An Affective Student Model To Assess Emotions
In An Educational Game
by
Xiaoming Zhou
B.A., Nanjing University, 1999

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
Master of Science
in
THE FACULTY OF GRADUATE STUDIES
(Department of Computer Science)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 2002
© Xiaoming Zhou, 2002
In presenting this thesis in partial fulfilment of the requirements for an advanced
degree at the University of British Columbia, I agree that the Library shall make it
freely available for reference and study. I further agree that permission for extensive
copying of this thesis for scholarly purposes may be granted by the head of my
department or by his or her representatives. It is understood that copying or
publication of this thesis for financial gain shall not be allowed without my written
permission.

Department of Computer Science
The University of British Columbia
Vancouver, Canada

Date 13/12/2002
Electronic educational games integrate the target domain knowledge in a game-like environment in order to help students learn. While, in general, these educational games are more engaging than the traditional computer-based educational software, they often do not necessarily trigger learning. One explanation is that many students play the games without actively reasoning about the underlying domain knowledge. To make learning more effective in educational games, we are designing intelligent pedagogical agents that can provide tailored interventions to students. However, in order not to compromise the high level of engagement that is the main advantage of educational games, it is important for these agents to consider students' emotional states in addition to their cognitive states (such as learning) to decide when and how to provide interventions.

This thesis focuses on the creation of an affective student model that assesses the students' emotional states while they are playing an educational game Prime Climb. The affective student model explicitly represents the cognitive appraisal process of emotions by implementing the cognitive theory of emotions (OCC Model). It relies on Dynamic Decision Networks (DDNs) to deal with the high level of uncertainty involved in affective user modeling.

The initial version of the model was built based on the observation of two preliminary Prime Climb user studies, our intuitions, and psychological findings. This model was then revised based on the results of a third Prime Climb study.
CONTENTS

ABSTRACT .................................................................................................................. ii

CONTENTS ................................................................................................................ iii

LIST OF TABLES ......................................................................................................... vii

LIST OF FIGURES ...................................................................................................... ix

ACKNOWLEDGEMENTS ............................................................................................. x

DEDICATION ............................................................................................................... xi

CHAPTER 1 INTRODUCTION ....................................................................................... 1
  1.1 Emotionally Intelligent Computers ................................................................. 2
  1.2 The Role of Emotion in Intelligent Learning Environments ......................... 3
    1.2.1 Adapting the System’s Behaviors by Considering Emotions ............... 3
    1.2.2 Emotion in the Design of Life-like Pedagogical Agents ..................... 4
  1.3 Affective User Modeling in Electronic Educational Games ......................... 6
  1.4 Difficulties in Affective User Modeling ........................................................... 8
  1.5 The Approach Used in this Thesis ................................................................ 9
  1.6 Thesis Outline ................................................................................................. 10

CHAPTER 2 BACKGROUND ....................................................................................... 12
  2.1 The OCC Model .............................................................................................. 12
  2.2 Dynamic Decision Networks (DDNs) ............................................................. 14

CHAPTER 3 RELATED WORK .................................................................................... 19
  3.1 Three Dimensions of User Assessment .......................................................... 19
  3.2 A Conversational Agent that Responds to Users’ Emotions ......................... 19
  3.3 Interface Adaptation by Assessing Users’ Affective and Belief States ........ 20
  3.4 The Affective Reasoner Framework ............................................................... 21
  3.5 Detecting Emotions from Physiological Signals ............................................ 22
  3.6 Motivation Diagnosis .................................................................................... 23
3.7 Probabilistic User Modeling

CHAPTER 4 PRIME CLIMB: AN ELECTRONIC EDUCATIONAL GAME

4.1 The Prime Climb Interface
4.2 The Pedagogical Agent

CHAPTER 5 THE AFFECTIVE MODEL: INITIAL VERSION

5.1 Uncertainty in Affective User Modeling For Prime Climb
5.2 Two Preliminary Prime Climb Studies
   5.2.1 The Study Goal
   5.2.2 The Study Participants
   5.2.3 The Study Design
5.3 The Results of the Preliminary Studies
   5.3.1 Students’ Goals in Playing the Game
   5.3.2 Interaction Patterns and Their Relationships to Goals
   5.3.3 The Emotion Eliciting Situations
5.4 High Level Description of The Initial Affective Model
5.5 Details on the Model Design
   5.5.1 Portion of the Network Assessing Students’ Goals
   5.5.2 Portion of the Network Assessing Students’ Emotions
   5.5.3 Conditional Probability Tables
5.6 Limitations of the Preliminary Studies
5.7 Limitations of the Initial Model

CHAPTER 6 THE AFFECTIVE MODEL: REVISED VERSION

6.1 The Third Prime Climb Study
   6.1.1 The Study Goal
   6.1.2 The Study Participants
   6.1.3 The Study Design
6.2 The Study Results and Interpretation
   6.2.1 Relationships between Students’ Personalities and Goals
APPENDIX B  PRE-TEST (FIRST VERSION).................................102

APPENDIX C  POST-QUESTIONNAIRE (FIRST VERSION)......................103

APPENDIX D  OBSERVATION SHEET........................................105

APPENDIX E  PRE-TEST (SECOND VERSION)................................107

APPENDIX F  POST-QUESTIONNAIRE (SECOND VERSION)......................109

APPENDIX G  FULL CROSS TABLE FOR INTERACTION PATTERNS..............111
Table 5.1  Number of Students Who Chose Each Goal in the Post-Questionnaire ......37
Table 5.2  CPT for Avoid Falling Given Neuroticism .................................................46
Table 5.3  CPT for Move Quickly Given Have Fun, Avoid Falling and Beat Partner ....46
Table 5.4  CPT for Quick Move Given Move Quickly ................................................47
Table 5.5  CPT for Succeed By Myself Satisfied When Agent Intervenes ..............47
Table 5.6  CPT for Emotion For Agent When Agent Intervenes ................................48
Table 6.1  Information Recorded in the Log Files ..........................................................53
Table 6.2  Correlations between Personality Traits and the Goals ...............................55
Table 6.3  Contingencies between Goals and Interaction Patterns with Fisher’s Score < 0.4 ...............................................................................................................................62
Table 6.4  Cross Table for Move Quickly and Beat Partner ..........................................62
Table 6.5  Cross Table for Move Quickly and Succeed By Myself ................................63
Table 6.6  Cross Table for Follow Advice and Succeed By Myself ...............................64
Table 6.7  Cross Table for Follow Advice and Avoid Falling .........................................64
Table 6.8  Cross Table for Ask Advice and Succeed By Myself .....................................65
Table 6.9  Cross Table for Use Magnifying Glass Often and Learn Math .....................65
Table 6.10 Cross Table for Use Magnifying Glass Often and Beat Partner ....................65
Table 6.11 Cross Table for Fall Often and Have Fun ....................................................66
Table 6.12 Cross Table for Fall Often and Learn Math ..................................................66
Table 6.13 Cross Table for Fall Often and Succeed By Myself .......................................67
Table 6.14 Cross Table for Avoid Falling and Neuroticism ..........................................72
Table 6.15 CPT for Avoid Falling Given Neuroticism .................................................72
Table 6.16 CPT for Ask Advice Given Succeed By Myself ..........................................74
Table 6.17 CPT for Quick Move Given Move Quickly ................................................75
Table G.1 Full Cross Table for Move Quickly ............................................................111
Table G.2 Full Cross Table for Follow Advice .............................................................111
Table G.3 Full Cross Table for Ask Advice .................................................................111
Table G.4 Full Cross Table for Use Magnifying Glass Often ........................................111
Table G.5 Full Cross Table for Fall Often
<table>
<thead>
<tr>
<th>Figure Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The OCC Cognitive Structure of Emotions (Ortony, Clore and Collins)</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>A Generic Structure of a Two-Slice Fragment of DDN</td>
<td>15</td>
</tr>
<tr>
<td>4.1</td>
<td>The Prime Climb Interface</td>
<td>28</td>
</tr>
<tr>
<td>4.2</td>
<td>Magnifying Glass</td>
<td>29</td>
</tr>
<tr>
<td>4.3</td>
<td>Help Box</td>
<td>29</td>
</tr>
<tr>
<td>4.4</td>
<td>Wizard of OZ Interface</td>
<td>31</td>
</tr>
<tr>
<td>5.1</td>
<td>A High Level Description of the Affective Model</td>
<td>38</td>
</tr>
<tr>
<td>5.2</td>
<td>Part of the Network Assessing Students’ Goals</td>
<td>40</td>
</tr>
<tr>
<td>5.3</td>
<td>Sample Nodes Involved in Assessing Students’ Emotions</td>
<td>43</td>
</tr>
<tr>
<td>6.1</td>
<td>Eight Possible Structures for Agreeableness, Conscientiousness, Extraversion, Have Fun, and Succeed By Myself</td>
<td>69</td>
</tr>
<tr>
<td>6.2</td>
<td>Cross Validation of the Eight Structures</td>
<td>70</td>
</tr>
<tr>
<td>6.3</td>
<td>Three Possible Structures for the Personality Traits and Goals, Structure 1</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>is the Final Selected Structure</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Cross Validation of Three Structures for Personality and Goals</td>
<td>70</td>
</tr>
<tr>
<td>6.5</td>
<td>The Revised Sub-Model to Assess Students’ Goals</td>
<td>73</td>
</tr>
<tr>
<td>6.6</td>
<td>Cross Validation of the Three Sub-Models Assessing Goals</td>
<td>73</td>
</tr>
<tr>
<td>6.7</td>
<td>Model the Goals as Dynamic Nodes</td>
<td>77</td>
</tr>
<tr>
<td>6.8</td>
<td>A Static Network Assessing Students’ Goals, after 3 Actions</td>
<td>77</td>
</tr>
<tr>
<td>6.9</td>
<td>A Dynamic Network Assessing Students’ Emotions</td>
<td>78</td>
</tr>
<tr>
<td>6.10</td>
<td>Time Spent to Make the Inference on the Goals</td>
<td>79</td>
</tr>
<tr>
<td>7.1</td>
<td>Assessment Without Personality Information</td>
<td>82</td>
</tr>
<tr>
<td>7.2</td>
<td>Assessment with Personality Information</td>
<td>84</td>
</tr>
</tbody>
</table>
First, I would like to thank my supervisor Dr. Cristina Conati for all of her valuable guidance. Her attitude toward research has inspired me to continuously challenge myself to reach new levels. I would also like to thank Dr. David Poole for being my second reader and providing me with many insightful comments.

Working with different people in the LCI lab at UBC was a great experience. In particular, I greatly appreciate those people in the lab who gave me enormous support and encouragement. They are Dr. Alan Mackworth, Valerie McRae, Andrea Bunt, Kasia Muldner, Fengguang Song, and Xiaohong Zhao.

I am deeply grateful to my friend Dr. Mircea Munteanu for his very helpful advice on the thesis, and his continuous encouragement and care during my life in Canada. I also like to thank my girlfriend Xin Wang who is always there to give me support and care and help me get through all the difficulties. I also thank my sister Yehong Zhou and my cousin Chun Zhu, for their always support. Finally, I would like to thank my parents, Jianhua Mao and Minggao Zhou for their great love and care in my life. This thesis would not be possible without their encouragement and calming words during my stressful periods of time.

Xiaoming Zhou
The University of British Columbia
October, 2002
To my mom and dad
Building intelligent machines, which behave logically and rationally, has long been the focus of interest in the Artificial Intelligence (AI) community. However, in recent years, scientists have begun to realize the importance of emotions in attention, planning, learning, memory, and decision-making, based on evidence from neuroscience, cognitive science, and psychology [65][68]. Thus, some researchers in the AI community have begun to explore the issue of building emotionally intelligent computers, “computers which are able to recognize and express emotions, respond intelligently to human emotions, and regulate and utilize its emotions” [56].

One of the challenging problems in building emotionally intelligent computers is recognizing users’ emotional states. Humans use different sources of information to assess a person’s emotional states. Such sources of information can be knowledge about that person’s background, personality traits, contextual information, and bodily expressions. However, the available information is often incomplete and even contradictory, which makes the assessment of emotions an uncertain task. The task of giving a computer the ability to recognize users’ emotions inevitably faces the same problem.

This thesis proposes a general framework for affective student modeling. This framework is based on an existing cognitive theory of emotions, and uses Dynamic Decision Networks (DDNs) to combine different sources of information in order to deal with the high level of uncertainty involved in the modeling task. The framework is applied to build an affective student model that is to be used by intelligent pedagogical agents in an electronic educational game.
1.1 Emotionally Intelligent Computers

Giving computers the ability to act logically, rationally, and predictably has been the pursuit of many computer scientists for a long time. Because emotions have been regarded as non-scientific and impairing the decision-making process, for a long time they have been excluded from the properties that scientists were giving to intelligent computers. However, recently researchers have begun to argue that emotions play too strong a role in motivation, cognition, coping, and creativity, to be ignored. Therefore, researchers in the AI community have started investigating the issue of building emotionally intelligent computers [4][19][20][32]. According to Picard [56], an emotionally intelligent computer should have the ability to recognize and express emotions, as well as the ability to regulate and utilize emotions.

In addition to recognizing its own emotions and being aware of how to express them, an emotionally intelligent computer should be able to assess others' emotional states by both recognizing their bodily expressions and reasoning about the given emotion-eliciting situation. The ability to recognize and express emotions is especially crucial in human-computer communication. Recognizing user emotions can help the computer better understand a user's cognitive states, such as goals and preferences, and can also give the computer information for adapting its behaviors. By expressing its own emotions, the computer can convey the affective information to the users, which is an important part in human-human communication.

Regulating emotions is also an important aspect of emotional intelligence. Once users' emotions are recognized, the computer should be able to respond appropriately in ways that positively influence those emotions, for instance, by providing emotional support to someone who is feeling frustrated in order to restore his or her confidence.

This thesis focuses on the task of providing the computer with the ability to recognize users' emotions.
1.2 The Role of Emotion in Intelligent Learning Environments

Intelligent Learning Environments (ILEs) are computer-based learning environments, which can adapt to individual learners through the use of artificial intelligence. According to Sleeman and Brown [63], an intelligent learning environment must have three components: knowledge of the target domain, knowledge of the learner, and knowledge of teaching strategies. Domain knowledge refers to the topic or curriculum being taught. Knowledge of the learner includes information such as the learner’s personality traits, level of expertise, cognitive states, affective states, and preferences. Teaching strategies are the methods of instruction and the way in which the material is presented. The goal for every ILE is to communicate its embedded domain knowledge effectively.

Two major trends have been seen in applying emotional intelligence to intelligent learning environments. One trend is to adapt the system’s behaviors by considering the users’ emotional states, for example, adapting its interface layout, the type of information displayed, and its tutorial actions [16][32][51]. The other trend is to enrich the system with life-like pedagogical agents that have emotional intelligence [20]. By generating and expressing their own emotions, these agents give the impression of having their own desires, interests, and personalities. By recognizing and responding appropriately to the users’ emotions, these agents show a human-like quality, caring about the feelings of others.

1.2.1 Adapting System’s Behaviors by Considering Emotions

When adapting the system’s behaviors, researchers in ILEs have paid attention mostly to users’ cognitive aspects, such as the users’ domain knowledge, their goals, beliefs, and preferences [11][14][48]. Less attention has been paid to affective or motivational states. Consequently, very few systems have been found to adapt their behaviors based upon the users’ affective or motivational states. However, motivation has long been recognized to be an important factor in learning. Furthermore, emotions play a key role in motivation. Expert human tutors pay at least the same amount of attention to the
students' affective and motivational states, as they do to the achievement of the students' cognitive goals [42]. In order to achieve the ultimate goal of ILEs, that is, to make learning more effective, researchers have started investigating the issue of adapting the system's behaviors by considering the users' affective states in addition to their cognitive states [16][32][51]. The Affect and Belief Adaptive Interface System (ABAIS) described in [32] is a training system that teaches pilots to perform combat operations. The system reacts to the users' anxiety level in addition to their performance-relevant beliefs by varying the interface layout and the type of information displayed on the cockpit instruments (e.g., radar display, or head up display). The Decision-Theoretic Tutor (DT Tutor) presented by Murray and Vanlehn [51] selects the optimal tutorial actions by considering the students' morale in addition to domain knowledge, focus of attention, and so forth. The Affective Dialogue (AFDI) [16] uses the students' motivational states, such as satisfaction, confidence, and interest, to guide the tutor's dialogue with the students.

1.2.2 Emotion in the Design of Life-like Pedagogical Agents

Life-like pedagogical agents have been the subject of increasing attention in ILEs over the past few years. Pedagogical agents with believable, life-like qualities can increase the bandwidth of communication channels between students and computers. By demonstrating complex tasks, using gestures to direct students' attentions, and expressing emotional responses to tutorial situations, the agents can make learning environments more engaging, and thus promote learning [36]. However, some critics argue that life-like pedagogical agents can distract students' attention, and therefore, prevent students from learning. As research in life-like pedagogical agents is still in its infancy, few empirical studies have been conducted so far on the effectiveness of life-like pedagogical agents in learning environments. Lester, Converse, Stone, Kahler, and Barlow [43] reported two encouraging results from a formal empirical study with the pedagogical agent "Herman the Bug" in the Design_A_Plant learning environment. In that environment, the students were instructed to design plants that could thrive in a given environment. The agent observed the students' performance and provided
explanations and hints to them. The authors found a statistically significant increase from pre-test scores to post-test scores for the students interacting with the learning environment embedded with the animated pedagogical agent. Another finding was that life-like pedagogical agents that provide multiple levels of advice with multiple modalities produced greater improvements in problem solving than did less expressive agents.

One way to achieve this believability is through the physical aspects of the agent, such as appearance, voice, locomotion, gaze, and gesture. It is crucial for the agent to be able to react to unexpected events and maintain coherent behavior. This requires a behavior engine to coordinate the agent’s behavior. For instance, the agents in [3][44][57] all have complex behavior engines to generate appropriate and coherent actions.

The other way to construct believable agents is to make agents emotionally intelligent [5]. This includes giving agents the ability to recognize the users’ emotions and respond appropriately to them, and the ability to generate and manifest their own emotions. Research shows that “people respond to computer personalities in the same way they would respond to human personalities” [50][52]. Thus, an agent that attends to a user’s emotions can make the user feel that the agent cares about what he or she is feeling. An agent that displays its own emotions can make a user feel that the agent has its own desires and emotional life, and is not just a lifeless machine. Therefore, giving agents emotional intelligence can not only increase the believability of these agents and make learning more engaging, but also open up a new communication channel between the students and agents. For instance, an agent can have a puzzled look to provide disapproving feedback to a student who just made a mistake. This channel can be useful in situations in which more direct agent interventions could interrupt the students’ problem solving.

Research on applying emotional intelligence to life-like pedagogical agents has just begun. Very few systems have been developed to date. Elliott, Lester, and Rickel [20] discuss the potential to use the Affective Reasoner [19] in building life-like pedagogical agents that are sensitive to students’ emotional states. The Affective
Reasoner is a rule-based framework for building agents that can generate and manifest their own emotions, which we discuss in detail in a later section.

1.3 Affective User Modeling in Electronic Educational Games

Some researchers have proposed to motivate students by using multimedia and electronic games as educational tools [46][62]. Electronic educational games integrate the target domain knowledge in a game-like environment in order to help students learn. The underlying hypothesis is that students will generally spend more time on the game and will be motivated by the game-like interactions, thus acquiring the target domain knowledge. However, empirical studies have shown that educational games do not necessarily trigger the learning as designers expected [13][40].

While, in general, electronic educational games are more engaging than the traditional computer-based educational software, they tend to generate different levels of motivation in different students. Some students with a low level of domain knowledge may find them too difficult to play, and thus feel extremely frustrated and lose motivation. In contrast, students with relatively high knowledge may find the games too easy and may lose their interest if there is no adaptation. Therefore, not all students will be motivated through the interactions with educational game environments. Furthermore, each student might have different emotional states and different levels of motivation during different stages of interaction with the game. According to [41], a learning experience in science, math, engineering, and technology usually involves a series of emotions. For instance, in the beginning, a typical student might be curious or fascinated with a new topic; then the student might be confused or disappointed when faced with failures; this may lead to frustration or hopelessness; eventually as the student consolidates his or her knowledge in the domain, the emotions might turn into hope or confidence. This limitation can be overcome by developing a student model in educational games, based on which we can track the students’ affective or motivational states and adapt the game environments accordingly in order to keep students at a high motivational level.
However, even for those students who are highly engaged in game environments, learning does not always occur. A possible explanation for this may be the fact that educational games are open learning environments which require meta-cognitive skills (i.e., skills related to how to reason and learn, which are independent of the underlying domain), such as self-explanation (generating explanation to oneself about the observations) and self-monitoring (monitoring one's own progress and understanding) [9][45][61][64]. Self-explanations are required in order to generalize the results of experimentations. The ability to self-monitor is also important for identifying one's defects in knowledge and thus performing the necessary explorations. Therefore, students who do not possess these skills will not be able to learn effectively from educational games. Solving this problem requires additional scaffolding, for example, appropriate interventions from a pedagogical agent to increase students' constructive reasoning and reflection. However, in order not to compromise the engaging aspect of educational games, which is a key factor that makes these games appealing for students, it is important to consider the students' emotional states when providing such scaffolding.

To make learning more effective in educational games, we are designing emotionally intelligent pedagogical agents that can provide tailored interventions to students. By taking into account the students' emotional states in addition to their cognitive states, we hypothesize that these pedagogical agents can, on the one hand, give the necessary scaffolding for students to explore the environment more effectively, and on the other hand, keep students engaged and entertained. Several problems need to be solved in order to build the emotionally intelligent pedagogical agents. One is to effectively assess the students' domain knowledge level, which requires a clear understanding of the underlying domain knowledge structure, as well as a complete analysis of the students' behaviors and their relationships to the knowledge structure. A second problem is to recognize the students' emotional states while they are playing the game. A third problem is to develop appropriate pedagogical strategies for the agents to decide when and how to act based on the information about the student's knowledge level and affective states.
The focus of this thesis is on recognizing the students' emotional states while they are playing an educational game. We introduce a probabilistic affective model, which the pedagogical agents can use along with the diagnosis of the students' domain knowledge, in order to generate tailored interventions that can increase learning and keep the students engaged.

1.4 Difficulties in Affective User Modeling

"Recognizing emotions is to infer an emotional state from observations of emotional expressions and behavior, and through the reasoning about an emotion-generating situation" [56]. This definition recognizes two aspects of emotions: physical and cognitive, which provide, respectively, symptomatic and causal information to the emotion recognition process.

Research on the physical aspects of emotions focuses on physical responses that are triggered by emotions and tries to match emotions to their physical expressions. Bodily expressions such as facial expressions, vocal intonation, and posture are visible physical means by which an emotional state is typically expressed [18]. However, the match between emotions and visible physical expressions is often complicated by a number of factors, such as the type of the emotion, the intensity of the emotion, the personality traits of the person who expresses the emotion, and the context in which the emotion arises. For example, while an extrovert usually tends to exaggerate the expression, an introvert might tend to control his or her display of emotions, especially if faced with unfamiliar people. Therefore, while emotions are often reflected in visible physical expressions, these expressions are highly user and context dependent, and not always discriminating enough to allow a precise diagnosis of the specific underlying emotion. There are also some unobservable physiological responses, such as changes in blood pressure, heart rate, and skin conductivity, which can be affected by emotional states [56]. For example, skin conductivity is a very good measure of a person’s overall level of arousal (intensity of emotion) while giving little information on the valence (whether the emotion is positive or negative) [56]. Negative emotions increase blood pressure more than positive emotions [8]. However, these physiological signals can be
very difficult to assess precisely. Furthermore, a single biometric measure usually provides little information to allow the assessment of specific emotions.

Research on the cognitive aspects of emotions concentrates on the understanding of the situations and the cognitive factors that cause emotions. The same person can have different emotions in different contexts, and different people can have different emotional responses in the same context. Several appraisal theories have been developed to explain how emotions are cognitively generated [54][58]. One of the theories is the so-called “OCC Model”, developed by Ortony, Clore and Collins [54]. According to this theory, emotions come from the appraisal of the current situation. The outcome of the appraisal depends upon how the current situation fits with one’s goals and preferences. Situations are specified by three categories of factors: consequences of events, actions of agents, and aspects of objects. Each category of factors can generate a different set of emotions. For instance, in the category of actions of agents, the emotions include: pride or shame (if the emotion eliciting situation is caused by one’s own action), and admiration or reproach (if the situation is caused by another agent). The OCC theory defines twenty-two emotion types that result from situation appraisal. However, it is often difficult to obtain information on the users’ goals and preferences to predict their emotions by applying the cognitive appraisal theory. For instance, in educational game environments, students can have different goals when they are playing the games. Goals are often influenced by one’s personality traits [15]. However, assessing these traits is not trivial, and the relationships between students’ personality traits and their goals are always uncertain.

The factors mentioned above complicate emotion recognition and make it a task permeated with uncertainty.

1.5 The Approach Used in this Thesis

To deal with the high level of uncertainty involved in the assessment of students’ emotional states during the interactions with an educational game, we propose an affective student model that relies on Dynamic Decision Networks (DDNs) to explicitly
represent the probabilistic relationships in the domain, and to leverage any evidence available to perform the probabilistic inferences.

Unlike other approaches that assess affective states from the effects of emotions, such as physiological signals, vocal intonation, and facial expressions [4][28][37], the approach used in this thesis explicitly models the cognitive appraisal process of emotions by implementing the OCC cognitive theory of emotion. Thus, our affective model can assess not only students' affective states, but also the causes of emotions.

The affective model is implemented in an electronic educational game called Prime Climb, developed by the Electronic Games for Education in Math and Science (EGEMS) group at the University of British Columbia. The model is a part of an intelligent pedagogical agent we are developing for Prime Climb, and should be used in conjunction with a model of student learning to help the agent decide when and what kind of intervention should be provided in order to help students learn effectively in the game, while maintaining a high level of engagement.

The advantages of the affective model we propose are as follows: (1) by using DDNs, the model explicitly represents the probabilistic relationships between situations, students' traits, and emotions, and is able to integrate any evidence available to make the probabilistic assessment; (2) the use of DDNs provides a theoretic basis for the decisions guiding the pedagogical agent's behavior; (3) the model is able to explain the causes of emotions by implementing the cognitive appraisal theory of emotions (OCC Model); and (4) the model can be easily extended to incorporate new evidence, for example, input from physiological signals, vocal intonation, or facial expressions. Thus, it can be modified to combine both causes and effects of emotions to perform both predictive and diagnostic assessment, as suggested by Conati [10].

### 1.6 Thesis Outline

Chapter 2 provides background information on the cognitive theory of emotions (OCC Model) and on Dynamic Decision Networks. Chapter 3 reviews the research literature on affective user modeling and probabilistic user modeling. Chapter 4 describes the Prime Climb educational game, including the interface and the pedagogical agent.
Chapter 5 presents the initial findings from two preliminary studies with Prime Climb, which served as the basis for the initial version of the affective student model. Chapter 6 provides the results from a third Prime Climb study, and describes in detail the revised affective student model based on those results. Chapter 7 includes a sample assessment based on the revised affective model. Finally, Chapter 8 addresses the contributions and limitations of the thesis and includes directions for future research.
The affective student model that is the object of this thesis explicitly represents the cognitive appraisal process of emotions by implementing the cognitive theory of emotions known as the OCC Model [54]. The model relies on Dynamic Decision Networks (DDNs) [59] to deal with the high level of uncertainty involved in the task. This chapter provides background information on the OCC Model and DDNs.

2.1 The OCC Model

Several emotion theories have been developed to date, which try to explain how emotions are cognitively generated [54][58]. Among them, the OCC model, which was developed by Ortony, Clore and Collins [54], has been the focus of attention in the AI community because of its clear structure and the relative easiness to implement in a computational model.

Instead of using sets of basic emotions or a dimensioned space (arousal and valence), the OCC model provides a classification scheme for common emotion labels based on one’s valenced (positive or negative) reaction to situations in light of one’s goals and preferences. According to this model, emotions come from the cognitive appraisal of the current situation consisting of events, agents, and objects. The outcome of the appraisal depends on how the situation fits with one’s goals and preferences. Based on this structure, the authors outlined 22 emotion labels (see Figure 2.1).

For instance, based on this model, if the current event fits with one’s goal, that person will feel joy toward the event; if the current satisfying event is caused by a third-party agent, that person will feel admiration toward the agent; if that agent is oneself, that person will feel proud. Otherwise, if the current event conflicts with one’s goal, that
person will feel distressed by the event; if this conflicting event is caused by a third-party agent, the individual will feel resentful toward that agent; if that agent is oneself, that person will feel ashamed.

However, the OCC theory only models the valenced (positive or negative) reactions to situations, and does not model the intensity of the resulting emotions. The OCC theory assumes that there is only one goal activated at a given time in the cognitive appraisal process and, thus, the resulting reaction is always deterministic. However, in a real situation, a person can have multiple goals or even conflicting goals, and the goals might have different weights. For instance, a student can have two or more goals at the same time when playing an educational game: he or she can have the goal of having fun and also wanting to learn math. The OCC Model does not provide any information on how to deal with multiple goals and conflicting goals, nor how to assess these goals in the first place.

The affective model in this thesis implements the OCC model applied to the interaction with the Prime Climb educational game, by explicitly representing the situations, the students’ goals, and the appraisal outcomes during the game interactions. In addition, our model includes ways to infer goals from a student’s interaction patterns and personality traits. Six out of twenty-two emotion types are currently implemented in the affective model because of their high relevance in our target domain, an educational game with a pedagogical agent. The six emotion types are joy/distress, admiration/reproach, and pride/shame. Joy and distress are two event-based emotions, which are directly linked to the appraisal of a current event. Admiration and reproach are emotions toward the pedagogical agent who caused the event. Pride and shame are emotions toward oneself if the current event is caused by the student.
Valenced reaction to

Consequences of events
- Pleased, displeased, etc.
  Focusing on

Consequences for other
- Desirable for other
  - happy-for
  - floating for
  - resentment
  - pity
- Undesirable for other
  - fortunat-of-others

Consequences for self
- Prospects relevant
  - prospects
- Prospects irrelevant

Actions of agents
- Approving, disapproving, etc.
  Focusing on

Aspects of objects
- Approving, disapproving, etc.
  Focusing on

Self agent
- Gratification
- Remorse
- Gratitude
- Anger

Other agent
- Pride
- Admiration
- Reproach

Well-being
- Joy
- Distress
- Pride
- Admiration
- Love
- Hate

Attribution
- Gratification
- Gratitude
- Remorse
- Anger

Attraction

Figure 2.1 The OCC Cognitive Structure of Emotions (Ortony, Clore and Collins)

2.2 Dynamic Decision Networks (DDNs)

In order to deal with the high level of uncertainty involved in the assessment of the students' emotional states while playing the educational game Prime Climb, our affective model relies on Dynamic Decision Networks (DDNs) [59] to explicitly represent the probabilistic relationships between the emotional states and their causes. DDNs are an extension of Bayesian Networks [55] that can model, in addition to random variables and probabilistic dependencies among the variables, (1) the action space of an agent, represented as decision variables; (2) an agent's preferences (utility) over the possible outcomes of its actions, and (3) the temporal evolution of variables.
Figure 2.2 shows the generic structure of a DDN modeling a sequential decision problem. This DDN models the agent’s behavior over two time slices and can be used to select the agent action that maximizes the agent’s expected utility at time slice t+1, based on the current available evidence (shaded nodes in Figure 2.2).

There are three types of nodes in a DDN. Chance nodes (shown as ovals in Figure 2.2) represent random variables or probabilistic events that an agent is uncertain about. The values of a chance node correspond to all the possible states of the event that the node represents. Decision nodes (shown as rectangles in Figure 2.2) represent an agent’s available actions, with values representing all possible agent actions. The utility node (shown as a diamond in Figure 2.2) represents an agent’s utility function. The utility function defines an agent’s preferences over the possible outcomes of its actions, which are often non-deterministic. It assigns a single number to express the desirability of each outcome. A decision theoretic agent always selects an action with the maximum expected utility, where the expected utility of an action is computed assuming the following:

- An agent’s action space is $A = \{a_1, a_2, \ldots, a_n\}$
- The possible outcome space for each action $a_i$ is $S_i = \{S_{ij}, S_{i2}, \ldots, S_{im}\}$
• The associated probability that action \( a_i \) will result in state \( S_{ij} \) is \( P(S_{ij}|E, a_i) \), where \( E \) represents the currently available evidence.

• The utility for each state \( S \) is \( U(S) \). It usually assumes a reward (the agent has to also be concerned about the future values).

The expected utility of each action \( a_i \) is defined as the probability-weighted sum of the utilities of each possible outcome, as follows:

\[
EU(a_i | E) = \sum_j P(S_{ij}|E, a_i) U(S_{ij})
\]

Most links in DDNs represent probabilistic dependencies among variables. Only the links to utility nodes do not represent probabilistic relationships. They represent which factors influence the agent’s preferences. For instance, in Figure 2.2, the link from node “Agent Action\(_i\)” to node “State\(_i\)” means that “Agent Action\(_i\)” has a direct influence on “State\(_i\)”.

In DDNs, links can also be used to model the temporal evolution of the environment (see Figure 2.2, the link from node “State\(_i\)” to node “State\(_{i+1}\)”). Each chance node has an associated conditional probability table (CPT) defining the exact probabilistic dependencies between the chance node and its parent nodes. The CPT in a node with no parents represents the prior probabilities of each possible value of the variable.

By relying on algorithms that explicitly exploit the probabilistic dependencies in the network, DDNs provide a sound and often efficient way of modeling and processing the uncertainty usually involved in user modeling tasks. DDNs also have the ability to model the environment that changes over time. This property is especially useful for some user modeling tasks that try to capture the evolution of the user’s states over time, for example, domain knowledge and emotional state. In addition, DDNs enable user models to integrate any evidence available to perform different kinds of inference, that is, diagnostic inference (from effects to causes) and causal inference (from causes to effects). Finally, DDNs also provide a theoretic basis for an agent’s decisions by allowing the representation of the utility functions and providing algorithms to efficiently compute the action with maximum expected utility.
However, both Bayesian Networks and DDNs have various limitations as user modeling tools. One of the main concerns with using these networks in a modeling application is the issue of how the networks are defined. This includes both the network structure and the CPTs.

One way to define the networks is to ask an expert in the domain to hand-build the network structure, define the CPTs using the expert’s informed estimates and refine these values through empirical evaluations [30]. However, if the networks grow in size, it becomes extremely time-consuming to define the networks with the help of a domain expert. Some researchers in student modeling have designed ways to build the network at run time. For instance, in the ANDES student model [11], the structure of its Bayesian Networks is built from problem solution graphs generated by a problem solver, and the CPTs are defined automatically through predefined Nosiy-AND and Leaky-OR gates. Although this process still relies on a knowledge base of concepts and parameters defined by hand, each situation dependent network is built automatically, thus highly reducing the construction effort.

Another approach is to use machine-learning techniques to learn the networks from the data. This includes learning the structure only, learning the parameters (CPTs) for a given structure, or learning both the structure and parameters [7]. For instance, Mayo and Mitrovic [49] use machine-learning techniques to learn both the structure and CPTs from the data, whereas Albrecht [1] uses the data to learn the CPTs, given a set of candidate network structures. However, when using standard learning algorithms to learn the CPTs of a Bayesian Network with hidden variables, the outcome can be a network that may fit the data well, but not exhibit the expected qualitative relationships. Wittig and Jameson [67] present a method for integrating qualitative knowledge into standard learning algorithms to learn the CPTs of Bayesian Networks from the data.

Another potential problem with using DDNs or Bayesian Networks in real-time environments is that they can become computationally expensive if they get too large. Although Bayesian Networks use conditional independence to simplify probabilistic inference, exact inference in an arbitrary Bayesian Network for discrete variables is NP-hard [33]. Among others, researchers in Intelligent Tutoring Systems (ITSs) have developed some techniques for improving the tractability in some applications. For
instance, the system Decision-Theoretic Tutor (DT tutor) [51] uses techniques such as
dynamic topology changes, filter nodes, and the tunnel network. The system Adaptive
Coach for Exploration (ACE) [6] uses two techniques to keep the Bayesian student
model in a manageable size, namely, dynamic topology changes and dividing the
network into two small networks. DDNs are particularly subject to inefficiency
problems because they can get extremely large as new time slices are added to the
network. In our model, we keep the size of the network under control by maintaining
only two time slices, as we describe in Chapter 5.
This chapter reviews previous work on affective user modeling. It also reviews work on probabilistic user modeling using Bayesian or Decision Networks.

3.1 Three Dimensions of User Assessment

This chapter compares related work on affective user modeling based on the following three dimensions of user assessment:

1. Which user characteristics does the model try to assess? The user characteristics can be cognitive states (e.g., domain knowledge level, goals, or beliefs), emotional states, and personality traits.

2. What kind of data does the model use to perform the assessment? The possible data include interaction patterns, performance, bodily reactions (e.g., physiological signals, facial expressions, and vocal intonation), and self-reports.

3. Which method(s) does the model use to combine the data to assess the users' characteristics? The methods include rule-based matching, probabilistic approaches (e.g., based on Bayesian or Decision Networks), case-based reasoning, and pattern recognition algorithms.

3.2 A Conversational Agent that Responds to Users' Emotions

In [4], Ball and Breese describe a reasoning architecture for a conversational agent that can generate appropriate behaviors when communicating with users. The agent can adapt its choice of paraphrases (semantically equivalent but emotionally diverse phrases), speech speed, and volume by diagnosing the users' emotions and personality
traits. However, instead of detecting individual emotional states, such as joy or distress, as our model does, the Ball and Bresso model focuses on two basic dimensions of emotional responses, arousal and valence. Since interpersonal relationships are the most critical in the domain of human-computer conversation, the Ball and Bresso model includes two personality traits, dominance and friendliness, which are the most related to social interactions. However, their model does not explicitly represent the relationships between emotional states and personality traits, as our model does. The data the Ball and Bresso model uses come from the users’ wording choices, speech characteristics (e.g., speech pace, rhythm, and pitch contour), and body language and movements, that is, data that represent the effects of different personalities and emotional states. Similar to our approach, their model relies on Bayesian Networks to combine the different sources of evidence on users’ emotions and personalities. However, while we model the causes of emotions by relying on the causal theory of emotions (OCC model), they model the effects of emotions. Another main difference is that our model represents the temporal evolution of emotional states by using DDN, while their model is static.

3.3 Interface Adaptation by Assessing Users’ Affective and Belief States

Hudlicka and McNeese [32] describe a system that adapts its interface format and content to the users’ affective state, select key personality traits, and situation-specific beliefs, in the context of an air force pilot combat task. In particular, the system assesses a pilot’s anxiety level, as well as the task-relevant beliefs that reflect the pilot’s situation assessment and awareness. For example, a pilot can believe that he is out of danger, about to be attacked, or under attack, and consequently be at different anxiety levels. The reason why only one affective state, anxiety level, is modeled is because it is the most relevant in the context of supporting pilot performance during combat, and also because the influences of anxiety on users’ performance are relatively well studied and understood. For instance, anxiety can narrow one’s attention and lead to perceptual biases, thus influencing the pilots’ judgment and causing dangerous actions (e.g.,
shooting an non-enemy plane). The model uses both static and dynamic factors to assess the pilots' affective state. The static factors include the level of task difficulty, the users' training and proficiency, the users' personality traits, and individual history. The dynamic factors include information about changing signals in the external environment (e.g., radar contacts, state of the aircraft, equipment failures, and team specific factors, such as the geometry of the intercept), as well as physiological signals, such as the pilot's heart rate. The model performs both predictive and diagnostic reasoning by considering both the causes and effects of the affective state. To combine the multiple factors influencing the users' affective state, it uses a fuzzy rule approach by specifying a weight for each relevant factor. The system also includes rules for predicting the possible influences of the level of anxiety on the users' belief states and performance, as well as rules that implement compensatory strategies to counteract the identified performance biases. Because of its highly constrained domain and exclusive focus on anxiety level, the model does not have to deal with the high level of uncertainty usually involved in affective user modeling and therefore does not need to resort to a formal method for reasoning under uncertainty.

### 3.4 The Affective Reasoner Framework

Elliott [19] has implemented a general platform called the Affective Reasoner to model a multi-agent world in which agents can reason about the world, generate appropriate emotions, and manifest their emotions through different channels. The agents also keep internal models of the concerns (goals, principles, and preferences) of other agents, which allow them to explain the behaviors of those agents. As in our approach, the Affective Reasoner is based on the OCC Model of emotion appraisal. However, it uses the OCC model only to generate agent emotions, and relies on a rule-based approach to implement the cognitive appraisal process because it assumes that all the information necessary for emotion generation is unequivocally available to the agents. While our model only includes six emotion types that are considered the most relevant in our domain, the Affective Reasoner includes twenty-four emotion types, twenty-two of which were originally specified by the OCC model. While the original Affective
Reasoner does not use the OCC theory to recognize other agents’ affective states, Elliott, Richel, and Lester [20] discuss the potential of applying the framework to affective user modeling. However, applying the framework to the assessment of users’ emotions requires knowledge of users’ goals, principles, and preferences. The authors do not discuss how to assess these factors.

In order to facilitate the assessment of the students’ goals while they are playing the Prime Climb game, our model includes the probabilistic relationships between the students’ personality traits, their interaction patterns, and their goals. Thus, our model can perform both diagnostic reasoning (from interaction patterns to goals) and predictive reasoning (from personality traits to goals) to assess the students’ goals.

3.5 Detecting Emotions from Physiological Signals

Researchers in the affective computing group at the MIT media lab have been actively investigating pattern recognition techniques to identify and combine features of physiological signals that can be used to detect emotional states. The assumption of the approach is that a particular emotional state has a corresponding unique physiological signal pattern. A variety of systems have been developed to demonstrate the feasibility of this approach. Vyzas and Picard [66] have developed an experimental system that can recognize eight emotional states (i.e., neutral, anger, hate, grief, platonic love, romantic love, joy, and reverence) with a success rate of 81.25%. The system uses the physiological signals coming from jaw clenching, blood volume pressure, skin conductivity, and respiration. However, the reported success rate was obtained in a restricted experimental environment, in which the eight emotions were intentionally expressed by an actress in the same sequence every time, and for a similar duration. Healey and Picard [28] use signals from electromyogram, electrocardiogram, respiration, and skin conductance sensors to assess the stress level in a car driver. Fernandez and Picard [21] discuss a system that uses signals of blood volume pressure and skin conductivity to detect the users’ frustration in human-computer interaction. Kapoor, Mota, and Picard [37] discuss preliminary work on assessing student affective states commonly seen in learning experiences (e.g., interest, boredom, confusion, and
excitement) by measuring their eye-gaze, eyebrow positions, postures, and head movements during the interaction with a computer-based learning companion. However, the authors do not present any evidence on the feasibility and effectiveness of their approach.

3.6 Motivation Diagnosis

An issue closely related to affective assessment is motivation diagnosis. Vicente and Pain [16] have developed a system called the Affective Dialoguer, which can generate educational dialogues by considering the students’ motivational states. The motivational state characteristics assessed by the system include confidence (which refers to the student’s belief in being able to perform the task correctly), sensory interest (which refers to the curiosity aroused from the interface presentation), cognitive interest (which refers to the curiosity aroused through cognitive characteristics of the task), effort (which refers to the amount of work the student puts into the task), and satisfaction (which refers to the overall feeling of goal accomplishment). The sources of data for the assessment are students’ behaviors and self-reported states. The system uses a rule-based approach to update the motivational model by considering the students’ replies during the educational dialogue process.

The rules in this system were originally drawn from theories of motivation and education, but Vicente and Pain also conducted an empirical study to further formalize them using an application system called MOODS for teaching Japanese numbers [17]. In that study, expert human tutors were asked to infer a student’s motivational state by watching the pre-recorded screen interaction of the student with the system. The experts used different sources of information to assess the students’ motivational states, such as the students’ performance, level of difficulty of the teaching materials, and the students’ traits. The knowledge from these experts was then formalized to create a final set of motivation diagnosis rules.

Another system considering motivational states is the Decision-Theoretic Tutor (DT Tutor) developed by Murray and Vanlehn [51]. The DT Tutor considers the students’ morale and independence, in addition to their cognitive states in order to decide
how to select the optimal tutorial actions. However, the authors do not discuss how to assess those motivational states in their system.

3.7 Probabilistic User Modeling

In user modeling tasks, the assessment of users’ characteristics, such as domain knowledge and cognitive or affective states, is often based upon limited and ambiguous information. Formalisms for reasoning under uncertainty, such as Bayesian Networks [55], Dempster-Shafer Theory [25], and fuzzy-logic [69], provide a sound way to represent and process the resulting uncertainty in the modeling process. Among these formalisms, Bayesian Networks have become especially popular in user modeling applications both because of their clear semantics and because of the large amount of research that has been devoted to increasing their computational efficiency and to learning them from the data. In addition to the ability to explicitly represent the probabilistic relationships in the modeling tasks, Bayesian and Decision Networks include sound and often efficient inference mechanisms, which can integrate any evidence available to compute the posterior probability of any given node in the network.

A number of systems using Bayesian and Decision Networks in user modeling have been developed over the past few years. The SQL-Tutor [48] is a system that teaches undergraduate students Structured Query Language in database courses. It uses Bayesian Networks to assess students’ knowledge of SQL, which the tutor takes into account when selecting tailored problems for the students. ANDES [11] is an Intelligent Tutoring System that teaches Newtonian physics via coached problem solving and example studying. The Bayesian student model in ANDES not only performs the assessment of students’ domain knowledge, but also performs the plan recognition (i.e., to infer which solution path the student is following) and prediction of students’ goals and actions. Besides goals and knowledge assessment, some researchers have applied Bayesian Networks to other kinds of assessment. For example, the student model in the Self-Explanation-Coach (SE-Coach) [14] extends the Bayesian model in ANDES to assess the effectiveness of students’ self-explanation from students’ reading and self-
explaining actions during example studying. In the ACE environment [6], the Bayesian student model is used to assess the effectiveness of students’ exploration in an open learning environment.

The pedagogical actions in all of the systems mentioned above are based on some heuristic decision rules, which take the posterior probabilities of their networks as input. Recently, there has been interest in using decision-theoretic methods combined with probabilistic reasoning for action selection. For instance, the DT Tutor [51] uses DDNs to select optimal tutorial actions during physics problem solving, given the assessment of users’ characteristics, such as domain knowledge, focus of attention, morale, and independence. The system developed by Horvitz and Barry [30] relies on Decision Networks to manage the complexity of information displayed in a time-critical decision making environment by modeling users’ beliefs and actions. The Lumiere system [31] also uses Decision Networks to decide when to provide help by considering users’ goals and needs when they are interacting with Excel application software. Mayo and Mitrovic [49] use DDNs to guide the behaviors (i.e., problem and error message selection) of a computer–based tutor based on the assessment of students’ knowledge in English capitalization and punctuation.

The approach taken by this thesis is to use DDNs to represent and reason about the uncertainty in affective user modeling and to provide theoretic guidance for the behaviors of the pedagogical agent we are designing. However, the focus of this thesis is not on the decision part of the network, but rather on the part of the network that assesses the students’ emotions. Before discussing the affective model, we first describe the educational game Prime Climb, which is the application domain for our research.
Prime Climb is an electronic educational game developed by the Electronic Games for Education in Math and Science (EGEMS) group at the University of British Columbia. The primary goal of the game is to help 6th or 7th grade students to learn number factorization. This includes how to factorize a number, and how to identify whether two numbers share any common factors. A secondary goal of Prime Climb is to foster cooperation among students. In the game, two players need to cooperate with each other and reason about the underlying math knowledge to climb a series of mountains (see Figure 4.1). The game also includes a tool to help players inspect the factorization of numbers and make moves.

While students are usually highly engaged when playing Prime Climb, initial studies show that students do not necessarily learn from it [12]. While the pedagogical tools in the game are helpful for some students, other students do not learn much from them. This happens mostly with students who have low math knowledge and/or with students who do not pause and reason about the underlying math knowledge. These students either keep falling or manage to get to the top of a mountain using heuristics unrelated to the knowledge of number factorization. This problem cannot be easily solved by changing the Prime Climb interface, which already underwent a careful process of iterative designs, and has been shown to work well for some students. The problem mainly derives from the fact that the effectiveness of educational games, and of open learning environments in general, does depend on a number of user dependent factors that cannot possibly be matched by a unique design [13][61]. Thus, to improve the effectiveness of Prime Climb, one approach is to design a pedagogical agent that can provide tailored support to help students learn from playing the game. Because engagement is the main advantage of educational games, it is required that the agent’s actions not compromise the students’ level of engagement. Therefore, it is important for
the agent to consider the students’ emotional states, in addition to their cognitive states (such as learning), when deciding when and how to provide its pedagogical interventions.

This chapter describes the interface of the educational game Prime Climb and the pedagogical agent. The affective model, which provides the pedagogical agent with the information on students’ emotional states, is discussed in detail in Chapter 5 and 6.

4.1 The Prime Climb Interface

Figure 4.1 is a screen shot of the Prime Climb interface. As mentioned earlier, Prime Climb is designed for 6th or 7th grade students to help them learn concepts related to number factorization, such as prime numbers, factors, common factors, and the factor tree. It is a two-player game, in which two players need to cooperate with each other to get to the top of a series of mountains. The mountains are divided in numbered sectors. Each player can only move to a numbered sector that does not share any factors with the numbered sector occupied by his partner. When a player moves to a number that shares a factor with the one the partner is on, the player will fall down the mountain. For example, in Figure 4.1, the player at the bottom fell down and is swinging on the mountain because she had tried to move to number 42, which shares factor 3 with number 9 occupied by the partner.

In addition to this main rule, there are also other rules that regulate the players’ moves. For instance, a player can only move to the numbers that are adjacent to his current position. There is a rope that ties two players together and the rope has a maximum length of three numbers. Therefore, each player can move no further than two sectors away from her partner. Sometimes one player has to wait for the partner to climb, or has to move backward if there are no “good” numbers for the partner to move to. The mountain is also covered with some obstacles, shown as rocks or forests (Figure 4.1). Players are not allowed to move to those obstacles.
The Prime Climb interface also includes two tools that help students during the climbing moves, the magnifying glass and help box. These tools are accessible by clicking the corresponding buttons in the PDA window shown in the top-right corner of Figure 4.1. The magnifying glass tool is for players to see the factor tree of the numbers on the mountain. To use it, one needs to click the magnifying glass button and then select a number on the mountain. The factor tree of that number will then be shown in the PDA window. Figure 4.2 shows a factor tree of number 40.
The tool "help box" is designed to allow a player to communicate with the pedagogical agent (see Figure 4.3). This tool is especially useful during a phase of Prime Climb called the "Practice Climb". During this phase, instead of playing with a peer, a student climbs with the pedagogical agent, which acts as a climbing instructor. The affective model that we have developed models the student behavior during this phase. We concentrated on this phase because modeling the student’s affect during the interaction with a peer would have entailed modeling emotions toward this peer in addition to the emotions toward the agent and self. Given the many challenges that already exist in modeling the interaction with the agent only, we decided to concentrate on solving them before moving to a more complex interaction.
4.2 The Pedagogical Agent

The character in the top-left corner of Figure 4.1 is the pedagogical agent we are designing. The main goal of the pedagogical agent is to provide tailored support to help students learn number factorization without compromising their level of engagement in the game. For example, in Figure 4.1, the pedagogical agent is trying to help the player who just fell down the mountain to recover from a fall by saying “Can you grab a hex that doesn’t have a common factor with 9”. The actions that the agent may perform include the followings:

- Stimulating a student to reason about the cause of a fall or a successful move
- Providing specific advice on where a student should move
- Reminding a student of the tools available to help make a move
- Deciding the level of difficulty of the climbing task
- Giving emotional support
- Reminding a student of the game-related rules.

Since this thesis research focuses on developing a model of students’ emotions, and not on developing an effective pedagogical model for the agent’s behavior, we have developed a Wizard of OZ interface (see Figure 4.4) that allows the agent to be controlled by an experimenter. We describe how we used this interface to collect information on the design of the affective model in the next chapter. Here we briefly describe how the interface works. When playing with a student, an experimenter uses the interface to control the agent’s behavior. For instance, clicking the “Hide” button makes the agent disappear (this can be useful if the experimenter finds that the student does not like the agent’s help at all); clicking the “ComFactor” button makes the agent remind the student of the common factor rule in the game; clicking “MagGlass” button makes the agent remind the student of the magnifying glass tool by saying “Do you know you can use the magnifying glass to see number factorization?”; clicking the “WhyFail” button lets the agent ask the student the question “Do you know why you fell this time?” in order to stimulate the student to reason about the fall; and typing a number in the text box (the first text box in the “Move” frame in Figure 4.4) and clicking
the "Move" button makes the agent give specific advice on where the student can move. Through the interface, the experimenter can also make the agent speak anything to the student, including answering the student’s questions, by typing the words in the text box (at the bottom of Figure 4.4) and clicking the "Speak" button.

Figure 4.4: Wizard of OZ Interface
While emotions can be recognized by analyzing both cause and effect, this thesis only focuses on a causal model of emotions that relies on the OCC cognitive theory. Conati [10] describes how the affective model can be extended by incorporating evidence coming from bodily expressions, such as heart rate, skin conductance, and eyebrow movements.

The initial version of the model was built based on the observations from two preliminary Prime Climb user studies, our intuitions, and psychological findings. This chapter first discusses the uncertainty involved in modeling user affect in the Prime Climb environment. Then it describes the two Prime Climb user studies that gave us the intuition for the initial model design, presents a high level description of the model, and finally describes the initial version of the affective model.

5.1 Uncertainty in Affective User Modeling For Prime Climb

As discussed in Section 1.4, affective user modeling usually involves a high level of uncertainty. Some existing work has tried to reduce this uncertainty by focusing on one specific emotion that is relevant to the target domain. For instance, Hudlicka and McNeese [32] describe a system that only assesses a combat pilot’s anxiety level in a combat pilot training environment, while the system designed by Healy and Picard [28] detects the stress level of a car driver. Other researchers [4] have tried to reduce the uncertainty in affective user modeling by assessing two basic dimensions of an emotional response, arousal and valence, instead of individual emotion types.

Neither of the above approaches is appropriate for modeling student emotions in Prime Climb because the game tends to elicit a variety of different emotions in different
students. For example, when falling down the mountain, a student whose goal is to have fun might find the animation and music accompanying the falling appealing, and may consequently feel joy, while a different student whose goal is to perform well and avoid falling might feel distressed and even ashamed of his performance. Students can also have different emotional responses to the agent's help. When the agent provides help, some students might feel admiration toward the agent, while other students may think it is interfering with their game playing and thus feel resentful toward the agent. Detecting a student's specific emotional responses can give the Prime Climb pedagogical agent additional information to take more tailored actions and improve the student's interaction with the game. For example, if the agent realizes that a student does not like to be helped, the agent can try to reduce the help, or find more subtle ways to provide it. If the agent detects that a student is continuously frustrated because she keeps making mistakes, the agent can decide to restore the student's confidence by lowering the difficulty level of the task.

Since Prime Climb does tend to elicit different emotions, and detecting individual emotions can allow the pedagogical agent to take more tailored actions, unlike the approaches described earlier, our affective model includes the six emotion types considered the most relevant to the domain. In order to deal with the higher level of uncertainty caused by the increased number of modeled emotions, the affective model relies on DDNs, which we describe after illustrating two preliminary user studies that we performed to gain more information on the interaction we need to model.

5.2 Two Preliminary Prime Climb Studies

In order to gain preliminary information on the affective student model design, two pilot Prime Climb user studies were conducted in October and December of 2001.
5.2.1 The Study Goal

The general goal of the two preliminary studies was to gain information on what variables and dependencies our affective student model should include. In particular, we wanted to identify the following:

- What goals students have while they are playing the game
- What interaction patterns they follow
- If there are any relationships between goals and interaction patterns
- What emotional reactions students have in playing the game and under what conditions

We would have also liked to identify any dependencies between personality traits and goals. However, we could not find an adequate personality test in time for the study, and therefore, we could not gather information on the students' personalities.

5.2.2 The Study Participants

The participants in both preliminary studies were 6th grade students. Before the studies, they had already been exposed to number factorization in class. The 13 participants in the first study were from University Hill Elementary School in Vancouver. The second study included 10 participants from False Creek Elementary School in Vancouver. All 23 participants volunteered for the studies, and parental permission for participation was obtained for all of them.

5.2.3 The Study Design

Both preliminary studies were based on the Wizard of Oz design. Before the studies, the participants were told that they would be playing with the computer agent called Merlin who was their climbing instructor. In fact, the participants played with an experimenter (Wizard) who controlled the agent from a separate room. At the end of each study, the participants were told that they had been playing with a human experimenter. Two different experimenters (both are MSc. students in the computer science department at
the university of British Columbia) acted as the wizard during the study, following the criteria given below to decide when and how to act:

- Advise a student to use the magnifying glass if the student seldom uses the tool and often makes mistakes
- Remind a student that he or she can use the help box to ask questions when the student does not seem to understand the rules of the game
- Give specific advice on where a student can move if the student is in difficulty
- Stimulate a student to reason about a fall or a successful move if the student does not seem to be taking the time to do so
- Provide positive feedback to a student if the student makes progress.

The two studies took place at the two elementary schools. Each student participated in one session, which lasted 30 minutes and consisted of a 10-minute introduction and pre-test phase, a 15-minute game playing phase, and a 5-minute post-questionnaire phase. In the introduction, an experimenter described the purpose of the study and the rules of the game (see Appendix A for the introduction text that was used to keep the introduction constant for all the students). The pre-test (see Appendix B) included three math questions assessing the students’ knowledge of number factorization and two other general questions about their attitudes toward math and computer games. The post-questionnaire included questions about the students’ goals while playing the game and about their attitudes toward the agent’s help (see Appendix C for the post-questionnaire).

An observer recorded the students’ behaviors during their game playing (see Appendix D for the observation sheet). We also collected videotapes of each student, and log files in the second study (10 students).

### 5.3 The Results of the Preliminary Studies

The preliminary studies led to several insights for the model construction. We found that students had different focuses and different interaction styles when playing the game. In addition, different students had different emotional responses and different
ways of expressing them. The following is a summary of the findings from the two preliminary studies.

5.3.1 Students' Goals in Playing the Game

Different students had different goals when playing the game. Through the studies, we were able to identify five high level goals: Have Fun, Avoid Falling, Learn Math, Beat Partner, and Succeed By Myself. Table 5.1 shows the number of students that chose each goal in the post-questionnaire. Because Prime Climb is an educational game which combines factorization knowledge with game activities, it is quite normal for the students to have both the goal Have Fun and Learn Math. The goal Avoid Falling is also consistent with the game, because the students must get to the top of the mountain in order to win the game.

A less intuitive goal is Beat Partner, because the game is designed to foster collaboration (in the introduction phase the students were instructed to work with their climbing instructor to get to the top). However, as many as eight students in total chose the goal Beat Partner in the post questionnaire. This can probably be explained by the fact that, in most commercial electronic games, players are normally instructed to be competitive or aggressive; consequently, some students displayed a tendency to behave competitively, even in a collaborative environment. This finding concerning the competitive goal of some participants echoes the findings in the psychological research which indicates that people who scores low on the personality trait Agreeableness tend to be more competitive rather than cooperative [15][35].

As Table 5.1 shows, eight students in total selected the goal Succeed By Myself. One possible explanation is that students with high self-esteem are more independent and have a high opinion of their own abilities, and therefore, prefer to work on their own.

We also found that students could have multiple goals. In the preliminary studies, sixteen out of twenty-three students selected two or more goals in the post questionnaire.
Table 5.1: Number of Students Who Chose Each Goal in the Post-Questionnaire

<table>
<thead>
<tr>
<th>Goals</th>
<th>Have Fun</th>
<th>Avoid Falling</th>
<th>Learn Math</th>
<th>Beat Partner</th>
<th>Succeed By Myself</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Students</td>
<td>14</td>
<td>15</td>
<td>12</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

5.3.2 Interaction Patterns and Their Relationships to Goals

The students showed different interaction patterns when playing the game. Five interaction patterns were identified from observations during the preliminary studies:

- **Move Quickly**: indicating whether a student tends to move quickly or slowly
- **Use Magnifying Glass Often**: indicating how often a student uses the magnifying glass
- **Ask Advice**: indicating how often a student uses the help box to ask the agent for help
- **Follow Advice**: indicating how often a student follows the agent’s advice
- **Fall Often**: indicating how often a student falls

We also found several qualitative dependencies between the students’ goals and their interaction patterns. For instance, we found that (1) the students who had the goal *Succeed By Myself* seldom asked the agent for help, (2) the students whose goal was to *Learn Math* were more likely to use the magnifying glass to see the number factorization, and (3) the students whose goal was to *Avoid Falling* tended to move more slowly than the students whose goal was to *Beat Partner*.

5.3.3 The Emotion Eliciting Situations

By analyzing the videotapes and the observation sheets, we found that the students had different emotional reactions when playing the game. For example, when falling down the mountain, some students seemed very frustrated, while other students seemed excited.
to see the animation accompanying the falling. When getting to the top of the mountain, some students seemed happy, while other students expressed their concern by asking who got to the top first or how much their partners scored. When the agent provided help, some students did not like it and frowned, while others attended to and followed the agent's advice.

5.4 High Level Description of The Initial Affective Model

Figure 5.1: A High Level Description of the Affective Model

Figure 5.1 shows a high level description of the model that we built, which takes into account the results of the two user studies described above, consisting of two time slices of the DDN. The nodes in Figure 5.1 represent classes of variables in the actual model. Details on individual nodes and dependencies are given in section 5.5.

In order to apply the OCC theory to the assessment of student emotions when playing the Prime Climb game, the model includes goal nodes (Goals nodes in Figure 5.1) representing the five high level goals (Have Fun, Avoid Falling, Learn Math, Beat Partner, and Succeed By Myself) identified from the two preliminary studies. Each goal
is represented as a different node in the actual model (see Figure 5.2). The emotion eliciting situations represented in the model are events resulting from the outcome of either a student’s action (Student Action Outcome node in Figure 5.1) or the agent’s action (Agent Action Outcome node in Figure 5.1). The Agent Action Outcome node is represented as a decision variable in the model indicating the different actions the agent can choose from. Because this thesis focuses on the affective modeling, the utility node describing the preferences of the pedagogical agent is not shown in Figure 5.1. The Goals Satisfied nodes in Figure 5.1 represent the outcome of the student’s situation appraisal, which is affected by both the student’s goals and the current situation, and that in turn affects the student’s emotions (Emotional States nodes in Figure 5.1), as specified by the OCC theory.

As students’ goals are a key element in the application of the OCC theory, the model also includes nodes to facilitate the assessment of these goals. The goals that students have can depend on both the students’ personality traits [15] and their knowledge of number factorization. This dependency is represented by the links from the Personality nodes and Factorization Knowledge nodes to the Goals nodes in Figure 5.1. Since the preliminary studies showed that students with different goals have different behaviors while playing the game, the links from the Goals nodes to the Interaction Patterns node in Figure 5.1 reflect the influence of the students’ goals on their game behaviors. We also believe that factorization knowledge can affect interaction patterns, and this influence is represented by the links from the Factorization Knowledge nodes to the Interaction Patterns nodes in Figure 5.1. Specific features of the students’ individual actions (represented by some of the student action outcome nodes) are used to infer their interaction patterns, as we describe in more detail in the next section.

The link between the Factorization Knowledge nodes in different time slices models the evolution of students’ knowledge as they play the game. The link between Emotional States nodes reflects the fact that a student’s emotional state at time $t_{i+1}$ can be influenced by his or her emotional state at time $t_i$. For instance, if a student was feeling joy previously, a satisfying event is more likely to have the student stay in that emotional state. Our model currently includes the assumption that the student high level goals do
not change over time, as indicated by the lack of a link between the Goals node at $t_i$ and the Goals node at $t_{i+1}$.

### 5.5 Details of the Model Design

In this section, we look in more detail at the nodes that form the classes described in the previous section, and at the corresponding probabilistic dependencies. Section 5.5.1 describes the part of the network that assesses students' goals (see Figure 5.2). Section 5.5.2 describes the part of the network that assesses students' emotions from these goals (see Figure 5.3).

![Figure 5.2: Part of the Network Assessing Students' Goals](image)

**Figure 5.2: Part of the Network Assessing Students' Goals**

#### 5.5.1 Portion of the Network Assessing Students' Goals

Five high level goals were identified from the preliminary studies, and they were not mutually exclusive, showing that students can have multiple goals when playing the game. Therefore, the model includes five goal nodes, each representing an individual
goal: *Have Fun, Avoid Falling, Beat Partner, Learn Math, and Succeed By Myself* (see Figure 5.2). Each goal node has two values, True and False, representing whether a student has the corresponding goal or not.

Personality traits affect one's goals and behaviors [15][35]. To represent the influence of the students' personality traits on their goals during the game, the model includes personality nodes and links from the personality nodes to the goal nodes. To represent the students' personality, our model uses the Five Factor Personality theory developed by Costa and McCrae [15] in which personality traits are classified under five domains: *Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness*. Of these, only four personality domains are represented in the model (see Figure 5.2), as the *Openness* domain does not seem to be directly relevant to our task.

The links between the students' personality traits and their goals follow the definitions of the four personality domains in the Five Factor Personality theory. For example, an agreeable person is defined as someone who is “sympathetic to others and eager to help them, and believes that others will be equally helpful in return”, whereas a disagreeable person is defined as “egocentric, skeptical of others’ intentions, and competitive rather than cooperative”. This definition indicates that *Agreeableness* can influence the two goals *Beat Partner* and *Succeed By Myself*. This influence is represented in the network by two links from the personality node *Agreeableness* to the two goal nodes *Beat Partner* and *Succeed By Myself* (see Figure 5.2). Furthermore, the CPTs for the two goal nodes are set to reflect that they are likely for a disagreeable player, and unlikely for an agreeable one. Similarly, the CPT for the node *Avoid Falling* is set to indicate that it is more likely for a neurotic player to have this goal, while the CPT for the node *Have Fun* represents that this goal is more likely for a player who is an extrovert.

Each personality node has two values, True and False, with the value True representing that a student belongs to one end of the corresponding personality domain (e.g., extraversion or agreeableness), and the value False representing that a student belongs to the opposite end (e.g., introversion or disagreeableness).

Students' goals may also be influenced by their domain knowledge, as it is represented in the model by the link between the *Factorization Knowledge* node and the
goal node Learn Math (see Figure 5.2). This link is defined based on the intuition that it is less likely for a student with high math knowledge to want to learn math through the game. However, this may not be true for a conscientious student, who tends to be a perfectionist or an overachiever. Although we did not have data on the students' personality traits and could not verify this dependency, we decided to keep the link in the initial model and refine it by running additional studies in which we have information on the students' personalities.

As the preliminary studies indicated, students with different goals could have different interaction patterns when playing the game. Thus, students' goals can also be inferred from their interaction patterns. The model includes the five interaction patterns identified from the two preliminary studies (see Figure 5.2). Some links from the goal nodes to the interaction pattern nodes are defined based on the observations from the preliminary studies. Those links are: (1) the link from Succeed By Myself to Ask Advice, (2) the link from Learn Math to Use Magnifying Glass Often, (3) the two links from Avoid Falling and Beat Partner to Move Quickly, and (4) the link from Have Fun to Fall Often. The other links are defined based on our intuitions. For instance, the interaction pattern Fall Often, is represented as being influenced by both the goal Have Fun and the students' Factorization Knowledge. While the link from Have Fun comes from the observation that falling down the mountain can happen because a student finds it appealing to see the animation accompanying the fall and thus deliberately tries to fall, the link from Factorization Knowledge comes from the intuition that the less knowledge a student has, the more likely he is to fall. The True and False values of each interaction pattern node represent whether a student follows the corresponding interaction pattern or not.

Features of the students' individual actions are used as evidence to infer their interaction patterns. Corresponding to each interaction pattern node, the model includes a node describing the students' individual action that supports that pattern (see Figure 5.2). Individual action nodes include: Magnifying Glass Used (indicating whether a student used the magnifying glass or not), Quick Move (indicating whether a student made a quick or slow move), Advice Asked (indicating whether a student used the help box to ask the agent for advice), Advice Followed (indicating whether a student followed
the advice from the agent or not), and Successful Move (indicating whether a student made a successful move or not). The node Successful Move also defines the emotion-eliciting situation triggered by a student's move, in addition to providing evidence to infer the Fall Often interaction pattern (see Figures 5.2 and 5.3).

Given the structure we have just described, the model can perform both predictive assessment of the student's goals when information on the student's personality traits and math knowledge is available, as well as diagnostic assessment based on evidence from the student's individual game actions.

### 5.5.2 Portion of the Network Assessing Students' Emotions

![Sample Nodes Involved in Assessing Students' Emotions](image)

Figure 5.3 Sample Nodes Involved in Assessing Students' Emotions

This part of the network implements the OCC theory. Two kinds of emotion-eliciting situations are modeled in the network: a student making a move (see Figure 5.3 slice $t_i$) and the agent providing an intervention (see Figure 5.3 slice $t_{i+1}$). We currently do not consider the student's use of the magnifying glass or help box as an emotion-eliciting situation, since we did not observe clear emotional reactions to these actions during the
studies. These actions are only used as evidence to update the assessment of the student's goals.

Every time a student makes a move or the agent performs an action, a new time slice is added to the DDN. The nodes describing the situation after a student makes a move are these: Successful Move (indicating whether a move is successful or not) and Ahead Of Partner (indicating whether the student is ahead of the partner or not after a move). We currently summarize the actions available to the agent in three types, represented as different values of the agent's decision node: (1) Hints (suggestions to help a student without telling exactly what to do, for example, asking a student if there is a common factor between number x and y); (2) Move to (specific advice telling a student where to move); and (3) Reflect (prompting a student to reason about the outcome of a move, for example, asking a student why she succeed at a given time). In our model, in order to keep the size of the network under control, only two time slices are maintained at a given time, because the influence of previous time slices can be summarized as prior probabilities in the first of the two active time slices.

The model includes a layer of Goals Satisfied nodes to represent the outcome of the student's situation appraisal (see Figure 5.3). Only a subset of the Goals Satisfied nodes is shown in the figure. For each goal, there is a corresponding Goals Satisfied node describing whether the goal is satisfied or not in the given emotion-eliciting situation. The Goals Satisfied nodes are influenced by both the student's goals and the outcome of the student's moves or the agent's actions (see Figure 5.3). For example, if a student has the goal Beat Partner and makes a move at time t, the value of the node Beat Partner Satisfied in that time slice depends on whether the student is ahead of his partner after that move (see Figure 5.3). Similarly, if a student has the goal Succeed By Myself, it is unlikely for that goal to be satisfied if the agent provides general advice about how to move, and even more unlikely if the agent tells the student exactly where to move.

According to the OCC theory, the outcome of situation appraisal in turns triggers emotions. Three pairs of emotions, which are considered the most relevant for this domain, are represented in the model by three separate nodes (see Figure 5.3): (1) Emotion For Event (representing the emotion toward the appraised event), with values joy and distress, (2) Emotion For Self (representing the emotion a student feels when she
caused the event), with values *pride* and *shame*, and (3) *Emotion For Agent* (representing a student’s emotion toward the agent when the agent caused the event), with values *admiration* and *reproach*.

The *Goals Satisfied* nodes are always linked to the *Emotion For Event* node in every time slice, that is, every time a student makes a move or the agent provides an intervention, representing the result of the corresponding event appraisal. The *Goals Satisfied* nodes are also linked to the *Emotion For Self* node in time slices in which the current emotion-eliciting situation is triggered by a student’s move (see time slice $t_i$ in Figure 5.3), or linked to the *Emotion For Agent* node in time slices when the current situation is triggered by the agent’s action instead (see time slice $t_{i+1}$ in Figure 5.3). The links between emotion nodes across time slices (see Figure 5.3) model the evolution of emotions, representing the fact that a previous emotion affects a student’s subsequent emotion.

For example, in Figure 5.3, when a student makes a move at time slice $t_i$, the *Goals Satisfied* nodes are linked to *Emotion For Event* and *Emotion For Self*, indicating that the student can have an emotion (*joy* or *distress*) related to the resulting situation (*Successful Move* and *Ahead of Partner*), as well as an emotion (*pride* or *shame*) toward himself. When the agent performs an action at time slice $t_{i+1}$, the *Goals Satisfied* nodes are linked to *Emotion For Event* and *Emotion For Agent*, indicating that the student can have an emotion (*admiration* or *reproach*) toward the agent intervention, in addition to an emotion related to the event itself. Because agent actions do not affect emotions toward oneself, the node *Emotion For Self* at time slice $t_{i+1}$ is only influenced by the same node at time slice $t_i$. In particular, the CPT for the node *Emotion For Self* at time slice $t_{i+1}$ models the fact that an emotion fades over time if no event revives it.

### 5.5.3 Conditional Probability Tables

The Conditional Probability Tables (CPTs) in the initial model were defined using our estimates. In Chapter 6, we describe how we refine them based on the results of a third Prime Climb study. This section gives some examples of the CPTs in the initial model.
For the \textit{Goals} nodes, the CPTs were constructed following the definitions of the personality domains. For example, the CPT for the node \textit{Avoid Falling} is set to indicate that the goal \textit{Avoid Falling} is likely for a neurotic player, and unlikely for a non-neurotic person (see Table 5.2).

\begin{table}
\centering
\caption{CPT for \textit{Avoid Falling} Given Neuroticism}
\begin{tabular}{|c|c|}
\hline
\textbf{Neuroticism} & \textbf{Avoid Falling} \\
\hline
True & 0.9 & False & 0.1 \\
\hline
False & 0.1 & False & 0.9 \\
\hline
\end{tabular}
\end{table}

Table 5.3 shows the CPT for the node \textit{Move Quickly} that depends on three goals: \textit{Have Fun}, \textit{Avoid Falling}, and \textit{Beat Partner}. It is set based on the following criteria:

- If a student has either the goal \textit{Have Fun} or \textit{Beat Partner}, it is likely for the student to move quickly.
- If a student has the goal \textit{Avoid Falling}, it is likely for the student to move slowly.
- If a student has the either the goal \textit{Have Fun} or \textit{Beat Partner}, as well as the goal \textit{Avoid Falling}, the network is uncertain about the \textit{Move Quickly} pattern.

\begin{table}
\centering
\caption{CPT for \textit{Move Quickly} Given \textit{Have Fun, Avoid Falling and Beat Partner}}
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{Have Fun} & \textbf{Avoid Falling} & \textbf{Beat Partner} & \textbf{Move Quickly} \\
\hline
\text{True} & True & True & 0.5 & 0.5 \\
& True & False & 0.5 & 0.5 \\
& False & True & 0.9 & 0.1 \\
& False & False & 0.7 & 0.3 \\
\hline
\text{False} & True & True & 0.5 & 0.5 \\
& True & False & 0.1 & 0.9 \\
& False & True & 0.7 & 0.3 \\
& False & False & 0.5 & 0.5 \\
\hline
\end{tabular}
\end{table}
As illustrated in table 5.4, the CPTs for the individual action nodes are set so that if a student follows a given interaction pattern, it is more likely for the student to make an action that supports that pattern.

**Table 5.4: CPT for Quick Move Given Move Quickly**

<table>
<thead>
<tr>
<th>Move Quickly</th>
<th>Quick Move</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True</td>
</tr>
<tr>
<td>True</td>
<td>0.7</td>
</tr>
<tr>
<td>False</td>
<td>0.3</td>
</tr>
</tbody>
</table>

The CPTs for the Goals Satisfied nodes are set so that:

- If a student does not have a given goal, the network is uncertain about whether the goal is satisfied,
- Otherwise, the probability of the goal being satisfied is high if the situation satisfies the goal. The probability is low if the situation conflicts with the goal.

Table 5.5 shows the CPT for the node Succeed By Myself Satisfied given the corresponding goal and the agent action.

**Table 5.5: CPT for Succeed By Myself Satisfied When Agent Intervenes**

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Agent Action</th>
<th>Succeed By Myself Satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>True</td>
</tr>
<tr>
<td>True</td>
<td>Hint</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>MoveTo</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Reflect</td>
<td>0.3</td>
</tr>
<tr>
<td>False</td>
<td>Hint</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>MoveTo</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Reflect</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The CPTs for the emotion nodes are set based on the following criteria:

- If an event satisfied more than half of the related goals, the event will trigger positive emotions. Otherwise, the event will trigger negative emotions.
If a student is in a given emotional state at the previous time slice and no event revives it, it is likely for the student to stay in that emotional state but with a slightly lower probability.

If a student is in a given emotional state at the previous time slice and the current event triggers the same emotion, the probability that the student has that emotion is high. If the current event triggers the opposite emotion, it is still likely for the student to stay in that emotion which she felt previously, but with a low probability.

Table 5.6 shows the CPT for the node *Emotion For Agent* when the agent intervenes.

**Table 5.6: CPT for Emotion For Agent When Agent Intervenes**

<table>
<thead>
<tr>
<th>Emotion For Agent (t_i)</th>
<th>Succeed By Myself Satisfied</th>
<th>Learn Math Satisfied</th>
<th>Have Fun Satisfied</th>
<th>Emotion For Agent (t_{i+1})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Admiration</td>
</tr>
<tr>
<td>Admiration</td>
<td>True</td>
<td>True</td>
<td>True</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>True</td>
<td>0.51</td>
</tr>
</tbody>
</table>

48
5.6 Limitations of the Preliminary Studies

The preliminary studies did have several limitations, which are reflected on the design of the initial student model:

- Since we could not find an adequate personality test in time for the study, we could not gather information on the students' personalities and their influences on the students' goals.

- The pre-test only included three questions to assess the students' factorization knowledge. Thus, it could not give reliable information on the students' factorization knowledge and its influence on other variables in the model.

- In the post questionnaire, the questions assessing student goals are True/False questions. This limits the students' choices and can force a student to choose a goal that he might have, but with low priority.

- We had log files for only 10 students. The other students' game behaviors were recorded on the observation sheet by different observers. Therefore, the results can be subjective, although we did train the observers to rate the students' game behaviors by using a set of predefined guidelines.

5.7 Limitations of the Initial Model

In the initial model, the links between the students' personalities and their goals when they play the educational game are defined following the general definitions of the personality domains in the Five Factor personality theory. Because of lack of data on the students' personalities, we could not verify these relationships. Further studies with access to the students' personalities are needed in order to verify the relationships.

The links between the students' goals and their interaction patterns are defined based on the observations from the two preliminary studies, as well as our intuition. However, because the observations were subjective, the resulting links are not quite reliable. Refining those links requires additional studies in which the students' game behaviors can be recorded more objectively to allow for a more reliable analysis.
The CPTs in the model are constructed using our estimates. As is the case with all human-designed CPTs, such estimates need to be refined empirically.

In the next chapter, we describe a third study that we conducted to overcome the limitations of the previous studies and the corresponding limitations of the initial model.
As discussed in Sections 5.6 and 5.7, the preliminary studies and the initial model have several limitations. This chapter discusses a new study and the changes made to the initial model to address these limitations. It first presents the results of the third Prime Climb study, and then describes the revised model based on the results of the study.

### 6.1 The Third Prime Climb Study

In order to refine the initial model, especially the part of the network that assesses student goals, we conducted a third Prime Climb study in April 2002.

#### 6.1.1 The Study Goal

The purpose of the study was to clarify (1) the relationships between the students' personality traits, number factorization knowledge, and their goals when playing the game; and (2) the relationships between the students' goals and their interaction patterns.

#### 6.1.2 The Study Participants

A total of nineteen participants volunteered for the study. All the participants were 6th grade students from Maple Grove Elementary School in Vancouver. The students had been exposed to number factorization in class before the study. Parental permission for participation was obtained for all the students.
6.1.3 The Study Design

The study was based on essentially the same design as the two preliminary studies (see section 5.2.3). However, we revised the pre-test and the post questionnaire in order to make a more accurate assessment of both the students' number factorization knowledge and their goals when playing the game.

Compared to the first version of the pre-test, which contained only three math questions, the revised pre-test (see Appendix E) consisted of fifteen math questions assessing the students' number factorization knowledge.

Compared to the first version of the post questionnaire in which the questions assessing the students' goals were True/False questions, the revised post questionnaire (see Appendix F) consisted of fifteen questions to be answered using a scale of 1 to 5. The questionnaire also included an open-ended question to see if the students had any other goals, in addition to the five goals we had already identified in the preliminary studies.

Before the study, we used a revised version of Goldberg's 100 standard markers [23] to measure the five personality domains of the Five Factor theory. This personality test consists of one hundred adjectives, twenty for each personality domain. The students were instructed to rate each adjective as an accurate description of themselves, from 1 (strongly disagree) to 5 (strongly agree). The personality test is appropriate for children as it contains the definitions of the adjectives, making it easier for young age students to understand [26]. Campbell and Graziano [35] used the same test to measure the personality traits of 6th, 7th, and 8th grade students.

An observer standing behind each student recorded informal observations on the same observation sheet used in the preliminary studies (see Appendix D). Each student's interactions with the game and facial expressions were also videotaped. We used a video mixer to combine video sources from both the camera (recording facial expressions) and the computer (game interactions). A log file was created for each student, recording the game events in detail. Table 6.1 summarizes the information recorded in the log files.
<table>
<thead>
<tr>
<th>Game Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move Action</td>
<td>Records the characteristics of the student’s move, including:</td>
</tr>
<tr>
<td></td>
<td>• The time spent between two successive moves</td>
</tr>
<tr>
<td></td>
<td>• Whether the move is successful or not</td>
</tr>
<tr>
<td></td>
<td>• Whether the student is ahead of partner after the move</td>
</tr>
<tr>
<td>Using Magnifying Glass</td>
<td>Records when the student uses the magnifying glass.</td>
</tr>
<tr>
<td>Using Help Box</td>
<td>Records when the student uses the help box and the content typed.</td>
</tr>
<tr>
<td>Following Advice</td>
<td>Records when the student follows the advice from the agent, including:</td>
</tr>
<tr>
<td></td>
<td>• Follows the advice to use the magnifying glass</td>
</tr>
<tr>
<td></td>
<td>• Follows the advice to use the help box</td>
</tr>
<tr>
<td></td>
<td>• Follows the advice to move to a specific number</td>
</tr>
<tr>
<td>Agent Intervention</td>
<td>Records the type of interventions provided by the agent, including:</td>
</tr>
<tr>
<td></td>
<td>• Advise the student to use the magnifying glass</td>
</tr>
<tr>
<td></td>
<td>• Advise the student to use the help box to ask questions</td>
</tr>
<tr>
<td></td>
<td>• Give specific advice on where the student can move</td>
</tr>
<tr>
<td></td>
<td>• Stimulate the student to reason about a fall or a successful move</td>
</tr>
<tr>
<td></td>
<td>• Provide acknowledge to the student</td>
</tr>
</tbody>
</table>
6.2 The Study Results and Interpretation

6.2.1 Relationships between Students' Personalities and Goals

In order to put in the model only the realistic dependencies between goals and personality traits, we ran a Pearson correlation analysis between the goal and personality scores, and added to the model only those correlations that are substantial and statistically significant. In this section, we first describe how the goals and personality traits were scored. Then, we discuss the correlations found and their influences on the model.

6.2.1.1 Scores of Personality Traits and Goals

For each personality domain, the personality test includes 10 positive items and 10 corresponding negative items. The students were instructed to rate themselves for each item on a scale ranging from 1 (strongly disagree) to 5 (strongly agree). The score for each domain is simply calculated as:

\[
\text{Score} = \frac{\text{Sum of the positive items}}{10} - \frac{\text{Sum of the negative items}}{10}
\]

The revised post questionnaire included 3 questions for each goal to increase the reliability of the students' answers. However, after the study, we realized that some of the questions where phrased so that the answers could depend on some additional factors other than the goals. For example, question #12 in the post questionnaire, "The agent was very helpful to me." was designed to probe the goal Succeed By Myself, but the answer to that question may depend more on the quality of the agent's help than on the student's tendency to work independently. Thus, we did not count these ambiguous questions when we computed the scores for the five goals (these questions are question #7, #9, #12, and #14 in Appendix F). Questions #8, #13, and #15 are the negative questions. We converted their scores before adding them to the corresponding final goal scores. For example, if a student scores 5 (strongly agree) on the negative question #8 "I don't care whether the game is funny or not", the score is converted into 1 (6 - 5),
representing a low score for the goal *Have Fun*. The following are the formulas used to compute the scores for the goals from the related items:

- **Have Fun**: \([\text{question \#1}] + 6 - \text{[question \#8]} + \text{[question \#11]}\)
- **Beat Partner**: \([\text{question \#2}] + \text{[question \#6]}\)
- **Avoid Falling**: \([\text{question \#3}] + 6 - \text{[question \#13]}\)
- **Learn Math**: \([\text{question \#5}] + \text{[question \#10]} + 6 - \text{[question \#15]}\)
- **Succeed By Myself**: \([\text{question \#4}]\)

After computing the scores for the goals and personality traits, we then used a Pearson correlation analysis to verify the relationships between the students’ personality traits and their goals when playing the game. The following subsection discusses the correlations we found.

### Table 6.2: Correlations between Personality Traits and Goals

<table>
<thead>
<tr>
<th></th>
<th>Extraversion</th>
<th>Agreeableness</th>
<th>Conscientiousness</th>
<th>Neuroticism</th>
<th>Openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have Fun</td>
<td>.107</td>
<td>.589*</td>
<td>.528*</td>
<td>-.097</td>
<td>.424</td>
</tr>
<tr>
<td>Learn Math</td>
<td>.650*</td>
<td>.634*</td>
<td>.550*</td>
<td>-.480*</td>
<td>.621*</td>
</tr>
<tr>
<td>Avoid Falling</td>
<td>-.315</td>
<td>-.008</td>
<td>-.087</td>
<td>.529*</td>
<td>-.120</td>
</tr>
<tr>
<td>Succeed By Myself</td>
<td>-.646*</td>
<td>-.482*</td>
<td>-.314</td>
<td>.434</td>
<td>-.312</td>
</tr>
<tr>
<td>Beat Partner</td>
<td>.154</td>
<td>.270</td>
<td>.262</td>
<td>-.317</td>
<td>.344</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).

### 6.2.1.2 Analysis of Correlations and Influences on the Model

In this subsection, we discuss the correlations we found between the students’ personality traits and goals, and their influences on the model for each goal. Table 6.2 summarizes the correlations.
Have Fun

Significant Correlation Found:

• **Agreeableness** was significantly correlated to the goal *Have Fun* \( (r = 0.589, p < 0.01) \).
• **Conscientiousness** was also significantly correlated to the goal *Have Fun* \( (r = 0.528, p < 0.05) \).

The correlation between **Conscientiousness** and *Have Fun* seemed somewhat surprising at first, but it may just be explained by the significant correlation between **Agreeableness** and **Conscientiousness** \( (r = 0.810, p < 0.01) \). This high correlation, as well as the other correlations between personality traits we discuss later in this section, was consistent with the previous findings reported by Graziano, Campbell, and Finch [27].

**Changes in the Model:**

Since there was no significant correlation between **Extroversion** and *Have Fun* (see table 6.2) as we had assumed in the initial version of the model, the link between the two nodes is removed in the revised model. The revised model should then represent the relationships between this goal and the two personality traits **Agreeableness** and **Conscientiousness**. However, because the two personality traits were correlated with each other, various structures are suitable to represent these correlations. We describe how we selected the most appropriate structure in section 6.3.1 after discussing all the other correlations.

Succeed By Myself

Significant Correlation Found:

• **Extroversion** was significantly correlated to *Succeed By Myself* \( (r = -0.646, p < 0.01) \).
• **Agreeableness** was significantly correlated to *Succeed By Myself* \( (r = -0.482, p < 0.05) \).

The negative correlation with **Agreeableness** supports the hypothesis we embedded in the initial model, and fits with the personality trait definition: an agreeable person is "sympathetic to others and eager to help them, and believes that others will be equally helpful in return", whereas a disagreeable person is defined as "egocentric, skeptical of
others' intentions" [15]. The fact that Extraversion, as Agreeableness, is related to interpersonal tendencies, and an extravert is “sociable, liking people, and preferring large groups and gatherings” provides an explanation for the negative correlation between Extraversion and Succeed By Myself. The positive correlation between Extraversion and Agreeableness (r = 0.567, p < 0.05) further supports that explanation.

Changes in the Model:
Since there was no significant correlation between Conscientiousness and Succeed By Myself, the link between these two nodes in the initial model is removed. The revised model should then represent the relationships between this goal and the two personality traits Agreeableness and Extraversion. Because Agreeableness and Extraversion were also correlated, this generates the same problem of multiple plausible structures existing for Have Fun, Agreeableness, and Conscientiousness. We will address it in section 6.3.1.

Avoid Falling

Significant Correlation Found:

- Neuroticism was significantly correlated to the goal Avoid Falling (r = 0.529, p < 0.05).

The interpretation of this result is that the more neurotic a student is, the more she is afraid of falling and making mistakes. This result fits with the general definition of the Neuroticism personality trait, which is “prone to worry, nervous, easy to become upset in face of stressful situations” [15].

Changes in the Model:
The correlation confirms the assumption we made in the first version of the model. Therefore, the link between Neuroticism and Avoid Falling is kept in the revised model.

Learn Math

Significant Correlation Found:
All the five personality traits were found to be significantly correlated with the goal Learn Math (see table 6.2). This result was quite surprising. It may be explained, however, by the correlations among the personality traits. In addition to the correlations
mentioned above among the personality traits Agreeableness, Conscientiousness, and Extraversion, Neuroticism was found to be negatively correlated with Conscientiousness, and Openness was correlated with the other four personality traits.

Changes in the Model:
Although openness was correlated with the goal Learn Math, this correlation can be explained by the correlations between openness and the other four personality traits. Furthermore, including the personality trait openness entails representing its correlations with the other personality traits, which could significantly increase the complexity of the model construction, without adding much information for goal assessment. Therefore, we decided not to include the personality trait Openness in the revised model. The revised model, thus, should represent the correlations between the goal Learn Math and the other four personality traits.

Beat Partner

Significant Correlation Found:
No significant correlation was found between the students’ personality traits and the goal Beat Partner. The initial model assumed that Agreeableness would be negatively correlated with Beat Partner, based on the general definition of the personality trait. A possible explanation is that, in the study, the students were playing with the agent acting as a climbing instructor, not with a student peer, and the agent always played cooperatively and tried to provide help to the students. Thus, even students with a more competitive personality might have felt compelled to cooperate with the climbing instructor. When the model will be extended to assess emotions during the regular play between two peers, a new study will be needed to clarify the relationship between personality and Beat Partner, in this context.

Changes in the Model:
The two links from Extraversion and Agreeableness to Beat Partner in the initial model are then removed.
In addition, we found no statistically significant correlation between factorization knowledge and the goals, and therefore, the link between Factorization Knowledge and Learn Math in the initial model is removed.

6.2.2 Relationships between Students' Goals and Interaction Patterns

In this section, we first describe how we defined the interaction patterns, and then present the results on the relationships we found between student goals and interaction patterns.

For each student, we calculated the following scores related to the corresponding interaction patterns, by analyzing the log files:

- **Percentage of Quick Moves**: The percentage of quick moves made by the student. We used 4-seconds as a threshold to determine whether a move is a quick move. If it takes more than 4 seconds to make a move, that move is considered as a slow move. Otherwise, that move is considered as a quick move. We chose this threshold because the average time spent to make a move by all the students is 4.4 seconds. We also tried 3, 4, and 5-seconds as thresholds, and 4-second generated the best match with the observers' judgment.

- **Percentage of Following Advice**: The percentage of the agent's advice that was followed by the student, given all the advice from the agent. A following advice event occurs when the student (1) uses the magnifying glass immediately after the agent advises to do so, (2) uses the help box immediately after the agent advises to do so, and (3) moves to a number the agent advises to move to.

- **Number of Magnifying Glass Usages**: The number of times the student used the magnifying glass.

- **Number of Help Box Usages**: The number of times the student used the help box.

- **Percentage of Falling**: The percentage of unsuccessful moves the student made.

Pearson correlation analysis was first tried to verify if there were any significant relationships between the goal scores and interaction pattern scores, and also between
factorization knowledge and interaction patterns. However, no statistically significant correlation was found. We also tried ANOVA. For each goal, we divided the students into two groups, depending on whether a student scored high on the corresponding goal or not. ANOVA was then performed to see if there were any statistically significant differences in the average interaction patterns scores between the two groups for each goal. Again, no statistically significant result was obtained.

We then decided to resort to simple cross tables to analyze the data. A cross table allows us to analyze the relationships among categorical variables by displaying the number of cases falling into each combination of the variable categories. In order to use cross tables to analyze the data, we first converted the goals and interaction patterns scores into two categorical values. Next, we scored the cross tables for all the possible two-way combinations of goals with interaction patterns¹, using the Fisher exact test. Fisher's exact test computes the strength of the association between two categorical variables (a chi-square test cannot be used because many cells in the cross tables have less than 5 expected elements). The Fisher's exact test returns a value between 0 and 1, where a value of 1 indicates the virtual absence of association and a value of 0 indicates the strongest possible association. We then included in the revised model only those links connecting goals and interaction patterns that had a cross table with a Fisher's exact value less than 0.4. Before we go into the details of the associations found, we describe how we converted our scores into categorical values.

The interaction pattern scores were converted into two categorical values (True and False) by using the corresponding average score for all the students as a threshold. The value True represents that a student showed the corresponding pattern, while the value False represents that a student did not show the pattern. A student was considered to show an interaction pattern if he scored more than the average on it. Otherwise, the student was considered to not show the corresponding pattern.

¹ Although we did the same for factorization knowledge and interaction patterns, we do not describe the results we found here, because we decided not to include any links between Factorization Knowledge and interaction pattern nodes in the revised model. This decision was made because including these links would not contribute to goal assessment, given the lack of dependencies between the goals and factorization knowledge.
• The average **Percentage of Quick Moves** for all the students was 60%. If more than 60% of a student's moves were quick moves, the student was considered to tend to move quickly. Otherwise, the student was considered to move slowly.

• The average **Percentage of Following Advice** for all the students was 50%. A student was considered to tend to follow advice often if he followed more than 50% of the advice from the agent.

• The average **Number of Magnifying Glass Usages** for all the students was 7.58. A student was considered to use the magnifying glass often if she used the magnifying glass more than 7 times while playing the game.

• The average **Number of Help Box Usages** for all the students was 0.84\(^2\). If a student used the help box at least one time, the student was considered to tend to ask advice often.

• The average **Percentage of Falling** for all the students was 20%. If a student fell more than 20% of all the moves, the student was considered to tend to fall often.

We also converted the goal scores into High and Low values\(^3\). The score of each goal question ranges from 1 to 5, with 3 representing that a student is neutral about the corresponding goal. Thus, if there are \(n\) questions related to a goal, \(3*n\) is used as a threshold to convert the scores into the two categories. For example, since there are three questions assessing the goal *Have Fun*, a score more than 9 on the goal was converted into the High value.

---

\(^2\) Note, although this average is very low, it is consistent with the findings from various studies indicating that students tend not to ask for help even when they need it [2][12].

\(^3\) Before choosing a two-valued categorization for both interaction patterns and goals, we first tried a finer-grained one based on three values, *low, medium, and high*. This would have been especially useful to have a representation of goal priority. However, given the low number of students, the three-valued categorization would have generated cross tables with too few elements per entity to hold any meaningful information.
6.2.2.1 Analysis of the Results and Influences on the Model

In this section, we discuss the associations we found between the students’ goals and interaction patterns, and their influences on the model. We included in the revised model only those links connecting goals and interaction patterns that had a cross table with a Fisher’s score less than 0.4. Table 6.3 summarizes the cross tables that had this property.

Table 6.3: Contingencies between Goals and Interaction Patterns with Fisher’s Score < 0.4

<table>
<thead>
<tr>
<th>Interaction Pattern</th>
<th>Goal</th>
<th>Fisher’s Score</th>
<th>Type of Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move Quickly</td>
<td>Beat Partner</td>
<td>0.38</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Succeed By Myself</td>
<td>0.352</td>
<td>Negative</td>
</tr>
<tr>
<td>Follow Advice</td>
<td>Avoid Falling</td>
<td>0.304</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Succeed By Myself</td>
<td>0.074</td>
<td>Negative</td>
</tr>
<tr>
<td>Ask Advice</td>
<td>Succeed By Myself</td>
<td>0.38</td>
<td>Negative</td>
</tr>
<tr>
<td>Use Magnifying Glass Often</td>
<td>Learn Math</td>
<td>0.35</td>
<td>Positive</td>
</tr>
<tr>
<td></td>
<td>Beat Partner</td>
<td>0.057</td>
<td>Positive</td>
</tr>
<tr>
<td>Fall Often</td>
<td>Have Fun</td>
<td>0.26</td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td>Learn Math</td>
<td>0.177</td>
<td>Negative</td>
</tr>
<tr>
<td></td>
<td>Succeed By Myself</td>
<td>0.181</td>
<td>Negative</td>
</tr>
</tbody>
</table>

**Move Quickly**

**Relationships with Goals:**

The interaction pattern Move Quickly was found to be influenced by two goals: Beat Partner and Succeed By Myself. Tables 6.4 and 6.5 show the corresponding cross tables.

Table 6.4: Cross Table for Move Quickly and Beat Partner

<table>
<thead>
<tr>
<th>Beat Partner</th>
<th>Move Quickly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 6.5: Cross Table for Move Quickly and Succeed By Myself

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Move Quickly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>2</td>
</tr>
<tr>
<td>High</td>
<td>6</td>
</tr>
</tbody>
</table>

As Table 6.4 shows, it was more likely for the students to move quickly if they had high score for the goal Beat Partner. Of the twelve students who scored high on the goal Beat Partner, 67% of them showed a tendency to move quickly. Table 6.5 shows that it was more likely for the students to move quickly if they scored low on the goal Succeed By Myself. Of the eight students who had low score for the goal Succeed By Myself, 75% of them tended to move quickly. Table G.1 (in Appendix G) is the full cross table for Move Quickly given Beat Partner and Succeed By Myself.

Changes in the Model:
The initial model assumed that Move Quickly would be influenced by the three goals: Have Fun, Beat Partner, and Avoid Falling. As the two goals Have Fun and Avoid Falling were not found to be associated with this pattern, the links from the two goals to the pattern are then removed. The revised model now includes links from Beat Partner and Succeed By Myself to Move Quickly.

Follow Advice
Relationships with Goals:
Two goals, Succeed By Myself and Avoid Falling, were found to have an influence on the interaction pattern Follow Advice. Tables 6.6 and 6.7 show the corresponding cross tables. Table 6.6 shows that it was more likely for the students to ignore the agent’s advice if they had high scores for the goal Succeed By Myself. Among the eleven students who scored high on the goal Succeed By Myself, 82% of them tended to ignore the agent’s advice. Furthermore, it was slightly more likely for the students to follow the agent’s advice if they scored low on the goal Succeed By Myself. Of the eight students who had low scores for that goal, 63% of them tended to follow the agent’s advice. Table 6.7 shows that the students were more likely to ignore the agent’s advice if they had low scores for the goal Avoid Falling. Of the fourteen students who scored
low on the goal Avoid Falling, 71% of them tended to ignore the agent’s advice. Table G.2 (in Appendix G) is the full cross table for Follow Advice given Succeed By Myself and Avoid Falling.

Table 6.6: Cross Table for Follow Advice and Succeed By Myself

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Follow Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>3</td>
</tr>
<tr>
<td>High</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 6.7: Cross Table for Follow Advice and Avoid Falling

<table>
<thead>
<tr>
<th>Avoid Falling</th>
<th>Follow Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>10</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
</tr>
</tbody>
</table>

Changes in the Model:
The initial model assumed that only the goal Succeed By Myself had an influence on the pattern Follow Advice. The link from Avoid Falling to Follow Advice is then added to the revised model.

Ask Advice

Relationships with Goals:
The initial model assumed that Ask Advice would be influenced by three goals: Avoid Falling, Learn math, and Succeed By Myself. However, only the relationship with Succeed By Myself was supported by the results. Table 6.8 shows that the students were unlikely to ask for advice if they scored high for the goal Succeed By Myself. Out of the eleven students who scored high on the goal Succeed By Myself, 73% of them did not ask the agent for advice. If the students scored low on that goal, they were equally likely to ask for advice as not.

Changes in the Model:
The revised model now only includes the link from Succeed By Myself to Ask Advice.
Table 6.8: Cross Table for *Ask Advice* and *Succeed By Myself*

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Ask Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>8</td>
</tr>
</tbody>
</table>

**Use Magnifying Glass Often**

*Relationships with Goals:*

The interaction pattern *Use Magnifying Glass Often* was found to be influenced by two goals *Learn Math* and *Beat Partner*. As Table 6.9 shows, the students who had high scores for the goal *Learn Math* were slightly more likely to access the magnifying glass than the students who had low scores for the goal. Of the six students who scored high on the goal *Learn Math*, 67% of them tended to use the magnifying glass often, whereas only 38% of the students who scored low on the goal displayed the pattern. Table 6.10 shows that it was more likely for the students who had high scores for the goal *Beat Partner* to use the magnifying glass than the students who had low scores for the goal. Out of the twelve students who scored high on the goal *Beat Partner*, 67% of them tended to use the magnifying glass often, whereas only 14% of the students who scored low on the goal showed the pattern. Table G.4 (in Appendix G) is the full cross table for *Use Magnifying Glass Often* given *Learn Math* and *Beat Partner*.

**Table 6.9: Cross Table for *Use Magnifying Glass Often* and *Learn Math***

<table>
<thead>
<tr>
<th></th>
<th>Use Magnifying Glass Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>8</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 6.10: Cross Table for *Use Magnifying Glass Often* and *Beat Partner***

<table>
<thead>
<tr>
<th>Beat Partner</th>
<th>Use Magnifying Glass Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>6</td>
</tr>
<tr>
<td>High</td>
<td>4</td>
</tr>
</tbody>
</table>

**Changes to the Model:**

Since there was no relationship between the goal *Avoid Falling* and the pattern *Use Magnifying Glass Often*, the link between the two nodes is removed in the revised
model. The revised model includes two links from Beat Partner and Succeed By Myself to Use Magnifying Glass Often.

**Fall Often:**

*Relationships with the Goals:*

The pattern Fall Often was found to be influenced by three goals: Have Fun, Learn Math, and Succeed By Myself. Tables 6.11, 6.12, and 6.13 show the corresponding cross tables. As Table 6.11 shows, it was less likely for the students who had high scores for the goal Have Fun to fall often than the students who had low scores for the goal. Of the fifteen students who scored high on the goal Have Fun, only 33% of them tended to fall often, whereas 75% of the students who scored low on the goal showed a tendency to fall often. As Table 6.12 shows, it was unlikely for the students to fall often if they had high scores for the goal Learn Math. Of the six students who scored high on the goal Learn Math, only 17% of them tended to fall often. Table 6.13 shows that it was less likely for the students with high scores for the goal Succeed By Myself to fall often than for the students who had low scores for the goal. Table G.5 (in Appendix G) is the full cross table for Fall Often given Have Fun, Learn Math, and Succeed By Myself.

**Table 6.11: Cross Table for Fall Often and Have Fun**

<table>
<thead>
<tr>
<th>Have Fun</th>
<th>Fall Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>10</td>
</tr>
</tbody>
</table>

**Table 6.12: Cross Table for Fall Often and Learn Math**

<table>
<thead>
<tr>
<th>Learn Math</th>
<th>Fall Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>6</td>
</tr>
<tr>
<td>High</td>
<td>5</td>
</tr>
</tbody>
</table>

---

4 This was the opposite of the relationship we had postulated in the initial model, after observing that a few students fell on purpose because they liked the resulting game animations.
Table 6.13: Cross Table for Fall Often and Succeed By Myself

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Fall Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Low</td>
<td>3</td>
</tr>
<tr>
<td>High</td>
<td>8</td>
</tr>
</tbody>
</table>

Changes to the Model:

As a result, the revised model includes links from the goals Have Fun, Learn Math, and Succeed By Myself to the pattern Fall Often. The semantics of the link between Have Fun and Fall Often has changed compared with the link in the initial model because of the negative relationship between the two variables we found in the third study.

In the next section, we describe in more detail how we revised the initial model based on the results discussed in this section.

6.3 Changes to the Network Assessing Students’ Goals

Based on the results of the third study, the part of the initial model that assesses the students’ goals was then revised. This includes both the structure and the conditional probability tables. Section 6.3.1 discusses the changes that represent the correlations we found between personality and goals. Section 6.3.2 discusses the changes that represent the relationships between goals and interaction patterns.

6.3.1 Personality and Goals

This section describes how the revised model represents the correlations between personality and goals. It first describes how the values of the personality nodes are set, then discusses the technique used to revise the structure, and finally presents some example CPTs.

We converted the personality scores into True and False values, with the value True representing that the student belonged to one end of the corresponding personality domain, and False representing that the student belonged to the opposite end (e.g., extrovert vs. introvert, agreeable vs. disagreeable). The score of each personality domain ranges from -40 to +40. We used zero as a threshold to convert the scores into...
the two categories. For a given personality domain, the positive scores were converted into True value, and non-positive scores were converted into False value. As described earlier in Section 6.2.2, goal scores were also converted into two values High and Low, representing whether a student scored high or low on the corresponding goal.

6.3.1.1 Structure

In order to find an appropriate structure that explains all the correlations found among the personality traits and goals, we proceeded as follows. We first enumerated the relevant structures that could explain the correlations among Agreeableness, Extraversion, Conscientiousness, Have Fun, and Succeed By Myself, and scored them using their log marginal likelihood [29]. Figure 6.1 shows eight of these structures that had the best scores, with “A” standing for Agreeableness, “E” standing for Extraversion, “C” standing for Conscientiousness, “HF” standing for Have Fun, and “SBM” standing for Succeed By Myself. Of all the relevant structures we scored, structures 1 and 2 in Figure 6.1 generated the highest score. We ran cross validations to verify how sensitive this result was to the data, using data subsets obtained by removing one case first, then two, and then three different cases. Figure 6.2 shows the cross validation result for the eight high score structures in Figure 6.1, obtained by removing two different cases each time (note that because structure 1 performs the same as structure 2, there is only one line representing both structure 1 and 2. The same applies to the other pairs). Structure 1 and 2 performed equally well and always outperformed the other six structures. We then decided to select structure 1 because the direct connection between Have Fun and Conscientiousness in structure 2 is not obviously supported by the definition of Conscientiousness in the Five Factor Model.

Based on structure 1 in Figure 6.1, we further generated and scored the relevant structures that include the additional correlations we found between the personality traits and the two goals Learn Math and Avoid Falling. Of the three structures that ranked highest (shown in Figure 6.3), structure 1 generated the highest score. As we did for the networks described above, we ran cross validations to check the sensitivity of the result to the data. Figure 6.4 shows the cross validation result obtained by removing two
different cases each time. Structure 1 in Figure 6.3 always outperformed the other two in all the runs.

We realize that this is a “greedy”, not completely sound approach to define the network structure from the data, because the log marginal likelihood measure is not additive over subparts of a given network. The correct approach would be to feed the complete data on goals, personality, and interaction patterns to an algorithm for structure learning [29]. However, 19 data points are clearly not enough to reliably learn the structure of a network with 14 variables, and therefore, a “greedy” approach that also exploits the semantics of the variables involved is the only way to make sensible use of our data.

Figure 6.1: Eight Possible Structures for Agreeableness, Conscientiousness, Extraversion, Have Fun, and Succeed By Myself
Figure 6.2: Cross Validation of the Eight Structures

![Diagram of eight structures with corresponding log marginal likelihood values: -28, -30, -32, -34, -36, -38, -40, -42.]

Figure 6.3: Three Possible Structures for Personality and Goals, Structure 1 is the Final Selected Structure

![Diagram of three structures with corresponding log marginal likelihood values: Score -8633 for Structure 1, Score -87.20 for Structure 2, Score -8683 for Structure 3.]

Figure 6.4: Cross Validation of Three Structures for Personality and Goals
6.3.1.2 CPTs for Goals

In this section, we describe how we computed the conditional probability tables (CPTs) for the goals given personality traits.

As described earlier, we converted the personality scores into two categorical values (True and False), with the value True representing that a student belongs to one end of the corresponding personality domain (e.g., extraversion or agreeableness), and False representing that a student belongs to the opposite end (e.g., introversion or disagreeableness). Goal scores were also converted into binary values (High and Low) as described earlier. We then constructed a cross table for each pair of linked personality traits and goals, which in turn was translated into the corresponding CPT by computing the relative frequencies.

However, because of the limited amount of available data, using pure frequencies to compute the CPTs would generate several very sparse tables. We used a simple add-one smoothing algorithm to address this problem. For those cross tables containing zero frequencies, we added one to each entry before converting them to the CPTs. The same technique was applied when we computed the CPTs for the interaction pattern nodes.

The add-one smoothing technique is to pretend each event occurs once more than it actually did, and thus, has the effect of giving some probability space to unseen events. It is actually the Bayesian estimator that one derives if one assumes a uniform prior on events. However, for very sparse data and a large number of unseen events, the add-one smoothing technique has a problem of overestimating the unseen events [22]. To overcome this overestimation problem, one solution is to add a smaller positive value instead of one. But it is always difficult to find an appropriate value [47]. There are also some other sophisticated smoothing techniques mostly used in the domain of language modeling, such as Good-Turing estimation [24], Absolute discounting [53], Simple linear interpolation [34], and Katz back-off smoothing [39]. One possible extension of the work is thus to apply different smoothing techniques and examine the sensitivity of our model's performance to them.

For a given goal (Goal), which is influenced by a personality trait (Trait), the CPT for that goal is computed as:
\[ P(\text{Goal} = x \mid \text{Trait} = y) = \frac{m}{n}; \]

Where \( x, y \) are the corresponding binary values; \( m \) is the number of students whose Goal value is \( x \) and Trait value is \( y \); and \( n \) is the number of students whose Trait value is \( y \). Note that, for those cross tables containing zero frequencies, \( m \) and \( n \) are the numbers after adding one to each entry.

For example, Table 6.14 shows the number of students falling into each combination of the categories of the two variables: Neuroticism and Avoid Falling. The number of students who were not neurotic is 16 (13+3), and among them 13 students scored low on the goal Avoid Falling. Therefore,

\[ P(\text{Avoid Falling} = \text{Low} \mid \text{Neuroticism} = \text{False}) = 0.81 \ (13/16). \]

Table 6.15 shows the complete CPT for the goal Avoid Falling given Neuroticism.

<table>
<thead>
<tr>
<th>Neuroticism</th>
<th>Avoid Falling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>False</td>
<td>13</td>
</tr>
<tr>
<td>True</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neuroticism</th>
<th>Avoid Falling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>False</td>
<td>0.81</td>
</tr>
<tr>
<td>True</td>
<td>0.33</td>
</tr>
</tbody>
</table>

### 6.3.2 Goals and Interaction Patterns

#### 6.3.2.1 Structure

As described in section 6.2.2.1, we included in the revised model only those links connecting goals and interaction patterns that had a cross table with a Fisher's score less than 0.4. To verify that the new links actually generate a better model, we compared the log marginal likelihood of the revised sub-model for goals, personality, and interaction patterns (see Figure 6.5), the initial model (see Figure 5.2), and a model we obtained from the revised model by removing some of the links that seem less intuitive to us (these were the two links from Succeed By Myself to Fall Often and Move Quickly). The
revised model with all the links suggested by the Fisher’s scores generated the highest score. To check the sensitivity of the results to the data, we also run cross validations following the same way described in section 6.3.1.1. Figure 6.6 shows the result of the cross validation by removing two different cases from the data set each time, with struct1 representing the initial model, struct2 representing the revised model based on the Fisher’s scores, and struct3 representing the model based on the revised model but without the less intuitive links. The revised model outperformed the other two models in all the runs.

Student's Personality Type

Student's Goals

Interaction Patterns

Individual Actions

Figure 6.5: The Revised Sub-model to Assess Students’ Goals

Figure 6.6: Cross Validation of the Three Sub-Models Assessing Goals
6.3.2.2 CPTs for Interaction Patterns

In this section, we describe how the CPTs for the interaction patterns given the goals were computed.

As described in section 6.2.2, the interaction pattern scores were converted into True and False values, representing whether a student displayed the corresponding interaction pattern or not. The CPTs for the interaction patterns were constructed following essentially the same method we used to build the CPTs for the goals. Suppose that a given interaction pattern (Pattern) is influenced by a goal (Goal). Then the CPT for that interaction pattern is computed as:

\[ P(\text{Pattern} = x \mid \text{Goal} = y) = \frac{m}{n}; \]

Where \( x, y \) are the corresponding binary values; \( m \) is the number of the students whose Pattern value is \( x \) and Goal value is \( y \); and \( n \) is the number of students whose Goal value is \( y \). Note, for those cross tables containing zero frequencies, \( m \) and \( n \) are the numbers after adding one to each entry.

For example, Table 6.8 shows the number of students falling into each combination of the categories of the two variables: Succeed By Myself and Ask Advice. The number of students who scored high on the goal Succeed By Myself is 11 (8+3), and among them 8 students tended not to ask advice. Therefore,

\[ P(\text{Ask advice} = \text{False} \mid \text{Succeed By Myself} = \text{High}) = 0.73 \text{ (8/11)}. \]

Table 6.16 shows the complete CPT for the interaction pattern Ask Advice given the goal Succeed By Myself.

Table 6.16: CPT for Ask Advice Given Succeed By Myself

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Ask Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>0.5</td>
</tr>
<tr>
<td>High</td>
<td>0.73</td>
</tr>
</tbody>
</table>
6.3.3 Interaction Pattern and Individual Actions

6.3.3.1 Structure

As the initial model, the revised model also includes an individual action node corresponding to each interaction pattern in each time slice (see Figure 6.5). Evidence on the individual actions can then be propagated upward to the network to update the probability of related interaction patterns and goals.

6.3.3.2 CPTs for Individual Actions

This section describes how we constructed the CPTs for the individual actions given the corresponding interaction patterns.

Each individual action node also has two values, True and False, representing whether a student performed the corresponding action or not. The CPT for an individual action (Action) given an interaction pattern (Pattern) is computed as:

\[ P(\text{Action} = x | \text{Pattern} = y) = \frac{m}{n}, \]

- \( x \) and \( y \) are the corresponding Boolean values;
- \( m \) is the total number of related actions performed by the group of students whose Pattern value is \( y \);
- \( n \) is the total number of moves for the group of students whose pattern value is \( y \), if Action is Quick Move, Magnifying Glass Used, or Advice Asked; or \( n \) is the total number of advice provided by the agent if Action is Advice Followed.

For example, Table 6.17 shows the CPT for the individual action node Quick Move. For the group of students who tended to move quickly, the total number of moves is 1937, and the total number of quick moves is 1475. Thus,

\[ P(\text{Quick Move} = \text{True} | \text{Move Quickly} = \text{True}) = 0.76 \left( \frac{1475}{1937} \right) \]

Table 6.17: CPT for Quick Move Given Move Quickly

<table>
<thead>
<tr>
<th>Move Quickly</th>
<th>Quick Move</th>
</tr>
</thead>
<tbody>
<tr>
<td>False</td>
<td>0.54</td>
</tr>
<tr>
<td>True</td>
<td>0.46</td>
</tr>
<tr>
<td>False</td>
<td>0.24</td>
</tr>
<tr>
<td>True</td>
<td>0.76</td>
</tr>
</tbody>
</table>
6.4 Changes to High Level Structure of the Affective Model

This section discusses the problem of the loss of information in the complete new model, and presents our solution to the problem.

6.4.1 Loss of Information

As in the old model, in the new model, a new time slice will be added to the network when either the student makes a move or the agent provides an intervention. To keep the network at a manageable size, only two time slices are maintained at any given time by applying the rollup mechanism [59]. Every time we make a rollup in the network, we set the prior probabilities of the root nodes at time slice \( t_{i+1} \) as the posterior probabilities of the corresponding root nodes at time slice \( t_i \) (see Figure 5.1). However, because we only have two root personality nodes and one root goal node in the network, the rollup will not keep the probabilities of the other personality and goal nodes at time slice \( t_{i+1} \) the same as the posterior probabilities of the corresponding nodes at time slice \( t_i \). Therefore, this will cause a problem of the loss of information on some of the personality nodes and on all but one goal node (Schafter and Weyrath [60] discuss the problem in detail). Because information on the Goals nodes is very important to assess student emotions in our model, such loss of information could make the model’s assessment of emotions significantly less accurate.

One possible solution to the problem is to model the static Goals nodes as dynamic nodes, as suggested by Schafter and Weyrath in [60]. The Goals nodes at time slice \( t_{i+1} \) can then be modeled as been influenced by the Personality nodes at time slice \( t_{i+1} \), and the Personality and Goals nodes at time slice \( t_i \) (see Figure 6.7). The links between Personality nodes at \( t_i \) and Goals nodes at \( t_{i+1} \) are included in order not to count the effects of Personality to Goals twice. However, the CPTs for defining such dependencies can be difficult to specify. We propose an alternative approach, discussed in the next section.
6.4.2 Two Phases of Assessment

To address the problem of the loss of information on the goals, we separate the final network for our affective model into two parts: a static network that assesses student goals (see Figure 6.8), and a dynamic network that assesses student emotions (see Figure 6.9). We can do so because we have the assumption that all the variables in the static network do not change over time.

The static network is a classic Bayesian network in which student action nodes are incrementally added and always kept around. Thus, there will be no loss of information on the student's goals. Every time a student makes an action, a new individual action node is added to the static network, and clamped to True or False. Figure 6.8 shows a snapshot of the high level structure of the static network at a given time. The network contains three Individual Action nodes, indicating that the student has made three actions so far.
In the dynamic network, every time a student makes a move or the agent provides an intervention, a new time slice will be added. The assessment of emotions thus involves two phases:

- Phase I: If the student performs an action, a new action node is added to the static network and clamped to True or False. The belief on the student’s goals is then updated.
- Phase II: If the student makes a move or the agent provides an intervention, a new time slice is added to the dynamic network. The relevant outcomes of the student move (e.g., if the student is ahead of the partner, if the move is correct) or the agent’s action are added as evidence nodes to the network. The priori probabilities of the Goals nodes in the new time slice are set as the posterior probabilities of the corresponding Goals nodes in the static network. The belief on the student’s emotions is then updated.

5 Only student moving actions are added to this network, because in our current model, these are the only actions that affect the student emotional states.
One drawback of this approach is that inference on the student’s goals could become costly as the more and more *Individual Action* nodes are added to the network and the network grows in size. For example, in a typical interaction with the game Prime Climb, a student can make as many as 200 moves. The static network can grow to have around 400 nodes. It can then become expensive to perform the assessment of student goals with an exact update algorithm. Figure 6.10 shows the result of an experiment to see the time spent to make the inference on student goals as the new action nodes are added to the network. The experiment was performed on a windows machine with an Intel Celeron 600MHZ Processor and 160M RAM, using the Microsoft MSBNx package [38]. The X-axis represents the number of moves made by the student so far. The Y-axis represents the time spent to make the inference on the goals. As shown in Figure 6.10, in the first 100 moves, the network spent less than 2 seconds performing the inference. However, as the network continued to grow in size, the time increased quite rapidly. In the last 20 moves, it took approximately 20-30 seconds to perform the assessment. The jump in Figure 6.10 (from 10 seconds to 19 seconds) was because the system started memory swapping at that point of time. Because online assessment is important in our domain, we plan to further investigate this issue in the future.

![Figure 6.10: Time Spent to Make the Inference on the Goals](image-url)
A second drawback of this approach is that it does not allow to soundly integrate downward assessment from goals to emotions and upward assessment from possible symptoms to emotions. This is not a problem at the moment, since our model does not actually take into account any evidence on emotion expressions. But the issue will need to be addressed when the model will be extended to take symptomatic information into account.

In the next chapter, we provide an example of the model's assessment, given the behavior of a simulated student.
As described earlier, our affective model can leverage any information available (interaction patterns, both interaction patterns and personality, or personality only) to perform the probabilistic assessment of student goals and consequently of student emotions. We show here a simple simulation to exemplify how the model works.

We first create a sequence of game interactions in which the simulated student: (i) always moves slowly, (ii) always moves successfully, and (iii) is always ahead of the partner. A few times during the interaction, the agent advises the student to use the magnifying glass to see the number factorization (these times are marked by the asterisks on the bottom axis in Figures 7.1 and 7.2) and the student always ignores the advice.

We then show how the model: (1) assesses the student’s goals and emotions from the interaction patterns only, (2) assesses the student’s goals and emotions from both the interaction patterns and personality traits.

7.1 Assessment from Actions Only

Figure 7.1 shows the model’s assessment of the student’s interaction patterns, goals, and emotional states by assuming that information on the student’s personality traits is not available (i.e., the priors of the root personality nodes are assigned to be 0.5 at the beginning of the interaction).
Figure 7.1: Assessment Without Personality Information
7.1.1 Assessment of Goals

The middle two charts in Figure 7.1 show the model's assessment of the student's goals. Because the student always moves slowly, the probability of the goal Beat Partner slowly decreases over time.

The goal Succeed By Myself has negative relationships with three interaction patterns Fall Often, Follow Advice, and Move Quickly (see Table 6.3). The probability of Move Quickly decreases because the student always moves slowly; the probability of Fall Often also decreases because the student always moves successfully; the probability of Follow Advice decreases every time the student ignores the agent's advice (see the top chart in Figure 7.1). Therefore, the goal Succeed By Myself receives evidence from all three interaction patterns. The probability of Succeed By Myself increases continuously.

The goal Avoid Falling has a positive relationship with the interaction pattern Follow Advice (see Table 6.3). Therefore, every time the student ignores the agent's advice, the probability of Avoid Falling decreases.

Both the goals Have Fun and Learn Math have negative relationships with the interaction pattern Fall Often (see Table 6.3). Since the probability of Fall Often decreases over time, the probability of the goal Have Fun and Learn Math slightly increases.

7.1.2 Assessment of Emotions

The assessment of the student's emotional states is shown in the bottom chart in Figure 7.1. The probability of Reproach increases every time the agent intervenes, because the intervention interferes with the goal Succeed By Myself (the model currently assumes that the agent's advice does not affect the goal Have Fun). The probability of Reproach slowly decreases over the time slices when the agent lets the student be.

Except when the agent intervenes, the goal Succeed By Myself is satisfied in most of the time slices because the student always moves successfully without following the agent's advice. Thus, the probability of Pride constantly increases, while the probability of Joy increases in all but those time slices when the agent provides advice to use the magnifying glass.
7.2 Assessment With Personality Information

Suppose now that we know the student is an introvert and neurotic person. Figure 7.2 shows how this information can influence the model’s assessment of the student’s goals (the top and middle charts in Figure 7.2) and emotions (the bottom chart in Figure 7.2).

Figure 7.2: Assessment With Personality Information
As Table 6.2 shows, Extraversion is negatively correlated with the goal Succeed By Myself. Since the student is an introvert, the network believes that it is likely for the student to have the goal Succeed By Myself. Furthermore, as described earlier, this goal also receives supporting evidence from the three interaction patterns Fall Often, Follow Advice, and Move Quickly. Therefore, the network becomes more certain that the student is likely to have this goal (see the top chart in Figure 7.2).

Since there is no personality trait influencing the goal Beat Partner, this goal only receives evidence from the interaction pattern Move Quickly. As a result, the assessment of this goal is the same as before (see the top chart in Figure 7.2). The probability of Beat Partner decreases over time because the student always moves slowly.

As Table 6.2 shows, Neuroticism is positively correlated with the goal Avoid Falling. Knowing that the student is neurotic makes the network first believe that the student is likely to have the goal Avoid Falling. However, because the student keeps ignoring the agent’s advice, which provides conflicting evidence of this goal, the network then becomes less certain that the student has this goal (see the middle chart in Figure 7.2).

Given that Agreeableness and Extraversion are positively correlated with each other, knowing that the student is an introvert makes the network believe that the student is unlikely to be an agreeable person. Furthermore, as Table 6.2 shows, Agreeableness is positively correlated with the goal Have Fun. Therefore, the network first believes that the student is unlikely to have the goal Have Fun. However, because the student always moves successfully, which provides supporting evidence of this goal, the network then becomes uncertain of whether the student has this goal or not (see the middle chart in Figure 7.2).

Given that it is unlikely for the student to be an agreeable person, and Agreeableness and Conscientiousness are also positively correlated, it is also unlikely for the student to be a conscientious person. Furthermore, Conscientiousness is positively correlated with the goal Learn Math, and Neuroticism is negatively correlated with this goal. Knowing that the student is a neurotic person and unlikely to be a conscientious person makes the network first believe that it is unlikely for the student to
have the goal *Learn Math*. However, moving successfully provides positive evidence for this goal. Therefore, the probability of the goal *Learn Math* slightly increases over time.

Because the network is now more certain that the student has the goal *Succeed By Myself*, the probability of *Reproach* increases slightly more in those time slices when the agent intervenes.

---

6 Note that because the conditional probability table for the goal *Learn Math* is one of the sparsest we had, the probability of this goal when the personality information is available (the middle chart in Figure 7.2) is actually slightly higher than the probability without personality information (the third chart in Figure 7.1). We plan to refine the conditional probability tables by running additional studies.
8.1 Conclusions

This thesis presents research on building an affective student model that assesses students' emotional states during the interaction with an educational game. Because assessing users' emotions generally involves a high level of uncertainty, the affective model illustrated in this thesis relies on Dynamic Decision Networks (DDNs) to explicitly represent the probabilistic relationships underlying the connections between situations, students' traits and emotions, as well as the temporal evolution of emotions. By relying on DDNs, the model leverages any evidence available to make the probabilistic assessment and can be extended to provide a decision theoretic basis for the pedagogical agent's behavior.

The affective model is based on the OCC cognitive theory of emotions [54], which models the cognitive appraisal process of emotions. According to the OCC model, emotions are the outcome of the cognitive appraisal of the current situation, which depends on how a situation fits with one's goals and preferences. By implementing the cognitive appraisal process of emotions, the affective model is able to detect the causes of emotions, in addition to assessing the emotions themselves. This is important if we want to use the model to adequately respond to and possibly improve the user's emotional state.

Since goals are a key element in the application of the OCC theory to the assessment of students' emotions during the interaction with the educational game, the affective model includes a structure to facilitate the assessment of students' goals. The model includes variables such as personality traits and interaction patterns, which provide evidence on students' goals.
Our affective model is the first model that uses the OCC theory to recognize emotions instead of generating them, and that explicitly takes into account the sources of uncertainty underlying the use of the OCC model for emotion recognition. Also, it is the first affective model that tries to recognize a variety of emotions in a quite unrestrictive setting.

Because of its modular structure, the affective model can be fairly easily extended to incorporate new variables describing the effects of emotions (e.g., physiological signals, vocal intonation, or facial expressions), thus combining causes and effects of emotions to perform both predictive and diagnostic assessment, depending on the available evidence.

Although the model was built specifically for the Prime Climb educational game environment, its high level design implementing the OCC theory can be applied to many other environments. This would require identifying users' goals, interaction patterns, the relationships between users' goals and interaction patterns, the relationships between users' traits and goals, and also the emotion eliciting situations.

The model was built and refined based on several user studies. The two preliminary user studies gave us information on the relevant students' goals, interaction patterns, and emotion eliciting situations. The initial version of the model was then built based on the information from the preliminary studies, our intuitions, and psychological findings. The third user study gave us further information on the relevant students' personality traits, the relationships between students' personality traits and goals, and the relationships between goals and interaction patterns. The part of the network that assesses student goals was then refined based on the results of the study.

8.2 Limitations and Future Work

Although the model was iteratively designed based on several user studies, more data are needed to further refine the structure and conditional probability tables in the model. For instance, due to the small sample size of the third study, we could not find any statistically significant correlations between the students' goals and interaction patterns, and we had to rely on weaker measures of association to identify the relationships that
we wanted to encode in the model. Therefore, the relationships between goals and interaction patterns are subjected to further investigation. In addition, the conditional probability tables we derived from the study can be inaccurate because of the small data set. The conditional probability tables can be learned incrementally. After a student has played the game, the system knows the student’s interaction patterns. The knowledge about the student's goals and personality traits can also be acquired by administering the questionnaires. Thus, we have one more data point. The conditional probability tables can then be refined after taking into account the new data point. Therefore, we can just put the game into classrooms and have students play the game. Potentially, the conditional probability tables can get more and more accurate as more students play the game.

A formal evaluation is needed to evaluate the effectiveness of the affective model. An evaluation that directly gauges the accuracy of the model would require having a measure of a student’s emotions during the interaction and then compare this measure with the model’s assessment. The measure of student emotions could be acquired either by using human judges, or by asking the students to input their emotional states at various times during game playing (a similar technique has been used to measure student motivation in [17]). The challenge of the first approach is that it may be difficult for the judges to provide an accurate assessment by relying only on visible student behaviors. The challenge of the second approach is to find a non-intrusive way for the students to report their own emotions. An indirect evaluation of the model could be performed after the pedagogical agent that uses the model’s assessment to guide its interventions is implemented. The evaluation would involve an experimental group that receives feedback tailored to the assessment of the affective model and a control group that receives feedback without considering students’ emotions. The effectiveness of the two types of feedback could then be compared using the differences in learning, game performance, and subject view of the agent’s help between the two groups.

The model currently does not consider the priorities among the goals. The goal nodes have only two values representing whether a student had a high or low score for the corresponding goal. However, students can have multiple goals and each goal can have its own weight. For instance, a student can have both the goal *Beat Partner* and the
goal *Avoid Falling*, but *Beat Partner* might be more important for that student. Multi-valued goal nodes would allow the model to represent goal priority. For instance, numerical values such as 1 to n can be used to represent different levels of importance of each goal for a student. This would also allow the model to represent the intensity of emotion, which is not currently modeled, as a function of goal priority. However, building a network with multi-valued nodes is difficult, as more data are needed to reliably set up the corresponding conditional probability tables.

The model assumes that students' goals are constant throughout the game, and thus it does not model the influence of students' emotions on these goals and on the situation appraisal processes. However, there is increasing evidence that emotions can affect one's cognitive states, as well as the situation appraisal process. For instance, it might be possible for a student to become more focused on the goal *Avoid Falling* and on the negative aspects of the game situations, if the student continuously makes mistakes and becomes increasingly frustrated. More user studies are needed to verify how emotions influence goals over time in Prime Climb, and whether this influence is strong enough to warrant being represented in the model.

It is important to consider both causes and effects of emotions to reduce the high level of uncertainty involved in affective user modeling. The model currently focuses on the causes of emotions by implementing the OCC cognitive theory of emotions. Extending the model to include evidence coming from bodily expressions, such as heart rate, skin conductance, and eyebrows position, will require user studies to better understand the relationships between students' emotional states and their bodily expressions in the game interactions, as well as research on how biometric sensors can detect these bodily expressions. One argument is that the Prime Climb game interaction might not be intense enough to elicit the detectable physiological signals such as heart rate and skin conductance. We plan to conduct user studies to further investigate this issue.

The user studies and the model we have presented target the Prime Climb practice phase in which students play with the climbing instructor instead of a peer student. The next step will be to develop a model for the phase in which students play together, and each student has his or her own agent to provide help. This model will be
considerably more complex, since it will need to represent the emotions toward the peer in addition to the emotions toward the agent and self.

The ultimate goal of this line of research is to build an intelligent pedagogical agent that decides when and how to act by considering both the students' emotional states and math knowledge. This requires integrating the affective model with (i) a model that tracks the evolution of the students' number factorization knowledge, (ii) a more sophisticated set of agent's actions, and (iii) the agent's utilities that take into account information on both student learning and emotional state.


[24] I.J. Good. The Population frequencies of Species and the Estimation of

G. Buchanan and E.H. Shortliffe (Eds), *Rule-based expert systems: The MYCIN
experiments of the Stanford Heuristic Programming Project*. Reading: MA:
Addison-Wesley, 1984.

words in self-reported personality: Lethargic and provincial in Texas. *Personality

Between Personality and Adjustment. *Journal of Personality and Social


Making. *Proceedings of Eleventh Conference on Uncertainty in Artificial

Project: Bayesian User Modeling for Inferring the Goals and Needs of Software
Users. *Proceedings of the Fourteenth Conference on Uncertainty in Artificial


Hi, firstly, welcome and thank you for helping us out today. My name is ____. I am from the E-GEMS Research group at UBC computer science department, and we make educational games to teach kids math and science. We are here today for you to help us with a game we are making called Prime Climb.

The purpose of the experiment is not to test you in any kind of way, we just need your help to make the game better, so maybe it could be used in your classroom one day. So, don’t worry about it too much, if you don’t know something.

Here are some rules of the game.
1. There is a mountain of numbers you have to climb, the object of the game is to climb to the top of the mountain with your partner. Today your partner will be the computer.
2. If you partner is on a number and you move to a number that has a shared common factor with the number your partner is on, you will fall.
3. You can only move to the neighbouring hexagons and you can only have a maximum of one number between you and your partner, so work with your partner.
4. You don’t have to wait for turns. When you want to move, just click (not drag) the number which you want to move to.
5. You are the highlighted player.

Also there are some tools to help you get to the top, which are located on the palm pilot.
1. Magnifying glass—if you click the magnifying glass, the cursor will change, if you then click a number, a factor tree will appear for that number, which will help you make decisions on where to move. When you don’t need it, click the button again to continue your move.
2. Help box—you can use the help box to talk to the computer. You can ask the computer anything you want.
3. Agent Merlin—who is the computer guide, will try to help you.

Questions? Enjoy the game! Before playing, there is a short questionnaire for you to fill out.
Login name:

1. Do you like Math?
   Not at all 1 2 3 4 5
   Not at all 1 2 3 4 5

2. Do you usually play computer games? 1 2 3 4 5

3.
   i) Do you know the factors of 25? If yes, what are they?
   Yes Don’t know No
   ii) Do 25 and 10 share a common factor? If yes, what is it?
   Yes Don’t know No
   iii) Do 25 and 14 share a common factor? If yes, what is it?
   Yes Don’t know No
Login name:

1. The game was
   Too Easy 1 2 3 4 5
   Ok

2. Did the game help you learn about common factor?
   Not at all Ok Completely
   1 2 3 4 5

3. Was Merlin Funny?
   Not at all Ok Completely
   1 2 3 4 5

4. Did you find Merlin’s advice helpful?
   Not at all Ok Completely
   1 2 3 4 5

5. Did Merlin help you when you needed it?
   Too soon Ok Too late
   1 2 3 4 5

6. How often did Merlin try to help you?
   Too rarely Ok Too Often
   1 2 3 4 5

7. I don’t like falling.
   True False

8. I want to beat the computer (my partner).
   True False

9. I want to learn math through the game.
   True False

10. I want to get to the top of the mountain.
    True False

11. I do want Merlin to help me.
    True False

12. The tool “Magnifying Glass” is helpful.
    True False

13. The tool “Help Box” is helpful.
    True False
14. I want to have fun. True False

15. It's funny to see swinging True False
<table>
<thead>
<tr>
<th></th>
<th>Slowly &amp; Successfully</th>
<th>Slowly &amp; Fall</th>
<th>Quickly &amp; Successfully</th>
<th>Quickly &amp; Fall</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How did the student climb?</td>
<td>Slowly &amp; Successfully</td>
<td>Slowly &amp; Fall</td>
<td>Quickly &amp; Successfully</td>
<td>Quickly &amp; Fall</td>
<td>Normal</td>
</tr>
<tr>
<td>2. How often did the student use the Magnifying Glass?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. How well did the student use the Magnifying Glass?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. How often did the student use the Help Box?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. How well did the student use the Help Box?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. How helpful was the agent to the student?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. What was the reaction of the student to the agent?</td>
<td>Never</td>
<td>Very Often</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Did the student show negative emotions when falling? (Facial exp. or speech)</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
10. Did the student show positive emotions when falling? (Facial exp. or speech)
   Yes  No

Note: (the interesting questions the student asked, the comments the student made, and
the interesting facial expressions and speech of the student during the playing)
APPENDIX E

PRE-TEST (SECOND VERSION)

Student#: ________________

1. Is 3 a factor of 18? Yes No Don’t know
2. Is 5 a factor of 20? Yes No Don’t know
3. Is 7 a factor of 42? Yes No Don’t know
4. Is 9 a factor of 33? Yes No Don’t know
5. Is 4 a factor of 48? Yes No Don’t know
6. Do 12 and 16 have a common factor? Yes No Don’t know
If yes, then what is it? __________
7. Do 9 and 15 have a common factor? Yes No Don’t know
If yes, then what is it? __________
8. Do 23 and 36 have a common factor? Yes No Don’t know
If yes, then what is it? __________
9. Do 10 and 25 have a common factor? Yes No Don’t know
If yes, then what is it? __________
10. Do 28 and 35 have a common factor? Yes No Don’t know
If yes, then what is it? __________
11. Do 13 and 21 have a common factor? Yes No Don’t know
If yes, then what is it? __________
12. Do 37 and 14 have a common factor? Yes No Don’t know
If yes, then what is it? __________
13. Do 22 and 38 have a common factor? Yes No Don’t know
If yes, then what is it? __________
14. Do 69 and 96 have a common factor? Yes No Don’t know
If yes, then what is it? __________
15. Do 25 and 52 have a common factor? | Yes | No | Don't know
If yes, then what is it? |__________|
POST-QUESTIONNAIRE (SECOND VERSION)

Student #: Name:
Gender: Boy / Girl

1. Please tell us what your goals were when playing the game:

2. Please circle the choice that best describes how you felt.

<table>
<thead>
<tr>
<th>No#</th>
<th>Question</th>
<th>Strongly Disagree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>I wanted to have fun when I was playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>I was eager to beat my partner (the agent) when I was playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>I felt upset when I fell down the mountain.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>I always like to win without anyone's help.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>I became curious about math (number factorization) by playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>I always wanted to be ahead of partner (the agent) when I was playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The agent was very helpful to me.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>I don't care whether the game is funny or not.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>When I fell down, the swinging made me laugh.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>I wanted to learn math by playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>It was important for me to have fun when I was playing the game.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>I liked getting help from the agent when I was stuck.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>I didn't mind falling down the mountain.</td>
<td>1 2 3 4 5</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>I liked working with my partner (the agent) to climb the mountain well.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>I didn’t want to think about math (number factorization) when I was playing the game.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table G.1: Full Cross Table for Move Quickly

<table>
<thead>
<tr>
<th>Beat Partner</th>
<th>Succeed By Myself</th>
<th>Move Quickly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>High</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Table G.2: Full Cross Table for Follow Advice

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Avoid Falling</th>
<th>Follow Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table G.3: Full Cross Table for Ask Advice

<table>
<thead>
<tr>
<th>Succeed By Myself</th>
<th>Ask Advice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>8</td>
</tr>
</tbody>
</table>

Table G.4: Full Cross Table for Use Magnifying Glass Often

<table>
<thead>
<tr>
<th>Learn Math</th>
<th>Beat Partner</th>
<th>Use Magnifying Glass Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>High</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
Table G.5: Full Cross Table for Fall Often

<table>
<thead>
<tr>
<th>Have Fun</th>
<th>Learn Math</th>
<th>Succeed By Myself</th>
<th>Fall Often</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>0</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High</td>
<td>4</td>
</tr>
</tbody>
</table>