a node i 's base activation is

$$B_i = \ln \left(\sum_{(k)} t_k^{-d} \right),$$

where t_k is the elapsed time since the k th activation of node i , and d represents a constant decay factor.

A.2.2. Unified Model of Process Cooperation

The blackboard architecture [126] is a basic computational model of distributed process cooperation that enables implementation of the unified models of cognition described by Newell and Kahneman. The architecture has as its basis the “blackboard,” a shared knowledge base, which in my system is the network of memory. The classical blackboard architecture also incorporates multiple specialist processes that provide partial solutions to a given problem on the blackboard. The blackboard architecture also contains a control shell or meta-module that moderates the specialists’ activities. In future versions of the system, that portion of the architecture could have several levels and components, such as the standard low-level processes of a modern digital computer, as well as the higher-level transactional control over asynchronous processes’ access to conceptual nodes in working memory.

A.3. Further Discussion of Emotional Theory

There are several psychological theories analyzing human emotions. In the 19th century, the psychologist William James and others theorized that emotions were brought on by physiological reactions to situations. James’s theory was a precursor to appraisal theory, whose proponents also view emotions as effects of reactions to situations, though with less of a focus on physiological reactions. Appraisal theory is dominant in the

community of computational emotional modeling, though other schools of thought have also made an impact in that arena. Within appraisal theory itself, there are several lines of subdivision.

Appraisal theory was developed as a means to predict individual human emotions given particular situations [127][128][129]. The basis of the theory is that a person can appraise (i.e., evaluate) an entity, concept, event, or situation with respect to the appraiser's beliefs, desires, and intentions. This evaluation can be organized by factors called appraisal variables, and a certain combination of appraisal variable values (collectively, an "appraisal frame") predictably gives rise to a distinct emotion. The mapping from appraisal variables to emotion has been termed "affect derivation" (often coupled with or subsuming a derivation of emotional intensity). For instance, a swimmer might appraise a shark encounter as likely to result in serious physical harm, and this appraisal would generate an intense emotion of fear in the person.

One way in which appraisal theories differ from one another is in the number, breakdown, and definition of appraisal variables accounted for by each theory, and the ways in which the variables combine to predictably generate labeled emotions. However, most appraisal theories share some basic variables: pleasure (a.k.a. valence, or a person's subjective positive/negative view of what is being appraised) is present or inferable among many theories' appraisal variables [129][130][131]. Also, arousal (intensity of

feeling) is measured as an appraisal variable in several theories or else assumed to be a factor in generating emotional response upon appraisal of a situation relevant to Emo and its goals [132]. Coping potential is another commonly used variable—the ability to deal with a situation. Coping itself may include taking direct action regarding the situation, or cognitive redefinition of one’s beliefs, desires, or intentions; for example, the “sour grapes” approach of reappraising a negative situation as a positive. Lazarus [128] developed the similar concept of primary vs. secondary appraisals: primary appraisal takes in a situation’s significance, and secondary appraisal (coping potential) assesses the ability to deal with the situation. Frijda [133] relates emotions to action tendencies, with emotional cues providing constraints on the next set of decisions or actions made by an agent. For instance, fear may limit action tendencies to avoidant behavior.

The OCC (Ortony, Clore, and Collins) appraisal theory [130] categorizes emotions based on appraisal of pleasure / displeasure and arousal. To more specifically predict emotion generation, OCC theory breaks down appraisal of pleasure / displeasure into three categories based on what is being appraised: desirability (of an event), praiseworthiness (of an action), and like/dislike (of an entity). Also, actions and events may be further differentiated by an attribution variable: for instance, was an action taken by (or did an event affect) oneself, or another? Appraisal across these variables defines different specific emotions; for instance, a positive-pleasure appraisal of an action attributable to oneself might create an emotion of pride in the appraiser, whereas the swimmer’s

appraisal of the shark encounter as above would be as a negative event attributable to the shark, with prospective negative consequences for the swimmer, thus producing fear of the shark.

Challenge and Threat theory [131] uses two main variables: demand and resources. Demand in turn is broken down into three variables: required effort, danger, and uncertainty. The resources variable is similar to the coping potential variable in other theories. According to Challenge and Threat theory, a swimmer’s appraisal of a shark encounter’s danger as “high”, combined with the appraisal of insufficient resources to deal with the shark, would produce a feeling of threat. However, if the swimmer’s evaluation of the situation identified sufficient resources like a nearby beach, the overall assessment would lead to a (positive) feeling of challenge.

Several researchers have devised high-variable-count appraisal theories that map specific configurations and values of appraisal variables to a range of generated emotions and a sequential series of checks. One such map [108] is summarized in Table 7 below.

Table 7. Map of appraisal dimensions to predicted emotion generation

Check	Seq	Joy	Fear	Anger	Sad- ness	Disgust	Shame	Guilt
Expected- ness	1	Open	Low	Open	Open	Open	Open	Open
Unpleasant- ness	1	Low	High	Open	Open	V.High	Open	Open

Goal Hindrance	1	V.Low	High	High	High	Open	Open	Low
External Causation?	2	Open	Ext.	Ext.	Open	Ext.	Int.	Int.
Coping Potential	3	Med.	V.Low	High	Low	Open	Open	Open
Immortality	4	Open	Open	High	Open	Open	Open	V.High
Self-Consistency	4	Open	Open	Low	Open	Open	V.Low	V.Low

For example, fear is the predicted result of an unexpected, unpleasant, goal-hindering, externally caused situation for which the appraiser has a low coping potential (resources): e.g., a shark attacking a swimmer. Immortality (defined loosely as the breaking of societal or personal norms or values) does not come into effect in that example; nor does the swimmer's sense of self-consistency (or integrity). Hence those two variables are generally categorized as "Open" in terms of generating fear.

Beyond divergent sets of variables, some appraisal theories postulate more than one appraisal level or sequence (related to primary and secondary appraisal as per Lazarus). Scherer [108] introduces the idea of "sequential checking" as an order of operations for appraisal. See "Seq." column in Table 9 for Scherer's organization of appraisal variables into a sequence of four ordered checks: 1) a check as to relevance of an event to the appraiser; 2) if relevant, a check for causality and other implications, followed by 3) a coping potential check if necessary and 4) a normative significance check (morality/integrity).

Other theories outline a distinction between non-cognitive and cognitive appraisal. One difference is that cognitive appraisal processes (like inferring the cause of a situation) are generally slower than appraisals based on direct sensory feedback (like physical pain) [10] expresses these different appraisal levels as somatic (primary) and recalled (secondary) “emotion inducers”. Some researchers use this distinction to define certain emotions as secondary— only able to arise following some cognitive processing[40]. For instance, anger at a person most likely stems from cognitively attributing an action or event (previously appraised as unpleasant) to that person, such as a “shark attack” revealed to be another swimmer playing a practical joke.

Part of the difference between theories that postulate two levels of appraisal and theories which only identify one (fast) appraisal level is semantic in nature: the two-level theories include non-evaluative cognitive processing (e.g., inference or recall) as part of “secondary appraisal,” whereas the one-level theories limit appraisal to quick situational evaluation of sensed situations and cognized situations alike [133a]. The one-level theories view the cognitive processing of events as coping, instead of as appraisal [134]. Also, according to Leventhal and Scherer [135], there is not a clear line between calling a given appraisal process cognitive or non-cognitive. The line can be further blurred in that routine appraisals of a particular situation can enable future appraisals (recognition response) of that situation to be quicker and more reflexive than an initial, purely cognitively based appraisal [136]. In keeping with the Associative subsystem /

Deliberative subsystem paradigm, primary appraisal in two-level theories would mainly be handled by Associative subsystem, while cognitive processing (whether described as coping or secondary appraisal) would be handled by Deliberative subsystem.

A.4. Further Effects to be Modeled

Aside from the studies mentioned in previous sections, there is a wealth of theoretical and experimental work, from ancient times onward, on how emotion affects decision-making behavior, which my model does not currently take into account. Some of these (whether or not they are emotional state-related) would be suitable for further exploration in the model once it is augmented as described above.

Several effects of emotion on deliberative cognitive processes fall under the heading of induced biases and heuristics. In one sense, these effects allow deliberative processes to assume continuity or predictability in the environment, given a priming emotional state or stimulus [137]. For instance, a person walking down a street at night might become fearful, and then evaluate ordinarily neutral concepts or percepts as threatening. However, this Affective Priming effect may be incorrect or misleading, as in the case of sad music causing shoppers to feel the need for short-term comfort and thus to buy more readily [138].

The Affective Infusion Model [139] provides a more structured paradigm for several emotional bias effects on decision than my model currently covers. The AIM posits four strategies, from least to most deliberation-intensive, and hence in order of probable use: 1) Direct-access processing: a previously stored, relevant decision is readily recalled. 2) “Motivated” processing: both recall and inference are guided by a strong motivating goal of the judge. 3) “Emotional state-heuristic” processing: the judge’s current emotional state guides decision without much further cognitive processing, and 4) “Substantive” processing: the judge performs integration of information into a complex concept, and emotional state may or may not affect any given recall or inference process used in that integration. The motivated and heuristic strategies align with the theory that at least some affective response precedes [140] and biases deliberative processing.

The Somatic Marker Hypothesis [141] predicts that an emotional heuristic aids the process of decision-making. Specifically, an emotionally charged state of mind is first stored in memory, associated with recollections of the situation that produced the state. Then, the recurrence of the situation (or a similar one) will produce a quick, reflexive response: “reliving” of the state and of the emotion behind it, before slower, more deliberative cognition can occur. Bechara demonstrated the effect (though not necessarily proving the Somatic Marker Hypothesis) using the Iowa Gambling Task experiment [10]. The experiment showed that emotion-linked reasoning causes people to avoid behavior that they have learned carries unnecessary risks. Specifically, subjects chose from two

sets of cards that gave monetary rewards and punishments; set 1 had higher individual rewards but had a long-term negative outcome potential, whereas set 2 would provide a long-term positive reward. The predicted result was that feelings of regret or loss would become associated with continued choosing of cards from set 1, and that set 1 would eventually be deemed less useful than set 2. Participants in the experiment were divided into two groups: a control group with normal brain function, and people with damage to the ventromedial sector of the prefrontal cortex, which enables emotional states to be stored in memory. The result showed that the control group built up a somatic “stress” reaction associated with choosing from set 1, and thus learned to minimize risk by favoring set 2. The group with cortical damage, as predicted, continued to use set 1’s riskier decks more often than not. According to some, the Somatic Marker Hypothesis is not necessarily proven by the Gambling Task results [25], as the same results could be attributable to deliberative processing of the losses accrued during the task.

Table 8. Further emotional effects for future work

Effect of Emotion	Computational Model Trigger	Computational Model Effects Summary
Learning under Cognitive Arousal	Arousal level of emotional state and of node(s) being attended to by deliberative subsystem processes	Scope of working memory increases, learned memories gain emotional state-congruent recall propensity; reverse effect begins to occur if arousal increases beyond a certain threshold
Affective Priming	Emotional state at certain level of arousal, pleasure, and dominance	Nodes attended to by cognitive processes have perceived pleasure, arousal, and dominance ratings closer to current emotional state values

Bias/Heuristic: Judgment Strategy Choice (Affective Infusion and Gambling Task effects)	Choice invoked	In decreasing order of priority: 1) Relevant judgment readily recalled (fast feature-matching); 2) Quick recall and/or inference congruent to the attended nodes' pleasure rating; 3) Emotional state-dependent decision 4) Cognitive judgment using emotional state-congruent recall (and inference)
--	----------------	---

Table 8 summarizes the how certain further emotional effects may be modeled. This will also necessitate more and finer-grained processes than currently exist in my model, among other enhancements (e.g., learning, inference, case-based planning, and episodic memory).

Appendix B: Implementation

B.1. Code specification Including Memory for ExperimentalDomains

<http://ncr.isi.edu/svn/Projects/Nuclear%20Plant%20Experiment>

Contact author for user id and password.

B.2. Sample Output from Nuclear Power Plant Decision-Maker

```
Got action check(coolantTemperature,)
level: 381.6
Result from performing check(coolantTemperature,) is 381
Got action check(waterPressure,)
level: 2.4
Result from performing check(waterPressure,) is 2
Got action checkSystem1(fixWaterPressure,)
EmoCog Loop initialized
Result from performing checkSystem1(fixWaterPressure,) is
```

nodeList(emergencyBypassPump, 0.6000000238418579,
nodeList(doNothing,
0.6000000238418579, nodeList(emergencyPump, 0.6000000238418579,
nodeList(emergencySealantSpray, 0.6000000238418579,
nodeList(pipeRupture, 0.6000000238418579, end))))))
asserting system1Fact(emergencyBypassPump,0.6000000238418579)
asserting system1Fact(doNothing,0.6000000238418579)
asserting system1Fact(emergencyPump,0.6000000238418579)
asserting system1Fact(emergencySealantSpray,0.6000000238418579)
asserting system1Fact(pipeRupture,0.6000000238418579)
comparing [emergencySealantSpray] and [emergencyPump] in model
correct
Project emergencySealantSpray on [] to
[[1.0,[emergencySealantSpray]]]Trigger correct on
[emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
created in simulation
Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
worlds created in simulation
Utility of [emergencySealantSpray] is 5
Utility of [emergencyPump,bad] is -10
Prefer [[1.0,[emergencySealantSpray]]] over
[[1.0,[emergencyPump,bad]]] since 5.0 > -10.0
comparing [emergencySealantSpray] and [emergencyPump] in model
correct
Project emergencySealantSpray on [] to
[[1.0,[emergencySealantSpray]]]Trigger correct on
[emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
created in simulation
Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
worlds created in simulation
Utility of [emergencySealantSpray] is 5
Utility of [emergencyPump,bad] is -10
Prefer [[1.0,[emergencySealantSpray]]] over
[[1.0,[emergencyPump,bad]]] since 5.0 > -10.0
Got action set(emergencyBypassPump,on,)
Running stub code to set a value in the simulator
[DatabaseLogger.logEvent] logging, type =

REACTOR_OPEN_BYPASS_VALVE_SUCCESS desc = opened bypass valve
 successfully
 Result from performing set(emergencyBypassPump,on,) is 1
 Got action checkSystem1(fixWaterPressure,)
 EmoCog Loop initialized
 Result from performing checkSystem1(fixWaterPressure,) is
 nodeList(emergencyBypassPump, 0.6000000238418579,
 nodeList(doNothing,
 0.6000000238418579, nodeList(emergencyPump, 0.6000000238418579,
 nodeList(emergencySealantSpray, 0.6000000238418579,
 nodeList(pipeRupture, 0.6000000238418579, end))))))
 asserting system1Fact(emergencyBypassPump,0.6000000238418579)
 asserting system1Fact(doNothing,0.6000000238418579)
 asserting system1Fact(emergencyPump,0.6000000238418579)
 asserting system1Fact(emergencySealantSpray,0.6000000238418579)
 asserting system1Fact(pipeRupture,0.6000000238418579)
 comparing [emergencySealantSpray] and [emergencyPump] in model
 correct
 Project emergencySealantSpray on [] to
 [[1.0,[emergencySealantSpray]]]Trigger correct on
 [emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
 created in simulation
 Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
 correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
 worlds created in simulation
 Utility of [emergencySealantSpray] is 5
 Utility of [emergencyPump,bad] is -10
 Prefer [[1.0,[emergencySealantSpray]]] over
 [[1.0,[emergencyPump,bad]]] since 5.0 > -10.0
 comparing [emergencySealantSpray] and [emergencyPump] in model
 correct
 Project emergencySealantSpray on [] to
 [[1.0,[emergencySealantSpray]]]Trigger correct on
 [emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
 created in simulation
 Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
 correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
 worlds created in simulation

Utility of [emergencySealantSpray] is 5
Utility of [emergencyPump,bad] is -10
Prefer [[1.0,[emergencySealantSpray]]] over
[[1.0,[emergencyPump,bad]]] since $5.0 > -10.0$
comparing [emergencySealantSpray] and [emergencyPump] in model
correct
Project emergencySealantSpray on [] to
[[1.0,[emergencySealantSpray]]]Trigger correct on
[emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
created in simulation
Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
worlds created in simulation
Utility of [emergencySealantSpray] is 5
Utility of [emergencyPump,bad] is -10
Prefer [[1.0,[emergencySealantSpray]]] over
[[1.0,[emergencyPump,bad]]] since $5.0 > -10.0$
comparing [emergencySealantSpray] and [emergencyPump] in model
correct
Project emergencySealantSpray on [] to
[[1.0,[emergencySealantSpray]]]Trigger correct on
[emergencySealantSpray] to [[1.0,[emergencySealantSpray]]]0 new worlds
created in simulation
Project emergencyPump on [] to [[1.0,[emergencyPump,bad]]]Trigger
correct on [emergencyPump,bad] to [[1.0,[emergencyPump,bad]]]0 new
worlds created in simulation
Utility of [emergencySealantSpray] is 5
Utility of [emergencyPump,bad] is -10
Prefer [[1.0,[emergencySealantSpray]]] over
[[1.0,[emergencyPump,bad]]] since $5.0 > -10.0$
Got action set(emergencySealantSpray,on,)
Running stub code to set a value in the simulator
[DatabaseLogger.logEvent] logging, type =
REACTOR_START_EMERGENCY_SEALANT_SPRAY_SUCCESS desc
= started emergency
sealant spray successfully
[DatabaseLogger.logEvent] logging, type =
REACTOR_PIPING_RUPTURE_DISABLED desc = piping rupture
disabled

[DatabaseLogger.logEvent] logging, type =
SYSTEM_RESET_WATER_PRESSURE_MID desc = reset water
pressure to mid
[DatabaseLogger.logEvent] logging, type =
SYSTEM_END_CONDITION_PIPE_FIXED desc = end condition: pipe
fixed
Result from performing set(emergencySealantSpray,on,) is 1

References

- [1] Madni, A.M., and Madni, C.C. “Intelligent Agents as Synthetic Role Players in Scenario-Based Training,” *Journal of Integrated Design and Process Science*, Vol. 12, No. 1, March 2008, pp. 41-56.
- [2] Madni, A.M. “The Role of Human Factors in Expert Systems Design and Acceptance,” *Human Factors Journal*, August, 1988, 30(4), 395-414.
- [3] Madni, A.M., Moini, A., and Madni, C. “Cross-Cultural Decision Making Training Using Behavioral Game Theoretic Framework” in *Advances in Cross-Cultural Decision Making*, Schmorrow, D. and Nichelson, D. (Eds.), CRC Press, 2010.
- [4] Madni, A.M., Moini, A., and Madni, C.C. “GameSim™: Game-based Simulation for Human, Social, and Cultural Training,” *Intelligent Systems Technology, Inc. Phase II Final Report, ISTI-FR-619-10/11, Contract # N00014-10-C-0426*, October 3, 2011.
- [5] Madni, A.M. and Madni, C.C. “MedGames™: Micro Games for Proactive Preventive Medicine,” *Intelligent Systems Technology, Inc. Phase I Final Report, ISTI-FR-608-8/09, Contract # FA8650-09-M-6933*, August 6, 2009.
- [6] A. Bechara, A. R. Damasio, H. Damasio, and S. W. Anderson, “Insensitivity to future consequences following damage to human prefrontal cortex,” *Cognition*, vol. 50, no. 1–3, pp. 7–15, 1994.
- [7] P. Slovic, M. Finucane, E. Peters, and D. G. Macgregor, “The affect heuristic,” in *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, pp. 397–420, 2007.
- [8] J. Gratch, “Why you should buy an emotional planner,” in *Proceedings of the Agents’99 Workshop on Emotion-based Agent Architectures*, pp. 99–107, 1996.
- [9] J. D. Yerkes and R. M. Dodson, “The relation of strength of stimulus to rapidity of habit-formation,” *Journal of Comparative Neurology and Psychology*, vol. 18, pp. 459–482, 1908.
- [10] A. Bechara, H. Damasio, and A. Damasio, “Emotion, decision making and the orbitofrontal cortex,” *Cerebral Cortex*, vol. 10, no. 3, pp. 295–307, 2000.

- [11] J. M. Harlow and E. Miller, "Recovery from the passage of an iron bar through the head," *History of Psychiatry*, vol. 4, no. 14, pp. 271–273, Jan. 1993.
- [12] P. Langley, J. E. Laird, and S. Rogers, "Cognitive architectures: Research issues and challenges," *Cognitive Systems Research*, vol. 10, no. 2, pp. 141–160, 2009.
- [13] P. J. Gmytrasiewicz and C. L. Lisetti, *Emotions and personality in agent design and modeling*, pp. 81–95, 2002.
- [14] B. G. Silverman, G. Bharathy, M. Johns, R. Eidelson, and T. Smith, "Socio-Cultural Games for Training and Analysis," *IEEE Transactions on Systems Man Cybernetics Part A (2006)*, vol. 37, no. 6, pp. 1113–1130, 2006.
- [15] M. D. Dickey, "Engaging By Design : How Engagement Strategies in Popular Computer and Video Games Can Inform Instructional Design," *Educational Technology, Research and Development*, vol. 53, no. 2, pp. 67–83, 2005.
- [16] M. Mateas, "Expressive AI : Games and Artificial Intelligence," in *Level Up: Digital Games Research Conference 1*, 2003.
- [17] R. L. Boring, D. I. Gertman, and K. Le Blanc, "HUMAN RELIABILITY ANALYSIS FOR COMPUTERIZED PROCEDURES," *Human Factors*, pp. 1720–1724, 2011.
- [18] D. Kahneman and S. Frederick, *Representativeness revisited: Attribute Substitution in intuitive Judgement*. New York, New York, USA: Princeton University Press, 2001.
- [19] M. L. Finucane, A. Alhakami, P. Slovic, and S. M. Johnson, "The affect heuristic in judgments of risks and benefits," *Journal of Behavioral Decision Making*, vol. 13, no. 1, pp. 1–17, Jan. 2000.
- [20] A. Hopkins, "Was Three Mile Island a 'Normal Accident'?", *Journal of Contingencies and Crisis Management*, vol. 9, no. 2, pp. 65–72, 2001.
- [21] A. Sloman, "Varieties of affect and the CogAff architecture schema," in *Proceedings of the AISB'01 Symposium on Emotion, Cognition, and Affective Computing*, pp. 39–48, 2001.
- [22] W. S. Reilly and J. Bates, "Building emotional agents," Pittsburgh, PA, 1992.
- [23] A. Staller and P. Petta, "Towards a Tractable Appraisal-Based Architecture for Situated Cognizers," *Appraisal*, pp. 1–19, 1998.
- [24] R. E. Cochran, F. J. Lee, and E. Chown, "Modeling emotion: Arousal's impact on memory," in *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pp. 1133–1138, 2006.
- [25] D. Fum and A. Stocco, "Memory, emotion, and rationality: An ACT-R interpretation for gambling task results," in *Proceedings of the Sixth International Conference on Cognitive Modeling*, pp. 106–111, 2004.
- [26] R. Belavkin, "The role of emotion in problem solving," in *Proceedings of the AISB'01 Symposium on Emotion, Cognition and Affective Computing*, pp. 49–57, 2001.

- [27] D. Watson and A. Tellegen, "Toward a consensual structure of emotional state," *Psychological bulletin*, vol. 98, no. 2, pp. 219–35, Sep. 1985.
- [28] A. Newell, *Unified theories of cognition*. Cambridge, MA: Harvard University Press, 1990.
- [29] J. R. Anderson and C. Lebiere, "The Newell Test for a theory of cognition.," *The Behavioral and Brain Sciences*, vol. 26, no. 5, pp. 587–601; discussion 601–48, Oct. 2003.
- [30] A. Sloman, "Damasio, Descartes, alarms and meta-management," *Systems, Man, and Cybernetics, 1998. 1998 IEEE*, 1998.
- [31] Madni, A.M., Moini, A., and Madni, C.C. "Cognitecture™: Cognition-Based Interactive Architecture for Next Generation Intelligent Agents," Intelligent Systems Technology, Inc. Phase II Final Technical Report, ISTI-FTR-567-8/06, Contract #W31P4Q-04-C-R222, August 10, 2006.
- [32] S. A. Sloman, "The empirical case for two systems of reasoning," *Psychological Bulletin*, vol. 119, no. 1, pp. 3–22, 1996.
- [33] P. Carruthers, "An Architecture for Dual Reasoning," *System*, pp. 1–28, 2004.
- [34] A. Sloman, R. Chrisley, and M. Scheutz, "The Architectural Basis of Affective States and Processes," in *Who Needs Emotions? The Brain Meets The Robot*, J.-M. Fellous and M. a. Arbib, Eds. 2005, pp. 203–244.
- [35] A. Ortony, D. Norman, and W. Revelle, "Affect and Proto-Affect in Effective Functioning," in *Who Needs Emotions? The Brain Meets The Robot*, ed. J.-M. Fellous and M. A. Arbib, 2005, pp. 173–202.
- [36] G. H. Bower, "Emotional state and memory," *American Psychologist*, vol. 36, no. 2, pp. 129–148, 1981.
- [37] S. C. Marsella and J. Gratch, "EMA: A process model of appraisal dynamics," *Cognitive Systems Research*, vol. 10, no. 1, pp. 70–90, Mar. 2009.
- [38] J. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.
- [39] A. Mehrabian and J. A. Russell, *An approach to environmental psychology*. Cambridge, MA: MIT Press, 1974.
- [40] C. Becker-Asano and I. Wachsmuth, "Affective computing with primary and secondary emotions in a virtual human," *Autonomous Agents and Multi-Agent Systems*, vol. 20, no. 1, pp. 32–49, May 2009.
- [41] H. Simon, "Motivational and emotional controls of cognition," *Psychological Review*, vol. 74, no. 1, pp. 29–39, 1967.
- [42] A. Ohman, A. Flykt, and F. Esteves, "Emotion drives attention: Detecting the snake in the grass," *Journal of Experimental Psychology: General*, vol. 130, no. 3, pp. 466–478, 2001.
- [43] J. Easterbrook, "The Effect of Emotion on Cue Utilization and the Organization of Behavior," *Psychological Review*, vol. 66, no. 3, pp. 183–201, 1959.

- [44] F. E. Ritter, A. Reifers, L. Klein, and M. Schoelles, "Lessons from Defining Theories of Stress," in *Integrated Models of Cognitive Systems*, W. Gray, Ed. New York, New York, USA: Oxford University Press, 2006.
- [45] K. Christoff, A. Gordon, and R. Smith, "The role of spontaneous thought in human cognition," in *Neuroscience of Decision Making*, O. Vartanian and D. R. Mandel, Eds. Psychology Press, 2007.
- [46] L. H. Phillips, L. Smith, and K. J. Gilhooly, "The effects of adult aging and induced positive and negative emotional state on planning.," *Emotion*, vol. 2, no. 3, pp. 263–272, 2002.
- [47] J. Gratch and S. Marsella, "Evaluating a computational model of emotion," *Autonomous Agents and Multi-Agent Systems*, vol. 11, no. 1, pp. 23–43, Jun. 2005.
- [48] C. J. Armitage, M. Conner, and P. Norman, "Differential effects of emotional state on information processing : evidence from the theories of reasoned action and planned behaviour," *European Journal of Social Psychology*, vol. 29, no. March 1998, pp. 419–433, 1999.
- [49] J. D. Mayer, "How Emotional state Influences Cognition," in *Advances in Cognitive Science*, N. E. Sharkey, Ed. Chichester, West Sussex: Ellis Horwood Limited, 1986, pp. 290–314.
- [50] J. Mayer and Y. Gaschke, "Emotional state-congruent judgment is a general effect," *Journal of Personality and Social Psychology*, vol. 63, no. 1, pp. 119–132, 1992.
- [51] G. H. Bower, "Affect and cognition," *Philosophical Transactions of the Royal Society of London*, vol. B, no. 302, pp. 387–402, 1983.
- [52] G. H. Bower, "How might emotions affect learning," in *The handbook of emotion and memory: Research and theory*, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, pp. 3–31.
- [53] K. Spies, F. Hesse, and C. Hummitzsch., "Emotional state and capacity in Baddeley's model of human memory," *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*, 1996.
- [54] F. W. Hesse and K. Spies, "Effects of negative emotional state on performance : reduced capacity or changed processing strategy ?," *European Journal of Social Psychology*, vol. 26, no. February 1995, pp. 163–168, 1996.
- [55] M. Oaksford, F. Morris, B. Grainger, and J. M. G. Williams, "Emotional state, reasoning, and central executive processes.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 22, no. 2, pp. 476–492, 1996.
- [56] K. Gasper, "When necessity is the mother of invention: Emotional state and problem solving," *Journal of Experimental Social Psychology*, vol. 39, no. 3, pp. 248–262, May 2003.
- [57] L. H. Phillips, R. Bull, E. Adams, and L. Fraser, "Positive emotional state and executive function: Evidence from Stroop and fluency tasks.," *Emotion*, vol. 2, no. 1, pp. 12–22, 2002.

- [58] G. Dreisbach and T. Goschke, "How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility.," *Journal of experimental psychology. Learning, memory, and cognition*, vol. 30, no. 2, pp. 343–53, Mar. 2004.
- [59] G. Dreisbach, "How positive affect modulates cognitive control: the costs and benefits of reduced maintenance capability.," *Brain and cognition*, vol. 60, no. 1, pp. 11–19, Feb. 2006.
- [60] R. L. C. Mitchell and L. H. Phillips, "The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative emotional state on executive functions.," *Neuropsychologia*, vol. 45, no. 4, pp. 617–29, Mar. 2007.
- [61] F. G. Ashby, a M. Isen, and a U. Turken, "A neuropsychological theory of positive affect and its influence on cognition.," *Psychological review*, vol. 106, no. 3, pp. 529–50, Jul. 1999.
- [62] F. G. Ashby, V. Valentin, and U. Turken, "The effects of positive affect and arousal on working memory and executive attention Neurobiology and computational models," pp. 245–287, 2002.
- [63] J. Gratch, S. Marsella, and P. Petta, "Modeling the cognitive antecedents and consequences of emotion," *Cognitive Systems Research*, vol. 10, no. 1, pp. 1–5, Mar. 2009.
- [64] C. Breazeal, "Toward sociable robots," *Robots and Autonomous Systems*, vol. 42, no. 3, pp. 167–175, 2003.
- [65] E. Bevacqua, M. Mancini, and C. Pelachaud, "Speaking with emotions," in *Proceedings of the AISB'04 Symposium on Motion, Emotion and Cognition*, 2004, pp. 197–214.
- [66] D. Moffat, N. H. Frijda, and R. H. Phaf, "Analysis of a model of emotions," in *Prospects for artificial intelligence*, A. Sloman, D. Hogg, G. Humphreys, and A. Ramsay, Eds. Amsterdam: IOS Press, 1993, pp. 219–228.
- [67] S. Rank, "Behaviour coordination for models of affective behaviour," Vienna University of Technology, Vienna, 2009.
- [68] H.-I. Ahn, "Modeling and analysis of affective influences on human experience, prediction, decision making, and behavior," Massachusetts Institute of Technology, Cambridge, MA, 2010.
- [69] S. Gadanho and L. Custodio, "Asynchronous learning by emotions and cognition," in *Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior*, 2003, pp. 224–225.
- [70] J. Gratch, "Émile: Marshalling passions in training and education," in *Proceedings of the Fourth International Conference on Autonomous Agents*, 2000, no. June, pp. 325–332.
- [71] P. Singh, "EM-ONE : An architecture for reflective commonsense thinking," Massachusetts Institute of Technology, Cambridge, MA, 2005.

- [72] J. Dias, S. Mascarenhas, and A. Paiva, "FAtiMA Modular: Towards an agent architecture with a generic appraisal framework," in *Proceedings of the International Workshop on Standards for Emotion Modeling*, 2011.
- [73] M. S. El-Nasr, J. Yen, and T. R. Ioerger, "FLAME - Fuzzy logic adaptive model of emotions," *Autonomous Agents and Multi-agent Systems*, vol. 3, pp. 219–257, 2000.
- [74] P. J. Gmytrasiewicz and C. L. Lisetti, "Using decision theory to formalize emotions in multi-agent systems," *Proceedings Fourth International Conference on MultiAgent Systems*, pp. 391–392, 2000.
- [75] M. Malfaz and M. A. Salichs, "Using emotions for behaviour-selection learning," in *Proceedings of the 2006 conference on ECAI 2006: 17th European Conference on Artificial Intelligence*, 2006, pp. 697–698.
- [76] E. Hudlicka, "Reasons for emotions: Modeling emotions in integrated cognitive systems," in *Integrated models of cognitive systems*, W. Gray, Ed. New York: Oxford University Press, 2007, pp. 1–37.
- [77] J.-J. C. Meyer, "Reasoning about emotional agents," *International Journal of Intelligent Systems*, vol. 21, no. 6, pp. 601–619, Jun. 2006.
- [78] R. Marinier, J. Laird, and R. Lewis, "A computational unification of cognitive behavior and emotion," *Cognitive Systems Research*, vol. 10, no. 1, pp. 48–69, 2009.
- [79] N. H. Frijda and J. Swagerman, "Can Computers Feel ? Theory and Design of an Emotional System," *Cognition*, no. 768400892, 1987.
- [79a] A. Dijksterhuis and T. Meurs, "Where creativity resides: the generative power of unconscious thought.," *Consciousness and cognition*, vol. 15, no. 1, pp. 135–46, Mar. 2006.
- [80] <http://www.masagroup.net/products/masa-life/>
- [81] <http://www.xaitment.com/>
- [82] A. Rizzo, J. Difede, B. Rothbaum, J. Daughtry, G. Reger. "Virtual Reality as a Tool for Delivering PTSD Exposure Therapy," in *Post-Traumatic Stress Disorder: Future Directions in Prevention, Diagnosis, and Treatment*, Springer, 2013
- [83] M. Eladhari. "The Pataphysic Institute," in *Proceedings of AIIDE*, 2010.
- [84] J. Panksepp, "The neuro-evolutionary cusp between emotions and cognitions: Implications for understanding consciousness and the emergence of a unified mind science," *Evolution and Cognition*, vol. 7, no. 2, pp. 141–163, 2001.
- [85] M. Quillian, "Semantic Memory," in *Semantic Information Processing*, M. Minsky, Ed. MIT Press, 1968, pp. 227–270.
- [86] A. Baddeley, "The episodic buffer: a new component of working memory?," *Trends in cognitive sciences*, vol. 4, no. 11, pp. 417–423, Dec. 2000.
- [87] J. R. Anderson, "A spreading activation theory of memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 22, no. 3, pp. 261–295, Jun. 1983.

- [87a] D. Watson and L. A. Clark, "Measurement and Mismeasurement of Emotional state : Recurrent and Emergent issues," no. August 2012, pp. 37–41, 2010.
- [88] D. K. Lee, L. Itti, C. Koch, and J. Braun, "Attention activates winner-take-all competition among visual filters.," *Nature neuroscience*, vol. 2, no. 4, pp. 375–81, Apr. 1999.
- [88a] E. Gray and D. Watson, "Emotion, Emotional state, and Temperament," in *Emotions at Work: Theory, Research and Applications for Management*, 2003, pp. 21–43.
- [89] G. Miller, "The Magic Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *The Psychological Review*, vol. 63, pp. 81–97, 1957.
- [90] B. Hayes-Roth and F. H.-R. Perrault, "A Cognitive Model of Planning," *Cognitive Science*, vol. 3, no. 4, pp. 275–310, Oct. 1979.
- [91] M. Minsky, *The Society of Mind*. Simon & Schuster, 1988, p. 339.
- [92] M. N. Huhns and L. M. Stephens, "Multiagent Systems and Societies of Agents," in *Multiagent systems*, 1999.
- [93] H. Einhorn and R. M. Hogarth, "Behavioral Decision Theory Processes of Judgment and Choice," vol. 19, no. 1, pp. 1–31, 1981.
- [94] T. G. Gill, "A Psychologically Plausible Goal-Based Utility Function," *Informing Science: the International Journal of an Emerging Transdiscipline*, vol. 11, 2008.
- [95] M. Minsky, "K-Lines : A Theory of Memory," *Cognitive Science*, vol. 4, pp. 117–133, 1980.
- [96] S. S. Choi, J. K. Park, J. H. Hong, H. G. Kim, S. H. Chang, and K. S. Kang, "Development Strategies of an Intelligent Human-Machine Interface for Next Generation Nuclear Power Plants," *IEEE Transactions on Nuclear Science*, vol. 43, no. 3, 1996.
- [97] IAEA, "Boiling Water Reactor Simulator with Active Safety Systems," *Atomic Energy*, October, 2009.
- [98] J. McCurry, [No Title]*Guardian*, 2012. [Online]. Available: <http://www.guardian.co.uk/environment/2012/jul/05/fukushima-meltdown-manmade-disaster>.
- [99] E. Hollnagel, A. Edland, and O. Svenson, "A Cognitive Task Analysis of the SGTR Scenario," Joint Report of Halden Reactor Project and University of Stockholm Dept. of Psychology, Stockholm, Sweden, April 1996.
- [100] M. Spraragen, P. M. Landwehr, B. Ranganathan, M. Zyda, K. M. Carley, Y.-H. Chang, and R. Maheswaran, "Cosmopolis: a massively multiplayer online game for social and behavioral research," *Journal of Artificial Societies and Social Simulation*, 2013.
- [101] K. Stanovich and R. West, "Individual differences in reasoning: Implications for the rationality debate?" *Behavioral and brain sciences*, 23(5), 645-665, 2000.

- [102] J. Cacioppo, B. Marshall-Goodell, L. Tassinari and R. Petty. “Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high.” *Journal of Experimental Social Psychology*, 28(3), 207-233, 1992.
- [103] J. Hammond, “CHEF: A Model of Case-based Planning,” in *AAAI-86 Proceedings.*, 1986, pp. 267–271.
- [104] G. Dreisbach, “Personal Correspondence.” 2013.
- [105] B. E. Kaufman, “Emotional arousal as a source of bounded rationality,” *Journal of Economic Behavior & Organization*, vol. 38, pp. 135–144, 1999.
- [106] L. J. Kleinsmith and S. Kaplan, “Paired-associate learning as a function of arousal and interpolated interval.,” *Journal of experimental psychology*, vol. 65, no. 2, pp. 190–3, Mar. 1963.
- [107] A. Mehrabian, “Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament,” *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.
- [108] K. R. Scherer, “Appraisal considered as a process of multi-level sequential checking,” in *Appraisal processes in emotion: Theory, Methods, Research*, K. R. Scherer, A. Schorr, and T. Johnstone, Eds. New York: Oxford University Press, 2001, pp. 92–120.
- [109] G. Ward, “Planning and the Executive Control of Thought and Action,” in *The Cognitive Psychology of Planning*, G. Ward and R. Morris, Eds. New York, New York, USA: Psychology Press, 2005, pp. 89–110.
- [110] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, New Jersey: Prentice Hall, 2009.
- [111] A. Newell and H. Simon, *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall, 1972.
- [112] A. Newell and H. Simon, “Elements of a theory of problem solving,” *Psychological Review*, vol. 65, pp. 151–166, 1958.
- [113] H. Simon, “Information Processing Theory of human problem solving,” in *Handbook of Learning & Cognitive Processes*, 1978, pp. 271–295.
- [114] K. J. Gilhooly, “Working Memory and Planning,” in *The Cognitive Psychology of Planning*, G. Ward and R. Morris, Eds. New York, New York, USA: Psychology Press, 2005, pp. 71–88.
- [115] G. Miller, A. Galanter, and K. Pribram, *Plans and the Structure of Behavior*. New York, New York, USA: Holt, Rinehart, & Winston, 1960.
- [116] M. Donald, “Human Cognitive Evolution : What We Were, What We Are Becoming,” *Social Research*, vol. 60, no. 1, pp. 143–170, 1993.
- [117] P. Burgess, J. Simons, L. Coates, and S. Channon, “The Search for Specific Planning Processes,” in *The Cognitive Psychology of Planning*, G. Ward and R. Morris, Eds. New York, New York, USA: Psychology Press, 2005, pp. 199–227.

- [118] J. Grafman, L. Spector, and M. Ratterman, "Planning and the Brain," in *The Cognitive Psychology of Planning*, R. Morris and G. Ward, Eds. New York, New York, USA: Psychology Press, 2005, pp. 181–198.
- [119] J. P. Das, B. Kar, and R. Parrila, *Cognitive Planning*. London: Sage Publications, 1996.
- [120] S. Davies, "Planning and Problem Solving in Well-Defined Domains," in *The Cognitive Psychology of Planning*, G. Ward and R. Morris, Eds. New York, New York, USA: Psychology Press, 2005, pp. 35–51.
- [121] J. Laird, A. Newell, and P. S. Rosenbloom, "SOAR: An architecture for general intelligence," *Artificial Intelligence*, vol. 33, no. 1, pp. 1–64, Sep. 1987.
- [122] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, Oct. 2004.
- [123] R. Sun, "The CLARION cognitive architecture: Extending cognitive modeling to social simulation," *Cognition and multi-agent interaction*, 2006.
- [124] R. Jensen and M. Veloso, "Interleaving Deliberative and Reactive Planning in Dynamic Multi-Agent Domains," in *Proceedings of the AAI Fall Symposium on on Integrated Planning for Autonomous Agent Architectures*. AAAI Press, October 1998.
- [125] J. E. Laird, *The Soar Cognitive Architecture*. Cambridge, MA: The MIT Press, 2012, p. 390.
- [126] B. Hayes-Roth, "A blackboard architecture for control," *Artificial Intelligence*, vol. 26, no. 3, pp. 251–321, Jul. 1985.
- [127] M. Arnold, *Emotion and personality*. New York: Columbia University Press, 1960.
- [128] R. S. Lazarus, *Psychological stress and the coping process*. New York: McGraw-Hill, 1966, p. 466.
- [129] K. R. Scherer, "Appraisal theory," in *Handbook of Cognition and Emotion*, T. Dalglish and M. Power, Eds. Chichester, UK: John Wiley & Sons Ltd, 1999, pp. 637–663.
- [130] A. Ortony, G. L. Clore, and A. Collins, *The cognitive structure of emotions*. Cambridge, MA: Cambridge University Press, 1988, p. 224.
- [131] J. Blascovich and W. B. Mendes, "Challenge and threat appraisals: The role of affective cues," in *Feeling and Thinking: The Role of Affect in Social Cognition*, J. Forgas, Ed. Cambridge University Press, 2000, pp. 59–82.
- [132] S. C. Marsella, J. Gratch, and P. Petta, "Computational models of emotion," in *A Blueprint for Affective Computing: A Sourcebook and Manual*, K. R. Scherer, T. Banziger, and E. Roesch, Eds. New York: Oxford University Press, 2010, pp. 21–45.
- [133] N. H. Frijda, "Emotion, cognitive structure, and action tendency," *Cognition & Emotion*, vol. 1, no. 2, pp. 115–143, Apr. 1987.

- [133a] A. Ryals and A. Cleary, “The Recognition without Cued Recall Phenomenon: Support for a Feature-Matching Theory over a Partial Recollection Account,” in *Journal of Memory and Language*, vol. 66, no. 4, pp. 747-762, 2012.
- [134] J. Gratch and S. Marsella, “A domain-independent framework for modeling emotion,” *Cognitive Systems Research*, vol. 5, no. 4, pp. 269–306, Dec. 2004.
- [135] S. Leventhal and K. R. Scherer, “The relationship of emotion to cognition: A functional approach to a semantic controversy,” *Cognition and Emotion*, vol. 1, no. 1, pp. 3–28, 1987.
- [136] D. Kahneman, *Thinking, Fast and Slow*. Farrar, Straus and Giroux, 2011.
- [137] C. MacLeod, a Mathews, and P. Tata, “Attentional bias in emotional disorders.,” *Journal of abnormal psychology*, vol. 95, no. 1, pp. 15–20, Feb. 1986.
- [138] J. S. Lerner, D. A. Small, and G. Loewenstein, “Heart strings and purse strings: Carryover effects of emotions on economic decisions,” *Psychological Science*, vol. 15, no. 5, pp. 337–341, May 2004.
- [139] J. P. Forgas, “Emotional state and judgment: the affect infusion model (AIM).,” *Psychological bulletin*, vol. 117, no. 1, pp. 39–66, Jan. 1995.
- [140] R. B. Zajonc, “Feeling and thinking: Preferences need no inferences.,” *American Psychologist*, vol. 35, no. 2, pp. 151–175, 1980.
- [141] A. Damasio, *Descartes’ error: Emotion, reason, and the human brain*. New York: Putnam, 1994, p. 336.