A Light-weight Emotion Model for Non-Player Characters in a Video Game

by

Yathirajan Brammadesam Manavalan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

University of Alberta

Abstract

Displaying believable emotional reactions in virtual characters is required in applications ranging from virtual-reality trainers to video games. Manual scripting is the most frequently used method and enables an arbitrarily high fidelity of the emotions displayed. However, scripting is labor intense and thus greatly reduces the scope of emotions displayed and emotionally affected behavior in virtual characters. As a result, only a few virtual characters can display believable emotions and only in pre-scripted encounters. In this thesis we implement and evaluate a light-weight algorithm for procedurally controlling both emotionally affected behavior and emotional appearance of a virtual character. The algorithm is based on two psychological models of emotions: conservation of resources and appraisal. The former component controls emotionally affected behavior of a virtual character whereas the latter generates explicit numeric descriptors of the character's emotions which can be used to drive the character's appearance. We implement the algorithm in a simple testbed and compare it to two baseline approaches via a user study. Human participants judged the emotions displayed by the algorithm to be more believable than those of the baselines.

Preface

The user study which is a part of this thesis, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "ACORE: Measuring Emotional Believability of virtual characters", No. Pro00051822, 29th January, 2015.

A version of Chapter 4 has been published as Yathirajan Brammadesam Manavalan and Vadim Bulitko, "Appraisal of emotions from resources", In *Proceedings of The Seventh International Conference on Interactive Digital Storytelling*, pages 224-227, Springer, 2014 and has contributions from Dr. Vadim Bulitko with concept formation and manuscript composition.

A version of Chapter 4 and Chapter 5 has been accepted as Yathirajan Manavalan, Vadim Bulitko, and Marcia Spetch, "A lightweight algorithm for procedural generation of emotionally affected behavior and appearance", In *Proceedings of the Eleventh Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, The AAAI Press, 2015.. Dr. Vadim Bulitko and Dr. Marcia Spetch provided me with guidance on running the user study and helped with the manuscript composition.

To my family and my closest friends.
I could not have done this without you.

"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not 'Eureka!' ('I found it!') but 'That's funny...' "

– Isaac Asimov

Acknowledgements

Although I organized and led the work on which this thesis is based, my efforts were not made alone. My supervisor, Vadim Bulitko, was always a constant source of support and guidance. Marcia Spetch provided insight into the methods for conducting the empirical evaluation and also provided me with the resources to do so. Mathew Fritze worked with me in the summer to setup the software which was used to conduct the evaluation. Pankaj, Nitya, Rohit, Karan and the actors from Die Nasty: The Live Improvised Soap Opera, volunteered their time and agreed to have their images used for this research. The IRCL group, Sergio, Thorey, David and Alejandro, have provided me with useful comments and insightful feedback. Alexandria Czeto worked with me during my second summer to verify further applications of the work presented here. This research was supported by the National Sciences and Engineering Research Council of Canada (NSERC).

I would also like to thank the people who were not directly involved in the thesis but I could not have done this work without them. My family, for their unending support and letting me choose my own path. Nishka, for her company, keeping me sane, the editing jobs, the lunches, coffees and the many laughs. Lokesh and Janani, for bringing me to Edmonton, making sure I never felt away from home and driving me around. Anu, for patiently listening to all my complaints and the delicious cakes. Pankaj and Nitya, for reminding me to not take myself too seriously and feeding me. Meera, for helping me discover Edmonton and being my model. Rohit and Abdul, for putting up with me as a roommate. Priya, for being there for me every time. Talat, for being my department senior. Sushant, for always making me feel welcome. Faiza and Adi, for helping me run INDSA and all the fun times. Karan, for putting up with me in my office and the cooking tips. Sergio, for all the geek trivia and teaching me on how to be gentle and mean at the same time. All these people have made my experience at the University of Alberta enjoyable and they are deserving of my many thanks.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Emotionally affected behavior of virtual characters in video games and training environments can significantly impact the user's experience. In movies and books characters are often remembered by their emotional reactions. We expect James Bond to be confident and suave, Master Yoda to be mysterious and wise and Sherlock Holmes to be calm and rational. Similarly, the witty and narcissistic antagonist GLaDOS (Figure 1.1) from the *Portal* video game series (Valve, 2007) stays with the player long after the game is over. The emotional reveal by Pagan Min at the climax of *Far Cry 4* (Ubisoft, 2014) brings a resolution to the player's struggle during the forty-some hour campaign. The cool and collected acknowledgment "Moving to position" of the non-player special-force companions in *Tom Clancy's Ghost Recon 2* (Ubisoft, 2004) becomes the strained "Move accomplished!" as the game's combat heats up. And how can one forget the gleeful "Yest, Kapitan!" of a battle station officer acknowledging a torpedo launch aboard a Russian attack submarine in *Sub Command* (Strategy First, 2001)?

In the early nineties, the US Navy IDCTT trainer used live acting recordings to create the emotionally charged atmosphere of damage control aboard a ship under attack (Johnson, 1994). Other simulators designed for military training also started researching into the influence of emotions on behavior (Hudlicka and Billingsley, 1999) and a survey conducted by the National Research Council concluded that a lack of "behavioral moderators" such as emotions could limit the capabilities of training simulators (Pew and Mavor, 1998). This survey is attributed in part to the creation of the Institute of Creative Technologies (Gratch, 2000). In the mid two thousands, an Iraqi police captain Farid, a non-player character in a virtual

Figure 1.1: Glados (Valve, 2007).

reality trainer by the Institute for Creative Technologies, would be offended by the player's inappropriate chit-chat (Solomon et al., 2008). These undertakings by various researchers demonstrate the growing need for a believable emotion model.

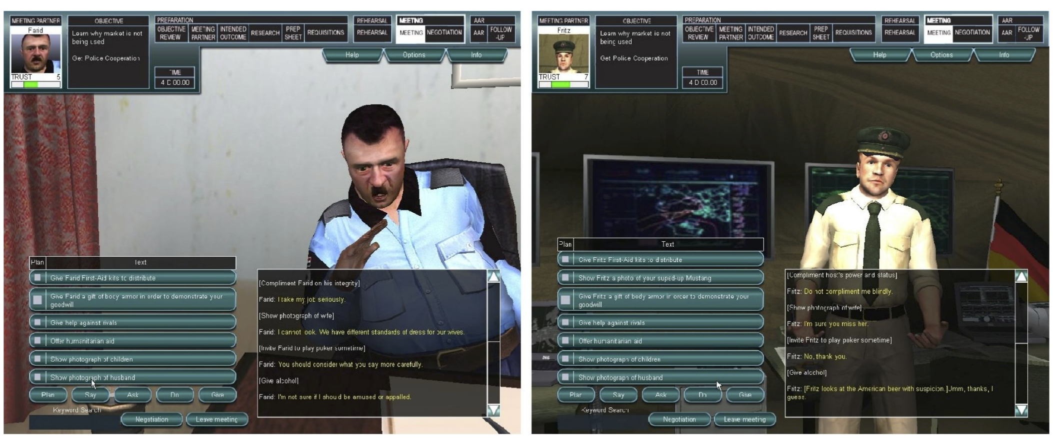


Figure 1.2: CAB: A virtual reality trainer (Solomon et al., 2008).

As the graphical fidelity of virtual characters has increased substantially, the gap between their realistic appearance and unrealistic behaviors has widened. This is in part due to two trends currently seen in the video game industry. First, modern video games are moving towards giving the player more agency. While most commercial video games still do not allow the player to free type like *Façade* (Mateas and Stern,

2003) or use their voice like *The Restaurant Game* (Orkin and Roy, 2007), more options are becoming available (e.g., the interaction wheel of the *Mass Effect* series, Figure 1.3). Second, even the traditionally story-oriented studios such as BioWare are making their games more open-world (Campbell, 2014) where the player is actively encouraged to roam massive worlds and interact with hundreds of AI-controlled virtual characters in them (Pramath, 2015).

Figure 1.3: The interaction wheel of the *Mass Effect* series (Bioware, 2007).

Combined, the two trends result in a progressively larger number of AI-controlled non-player characters (NPCs) the player can interact with and a larger number of ways to do so. As a result, even with tens of millions of dollars in development budgets, it is impossible to hand-script emotionally affected reactions and appearances of each character in response to each possible way the player may interact with them. Thus, a handful of characters may be fully scripted for the specific ways the player is allowed to interact with them whereas hundreds of others will deliver generic one-line reactions to actions the player takes toward them. The discrepancy breaks immersion and reminds the player that she is playing a game. For instance, Amita (Figure 1.4), a key story character in *Far Cry 4*, becomes a generic in-world NPC after the single-player champaign ends and the game transitions to an endless open-world simulator. Previously talkative and emotional, Amita now idly walks around the world, cycling through generic animations, mute and oblivious of her prior interactions with the player. While merely immersion-breaking in a video game, such lack of realism may be intolerable in a VR trainer (Traum et al., 2003).

Figure 1.4: Amita from Far Cry 4 (Ubisoft, 2014).

We propose to address the content bottleneck by procedurally generating both emotionally affected actions and emotional appearance descriptors of computer-controlled characters in video games and virtual reality (VR) trainers. Procedural generation is a promising approach with advances made in level generation (Valtchanov and Brown, 2012), NPC population (Booth, 2009) and gameplay difficulty adjustment (Hunicke and Chapman, 2004). Furthermore, work exists on visualizing emotions via procedurally animated facial expressions (Nvidia, 2013) and body language (NaturalMotion, 2014).

In this thesis we tackle the problem of procedurally generating believable emotionally charged NPC behavior using a resource based approach. We do this by adapting a recent resource-based model for NPC behavior and combining it with a simple appraisal model. While procedural models of emotional reactions have been explored over the past few years, our algorithm is intentionally light-weight for an easy integration in video games and trainers, especially running on mobile hardware and consoles.

1.2 Contributions

We briefly describe the contributions this thesis makes to existing work in the field of procedural generation of NPC emotion and behavior. Firstly, we build on an existing

Figure 1.5: Walk styles portraying different emotions (Desai and Szafron, 2012).

light-weight model for NPC behavior and add the ability to generate emotions to it. Second, we make improvements on the manner of behavior generation used in the earlier model to accommodate more diverse behavior. Lastly, we run a user study to verify the improvements over previously published work. The results have been disseminated in the field (Manavalan et al., 2015; Manavalan and Bulitko, 2014).

The rest of the thesis is organized as follows. In Chapter 2 we describe the problem of procedural NPC emotion and behavior generation which we are attempting to solve in this paper. Chapter 3 presents the existing work in this field and why it is insufficient to solve our problem. We describe our work in Chapter 4 followed by a description of the user study we conducted to verify our claims in Chapter 5. Chapter 6 discusses possible future work.

1.3 Summary

In this chapter we provided motivation as to why procedurally generated NPC characters are needed for video games and for training environments. We also briefly overviewed the contributions of this thesis to the field of procedural generation of NPC emotion and behavior.

Chapter 2

Problem Definition

2.1 What is a Believable NPC?

The believability of a non-player character is determined by the extent to which an observer or player engages and empathizes with the character (El-Nasr et al., 2009). The NPC should interact with the human player and with other NPCs in a natural manner as perceived by a human observer. It should also be reactive to changes in player actions and the environment around it. It has been shown that an NPC character or an artificial agent which always knows the best possible response to a problem is not considered human-like or believable (Traum et al., 2003). This often comes into play, for example, in a first-person shooter where an NPC is expected to reason where the enemy player is using the type of information a human would access (Hladky and Bulitko, 2008).

2.2 Developing Emotionally Believable NPCs

We are interested in a subset of the problem of creating a believable NPC: creating an emotionally believable NPC. It means that the NPC's actions, utterings and body language should be interpreted by an observer as caused by an emotion. An NPC's action is any action the NPC performs in a game world (e.g., fighting, standing, talking). Utterings are the lines spoken by an NPC capable of speech or sounds produced by an NPC to communicate. The body language of an NPC is the manner in which an NPC performs an action (e.g., an NPCs walk could be slouched or springy or aggressive).

2.3 Evaluating Emotional Believability of NPCs

One of the major open issues in the field of creating believable artificial agents is the lack of a standard method for model evaluation (Lin et al., 2012). Ideally, we would like to evaluate a model by assessing its ability to demonstrate human-like emotion. Various researchers are exploring the option of developing a standard test scenario to evaluate cognitive models (Adams et al., 2012) but to the best of our knowledge this is still an open problem. Thus, we use the currently accepted method which is conducting user studies to compare the believability of our emotion model to a baseline approach. Such studies involve a human participant interacting with several AI systems in a game-like setting (Figure 2.1). The participants are unaware of the difference between the systems and they are asked to rate their experiences along different measures. The collected response is then used to evaluate the new system with respect to an established or known system.

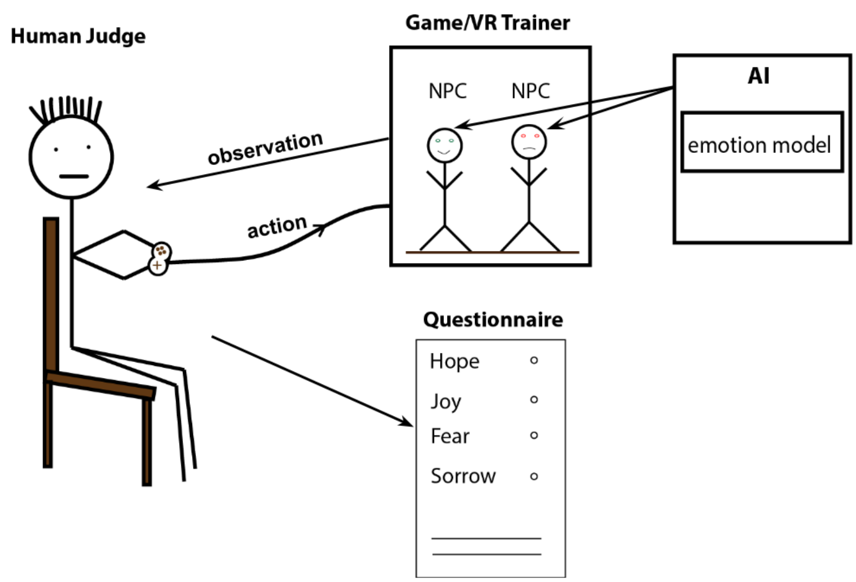


Figure 2.1: User study setup.

Chapter 3

Related Work

There are three primary approaches to achieving emotionally believable artificial characters in video games and other settings. The first is manual scripting, where the developers script the actions, facial expressions, utterings and body language of a virtual character for every possible in-game interaction. The second approach replaces hand-scripted emotional NPC responses with a procedural model that takes in the context and the player's actions and outputs a numerical descriptor of the NPC emotions. The third approach forgoes procedural generation of such emotion descriptors and instead generates emotion-related actions with the hope that the observer will infer the emotions from them.

In this chapter we discuss these approaches and some emotion models based on these approaches. We also discuss why these models do not adequately solve our problem formulated in Chapter 2. For clarity and uniformity we choose a running example with which to illustrate the emotion models discussed in this chapter. Consider a scenario where agents[1] are standing in a line to purchase a movie ticket (Figure 3.1). At each time step, an agent could wait in the line, pass another agent or protest against an agent passing it.

3.1 Manual Scripting

Manual scripting is the current industry-standard method to develop emotionally believable artificial agents where the developer has complete control over the emotion and behavior of the agents in each scenario. Manual scripting uses voice acting, hand-coded animations, dialog scripts and hand-coded behaviors to control non-player characters in a virtual environment. This can yield realistic movie-like experience

[1]the terms "artificial agents" and "NPCs" are used interchangeably throughout the thesis.

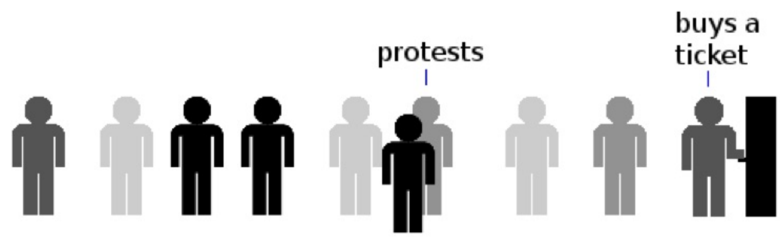


Figure 3.1: People standing in a line (Campano et al., 2013).

with highly detailed characters but requires a substantial amount of time and effort on the part of the designers. Consequently, most video games and virtual training environments necessarily limit the number of characters the player/trainee can interact with and/or the scope of the interactions available. For instance, BioWare's *Mass Effect* series (Bioware, 2007) constrains the player's actions with a dialog wheel within a conversation and quicktime events in certain situations and only fleshes out the primary characters in specific encounters. Even the primary characters exhibit emotionally unbelievable behavior beyond pre-scripted encounters: the implausibly repetitive "Wrex. Shepard." dialogue from *Mass Effect* (Figure 3.3) has become an Internet meme (Linguica, 2008). Other characters tend to utter generic one-liners and behave less believably when the player attempts to interact with them. For example, in *Guild Wars 2* (ArenaNet, 2012), the player could jump on a table and dance and the NPC characters surrounding him will still greet him (Figure 3.2).

3.1.1 Façade

Façade is a "fully-realized, one-act interactive drama", developed by Mateas and Stern (2003) as a research experiment in interactive storytelling and AI. In *Façade*, author knowledge is represented in the form of a set of dramatic beats each of which comes with preconditions. A beat is the smallest unit of dramatic action (e.g., looking away), and they can occur in many different orders while still maintaining a coherent narrative. *Façade* consists of about 200 beats which result in thousands of possible orderings. This provides the player with about 15-20 minutes of gameplay with emotional characters. The story is conveyed in a small 3D virtual world, in which the player character interacts with two NPC characters who are married to each other. The system's explicit design goal was to provide the player with local and global agency over the evolution of the dramatic experience.

Figure 3.2: NPCs greeting the PC as he dances on a table (ArenaNet, 2012).

However, looking at the code for *Façade* (Figure 3.4) we note that while the system uses a planner to order the different dramatic beats, it still required substantial manual effort to develop. *Façade* took a total of 10 man-years of development time for up to 20 minutes of gameplay involving two characters. This does not scale to AAA video games titles that routinely feature 30 or more hours of gameplay.

Thus manual scripting is highly resource intensive on the developer side. As the number of characters the player can interact with in a game increases, the cost of scripting every agent becomes prohibitive. One solution to this problem is procedural generation of emotions as we survey in the next section.

3.2 Procedural Emotion Modeling

The second approach to creating believable artificial agents replaces hand-scripting of emotional NPC responses with a procedural model that takes in the context and the player's actions and outputs a numerical descriptor of the NPC emotions. A few representatives of this approach use the appraisal theory of emotions (Scherer et al., 2001). The theory argues that emotion arises from two basic processes: appraisal

Figure 3.3: The "Wrex. Shephard." interaction in *Mass Effect* (Bioware, 2007).

and coping. Appraisal is the process by which a person or a virtual agent evaluates its current state and the relationship with the environment. It consists of not just the current state but of all the past events which led to the current state as well as the possible future states. The variables involved in this evaluation of the agent state are called appraisal variables and characterize each event from the individual's perspective. These concepts were made popular in the field of emotion research by Ortony et al. (1990) in their emotion model: OCC. The model describes a hierarchy that classifies 22 distinct emotion descriptors[2]. These emotion descriptors have numeric values corresponding to the intensity of the emotion associated with them. The hierarchy of emotion descriptors contains three branches, namely emotions concerning consequences of events (e.g., joy, pity), actions of agents (e.g., admiration, shame) and aspects of objects (e.g., love, hate). Some emotion descriptors are combined to form compound emotions, such as emotions concerning consequences of events caused by action of agents (e.g., gratitude, anger). Although the OCC model

[2]emotion descriptors are called emotion types by Ortony et al.

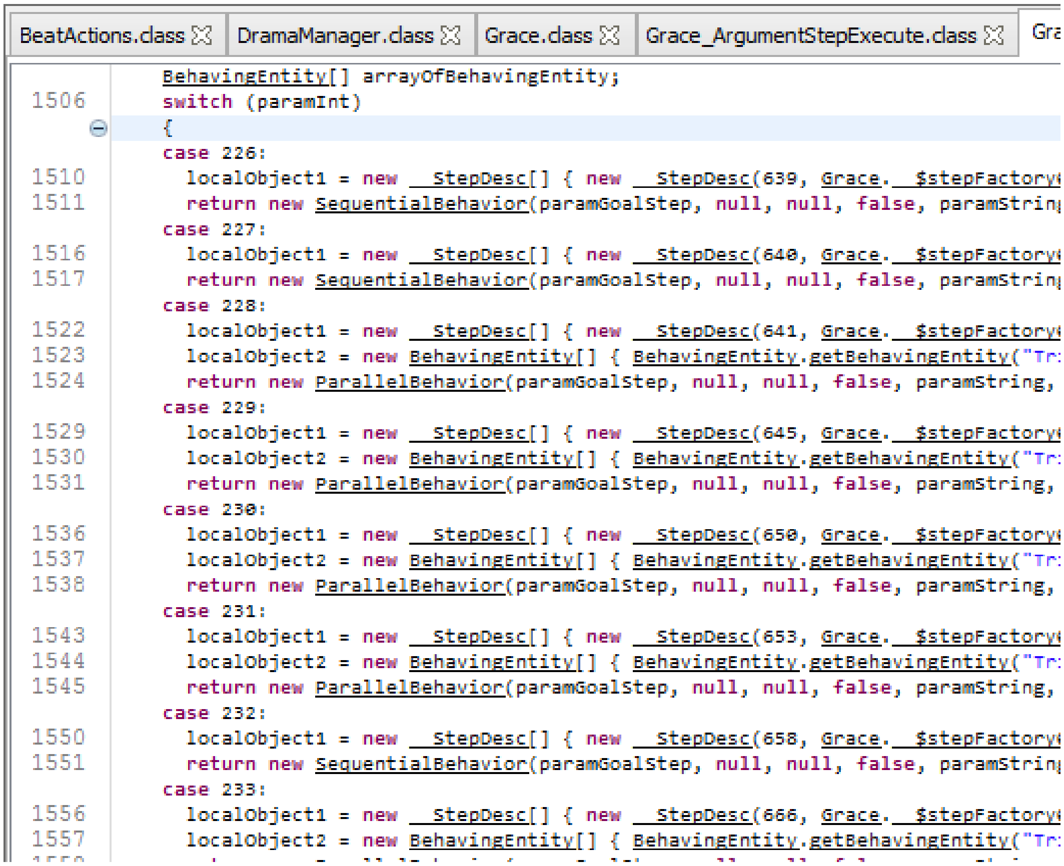


Figure 3.4: A fragment of *Façade* source code.

describes 22 numerical emotion descriptors, most emotion models use only a subset of these emotion descriptors.

3.2.1 Émile

Émile is an emotion model developed by Gratch (2000) primarily for virtual reality trainers and educational games. It aims to create autonomous agents to populate a "constructive simulation" where the interactions with these agents would be used for training. This is a plan-based model which uses an appraisal planner inspired by the OCC model described above. Each agent has a goal and the model determines a plan based on possible subgoals which may lead to the achievement of this goal. Each subgoal has probability of success and the joint probability of the subgoals is used as the probability of achieving the goal. Emotions are determined based on the importance of the goal to an agent and the probability of achieving the said goal. In *Émile* the probability of achieving a subgoals is handcoded by the designers of the system.

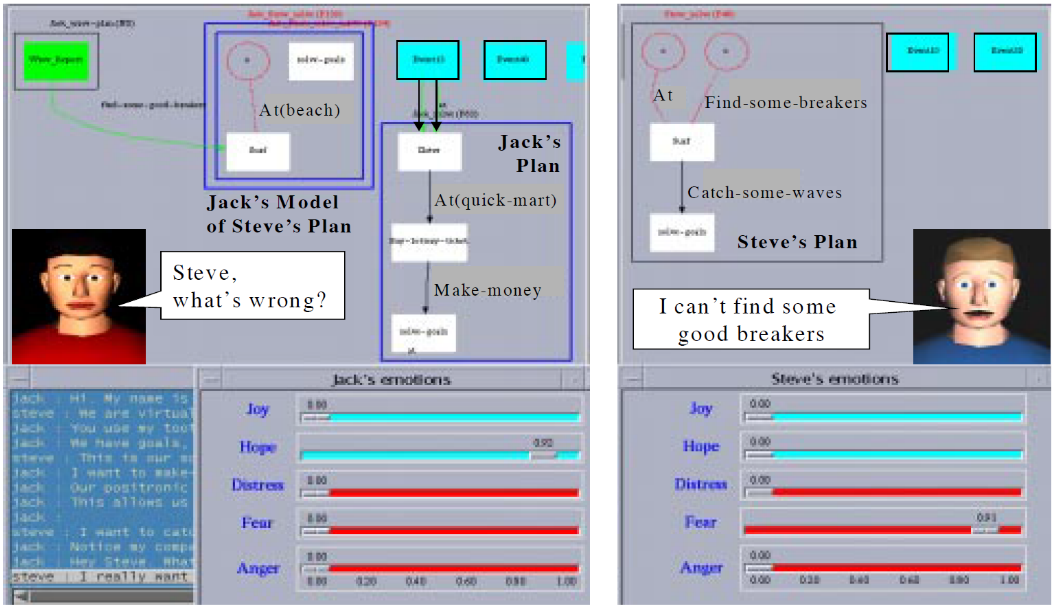


Figure 3.5: An interaction between two *Émile* agents (Gratch, 2000).

Here we illustrate the operation of *Émile* on our running example. When implemented in *Émile*, the agent's goal would be to get to the front of the line to purchase the movie ticket. All positions in front of an agent till the front of the line are the agent's subgoals. Passing another agent, protesting or waiting in line are actions which the agent might take to achieve the overall goal. Based on the probability values handcoded by the designers, if the agent computes that the probability of passing an agent successfully is higher than waiting in line, then the agent will experience the emotion of hope. *Émile* has a certain constraint incorporated by its authors: an agent should not introduce threat into other agent's plan. This implies that an agent would not take any action which might reduce another agent's probability of reaching its goal. Thus the agents in *Émile* will always stay in line and never pass another agent. Hence, the probability of achieving the goal would remain constant and the agents would not display any change in their emotions.

3.2.2 EMotion and Adaptation (EMA)

Another computational implementation of the appraisal theory of emotions is *EMotion and Adaptation* (EMA) (Gratch and Marsella, 2001, 2004a,b; Marsella and Gratch, 2009) which is reviewed in this section. In EMA, an agent's relationship with its environment is represented by beliefs, desires, intentions, plans and probabilities. The agent's interpretation of this relationship between itself and its environment is referred to as the causal interpretation of the agent.

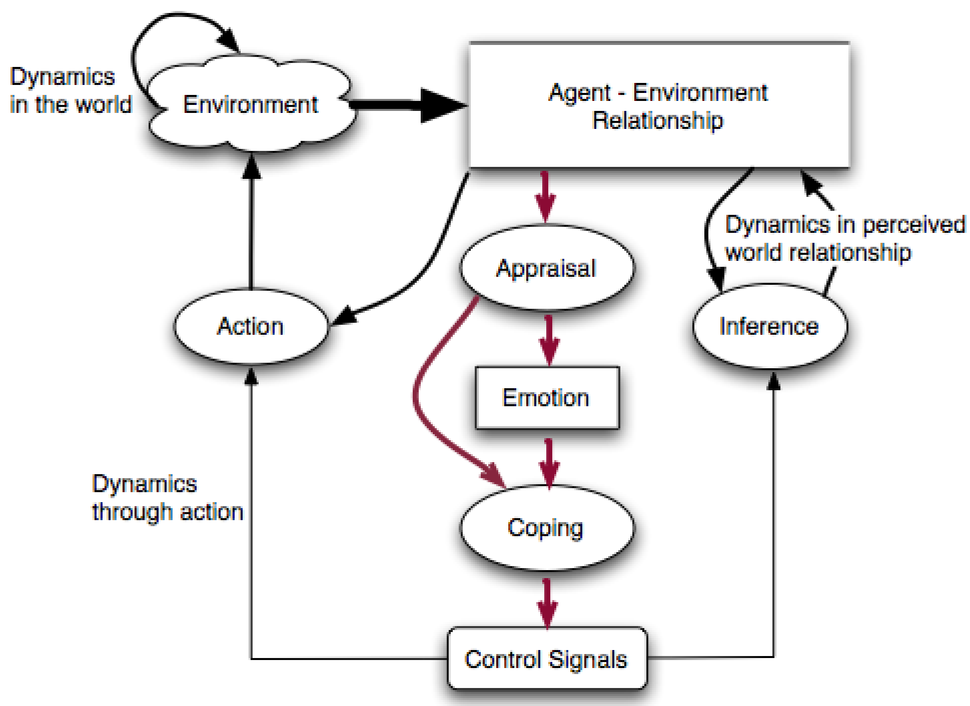


Figure 3.6: Flowchart of EMA (Marsella and Gratch, 2009).

The flowchart of the decision process in EMA is reproduced in Figure 3.6. EMA starts with the appraisal of the agent-environment relationship and an emotion is generated based on this appraisal. For example, if a situation is appraised as undesirable and the agent has low control over the situation, EMA will compute the emotion of fear. On the other hand, if the situation is appraised as desirable and certain, EMA will compute the emotion of joy. Coping is the decision making process which determines how an agent responds to the changes in the environment. Coping can either cause the agent to take an action which affects its environment which in turn changes the agent-environment relationship or the agent could change its interpretation or belief of the agent-environment relationship (e.g., reassess an ally as an adversary). Once the agent takes one of these two courses of action, the agent then re-evaluates the agent-environment relationship.

We now describe how EMA would operate in our running example. The goal of each agent is to reach the front of the line. A possible plan for an agent would be to pass another agent standing in line by executing the pass action. Each agent would have a certain probability of successfully completing the pass which may be supplied by the authors of the system. An agent could be waiting in line and EMA could determine that passing another agent could make the NPC achieve its goal earlier. This elicits the emotion of hope. As the NPC is passing, another agent could protest which has a possibility of changing our current NPC state to an undesirable state: being injured. This elicits the emotion of fear. The possible actions for the passing NPC now would be to go back to original position in line or ignore the protest. The action selection is done based on the agent's desirability for each state and the probability of reaching the state, both of which are supplied by the designer.

3.2.3 Culture-Emotion MAtrix

Culture-Emotion MAtrix (CEMA) (Bulitko et al., 2008) is a computationally light-weight combination of *Culturally Affected Behaviour* (CAB) and *EMotion and Adaptation* (EMA). CEMA reformulates EMA's appraisal mechanism as a matrix computation. The reformulation however, does not extend to the action selection mechanism implemented in EMA. This model is not demanding of computational resources and is easier to implement than EMA but it can only produce emotion descriptors (e.g., the amount of fear an NPC is feeling) and not the related coping actions (e.g., run away from source of fear).

3.3 Models without Explicit Emotion Descriptors

The third approach forgoes procedural generation of emotion descriptors and instead generates emotion-related actions. The observer is then expected to infer the underlying emotions. For instance, in *Halo 2* (Bungie, 2004) when the player encounters enemy soldier that retreat, the player thinks: "oh, the enemy just ran away screaming because I pulled out my energy sword and it was scared" (Isla, 2005). We believe the player might come to such a conclusion because some enemies in *Halo 2* have a large number of retreat impulses, which results in the player categorizing these enemies as cowardly and associating their retreat with the emotion of fear. Thus, the addition of retreat impulses to an NPC could make the player believe that the NPC is fearful without explicitly modeling the emotion of fear. In this section we illustrate two models which generate emotion-related actions but do not compute emotion descriptors.

3.3.1 Conservation of Resources Engine

The *Conservation of Resources Engine* (COR-E) is a light-weight emotion model for believable emotional responses from the artificial characters in a video game without having to generate an emotion descriptor. This model is based on the psychological theory of Conservation of Resources (Hobfoll, 1989). The theory postulates that an agent's actions emerge from protecting and gaining resources that the character deems valuable.

COR-E assumes that every agent has a set of discrete valued resources. The agent assigns a unique resource weight to each of its resource based on a ranking relation. The resource weights determine the agent's preferences. Any change in the environment which has a possibility of reducing an agent's resource will elicit a protective action from the agent. Otherwise, if no resource is threatened, the agent takes an action which would increase a desired resource.

To evaluate the effectiveness of their model, the authors of COR-E conducted a user study with a scenario similar to our running example. Participants were shown COR-E agents standing in line to purchase a movie ticket (Figure 3.1). Each agent had a set of reputation, health resources and rank resources. If an agent values a rank resource higher than all its reputation resource, then the agent will try to pass. And if a rank resource is threatened by another agent, the agent in question might protest by threatening the passing agent's reputation or health. The participants were asked if they found the behavior of the agents believable and if these behaviors were related to the characters' emotions. The participants were shown three video clips. In the first video clip, acquisitive (e.g., passing) and protective (e.g., protesting) behaviors were disabled. In the second clip, acquisitive behaviors were activated, but not protective. In the third clip which was also the experiment condition, acquisitive and protective behaviors were activated. According to Student's t-test, the believability scores for the normal configuration of COR-E, were significantly higher than an average score of 4. The result of the user study supports the claim that it is possible for an observer to associate emotions to a character's actions even though the actions are generated without explicitly generating any emotion descriptor at any stage of the process.

While computationally light-weight and less daunting than an emotion model like EMA to embed into a game, COR-E does not generate explicit emotion descriptors and therefore is unable to drive facial (MOVA, 2009; Nvidia, 2013) and non-facial (Desai and Szafron, 2012) NPC emotional displays or modulate the NPC voice (Black and Lenzo, 2003). The actions could be interpreted as resulting from an emotion but if the characters maintain a neutral expression, speak in a monotone or have a neutral walk the NPC's emotion believability may be reduced.

3.3.2 Thespian

As a framework for realizing interactive drama, *Thespian* takes author generated scripts as constraints on agent's behaviors while the agents themselves remain autonomous. Each *Thespian* agent has a view of the world around it and its relationship with the various entities in this world. The agent also maintains its guess of the relationships other agents have with the entities in the world them. *Thespian* replaces the concepts of resource and resource weights used in COR-E with goal states and goal weights respectively. It also make these values continuous instead of discreet values used in COR-E. *Thespian* is built on *PsychSim* (Si et al., 2005), a multi-agent system for social simulation based on Partially Observable Markov Decision Problems (POMDPs) (Smallwood and Sondik, 1973) and the dependency on the underlying *PsychSim* architecture suggests that the game developer or the author of the VR trainer would have to integrate *PsychSim* to make *Thespian* work. Like COR-E, *Thespian* does not compute explicit emotion descriptors and hence suffers from similar drawbacks such as an inability to modulate a character's voice or control facial expressions and body language.

We describe how our running example could be implemented in *Thespian*. Instead of resources used in COR-E, here we have goal states as: healthy, at-the-head-of-the-line, maintain-good-reputation. An agent standing in line could have different goal weights for each of these states. If the goal weight of at-the-head-of-the-line is higher than the weight of maintain-good-reputation and healthy, then the agent will try to pass the agent ahead of it in line. Otherwise if another agent tries to pass the agent in consideration, the agent updates its belief about the passing agent's state. Then, based on the goal weights of the agent being passed, *Thespian* will determine whether the agent should protest the passing action or let the other agent pass. If the agent's weight for the maintain-good-reputation goal is higher than the weight for at-the-head-of-the-line goal, the agent will let the other agent pass. Otherwise, it

Figure 3.7: Tactical Language Training System running *Thespian* agents (Si et al., 2006).

will protest the pass.

3.4 Summary

We discussed three approaches for generating emotionally believable NPC characters. The most widely used approach is manual scripting and it gives the designer a complete control on the emotion behavior of the NPC character. It requires however, substantial human effort and becomes intractable as the number of NPCs increases. The second approach is based on the procedural generation of emotions informed by the appraisal theory. Appraisal models require specifying NPC plans, tasks, goals and the probabilities of reaching the goals. The third approach is based on generating believable behavior without explicitly computing an emotion descriptor. The observer is expected to infer the accompanying emotions which may not be sufficient in the situations where the NPC's voice or facial expression needs to be colored with emotion to make them believable.

Chapter 4

Our Approach

In this chapter we present our approach to solving the problem formulated in Chapter 2. Thus, we develop a computational model of emotions to be implemented in a video game or a virtual reality simulator. We adopt COR-E (Campano et al., 2012) and extend it while maintaining its light computational requirements. We do so by adding the appraisal model of CEMA (Bulitko et al., 2008).

We start this chapter with a high-level introduction to our model, the Appraisal of Conservation of Resources Engine (ACORE) in Section 4.1. We explain the concept of resource in Section 4.2 followed by a detailed analysis of how ACORE computes NPC behavior and how it differs from COR-E (Section 4.3). Finally we describe how we use the idea of resources to compute emotion descriptors in ACORE (Section 4.4).

4.1 Introduction to ACORE

Our system uses resources similarly to COR-E, to generate actions that an observer can interpret as the NPC's display of emotions (e.g., fleeing from danger may be attributed to fear). Unlike COR-E, ACORE additionally uses an appraisal model of emotions to compute explicit emotion descriptors. For instance, if a valuable resource (e.g., health) is threatened, ACORE might make the NPC initiate a protective action (e.g., fleeing from danger) but at the same time it will also generate a numeric value for the emotion of fear which can be used to drive the NPC appearance and even affect a display of the protective action (e.g., the way the NPC flees from the scene).

Under ACORE, the in-game NPCs or the virtual agents in a VR trainer are controlled by Algorithm 1. Lines 1 through 3 initialize the NPCs' data. Then, as long as the game is not over (line 5), we model each NPC's action selection by invoking the resource conservation module $\text{ACORE}_{\text{action}}$ (line 7, detailed in Section 4.3). We then compute NPC's emotion descriptors by invoking the appraisal module $\text{ACORE}_{\text{emotion}}$ (line 8, detailed in Section 4.4). Resource values are updated using the actions of this and other NPCs (line 9). Emotion descriptor values decay over time (line 10).

Algorithm 1: ACORE

```
1  for each NPC do
2      initialize: resources r̄_1, resource weights w̄_1
3      set ē_1 ← (0, 0, 0, 0)
4  t ← 1
5  while simulation running do
6      for each NPC do
7          act: a_t ← ACORE_action(r̄_t, w̄_t, A_t)
8          display emotions: ē_{t+1} ← ACORE_emotion(r̄_t, w̄_t, ē_t)
9          update resources: r̄_{t+1} from r̄_t, a_t and other NPCs' actions
10         decay the emotion values: ē_{t+1} ← ē_{t+1}/2
11     t ← t + 1
```

4.2 Resources

ACORE uses the concept of resources similar to COR-E. A resource is a quantifiable property which the agent deems valuable. It could be a simulated physical object that the agent possesses (e.g., food, weapons) or a simulated physiological property (e.g., health). ACORE associates a set of valuable resources with each NPC:

$$\bar{r}_t = (r_t^1, \ldots, r_t^N)$$

where each r_t^i is a scalar representing the value of the i-th resource at time t. The NPC assigns different importance or weight to each resource. This is represented by the N-dimensional weight vector $\bar{w}_t$:

$$\bar{w}_t = (w_t^1, \ldots, w_t^N) \in (0, 1]^N$$

. A higher weight indicates higher importance of the resource to the NPC. The weights may change over time but remain in $(0, 1]$.

We will now walk through Algorithm 1 with our running example. A line of NPCs is shown in Figure 4.1[1]. Each of the NPCs has three resources: health, reputation and rank, which is inversely proportional to the NPC's place in line. Assume that at

[1] They are now waiting to buy a video game, "Destiny", instead of a movie ticket

some time t, the resources of our NPC are $\bar{r}_t = (1, 1, 0.5)$ in the aforementioned order. Suppose the NPC values the health resource at 0.26, the reputation resource at 0.17 and the rank resource at 0.44: $\bar{w}_t = (0.26, 0.17, 0.44)$. In the following sections we describe what actions this ACORE-controlled NPC might take while standing in the line and how its emotions change accordingly.

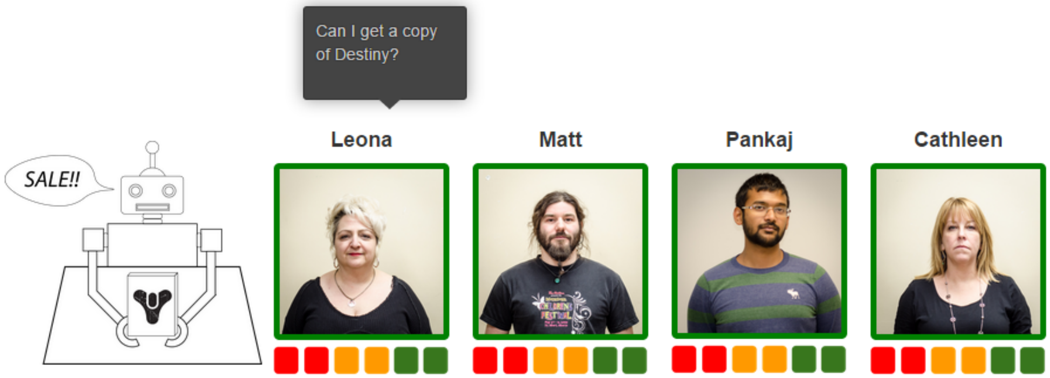


Figure 4.1: NPCs standing in line to purchase a video game.

4.3 NPC Actions

Each NPC has a set of actions (called *behaviors* in COR-E) available to it at time t, denoted by A_t. The actions can be primitive or complex (defined by a control policy). Each action can affect the resources the NPC holds (e.g., getting into a fight may negatively affect the NPC's health). An NPC selects the action that is expected to increase its cumulative weighted resource value (lines 2 and 3 of Algorithm 2).

Algorithm 2: ACORE$_{\text{action}}$

inputs : current resources $\bar{r}_t$, resource weights $\bar{w}_t$, available actions A_t
outputs : selected action: a_t

1 **for** $a \in A$ **do**
2   compute action value: $V(a) \leftarrow \bar{w}_t \times (E[\bar{r}_{t+1}|a] - \bar{r}_t)^T$
3 select action: $a_t \leftarrow \arg\max_{a \in A_t} V(a)$

The value of an action is the sum of the resource value deltas weighted by the resource weights, conditional on the action. In our example, an NPC standing in line has two actions available to it: $A_t = \{a^{\text{pass}}, a^{\text{wait}}\}$. By taking the a^{pass} action, the NPC will attempt to pass the NPC standing in front of it, thus improving its rank in the line but possibly losing reputation or even health (if the NPC being passed physically protests the pass). Suppose that by passing the NPC expects its resources to change from $\bar{r}_t = (1, 1, 0.5)$ to $E[\bar{r}_{t+1}|a^{\text{pass}}] = (0.95, 0.6, 1)$. Weighting the expected delta $E[\bar{r}_{t+1}|a^{\text{pass}}] - \bar{r}_t$ by $\bar{w}_t = (0.26, 0.17, 0.44)$ the NPC computes the value of passing as $V(a^{\text{pass}}) = 0.139$ (line 2).

The alternative action is to wait in line which does not immediately change the resource vector. Hence $V(a^{\text{wait}}) = 0$. Then, in line 3, the NPC will select the action with the maximum value: $a_t = a^{\text{pass}}$. Note that a different set of resource weights (e.g., $\bar{w}_t = (0.5, 0.5, 0)$ for an NPC that cares about its health and reputation but not about its place in the line) would have made the value of the passing action negative and led to the NPC selecting the wait action instead.

4.4 NPC Emotions

ACORE simplifies the emotion model of EMA (Gratch and Marsella, 2001, 2004a) by assuming that each NPC has only a single goal: to improve its weighted resource values. The appraisal process is thus limited to considering changes in the resource vector. In line with CEMA (Bulitko et al., 2008), ACORE models four emotions:

- *Hope* is elicited when there is a possible increase in a resource but it is uncertain.
- *Joy* is elicited when there is a certain increase in a resource.
- *Fear* is elicited when there is a possible decrease in a resource but it is uncertain.
- *Distress* is elicited when there is a certain decrease a in resource.

At time t, the emotion descriptor of an NPC is represented as a four-dimensional vector $\bar{e}_t = \left(e_t^{\text{joy}}, e_t^{\text{hope}}, e_t^{\text{distress}}, e_t^{\text{fear}}\right)$ where each e_t represents the intensity of the corresponding emotion. The intensity of each emotion with respect to a resource is computed as the product of the NPC's desirability α of the change in the resource caused by the action selected by $\text{ACORE}_{\text{action}}$ (Section 4.3) and the certainty of the change β. The desirability α is the product of the resource weight and the expected change in the value of a resource caused by the action (line 3, Algorithm 3). This is a measure of the value the NPC agent assigns to the change in the value of the resource. This could be positive which denotes an increase in the value of the resource and is favorable to the agent eliciting positive emotions of hope or joy. Negative values of α denote a decrease in the overall value of the resource which is unfavorable to the agent and elicits negative emotions of distress or fear. A desirability of zero leads to no changes to the agent's emotions.

The certainty of the change of a resource is the probability of the change in the direction specified by α (lines 5 and 7, Algorithm 3). The certainty of the change can either be handcoded by a domain expert or it could be computed by the system. For the user study presented in this thesis, we hand-code the probability values in to the system (Table 4.1).

The total intensity of each emotion is the sum of its intensity for each resource (the loop over i in line 2, Algorithm 3). Which emotion gets the update depends on the relation between the desirability and certainty. Desirable but uncertain changes in a resource contribute to hope (lines 10 and 11), desirable and certain changes contribute to joy (lines 8 and 9). Likewise, undesirable resource changes contribute to fear (when they are uncertain, lines 14 and 15) or distress (when they are certain, lines 12 and 13).

Algorithm 3: ACORE$_{\text{emotion}}$

---

inputs : current resources $\bar{r}_t$, resource weights $\bar{w}_t$, selected action a_t, current emotion $\bar{e}_t$

output : emotion $\bar{e}_{t+1} = \left(e^{\text{joy}}_{t+1}, e^{\text{hope}}_{t+1}, e^{\text{distress}}_{t+1}, e^{\text{fear}}_{t+1}\right)$

1 start with the current emotion $\bar{e}_{t+1} \leftarrow \bar{e}_t$
2 **for** resource index $i = 1, \ldots, N$ **do**
3   compute desirability: $\alpha \leftarrow w^i_t \cdot (E[r^i_{t+1}|a_t] - r^i_t)$
4   **if** $\alpha > 0$ **then**
5   compute certainty: $\beta \leftarrow \Pr(r^i_{t+1} > r^i_t \mid a_t)$
6   **else**
7   compute certainty: $\beta \leftarrow \Pr(r^i_{t+1} < r^i_t \mid a_t)$
8   **if** $\alpha > 0$ & $\beta = 1$ **then**
9   compute joy: $e^{\text{joy}}_{t+1} \leftarrow e^{\text{joy}}_{t+1} + \alpha$
10   **else if** $\alpha > 0$ & $\beta < 1$ **then**
11   compute hope: $e^{\text{hope}}_{t+1} \leftarrow e^{\text{hope}}_{t+1} + \alpha \cdot \beta$
12   **else if** $\alpha < 0$ & $\beta = 1$ **then**
13   compute distress: $e^{\text{distress}}_{t+1} \leftarrow e^{\text{distress}}_{t+1} - \alpha$
14   **else if** $\alpha < 0$ & $\beta < 1$ **then**
15   compute fear: $e^{\text{fear}}_{t+1} \leftarrow e^{\text{fear}}_{t+1} - \alpha \cdot \beta$

We will now illustrate the operation of ACORE$_{\text{emotion}}$ with concrete numbers, continuing our example from the previous section. Starting with the current emotion values $\bar{e}_t$ (line 1), our NPC computes an update to its new emotional state $\bar{e}_{t+1}$ for each resource in the loop in line 2. Suppose the current value of the NPC's health is $r^{\text{health}}_t = 1$. Under the action $a_t = a^{\text{pass}}$ selected by ACORE$_{\text{action}}$ in Section 4.3,

Table 4.1: Domain dynamics while an action is being performed.

Resource r_t	**Action** a_t	**Expected resource change** $E[r_{t+1}\|a_t] - r_t$	**The other NPC's action**	**Certainty of change** β
health	pass	-0.05	wait	0.95
reputation	pass	-0.4	wait	0.95
rank	pass	$1/(1/r_t^{\text{rank}} - 1) - r_t^{\text{rank}}$	wait	0.95
health	pass	-0.05	protest	0.5
reputation	pass	-0.4	protest	0.5
rank	pass	$1/(1/r_t^{\text{rank}} - 1) - r_t^{\text{rank}}$	protest	0.5
health	protest	-0.05	pass	0.5
reputation	protest	-0.1	pass	0.5
rank	protest	$1/(1/r_t^{\text{rank}} + 1) - r_t^{\text{rank}}$	pass	0.5
health	wait	0	wait	0.95
reputation	wait	0	wait	0.95
rank	wait	0	wait	0.95
health	wait	0	pass	0.5
reputation	wait	0	pass	0.5
rank	wait	$1/(1/r_t^{\text{rank}} + 1) - r_t^{\text{rank}}$	pass	0.5

the health resource is expected to decrease ($E[r_{t+1}^{\text{health}}|a^{\text{pass}}] = 0.95$) due to a possible physical opposition from the NPC being passed. The decrease in health in undesirable because health is positively weighted ($w = 0.26$). Thus,

$$\alpha = w \cdot \left(E[r_{t+1}^{\text{health}}|a^{\text{pass}}] - r_t^{\text{health}}\right) = 0.26 \cdot (0.95 - 1) = -0.013.$$

The certainty β of the undesirable health decrease is then computed as $\Pr\left(r_{t+1}^{\text{health}} < r_t^{\text{health}} \mid a^{\text{pass}}\right)$ (line 7) which, generally speaking, depends on the action the NPC being passed will take. We model this in two stages. At the first stage we assume that our NPC has not yet observed the actions of other NPCs and thus uses a prior for the certainty values. We compute the resulting changes to the emotion model which can then be visualized with the NPC's appearance. In stage two, the NPC has observed the actions of the affected NPCs and updates its emotion model accordingly. The emotional appearance can then be visualized once again.

For the health resource before the NPC is able to observe actions of the NPC being passed, it uses the prior $\beta = 0.95$. With this value, the conditions in line 14 will be satisfied and the intensity of emotion fear will be updated as

$$e_{t+1}^{\text{fear}} \leftarrow e_{t+1}^{\text{fear}} - \alpha \cdot \beta = e_{t+1}^{\text{fear}} - (-0.013) \cdot 0.95 = e_{t+1}^{\text{fear}} + 0.0123$$

in line 15. In other words, our NPC is now slightly more afraid of the prospects of losing some of its health.

The for loop in Algorithm 3 will then consider the next resource, reputation. The reputation is also predicted to decrease under the pass action: $E[r_{t+1}^{\text{reputation}}|a^{\text{pass}}] = 0.6$ because cutting in front of other people in line is disreputable. Our NPC cares about its reputation ($w = 0.17$) and so

$$\alpha = w \cdot \left(E[r_{t+1}^{\text{reputation}}|a^{\text{pass}}] - r_t^{\text{reputation}} \right) = 0.17 \cdot (0.6 - 1) = -0.068.$$

The prior probability of this undesirable loss of reputation is $\beta = 0.95$ which adds an extra $-\alpha \cdot \beta = -(-0.068) \cdot 0.95 = 0.0646$ to the emotion of fear.

The final resource is the rank in line whose current value is $r_t^{\text{rank}} = 1/2$ (i.e., the NPC is second in line). Under the pass action, the new value is expected to be $E[r_{t+1}^{\text{rank}}|a^{\text{pass}}] = 1$. Our NPC cares about its rank ($w = 0.44$) which means that

$$\alpha = w \cdot \left(E[r_{t+1}^{\text{rank}}|a^{\text{pass}}] - r_t^{\text{rank}} \right) = 0.44 \cdot (1 - 0.5) = 0.22.$$

The prior probability of this desirable gain of rank is $\beta = 0.95$ which adds $\alpha \cdot \beta = 0.22 \cdot 0.95 = 0.209$ to the emotion of hope (lines 10 and 11). Thus, at the end of the first stage, the NPC's fear is increased by $0.0123 + 0.0646 = 0.0769$ and the hope is increased by 0.209. The new emotion descriptor is then visualized via the NPC appearance (e.g., Cathleen in Figure 4.2).

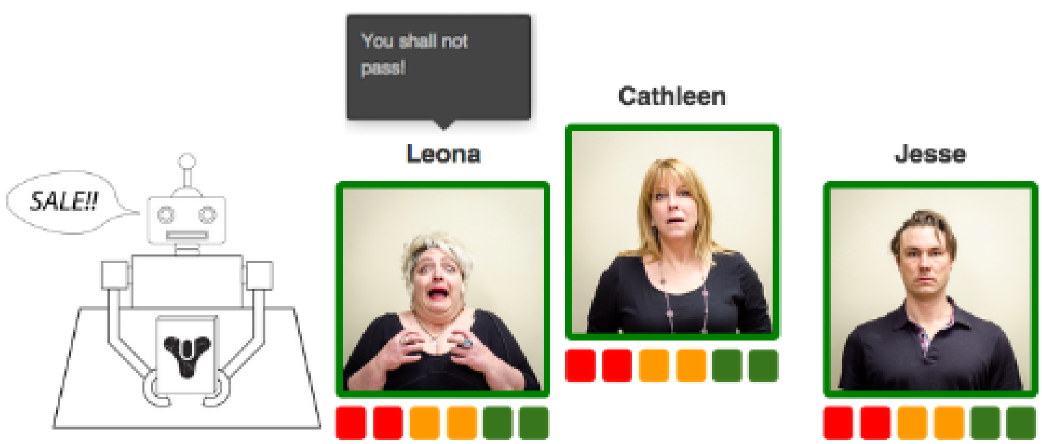


Figure 4.2: Screenshot showing Cathleen trying to pass Leona.

Suppose the NPC being passed protests the pass (e.g., Leona in Figure 4.2). Thus, in stage two, the passing NPC observes the protest and updates its α and β values. Retrieving the values from Table 4.1, ACORE computes the new increase of fear as $-\alpha \cdot \beta = -(-0.013) \cdot 0.5 = 0.0065$ which brings the intensity of fear to $0 + 0.0065 = 0.0065$.

In stage two, the change in the reputation resource remains the same leading to $\alpha = 0.17$. However, since the NPC is now aware of its pass action being protested, the expectation that the reputation will decrease is reduced to $\beta = 0.5$. Thus, the stage-two increase in fear due to possible change in reputation is $-\alpha \cdot \beta = -(-0.068) \cdot 0.5 = 0.034$.

For the rank resource, the expected value of the resource change remains the same as in stage one which leads to $\alpha = 0.22$. The certainty is now reduced to $\beta = 0.5$. Thus in stage two the NPC's hope is increased by $\alpha \cdot \beta = 0.22 \cdot 0.5 = 0.11$. Thus, in the end of the second stage, the NPC's the NPC's fear is increased by $0.0065 + 0.034 = 0.0405$ and its hope is increased by 0.11. The new emotion descriptor is visualized.

It should be noted that the β values for certainty of change in Table 4.1 are for the moment in time the actions are taken. Once the outcome of an action becomes known β becomes 1 or 0. For instance, once an NPC successfully passed another NPC, it will experience joy since its position in line (the rank resource) has improved (e.g., Cathleen in Figure 4.3).

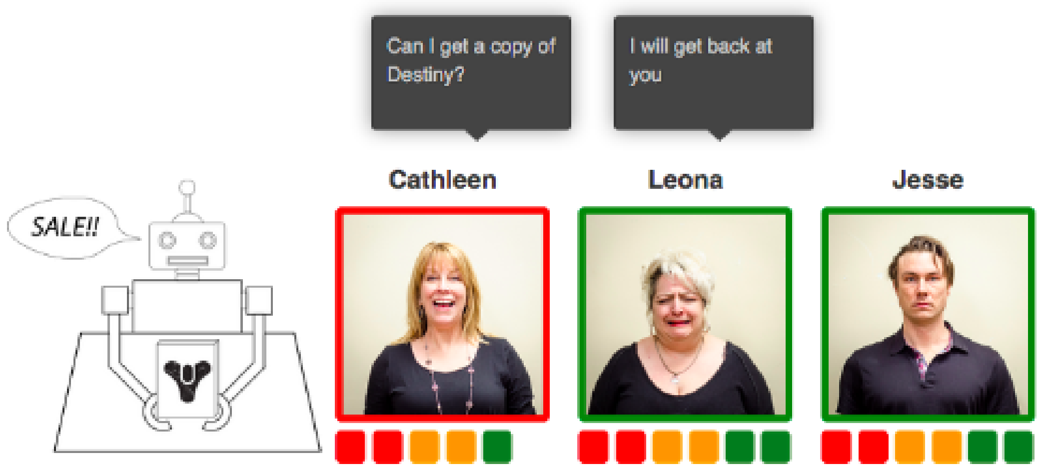


Figure 4.3: Screenshot showing Cathleen after she has passed Leona.

4.5 Summary

In this chapter, we presented ACORE (Appraisal of Conservation of Resources Engine), the NPC emotion model we had developed to generate believable NPCs in a video game or a training simulation. We provided an overview of ACORE's operation, including the concept of resources, the algorithm used to compute an NPC's action and the algorithm used to generate the NPC's emotion. We also walked through our running example to detail ACORE's operation.

Chapter 5

Empirical Study

We evaluated our implementation of ACORE with a user study in which we compared ACORE to two baselines.

5.1 Experimental Testbed

We consider our running example again which is an adaptation of the scenario from COR-E studies (Campano et al., 2013). For the reader's convenience we review the scenario. We considered a line of six people waiting to purchase a video game (Figure 5.1). Each person in line is represented by an NPC controlled by ACORE and has three resources: health, reputation and rank. They also have two basic actions available to them at each time step: pass the person in front of them or wait in line. Additionally, a person being passed can choose to protest the passing action or allow it to happen (i.e., continuing waiting in line).

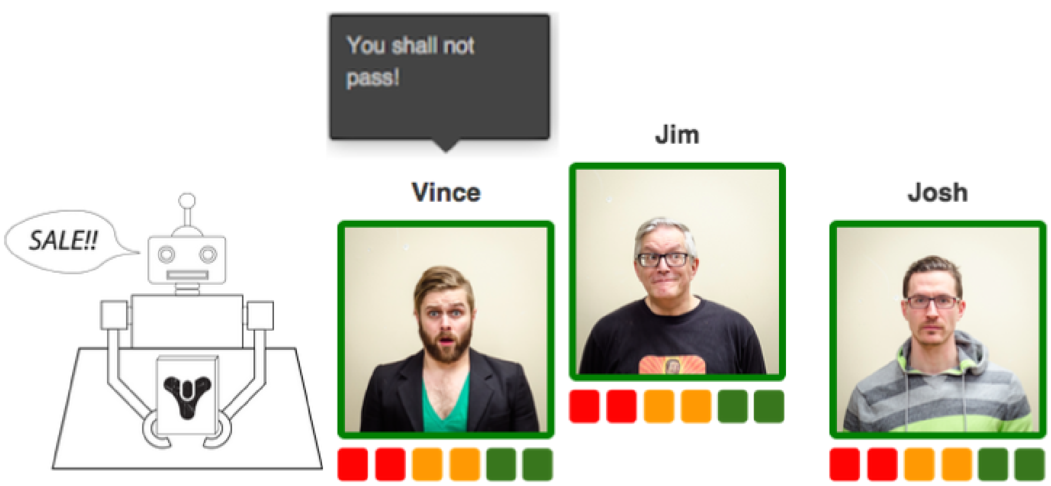


Figure 5.1: Individuals standing in line to purchase a video game. Jim is attempting to pass Vince.

Figure 5.2: An individual showing hope (top left), joy (top right), fear (bottom left) and distress (bottom right).

At every time step, the person at the head of the line purchases the video game and leaves the line. The simulation was stopped when the line became empty. Visually, each person in line was represented with a photograph showing their facial expression (only the highest intensity emotion was shown; Figure 5.2), their name and the three resources. Health was visualized with a bar underneath the image. The reputation was shown by the color of frame around their portrait (Figure 5.3). The rank was shown by the position of the person in line (Figure 5.1). Additionally, people in line uttered one-line remarks shown as text above their heads (Table 5.1). There were 18 individuals (Figure 5.4) from which a line of six was randomly populated (without repetition).

5.1.1 Implementation Details

ACORE was originally implemented as a text-based demonstration and the algorithm was written in Python. We then decided to show the emotions as facial expressions on a webpage. The webpage was created using Django which could directly run

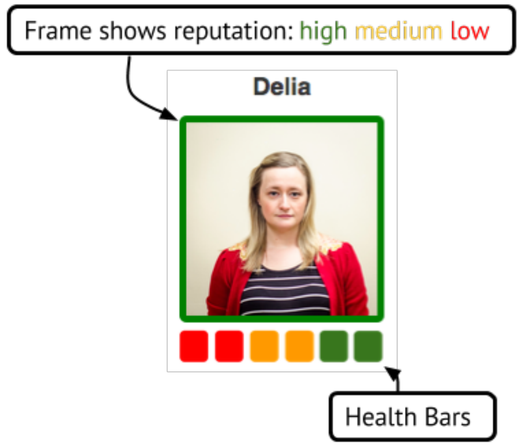


Figure 5.3: The visual representation of resources.

Table 5.1: One-line utterances.

Condition	Utterance
At the head of the line	Can I get a copy of Destiny?
Being passed	Stop! Where are you going? You shall not pass!
Having been just passed	Not fair! I will get back at you!

the Python code in the backend. For the facial expressions we initially used *Face Plus* (Mixamo, 2013) (Figure 5.5). The number of characters available in *Face Plus* however was very limited and to incorporate more variety we decided to use human faces (Figure 5.4). To run the user studies we obtained the domain `http://acore.cs.ualberta.ca/` but we ran into issue when trying to implement the Django server on the website. As a result, we rewrote the ACORE algorithm and the user interface in client side Javascript and ran the algorithm from a webpage in our user study.

5.1.2 Participants

For the user study we recruited 94 participants (30 males, 64 females; mean age 20). The participants came from the research pool at the Department of Psychology at the University of Alberta. The students participated for a partial credit and were also given the option of taking an alternate assignment (Appendix A.5) if they choose not to participate in the experiment. None of the participants opted for the alternate assignment.

Figure 5.4: Neutral expressions of the 18 individuals whose portraits were used in the user study.

Figure 5.5: Facial expressions generated using *Face Plus* (Mixamo, 2013).

No personal information was collected from the participants other than their age, gender and the numbers of reading and gaming hours per week. Research ethics approval granted was by the Research Ethics Board 1 at the University of Alberta; a complete copy of the approved application can be found in Appendix A.4.

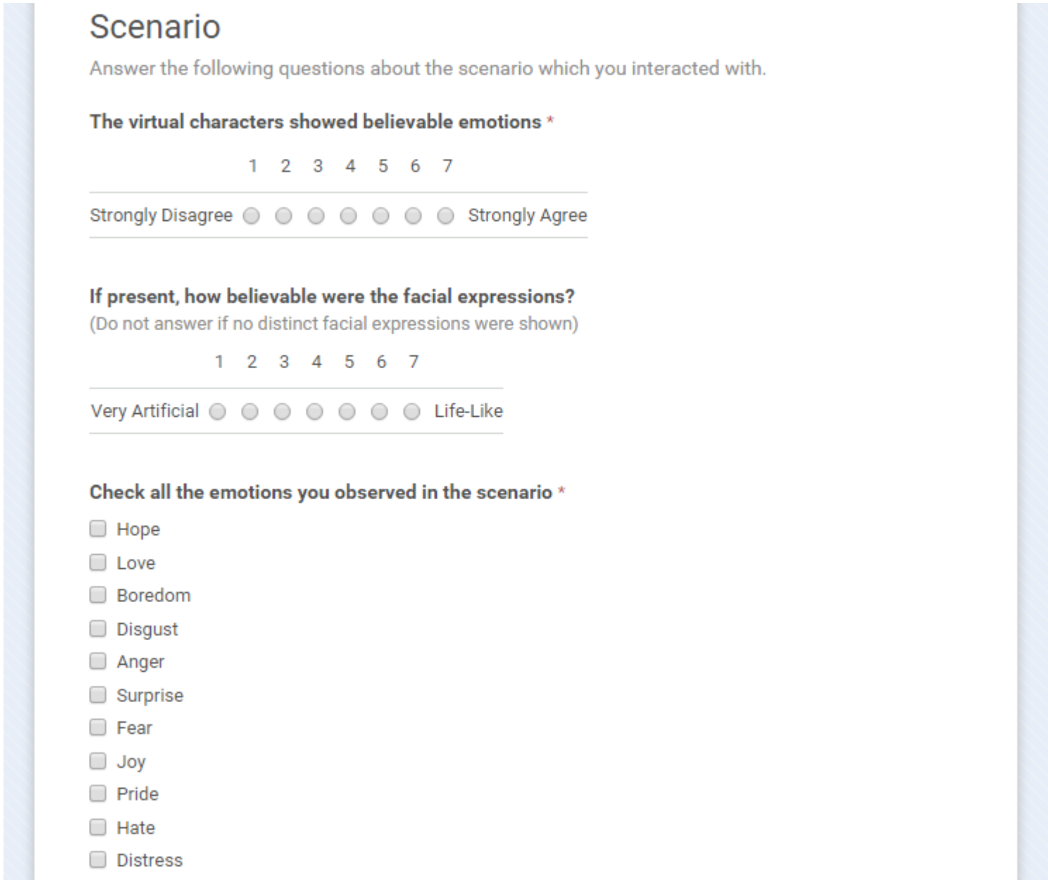
5.1.3 Experimental Setup

Multiple one-hour sessions were held in a computer laboratory at the Department of Computing Science, University of Alberta. Each session had an average completion time of 25 minutes and we had 5-15 participants per session. Each participant was assigned to a computer and was given a briefing and consent form. A short oral briefing was also provided to the participants. Afterwards, the participants followed the instructions on the screen and observe the six scenarios one by one. After the completion of the survey participants were given a debriefing form and were allowed to leave. All of the forms can be found in Appendix A.

5.2 The Experiment

Each participant was exposed to each of the following conditions.

Experimental condition (E) presented a participant with a line of people whose actions and appearances were controlled by the ACORE algorithm (Chapter 4).


Scenario

Answer the following questions about the scenario which you interacted with.

The virtual characters showed believable emotions *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

If present, how believable were the facial expressions?

(Do not answer if no distinct facial expressions were shown)

	1	2	3	4	5	6	7	
Very Artificial	○	○	○	○	○	○	○	Life-Like

Check all the emotions you observed in the scenario *

- Hope
- Love
- Boredom
- Disgust
- Anger
- Surprise
- Fear
- Joy
- Pride
- Hate
- Distress

Figure 5.6: A part of the questionnaire.

Control condition (C) was identical to E except each person in line maintained their neutral expression throughout the simulation. This condition approximates the existing model, COR-E, which models NPC actions but not their emotional appearances.

Random condition (R) was identical to E except the facial expressions were uniformly randomly selected from the four images we had for each character.

Each participant saw one of the six possible permutations of the conditions (i.e., ECR, ERC, CER, CRE, REC, RCE). Then he/she saw the same permutation again. For instance, one participant may have seen ECRECR whereas another may have seen RCERCE. The permutation order was assigned randomly to participants. After each of the six conditions, he/she was required to fill out a questionnaire (Figure 5.6, Appendix A.3). The primary statement the participants were asked to respond to was *"The virtual characters showed believable emotions"*. The answers were on a 1 (strongly disagree) to 7 (strongly agree) point scale (Figure 5.6).

5.3 Results

We present the results from each permutation first. As mentioned above, in each permutation we ran each of the three conditions twice. The two runs of each condition were treated independently for the statistical tests. The permutations were named in the order in which the conditions were shown to the participants. For example, RCE condition denote that the participant observed the Random condition first, followed by Control and then he or she observed the Experiment condition. The results for each of the six permutations are shown in Tables 5.2 to 5.7.

A one-way repeated-measures analysis of variance (ANOVA), with Believability as the dependent measure and Condition (E, C, R) as the independent variable revealed a significant effect of Condition, $F(2, 358) = 39.51$, $p < 0.001$, $\eta_p^2 = 0.181$.

Table 5.2: Results for RCE permutation.

Condition	**Average believability**	**Variance**	**Standard deviation (SD)**	**SD of the mean**
Experimental	5.11	1.32	0.27	0.06
Control	3.06	3.27	0.43	0.10
Random	3.61	2.79	0.39	0.09

Table 5.3: Results for CER permutation.

Condition	**Average believability**	**Variance**	**Standard deviation (SD)**	**SD of the mean**
Experimental	4.66	2.41	1.55	0.24
Control	3.71	2.84	1.69	0.26
Random	3.85	2.51	1.58	0.25

Table 5.4: Results for CRE permutation.

Condition	**Average believability**	**Variance**	**Standard deviation (SD)**	**SD of the mean**
Experimental	5.47	2.32	1.52	0.28
Control	3.93	3.80	1.95	0.36
Random	4.90	3.29	1.81	0.33

Table 5.5: Results for ECR permutation.

Condition	**Average believability**	**Variance**	**Standard deviation (SD)**	**SD of the mean**
Experimental	4.82	1.24	1.11	0.24
Control	2.77	2.36	1.53	0.33
Random	3.68	2.40	1.55	0.33

Table 5.6: Results for ERC permutation.

Condition	Average believability	Variance	Standard deviation (SD)	SD of the mean
Experimental	3.75	2.52	1.59	0.46
Control	3.00	3.5	1.87	0.54
Random	3.33	3.05	1.75	0.50

Table 5.7: Results for REC permutation.

Condition	Average believability	Variance	Standard deviation (SD)	SD of the mean
Experimental	5.27	2.3	1.43	0.23
Control	2.97	4.08	2.02	0.35
Random	4.27	2.95	1.72	0.28

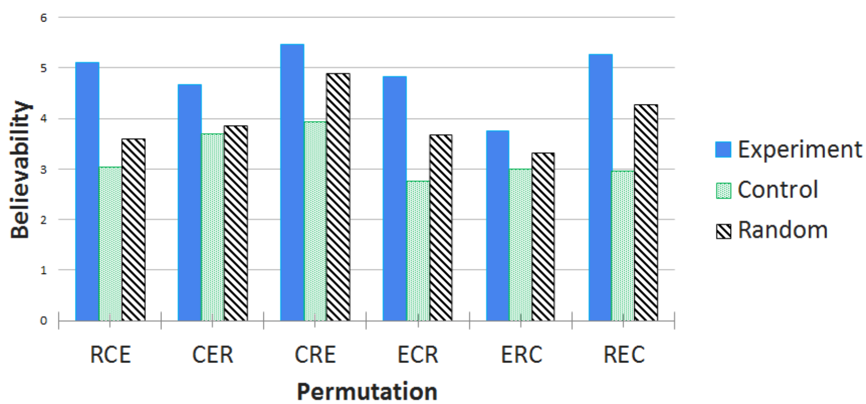


Figure 5.7: Average believability over each permutation.

Table 5.8: Mean believability of the conditions.

Condition	Mean $\pm$ standard error
Experimental	4.87 ± 0.12
Control	3.28 ± 0.14
Random	4.09 ± 0.13

Bonferroni-corrected pair-wise comparisons confirmed that all three conditions were significantly different from each other. The data is presented in Table 5.9.

5.4 Discussion

The participants found the experimental condition to be more believable than the control or the random condition (Table 5.8, Figure 5.8). Even when we consider each permutation individually, we see that the believability rating of the experiment

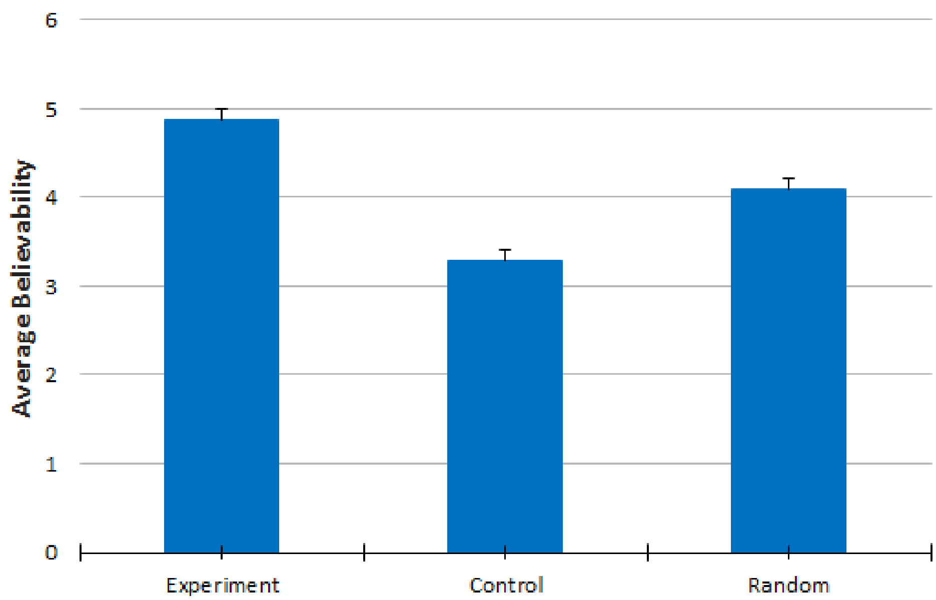


Figure 5.8: Average believability over all the runs.

Table 5.9: ANOVA results for the pairwise comparison between the different factors.

(I) Believ-ability	(J) Believ-ability	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval of Difference	
					Lower Bound	Upper Bound
1	2	−1.589	.179	.000	−2.023	−1.155
	3	−.811	.181	.000	−1.248	−.375
2	1	1.589	.179	.000	1.155	2.023
	3	.778	.176	.000	.352	1.203
3	1	.811	.181	.000	.375	1.248
	2	−.778	.176	.000	−1.203	−.352

condition appears higher than the other two conditions in every case (Figure 5.7). Thus, explicitly computing (via an appraisal model) and visualizing (via facial expressions) emotions appears to add to the believability over the control condition that approximated the previous algorithm (COR-E).

Interestingly, the participants rated the random condition as more believable than the control condition which appears to suggest that even random facial expressions, unrelated to the actions and utterings of people in the simulation, are better than a constant neutral expression. On the other hand, ACORE was consistently rated as more believable than random which suggests that actions and utterings based on the principle of conservation of resources can lead to more emotionally believable NPCs.

Chapter 6

Conclusion and Future Work

6.1 Contributions of the Thesis

The emotion model, ACORE, presented in this thesis makes three extensions to the previously published approach (COR-E). First, whereas COR-E did not explicitly compute emotion descriptors for a non-player character, we do so by adopting an appraisal-style model of emotions from CEMA. Second, whereas COR-E represented the NPCs preference over resources as a ranking relation, we do so with resource weights which allows us to represent multiple resources equally important to the NPC. Third, whereas an NPC in COR-E either had a resource or did not, an NPC in ACORE holds resource at a continuous value. The appraisal is done over resources which are already defined in the COR-E system and thus does not require explicitly computing the NPCs goals and plans to achieve them and is expected to maintain a light computational footprint.

We were able to rewrite the ACORE implementation from Django to a version implemented in Javascript running on a webpage in a short span of two weeks. This suggests that the algorithm is easily implementable in an object-oriented programming language.

6.2 Directions for Future Work

Our implementation of ACORE used *ad hoc* hand-coded values for expected resource changes as well as the associated certainty values. Since we use the same values for the COR-E and the random implementations, we do not suspect that they influenced the performance of ACORE with respect to the other conditions. If we had validated values for these changes it may have made all the three conditions better. Future work will generate such values procedurally via stochastic simulation.

Second, we used static images for the facial expressions which were photographs of actors acting out the emotions. Future work will investigate the effectiveness of procedurally generated facial expressions (MOVA, 2009; Nvidia, 2013), driven by ACORE.

Finally, our simulations were non-interactive as the participants watched the simulation unfold on a computer screen. We are currently working on incorporating ACORE into a video game where it will control actions and appearances of NPCs interacting with the player's avatar.

Bibliography

Sam Adams, Itmar Arel, Joscha Bach, Robert Coop, Rod Furlan, Ben Goertzel, J Storrs Hall, Alexei Samsonovich, Matthias Scheutz, Matthew Schlesinger, et al. Mapping the landscape of human-level artificial general intelligence. *AI Magazine*, 33(1):25–42, 2012.

ArenaNet. Guild Wars 2. `https://www.guildwars2.com/en/`, 2012.

Bioware. Mass Effect Series. `http://masseffect.bioware.com/`, 2007.

Alan W Black and Kevin A Lenzo. Building synthetic voices. *Language Technologies Institute, Carnegie Mellon University and Cepstral LLC*, 2003.

Michael Booth. The AI Systems of Left 4 Dead. `http://www.valvesoftware.com/publications/2009/ai_systems_of_l4d_mike_booth.pdf`, 2009.

Vadim Bulitko, Steven Solomon, Jonathan Gratch, and Michael van Lent. Modeling culturally and emotionally affected behavior. In *Proceedings of The Fourth Artificial Intelligence for Interactive Digital Entertainment Conference*, pages 10–15. The AAAI Press, 2008.

Bungie. Halo 2. `http://halo.bungie.net/Projects/Halo2/default.aspx`, 2004.

Sabrina Campano, Etienne de Sevin, Vincent Corruble, and Nicolas Sabouret. Evaluation of an affective model: Cor-e. In *Proceedings of the 12th International Conference on Intelligent Virtual Agents*, pages 508–510. Springer, 2012.

Sabrina Campano, Nicolas Sabouret, Etienne de Sevin, and Vincent Corruble. An Evaluation of the COR-E Computational Model for Affective Behaviors. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 745–752. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

Colin Campbell. How Skyrim is helping to shape Dragon Age: Inquisition. `http://www.polygon.com/2014/11/3/7151567/how-skyrim-is-helping-to-shape-dragon-age-inquisition`, November 2014.

Neesha Desai and Duane Szafron. Enhancing the believability of character behaviors using non-verbal cues. In *Proceedings of the Eighth Artificial Intelligence for Interactive Digital Entertainment Conference*, pages 130–135. AAAI Press, 2012.

Magy Seif El-Nasr, Leslie Bishko, Veronica Zammitto, Michael Nixon, Athanasios V Vasiliakos, and Huaxin Wei. Believable characters. In *Handbook of Multimedia for Digital Entertainment and Arts*, pages 497–528. Springer, 2009.

Jonathan Gratch. Emile: Marshalling passions in training and education. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 325–332. ACM, 2000.

Jonathan Gratch and Stacy Marsella. Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. In *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 278–285. ACM, 2001.

Jonathan Gratch and Stacy Marsella. A domain-independent framework for modeling emotion. *Journal of Cognitive Systems Research*, 5(4):269–306, 2004a.

Jonathan Gratch and Stacy Marsella. Evaluating the modeling and use of emotion in virtual humans. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 320–327. IEEE Computer Society, 2004b.

Stephen Hladky and Vadim Bulitko. An evaluation of models for predicting opponent positions in first-person shooter video games. In *Proceedings of the Fourth Computational Intelligence and Games, CIG'08.*, pages 39–46. IEEE, 2008.

Stevan E Hobfoll. Conservation of resources: A new attempt at conceptualizing stress. *American psychologist*, 44(3):513, 1989.

Eva Hudlicka and John Billingsley. Representing behavior moderators in military human performance models. In *Proceedings of the Eighth Conference on Computer Generated Forces and Behavioral Representation*, 1999.

Robin Hunicke and Vernell Chapman. AI for dynamic difficulty adjustment in games. In *Challenges in Game Artificial Intelligence AAAI Workshop*, pages 91–96. sn, 2004.

Damian Isla. Handling complexity in the Halo 2 AI. In *Proceedings of Game Developers Conference*, volume 12, 2005. URL `http://www.gamasutra.com/view/feature/130663/gdc_2005_proceeding_handling_.php?print=1`.

Mark S Johnson. Validation of an active multimedia courseware package for the integrated damage control training technology (IDCTT) trainer. Master's thesis, Naval Postgraduate School, Monterey, California, 1994.

Jerry Lin, Marc Spraragen, and Michael Zyda. Computational models of emotion and cognition. *Advances in Cognitive Systems*, 2:59–76, 2012.

Linguica. Wrex. Shepard. `https://www.youtube.com/watch?v=u5Fjc1hBwuE`, 2008.

Yathirajan Manavalan, Vadim Bulitko, and Marcia Spetch. A lightweight algorithm for procedural generation of emotionally affected behavior and appearance. In *Proceedings of the Eleventh Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. The AAAI Press, 2015.

Yathirajan Brammadesam Manavalan and Vadim Bulitko. Appraisal of emotions from resources. In *Proceedings of The Seventh International Conference on Interactive Digital Storytelling*, pages 224–227. Springer, 2014.

Stacy Marsella and Jonathan Gratch. EMA: A process model of appraisal dynamics. *Journal of Cognitive Systems Research*, 10(1):70–90, March 2009.

Michael Mateas and Andrew Stern. Façade: An experiment in building a fully-realized interactive drama. In *Proceedings of Game Developers Conference*, volume 2, 2003.

Mixamo. Face plus. `https://www.mixamo.com/faceplus`, 2013.

OnLive's MOVA. Geni4. `https://www.youtube.com/watch?v=0fF2pAsaaiw`, 2009.

NaturalMotion. Euphoria. http://www.naturalmotion.com/middleware/euphoria/, June 2014.

Nvidia. Nvidia Faceworks. https://www.youtube.com/watch?v=CvaGd4KqlvQ, 2013.

Jeff Orkin and Deb Roy. The restaurant game: Learning social behavior and language from thousands of players online. *Journal of Game Development*, 3(1):39–60, 2007.

Andrew Ortony, Gerald L Clore, and Allan Collins. *The cognitive structure of emotions*. Cambridge university press, 1990.

Richard Pew and Anne Mavor. Modeling human and organizational behavior, 1998.

Pramath. This is the full world map for the Witcher 3: Wild Hunt. http://gamingbolt.com/this-is-the-full-world-map-for-the-witcher-3-wild-hunt, 2015.

Klaus R Scherer, Angela Ed Schorr, and Tom Ed Johnstone. *Appraisal Processes in Emotion: Theory, Methods, Research.* Oxford University Press, 2001.

Mei Si, Stacy Marsella, and David V. Pynadath. Thespian: Using multi-agent fitting to craft interactive drama. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 21–28. ACM, 2005.

Mei Si, Stacy C Marsella, and David V Pynadath. Thespian: Modeling socially normative behavior in a decision-theoretic framework. In *Proceedings of the Sixth International Conference on Intelligent Virtual Agents*, pages 369–382. Springer, 2006.

Richard D Smallwood and Edward J Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5): 1071–1088, 1973.

Steven Solomon, Michael van Lent, Mark Core, Paul Carpenter, and Milton Rosenberg. A language for modeling cultural norms, biases and stereotypes for human behavior models. In *Proceedings of the 17th Behavior Representation in Modeling & Simulation (BRIMS) Conference, Providence, RI*, 2008.

Strategy First. Sub Command. http://en.wikipedia.org/wiki/Sub_Command, 2001.

David Traum, Jeff Rickel, Jonathan Gratch, and Stacy Marsella. Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 441–448. ACM, 2003.

Ubisoft. Tom Clancys Ghost Recon 2. http://en.wikipedia.org/wiki/Tom_Clancy's_Ghost_Recon_2, 2004.

Ubisoft. Far Cry 4. https://www.ubisoft.com/en-GB/game/far-cry-4/, 2014.

Valtchan Valtchanov and Joseph Alexander Brown. Evolving dungeon crawler levels with relative placement. In *Proceedings of the Fifth International C* Conference on Computer Science and Software Engineering*, pages 27–35. ACM, 2012.

Valve. Portal. http://www.valvesoftware.com/games/portal.html, 2007.

Appendix A

User Study Materials

In this appendix, we include the different support materials used during the empirical evaluation (user study).

A.1 Briefing and Consent Form

Computing Science Study: Evaluating Virtual Characters Generated by ACORE

Introduction and Briefing

Introduction
Welcome! You are invited to participate in a research study being conducted by Yathirajan Manavalan and Prof. Vadim Bulitko of the Department of Computing Science and Prof. Marcia Spetch of the Department of Psychology, from the University of Alberta. The purpose of this study is to determine the believability of the a set of virtual characters standing in a line. The results of this study will be used in subsequent research including dissemination and thesis defense of Yathirajan Manavalan. Each session of this study will last less than an one hour, typically about half an hour.

Your participation
Your participation in this study involves observing a set of virtual characters standing in a line to purchase a copy of a video game. You would have to observe their actions and how they go about purchasing the item over multiple runs which will last for approximately 30 minutes. Before beginning the study, you will be presented with a short set of instructions for interacting with the environment. Following each scenario, you will be asked to fill out a survey ranking the game across several measures. Your participation in this study is worth 2% toward your course mark.

Your rights
Your decision to participate in this study is entirely voluntary and you may decide at any time to withdraw. If you choose not to participate or withdraw after you have begun, but would like your 2% credit for participation, you may complete an alternative educational activity. In this case, you will be given a short article to read on the decisions that storytellers make while telling a story. You will be asked to answer on paper a few questions about the article. The time it takes to complete this assignment will be no longer than the time it takes to participate in this study. If you choose to withdraw your data after the completion of the study, you may do so within 3 days of the study date. You will need to provide your participant code number, which will be provided to you at the start of the study to withdraw your data. Your decision not to participate will not affect access to services from the University of Alberta. Your survey responses will remain confidential and anonymous, and our data file will NOT contain any personal identifiers (i.e., names or student ID numbers). Survey forms will be identified only by a researcher-assigned code number, for the purpose of associating them with the particular study run that the participant experienced. Only researchers associated with the project will have access to the questionnaires. The results of this study may be presented at scholarly conferences, published in professional journals or books, or presented in class lectures. All data presented will be anonymous. The survey is hosted by Google and the survey data is hosted on Google servers which can be located outside of Canada. As such, you should in the United States under the USA Patriot Act, the government has the right to access all information held in electronic databases. The data will also be securely stored by the Department of Computing Science (Yathirajan Manavalan) for a minimum of five years.

Benefits and risks

There are no major risks in this study; in general, the risks associated with this project are expected to be similar to those that are associated with reading a book or watching a movie. There is a minimal risk of fatigue and tension, as in any situation that involves observing animations on a computer screen If any risks should arise, the researcher will inform the participants immediately. If you should experience any adverse effects, please contact Yathirajan Manavalan and/or Prof. Vadim Bulitko.

Contact information

The plan for this study has been reviewed for its adherence to ethical guidelines and approved by Research Ethics Board 2 at the University of Alberta. For questions regarding participant rights and ethical conduct of research, contact the Research Ethics Office at (780) 492-2615. If you have any questions or comments on the study, or if you wish a clarification of rights as a research participant, please contact Yathirajan Manavalan.

Yathirajan Manavalan
M.Sc. Student
Department of Computing Science
University of Alberta
Edmonton, AB T6G 2E8
(780) 952-2425
yathi.bm@ualberta.ca

Vadim Bulitko, Ph.D.
Associate Professor
Department of Computing Science
University of Alberta
Edmonton, AB T6G 2E8
(780) 492-3854
bulitko@ualberta.ca

Marcia Spetch
Professor
Department of Psychology
University of Alberta
Edmonton, AB T6G 2E8
(780) 492-7548
mspetch@ualberta.ca

Signatures

Please sign below to indicate that you have read and understood the nature and purpose of the study. Your signature acknowledges the receipt of a copy of the consent form, as well as indicates your willingness to participate in this study.

Participant’s Signature: Date:

Researcher’s Signature: Date:

A.2 Debriefing Form

Evaluating Virtual Characters Generated by ACORE

Thank you for participating in this study! Your time and effort have been very valuable to us. Video games are an increasingly popular form of entertainment, but often most of the characters in a videogame do not behave as the characters would in real life. Our research in this study investigates the believability of game-like agents whose emotions and actions are controlled by artificial intelligence.

Specifically, we want to study how different AI algorithms affect perceived emotions and their believability. The algorithms use a concept of resources such as, the rank in the line, their reputation and their health. If a character tries to cut in front of another character in the line, he/she loses reputation, risks losing health and gains a rank. Different characters may assign different importance to their reputation, their health and rank in the line. Some might want to get the desired item as early as they can whereas others regard their reputation highly and will not want to tarnish it.

It was necessary to withhold this information in order to avoid any bias between the various runs. If you would like to withdraw your data from this study for any reason, you have 3 days to do so. Please contact Yathirajan Manavalan with your participant code if you wish to do so.

Thanks very much for participating. Do you have any questions that I can answer right now? If you have any questions, later on, about the study, please contact Yathirajan Manavalan via either phone (780-952-2425) or email (yathi.bm@ualberta.ca) or if you have general questions, contact Research Participation Coordinator at rescred@ualberta.ca or 780-492-5689. Please do not tell others about what we had you do here to avoid biasing them in case they participate in this study as well.

A.3 Questionnaire

User Study

Thanks for your participation. Please fill out this form below. Your time is appreciated.

* Required

Enter your participant code *

Enter the scenario number: *

This is shown above the survey form

Scenario

Answer the following questions about the scenario which you interacted with.

The virtual characters showed believable emotions *

	1	2	3	4	5	6	7	
Strongly Disagree	○	○	○	○	○	○	○	Strongly Agree

If present, how believable were the facial expressions?

(Do not answer if no distinct facial expressions were shown)

	1	2	3	4	5	6	7	
Very Artificial	○	○	○	○	○	○	○	Life-Like

Check all the emotions you observed in the scenario *

- ☐ Hate
- ☐ Hope
- ☐ Anger
- ☐ Disgust
- ☐ Distress
- ☐ Fear
- ☐ Joy
- ☐ Boredom

- [ ] Surprise
- [ ] Love
- [ ] Pride

Submit

Never submit passwords through Google Forms.

Powered by

This form was created inside of University of Alberta.

Report Abuse - Terms of Service - Additional Terms

A.4 Ethics Approval

1.1 Study Identification

*All questions marked by a red asterisk * are required fields. However, because the mandatory fields have been kept to a minimum, answering only the required fields may not be sufficient for the REB to review your application.*

Please answer all relevant questions that will reasonably help to describe your study or proposed research.

1.0 * **Short Study Title** *(restricted to 250 characters)***:**
ACORE: Measuring Emotional Believability of virtual characters

2.0 * **Complete Study Title** *(can be exactly the same as short title)***:**

ACORE: Measuring Emotional Believability of virtual characters

3.0 * **Select the appropriate Research Ethics Board (***Detailed descriptions are available by clicking the HELP link in the upper right hand corner of your screen***):**
REB 2

4.0 * **Is the proposed research:**
Funded (Grant, subgrant, contract, internal funds, donation or some other source of funding)

5.0 * **Name of Principal Investigator** *(at the University of Alberta, Covenant Health, or Alberta Health Services)***:**
Yathirajan Brammadesam Manavalan

6.0 **Investigator's Supervisor** *(required for applications from undergraduate students, graduate students, post-doctoral fellows and medical residents to Boards 1, 2, 3. HREB does not accept applications from student PIs)*

Vadim Bulitko

7.0 * **Type of research/study:**
Graduate Student - Thesis, Dissertation, Capping Project

8.0 **Study Coordinators or Research Assistants:** People listed here can edit this application and will receive all HERO notifications for the study:

Name	Employer
There are no items to display	

9.0 **Co-Investigators:** People listed here can edit this application but do not receive HERO notifications unless they are added to the study email list:

Name	Employer
Vadim Bulitko	SC Computing Science
Marcia Spetch	SC Psychology Science

10.0 **Study Team** *(Co-investigators, supervising team, other study team members):* People listed here cannot edit this application and do not receive HERO notifications:

Last Name	First Name	Organization	Role/Area of Responsibility	Phone	Email
There are no items to display					

1.3 Study Funding Information

1.0

* **Type of Funding:**

Grant (external)

If OTHER, provide details:

2.0

* **Indicate which office administers your award.** *(It is the PI's responsibility to provide ethics approval notification to any office other than the ones listed below)*
Other

If OTHER, provide details:

3.0 * **Funding Source**

3.1 Select all sources of funding from the list below:

NSERC - Natural Sciences And Engineering Research Council	NSERC

3.2 If not available in the list above, write the Sponsor/Agency name(s) in full (you may add multiple funding sources):
There are no items to display

4.0

* **Indicate if this research sponsored or monitored by any of the following:**

Not applicable

If applicable, indicate whether or not the FDA Investigational New Drug number or FDA Investigational Device Exception is required:

The researcher is responsible for ensuring that the study complies with the applicable US regulations. The REB must also meet particular review criteria and this application will likely receive full board review, regardless of level risk.

1.5 Conflict of Interest

1.0

* **Are any of the investigators or their immediate family receiving any personal remuneration (including investigator payments and recruitment incentives but excluding trainee remuneration or graduate student stipends) from the funding of this study that is not accounted for in the study budget?**

◯ Yes ◉ No

If YES, explain:

2.0 * **Do any of investigators or their immediate family have any proprietary interests in the product under study or the outcome of the research including patents, trademarks, copyrights, and licensing agreements?**
◯ Yes ◉ No

3.0 * **Is there any compensation for this study that is affected by the study outcome?**
◯ Yes ◉ No

4.0 * **Do any of the investigators or their immediate family have equity interest in the sponsoring company? (This does not include Mutual Funds)**
◯ Yes ◉ No

5.0 * **Do any of the investigators or their immediate family receive payments of other sorts, from this sponsor (i.e. grants, compensation in the form of equipment or supplies, retainers for ongoing consultation and honoraria)?**
◯ Yes ◉ No

6.0 * **Are any of the investigators or their immediate family, members of the sponsor's Board of Directors, Scientific Advisory Panel or comparable body?**
◯ Yes ◉ No

7.0 * **Do you have any other relationship, financial or non-financial, that, if not disclosed, could be construed as a conflict of interest?**
◯ Yes ◉ No

If YES, explain:

Important
If you answered YES to any of the questions above, you may be contacted by the REB for more information or asked to submit a Conflict of Interest Declaration.

1.6 Research Locations and Other Approval

1.0 * **List the locations of the proposed research, including recruitment activities. Provide name of institution or organization, town, or province as applicable**
Department of Psychology Lab, University of Alberta
Undergrad Laboratories, Department of Computing Science, University of Alberta

2.0 * **Indicate if the study will use or access facilities, programmes, resources, staff, students, specimens, patients or their records, at any of the sites affiliated with the following** *(select all that apply):*
Not applicable

List all facilities or institutions as applicable:

3.0 **Multi-Institution Review**

* **3.1 Has this study already received approval from another REB?**
○ **Yes** ● **No**

4.0 **Does this study involve pandemic or similar emergency health research?**
○ Yes ● **No**

If YES, are you the lead investigator for this pandemic study?
○ Yes ○ No

5.0 **If this application is closely linked to research previously approved by one of the University of Alberta REBs or has already received ethics approval from an external ethics review board***(s)*, **provide the HERO study number, REB name or other identifying information. Attach any external REB application and approval letter in Section 7.1.11 – Other Documents.**

2.1 Study Objectives and Design

1.0 **Date that you expect to start working with human participants:**
1/5/2015

2.0 **Date that you expect to finish working with human participants, in other words, you will no longer be in contact with the research participants, including data verification and reporting back to the group or community:**
12/31/2015

3.0 * **Provide a lay summary of your proposed research suitable for the general public** *(restricted to 300 words)*. **If the PI is not affiliated with the University of Alberta, Alberta Health Services or Covenant Health, please include institutional affiliation.**

This research primarily deals with generating autonomous emotionally believable virtual characters which could be used in a video game or a training environment. We use the concept of resources where every virtual character has certain resources important to it. The virtual character also has a set of actions which can help it protect its resources or acquire more resources. Our model decides on the actions of the virtual characters based on how its resources change in the virtual environment.

In this research, we will test whether the current model for virtual characters with facial expressions leads to characters which are deemed believable by a human observer.

4.0 * **Provide a description of your research proposal including study objectives, background, scope, methods, procedures, etc)** *(restricted to 1000 words)*. **Footnotes and references are not required and best not included here. Research methods questions in Section 5 will prompt additional questions and information.**

- Study Objective -

We wish to test if the addition of emotions and facial expressions to characters based on the conservation of resources leads to a more believable virtual character.

- Background -

Believable virtual characters are critical to creating an immersive character-rich experience whether it is for video game titles or training simulations. Emotionally plausible characters make the fictional world come alive. Conversely, emotionally implausible interactions break the player's immersion and remind her that she is merely playing a video game. While it is possible to manually script primary characters to display emotionally plausible responses, the procedure is expensive. Other, non-scripted, characters are left to utter repeated one-liners, ignore the player or warmly greet her after she walked on their dinner table during a meal. In this study we propose a step towards a more realistic story-telling experience by developing a light-weight computational model that drives character appearance and actions.

- Scope -

We expect this research to have an impact on virtual characters in video games and computer training environments.

- Methods and Procedures -

We intend to conduct a user study to test the aforementioned hypothesis, wherein human participants will observe a scenario where a small number of virtual characters are standing in line to purchase a valuable item. This scenario will be played out on a webpage which is hosted by the researchers. The participants will open the Web URL and they will be presented with a screen which will have 6 Faces/Silhouettes. These are the six virtual characters standing in line. Once the participant is ready and comfortable, they can press a green start button, also on the screen. Once the scenario starts, the virtual characters will show various expressions and utter dialogues as they try to go from right to left in the line. Some virtual characters will attempt to cut in front of others in the line. Various reactions will be observed. The most violent type of reaction in our system will be of the likes of "Stop" or "You shall not pass"; with no gore or detailed description of violence presented.

When a participant is done with observing two runs of a scenario of virtual characters standing in line, the webpage automatically presents them with a questionnaire. In the questionnaire the participants are asked to rate the believability of the emotions of the virtual characters they observed on a scale of 1 to 7. We also ask the participants to rate believability of the facial expressions, if present, on a scale of 1 to 7. Some scenarios have just silhouettes and they don't have facial expressions. We also ask the participants to select all the emotions they observed in the scenario from a list of 11 emotions. The questionnaire is attached with this application.

From this data we try to analyse if the addition of facial expressions to various actions and dialogue help make the virtual characters more believable.

5.0 **Describe procedures, treatment, or activities that are above or in addition to standard practices in this study area** *(eg. extra medical or*

health-related procedures, curriculum enhancements, extra follow-up, etc):
None.

6.0 **If the proposed research is above minimal risk and is not funded via a competitive peer review grant or industry-sponsored clinical trial, the REB will require evidence of scientific review. Provide information about the review process and its results if appropriate.**
None.

7.0 **For clinical research only, describe any sub-studies associated with this application.**

3.1 Risk Assessment

1.0 * **Provide your assessment of the risks that may be associated with this research:**
Minimal Risk - research in which the probability and magnitude of possible harms implied by participation is no greater than those encountered by participants in those aspects of their everyday life that relate to the research (TCPS2)

2.0 * **Select all that might apply:**

	Description of Potential Physical Risks and Discomforts
Possibly	Participants might feel physical fatigue, e.g. sleep deprivation
No	Participants might feel physical stress, e.g. cardiovascular stress tests
No	Participants might sustain injury, infection, and intervention side-effects or complications
No	The physical risks will be greater than those encountered by the participants in everyday life

	Potential Psychological, Emotional, Social and Other Risks and Discomforts
Possibly	Participants might feel psychologically or emotionally stressed, demeaned, embarrassed, worried, anxious, scared or distressed, e.g. description of painful or traumatic events
Possibly	Participants might feel psychological or mental fatigue, e.g intense concentration required
No	Participants might experience cultural or social risk, e.g. loss of privacy or status or damage to reputation
No	Participants might be exposed to economic or legal risk, for instance non-anonymized workplace surveys
No	The risks will be greater than those encountered by the participants in everyday life

3.0 * **Provide details of the risks and discomforts associated with the research, for instance, health cognitive or emotional factors, socio-economic status or physiological or health conditions:**
Participating in this user study can be somewhat fatiguing or stressful, in a similar way that watching television can strain one's eyes or elicit excitement and concern. In general, the risks associated with this project are expected to be similar to those that are associated with reading

literature or viewing a movie.

4.0 * **Describe how you will manage and minimize risks and discomforts, as well as mitigate harm:**
Participants will be informed that they may stop participating at any point during their study session, and be allowed to take breaks if they feel the need to do so.

5.0 * **If your study has the potential to identify individuals that are upset, distressed, or disturbed, or individuals warranting medical attention, describe the arrangements made to try to assist these individuals. Explain if no arrangements have been made:**
No arrangements have been made as the study will not identify these individuals.

3.2 Benefits Analysis

1.0 * **Describe any potential benefits of the proposed research to the participants. If there are no benefits, state this explicitly:**
Participants may get to experience how computer generated characters can behave in a life-like interesting manner.

2.0 * **Describe the scientific and/or scholarly benefits of the proposed research:**
The research would help in the improvement of virtual characters which could lead to better training as well as entertainment applications.

3.0 **Benefits/Risks Analysis: Describe the relationship of benefits to risk of participation in the research:**
The risk involved with this study is expected to be minimal, isolated, and short-term, while the benefits have the potential to improve a wide range of entertainment and training applications.

4.1 Participant Information

1.0 * **Who are you studying? Describe the population that will be included in this study.**
Undergraduate students obtained through the Psychology 104/105 Research Participation Pool.

2.0 * **Describe the inclusion criteria for participants (e.g. age range, health status, gender, etc.). Justify the inclusion criteria (e.g. safety, uniformity, research methodology, statistical requirement, etc)**

Participants must have normal or corrected vision, and be able to use a mouse and keyboard.
These restrictions are necessary to ensure that participants will be able to play through the experiences that our software system creates.

3.0 **Describe and justify the exclusion criteria for participants:**

Only participants who do not meet the criteria above will be excluded, for the reasons given above.

4.0

* **Will you be interacting with human subjects, will there be direct contact with human participants, for this study?**

(●) **Yes** ( ) No

Note: No means no direct contact with participants, chart reviews, secondary data, interaction, etc.

If NO, is this project a chart review or is a chart review part of this research project?

( ) Yes ( ) No

5.0

Participants

How many participants do you hope to recruit *(including controls, if applicable)*
200
Of these how many are controls, if applicable *(Possible answer: Half, Random, Unknown, or an estimate in numbers, etc).*

If this is a multi-site study, for instance a clinical trial, how many participants *(including controls, if applicable)* **are expected to be enrolled by all investigators at all sites in the entire study?**

6.0 **Justification for sample size:**
Similar previous studies achieved statistical significance with this amount .

7.0 **Does the research specifically target aboriginal groups or communities?**

( ) Yes (●) **No**

4.3 Recruit Potential Participants

1.0

Recruitment

* **1.1 Describe how you will identify potential participants (please be specific as to how you will find potentially eligible participants i.e. will you be screening AHS paper or electronic records, will you be looking at e-clinician, will you be asking staff from a particular area to let you know when a patient fits criteria, will you be sitting in the emergency department waiting room, etc.)**
Undergraduate students obtained through the Psychology 104/105 Research Participation Pool

1.2 Once you have identified a list of potentially eligible participants, indicate how the potential participants' names will be passed on to the researchers AND how will the potential participants be approached about the research.

1.3 How will people obtain details about the research in order to make a decision about participating? Select all that apply:

There are no items to display

1.4 If appropriate, provide the locations where recruitment will occur *(e.g schools, shopping malls, clinics, etc.)*

2.0

Pre-Existing Relationships

2.1 Will potential participants be recruited through pre-existing relationships with researchers *(e.g. Will an instructor recruit students from his classes, or a physician recruit patients from her practice? Other examples may be employees, acquaintances, own children or family members, etc)***?**

○ Yes ◉ **No**

2.2 If YES, identify the relationship between the researchers and participants that could compromise the freedom to decline *(e.g. professor-student)***. How will you ensure that there is no undue pressure on the potential participants to agree to the study?**

3.0 **Outline any other means by which participants could be identified, should additional participants be needed** *(e.g. response to advertising such as flyers, posters, ads in newspapers, websites, email, listservs; pre-existing records or existing registries; physician or community organization referrals; longitudinal study, etc)*

4.0 **Will your study involve any of the following** *(select all that apply)***?**

Payment or incentives, e.g. honorarium or gifts for participating in this study

4.5 Informed Consent Determination

1.0

* **Describe who will provide informed consent for this study (select all that apply). Additional information on the informed consent process is available at: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter3-chapitre3/#toc03-intro**

All participants have capacity to give free and informed consent

Provide justification for requesting a Waiver of Consent (Minimal risk only, additional guidance available at: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter3-chapitre3/#toc03-1b

2.0

How is participant consent to be indicated and documented? Select all that apply:

Signed consent form

Except for "Signed consent form" use only, explain how the study information will be communicated and participant consent will be documented. Provide details for EACH of the option selected above:

3.0

Authorized Representative, Third Party Consent, Assent

3.1 Explain why participants lack capacity to give informed consent *(e.g. age, mental or physical condition, etc.).*

3.2 Will participants who lack capacity to give full informed consent be asked to give assent?
◯ Yes ◯ No

Provide details. IF applicable, attach a copy of assent form(s) in the Documentation section.

3.3 In cases where participants (re)gain capacity to give informed consent during the study, how will they be asked to provide consent on their own behalf?

4.0 **What assistance will be provided to participants, or those consenting on their behalf, who have special needs?** *(E.g. non-English speakers, visually impaired, etc)***:**

5.0 * **If at any time a participant wishes to withdraw, end, or modify their participation in the research or certain aspects of the research, describe how their participation would be ended or changed.**
They will be asked to complete a different activity, such as reading a related article and answering some questions, or withdrawn completely if they do not wish to participate at all.

6.0
Describe the circumstances and limitations of data withdrawal from the study, including the last point at which it can be done:
If the participants wishes to withdraw his or her data from the study, they may do so within 3 days of the completion of the study. The participant needs to provide his or her unique participant code to withdraw their data from the study.

7.0 **Will this study involve any group(s) where non-participants are present? For example, classroom research might involve groups which include participants and non-participants.**
◯ Yes ◯ No

4.6 Reimbursements and Incentives

1.0 **IF you are providing expense reimbursements, describe in detail the expenses for which participants will be reimbursed, the value of the reimbursements and the process** *(e.g. participants will receive a cash reimbursement for parking, at the rate of $12.00 per visit for up to three visits for a total value of $36.00)***.**

2.0 **IF you will be collecting personal information to reimburse or pay participants, describe the information to be collected and how privacy will be maintained.**

3.0
Will participants receive any incentives for participating in this research? Select all that apply.

There are no items to display

Provide details of the value, including the likelihood (odds) of winning for prize draws and lotteries:

4.0 **Excluding prize draws, what is the maximum value of the incentives offered to an individual throughout the research?**

5.0 **IF incentives are offered to participants, they should not be so large or attractive as to constitute coercion. Justify the value of the incentives you are offering relative to your study population.**
Participation in the study is worth 2% towards their course mark.

5.1 Research Methods and Procedures

Some research methods prompt specific ethic issues. The methods listed below have additional questions associated with them in this application. If your research does not involve any of the methods listed below, ensure that your proposed research is adequately described in Section 2.0: Study Objectives and Design or attach documents in Section 7.0 if necessary.

1.0 * **This study will involve the following** *(select all that apply)*
The list only includes categories that trigger additional page(s) for an online application. For any other methods or procedures, please indicate and describe in your research proposal in the Study Summary, or provide in an attachment:
Surveys and Questionnaires (including internet surveys)
Use of Deception or Partial Disclosure (not including double-blind)

2.0 * **Is this study a Clinical trial? (Any investigation involving participants that evaluates the effects of one or more health-related interventions on health outcomes?**
○ Yes ● **No**

3.0 **If you are using any tests in this study diagnostically, indicate the member(s) of the study team who will administer the measures/instruments:**

Test Name	Test Administrator	Organization	Administrator's Qualification
There are no items to display			

4.0 **If any test results could be interpreted diagnostically, how will these be reported back to the participants?**

5.5 Use of Deception or Partial Disclosure

1.0 * **Describe the information that will be withheld from, or the misinformation that will be provided to, the participants:**
The participants will not be informed about how the actions or the emotions shown by the virtual agents are generated.

2.0 **Provide a rationale for withholding information:**
We wish to avoid any bias.

3.0 **Indicate how and when participants will be informed of the concealment and/or deception. Describe the plans for debriefing the participants. Indicate when the participants will be debriefed, and describe the nature and extent of debriefing:**

Participants will receive a textual debriefing upon completing their participation, and this debriefing will make clear that a software system was used to control the virtual characters, as well as explain our rationale for doing so.

4.0 **Describe the procedure for giving the participants a second opportunity to consent to participate after debriefing. Explain if debriefing and re-consent are not viable:**
The participants are given 3 days from the time of debriefing to withdraw from the study if they wish to do so.

5.0 **Indicate how participants may follow-up with researchers to ask questions or obtain information about the study:**
Full contact information (name, phone number, email address) for the principal researcher will be provided to participants during their study session.

5.7 Interviews, Focus Groups, Surveys and Questionnaires

1.0 **Are any of the questions potentially of a sensitive nature?**
○ Yes ● **No**

If YES, provide details:

2.0 **If any data were released, could it reasonably place participants at risk of criminal or civil law suits?**
○ Yes ● **No**

If YES, provide the justification for including such information in the study:

3.0 **Will you be using audio/video recording equipment and/or other capture of sound or images for the study?**
○ Yes ● **No**

If YES, provide details:

6.1 Data Collection

1.0 * **Will the researcher or study team be able to identify any of the participants at any stage of the study?**
● **Yes** ○ No

2.0 **Will participants be recruited or their data be collected from Alberta Health Services or Covenant Health or data custodian as defined in the Alberta Health Information Act?**
○ Yes ● **No**

Important: Research involving health information must be reviewed by the Health Research Ethics Board.

3.0 **Primary/raw data collected will be** *(check all that apply)***:**
Anonymous - the information **NEVER** had identifiers associated with it (eg anonymous surveys) and risk of identification of individuals is low or very low

4.0 **If this study involves secondary use of data, list all original sources:**

5.0 **In research where total anonymity and confidentiality is sought but cannot be guaranteed** *(eg. where participants talk in a group)* **how will confidentiality be achieved?**
After participation, each participant's data will be stored on Google's servers and securely in a locked lab for a minimum of 5 years.

6.2 Data Identifiers

1.0
* **Personal Identifiers:** will you be collecting - at any time during the study, including recruitment - any of the following *(check all that apply)***:**

Age at time of data collection
Other

If OTHER, please describe:
gender, prior experience playing video games (e.g. none at all, 1 hour per week, 5 hours per week, etc.), how many hours they spend on reading novels and stories, what kind of a gamer they think they are.

2.0
Will you be collecting - at any time of the study, including recruitment of participants - any of the following *(check all that apply)***:**

There are no items to display

If OTHER, please describe:

3.0 * **If you are collecting any of the above, provide a comprehensive rationale to explain why it is necessary to collect this information:**
Age, gender, and prior gaming experience are all factors that may bias a player's perception of the virtual characters. Collecting this information is necessary to control for any potential biases that may be introduced as a result.

4.0 **If identifying information will be removed at some point, when and how will this be done?**
The only potentially identifying information that will be recorded during our study is the age, gender, how many hours they spend reading novels and stories, what kind of a gamer they consider themselves and prior experience playing video games of each participant. No names or other identifying information will be recorded.

5.0 * **Specify what <u>identifiable</u> information will be RETAINED once data collection is complete, and explain why retention is necessary. Include the retention of master lists that link participant identifiers with de-identified data:**

The age, gender, how many hours they spend reading novels and stories, what kind of a gamer they consider themselves and prior gaming experience of each player will be retained along with their answers to our survey's questions. Doing so is necessary to perform a minimally biased analysis on the data that we acquire.

6.0 **If applicable, describe your plans to link the data in this study with data associated with other studies (e.g within a data repository) or with data belongong to another organization:**

6.3 Data Confidentiality and Privacy

1.0 * **How will confidentiality of the data be maintained? Describe how the identity of participants will be protected both during and after research.**
All data will be stored in a locked computer lab and/or on servers hosted by Google. These servers maybe located in the US and they may be subject to the Patriot Act.

2.0 **How will the principal investigator ensure that all study personnel are aware of their responsibilities concerning participants' privacy and the confidentiality of their information?**
As no personally identifiable data concerning our participants is retained, no special privacy training is necessary to have access to the study's data.

3.0

External Data Access

* **3.1 Will identifiable data be transferred or made available to persons or agencies outside the research team?**
○ Yes ● **No**

3.2 If YES, describe in detail what identifiable information will be released, to whom, why they need access, and under what conditions? What safeguards will be used to protect the identity of subjects and the privacy of their data.

3.3 Provide details if identifiable data will be leaving the institution, province, or country *(eg. member of research team is located in another institution or country, etc.)*

6.4 Data Storage, Retention, and Disposal

1.0 * **Describe how research data will be stored, e.g. digital files, hard copies, audio recordings, other. Specify the physical location and how it will be secured to protect confidentiality and privacy.** *(For example, study documents must be kept in a locked filing cabinet and computer files are encrypted, etc. Write N/A if not applicable to your research)*
The data will be collected using Google forms and stored on Google's servers which maybe hosted outside of Canada. All physical data will be stored in a locked computer lab managed by the Department of Computing Science in Athabasca Hall and the Computing Science Centre.

2.0 * **University policy requires that you keep your data for a minimum of 5 years following completion of the study but there is no limit on data retention. Specify any plans for future use of the data. If the data will become part of a data repository or if this study involves the creation of a research database or registry for future research use, please provide details.** *(Write N/A if not applicable to your research)*
The data will be retained for the purposes of replicating and verifying our results.

3.0 **If you plan to destroy your data, describe when and how this will be done? Indicate your plans for the destruction of the identifiers at the earliest opportunity consistent with the conduct of the research and/or clinical needs:**
No identifying information will be recorded, and all survey data will be stored securely for at least five years.

7.1 Documentation

Add documents in this section according to the headers. Use Item 11.0 "Other Documents" for any material not specifically mentioned below.

Sample templates are available in the REMO Home Page in the **Forms and Templates**, or by clicking HERE.

1.0 **Recruitment Materials:**

Document Name	Version	Date	Description
There are no items to display			

2.0 **Letter of Initial Contact:**

Document Name	Version	Date	Description
There are no items to display			

3.0 **Informed Consent / Information Document(s):**

3.1 What is the reading level of the Informed Consent Form(s):

3.2 Informed Consent Form(s)/Information Document(s):

Document Name	Version	Date	Description
Briefing and Consent Form | History	0.06	1/28/2015 1:55 PM	

4.0 **Assent Forms:**

Document Name	Version	Date	Description
There are no items to display			

5.0 **Questionnaires, Cover Letters, Surveys, Tests, Interview Scripts, etc.:**

Document Name	Version	Date	Description
Debriefing | History	0.07	1/28/2015 1:55 PM	
Online Survey | History	0.06	11/24/2014 8:49 AM	

6.0 **Protocol:**

Document Name	Version	Date	Description
There are no items to display			

7.0 **Investigator Brochures/Product Monographs** *(Clinical Applications only)*:

Document Name	Version	Date	Description
There are no items to display			

8.0 **Health Canada No Objection Letter** *(NOL)*:

Document Name	Version	Date	Description
There are no items to display			

9.0 **Confidentiality Agreement:**

Document Name	Version	Date	Description
There are no items to display			

10.0 **Conflict of Interest:**

Document Name	Version	Date	Description
There are no items to display			

11.0 **Other Documents:**
For example, Study Budget, Course Outline, or other documents not mentioned above

Document Name	Version	Date	Description
There are no items to display			

Final Page

You have completed your ethics application! Please select "Exit" to go to your study workspace.

This action will NOT SUBMIT the application for review.

Only the Study Investigator can submit an application to the REB by selecting the "SUBMIT STUDY" button in My Activities for this Study ID: Pro00051822.

You may track the ongoing status of this application via the study workspace.

Please contact the REB Coordinator with any questions or concerns.

A.5 Alternate Assignment

<u>Alternate Assignment: Judgements of Control</u>

Instructions

Read the research paper "*How Do We Judge Personal Control? Unconfounding Contingency and Reinforcement in Contr Judgments*" by S.C. Thompson et al., and answer the questions below.

Definitions

1) Heuristic:

2) Contingent Situation:

3) Reinforcement:

Short Answer Questions

1) What are two factors that can influence an individual's perception of control?

2) In Alloy and Abramson's llight-onset studies, is reinforcement treated as a dependent variable, or an independent variab

3) The authors describe two ways in which people seem to be prone to judging contingency inaccurately. What are they?

4) According to Arkes & Harkness's findings in1983, what often forms the basis of an individual's judgement of contingency between antecedent and consequent variables?

5) The participants in the authors' experiment ranged from 17 to 26 years of age, with the average age being 18.8 years, a the standard deviation being 1.2 years. What does this information suggest about how many 26-year-olds participated?

6) List six steps of the procedure that the authors used to administer their experiment.

Long Answer Questions

1) In your own words, explain the flaw that the authors identify in Alloy and Abramson's Experiment 1 from 1979. How doe the authors' presented work address the flaw?

2) Considering experimental design described by Table 1, explain why pressing the space bar makes the target appear 2 more often in the High Control (50%), High Reinforcement (70%) condition than it does in the No Control (0%), H Reinforcement (70%) condition. Why not 27% more often, or 23%?

3) Describe the authors' hypotheses in your own words.

4) The authors state that "[it] appears that under certain conditions of contingency, illusions of control are found." Wh "certain conditions" are they referring to?

Essay Questions

1) Are Thompson et al.’s findings extensible to the context of video games? Why, or why not?

Essay Questions (continued)

2) Imagine that you're a designer of a commercial video game whose primary focus is to give its players a sense of havi influence over a story. You've already done some development, and you've created enough content to tell a nice, lir story. As nice as it is, however, it doesn't actually provide the player with any opportunity to control how it turns out; ev player gets the same story, no matter what they do. Providing opportunities for players to control stories is expensi because each different thing that can happen in the game requires extra effort to create. Given that people seem overestimate their control even in no-control situations (see Figure 1), would you expect your game to be any better if y did the extra work, and gave players actual opportunities to control the course of the story? Why, or why not?

Appendix B

Actor Faces

In this appendix, we include the expressions shown by the actors for this user study. Each actor's neutral expression and the four emotions we use are displayed.

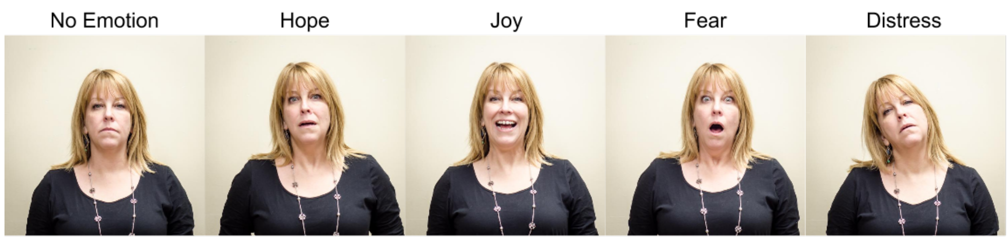


Figure B.1: Cathleen.

Figure B.2: Chris.

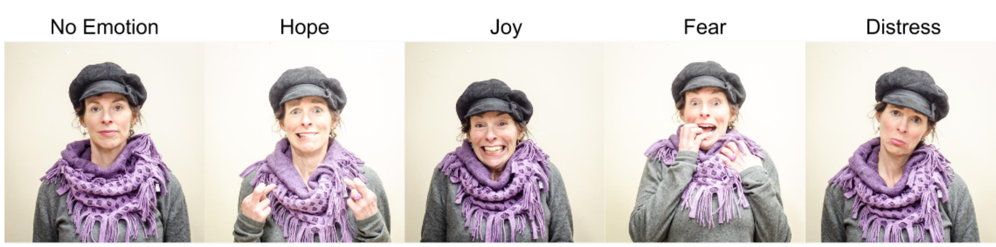


Figure B.3: Davina.

Figure B.4: Delia.

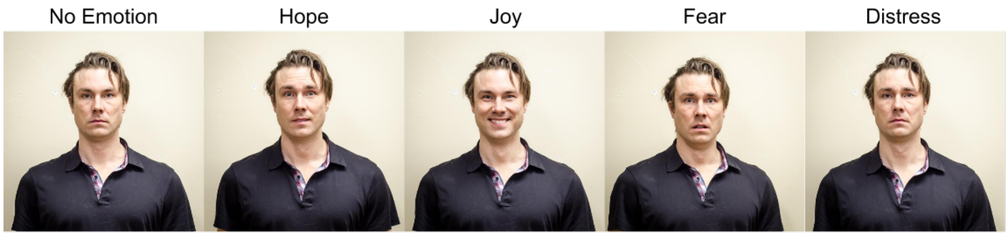


Figure B.5: Jesse.

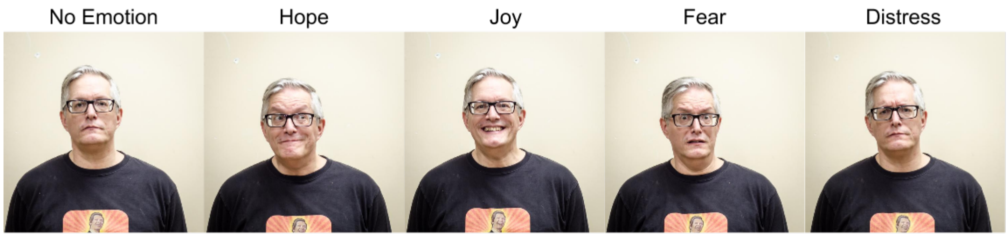


Figure B.6: Jim.

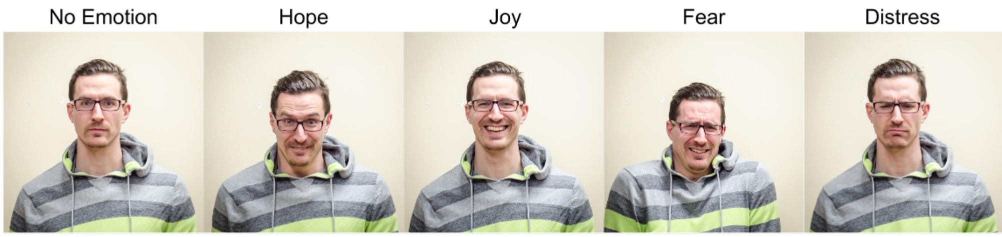


Figure B.7: Josh.

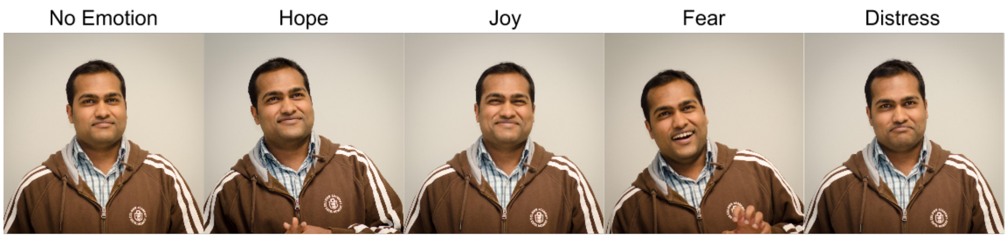


Figure B.8: Karan.

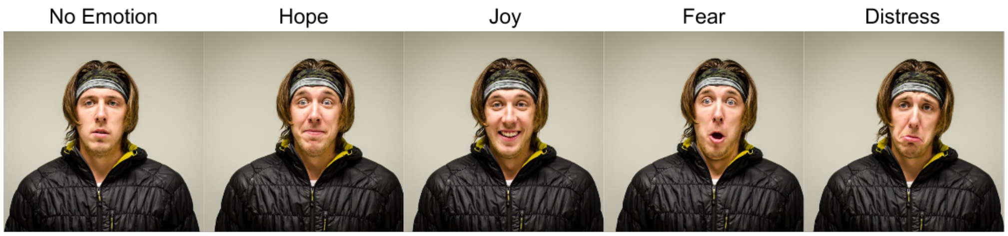


Figure B.9: Kory.

Figure B.10: Leona.

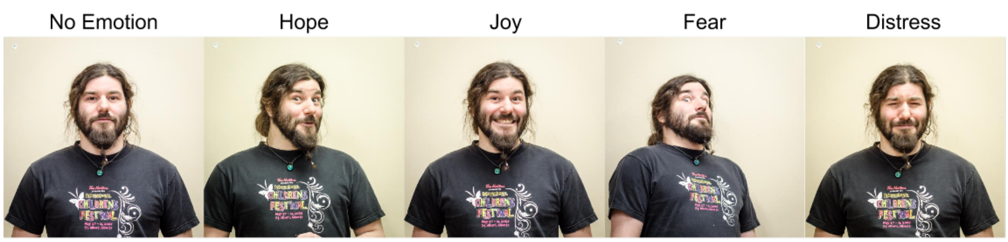


Figure B.11: Matt.

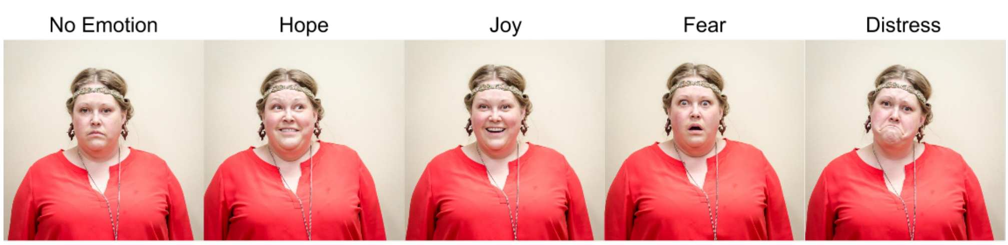


Figure B.12: Nicole.

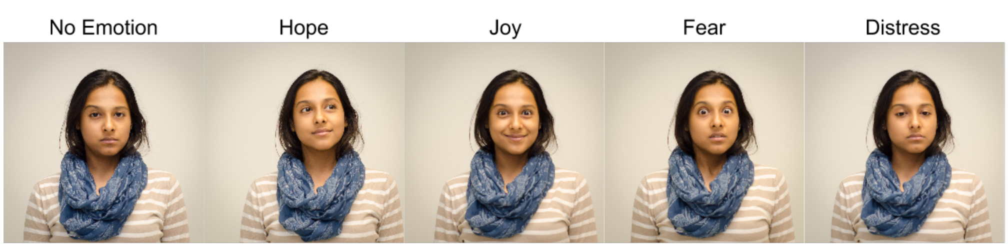


Figure B.13: Nitya.

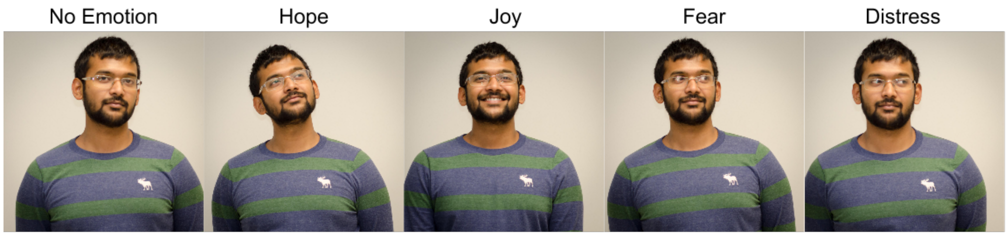


Figure B.14: Pankaj.

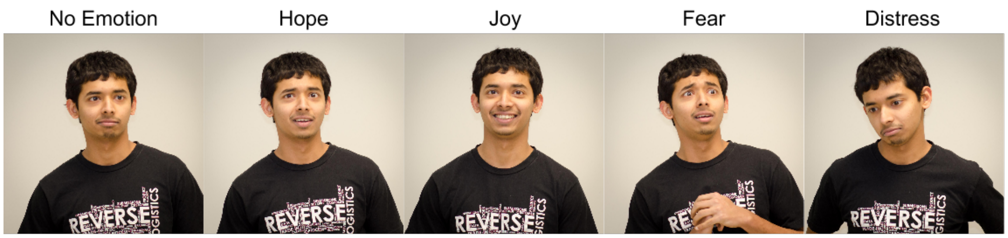


Figure B.15: Rohit.

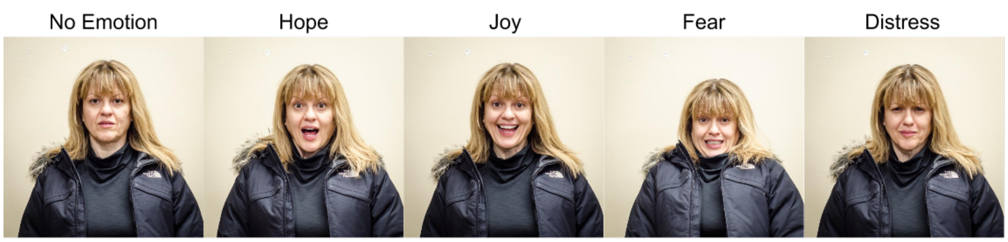


Figure B.16: Stephanie.

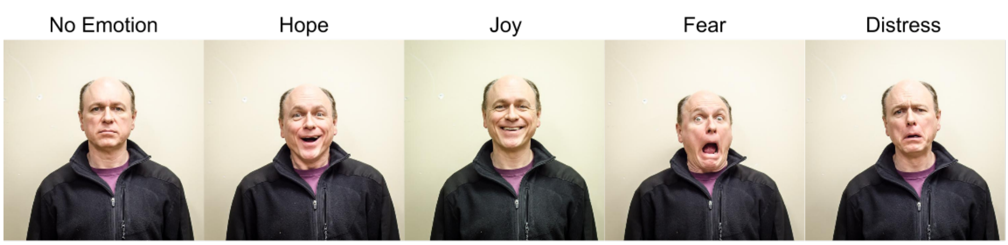


Figure B.17: Tom.

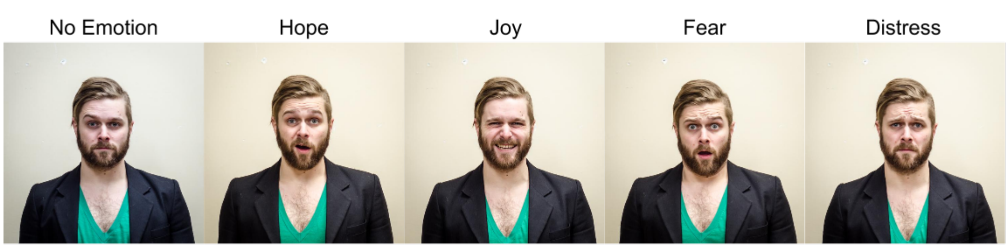


Figure B.18: Vince.