
Modelling affect expression and recognition in an interactive learning environment

Scott W. McQuiggan and James C. Lester*

Department of Computer Science
North Carolina State University
Raleigh, NC, USA
E-mail: swmcquig@ncsu.edu
E-mail: lester@ncsu.edu
*Corresponding author

Abstract: Affective reasoning holds significant potential for intelligent tutoring systems. Incorporating affective reasoning into pedagogical decision-making capabilities could enable learning environments to create customised experiences that are dynamically tailored to individual students' ever-changing levels of engagement, interest, motivation and self-efficacy. Because physiological responses are directly triggered by changes in affect, biofeedback data such as heart rate and galvanic skin response can be used to infer affective changes in conjunction with the situational context. This article explores an approach to inducing affect models for a learning environment. The inductive approach is examined for the task of modelling students' self-efficacy and empathy for companion agents. Together, these studies on affect in a narrative learning environment suggest that it is possible to build models of affective constructs from observations of the situational context and students' physiological response.

Keywords: affective student modelling; self-efficacy; intelligent tutoring systems; ITSs; inductive learning; human-computer interaction.

Reference to this paper should be made as follows: McQuiggan, S.W. and Lester, J.C. (2009) 'Modelling affect expression and recognition in an interactive learning environment', *Int. J. Learning Technology*, Vol. 4, Nos. 3/4, pp.216–233.

Biographical notes: Scott W. McQuiggan received his BS degree in Computer Science from Susquehanna University, USA in 2003, his MS degree in Computer Science from North Carolina State University, USA in 2005, and his PhD degree in Computer Science from North Carolina State University in 2009. In his research he pursues an interdisciplinary agenda combining artificial intelligence, education, psychology, human factors and human-computer interaction. His primary interests centre on investigating student interaction in intelligent tutoring systems, game-based learning environments and narrative-centred learning environments. He has extensively published his research on affective computing in intelligent tutoring systems, which has yielded computational models of empathy for companion agents, computational models of student affect in narrative-centred learning environments and computational models of self-efficacy in intelligent tutoring systems. He is currently with the Education Practice of SAS Institute, Inc.

James C. Lester is a Professor of Computer Science at North Carolina State University, USA. He received BA (Highest Honours), MSCS and PhD degrees in Computer Science from the University of Texas at Austin, and a BA

degree in History from Baylor University. A member of Phi Beta Kappa, he has served as Programme Chair for the ACM International Conference on Intelligent User Interfaces (2001), Programme Chair for the International Conference on Intelligent Tutoring Systems (2004), Conference Co-Chair for the International Conference on Intelligent Virtual Agents (2008) and on the editorial board of *Autonomous Agents and Multi-Agent Systems* (1997–2007). His research focuses on intelligent tutoring systems, computational linguistics and intelligent user interfaces. It has been recognised with several Best Paper awards. His current interests include intelligent game-based learning environments, affective computing, creativity-enhancing technologies, computational models of narrative and tutorial dialogue. He is Editor-in-Chief of the *International Journal of Artificial Intelligence in Education*.

1 Introduction

Affect has begun to play an increasingly important role in intelligent tutoring systems. Recent years have seen the emergence of work on affective student modelling (Conati and McLaren, 2005), detecting frustration and stress (Burlison, 2006; Burlison and Picard, 2004; Prendinger and Ishizuka, 2005), modelling agents' emotional states (André and Mueller, 2003; Gratch and Marsella, 2004; Lester *et al.*, 1999), devising affectively informed models of social interaction (Johnson and Rizzo, 2004; Paiva *et al.*, 2005; Porayska-Pomsta and Pain, 2004; Wang *et al.*, 2008), analysing student affective trajectories (Baker *et al.*, 2007; D'Mello *et al.*, 2007) and detecting student motivation (de Vicente and Pain, 2002). All of this work seeks to increase the fidelity with which affective and motivational processes are modelled in intelligent tutoring systems in an effort to increase the effectiveness of tutorial interactions and, ultimately, learning.

Self-efficacy is an affective construct that has been found to be a highly accurate predictor of students' motivational state and their learning effectiveness (Zimmerman, 2000). Defined as "the belief in one's capabilities to organize and execute the courses of action required to manage prospective situations" (Bandura, 1995), self-efficacy has been repeatedly demonstrated to directly influence students' affective, cognitive, and motivational processes (Bandura, 1997). Self-efficacy holds much promise for Intelligent Tutoring Systems (ITSs). Foundational work has begun on using models of self-efficacy for tutorial action selection (Beal and Lee, 2005) and investigating the impact of pedagogical agents on students' self-efficacy (Kim, 2005). Self-efficacy is useful for predicting what problems and sub-problems a student will select to solve, how long a student will persist on a problem, how much overall effort they will expend, as well as motivational traits such as level of engagement (Schunk and Pajares, 2002; Zimmerman, 2000). Thus, if an ITS could increase a student's self-efficacy, then it could perhaps enable the student to be more actively involved in learning, expend more effort, and be more persistent; it could also enable them to successfully cope in situations where they experience learning impasses (Bandura, 1997).

Because self-efficacious students are effective learners, it is important to incorporate mechanisms for diagnosing student affect, self-efficacy (Bandura, 1997), emotional state (Picard, 1997), and motivation (Malone and Lepper, 1987; Schunk *et al.*, 2007) to effectively inform pedagogical decision making. Self-efficacy influences students' reasoning, their level of effort, their persistence, and how they likely feel; it shapes how they make choices, how much resilience they exhibit when confronted with failure, and what level of success they are likely to achieve (Bandura, 1995; Schunk and Pajares, 2002; Zimmerman, 2000).

Emotional state is often an indication of how a student feels she is performing on a given task. For example, students often take pleasure in successfully completing tasks, while negative emotions, such as frustration, often accompany learning impasses. Motivation is an internal state that influences the activities students engage in and their persistence in such activities. It can also increase activity levels in students and guide students in the direction of particular goals (Maehr and Meyer, 1997). Adapting tutorial strategies that foster positive affective states affords a broad range of potential learning benefits, such as effectiveness, efficiency and transfer, because students tend to persist longer and put forth more effort in problem-solving activities when they enjoy what they are learning and believe in their abilities to succeed (Zimmerman, 2000). For these reasons, it is essential to coordinate affective reasoning with pedagogical and tutorial strategies.

We are currently investigating a variety of information channels for reasoning about affect, and, in particular self-efficacy. For example, monitoring physical changes (such as physiological response) in the student, observing her behaviours and relevant events occurring in the virtual world, tracking both narrative and tutorial planning actions, and considering internal models of believable characters in a virtual world are a few of the potential sources of information available. However, recognising student affective states is a challenging task and many approaches to date have required the use of invasive or distracting technologies. For practical purposes, deployable systems will call for affective reasoning techniques that limit or eliminate learning disruption. The ability to reason about student affect could provide potentially useful information about how the student feels about her learning experience. This would allow narrative and tutorial planning to consider refinements that may increase student interest, engagement, motivation, and self-efficacy by selecting specific tasks and providing directed guidance given the student's situation.

Additionally, we have begun exploiting similar information channels for affect expression tasks, such as inducing models of empathy for companion agents. Empathy is the expression of emotion based on another's situation and not merely one's own (Davis, 1994; Hoffman, 2000; Ickes, 1997). Its expression can demonstrate that the target's (the recipient of empathetic expression) feelings are understood or shared. To model empathy we employ a Wizard-of-Oz study design in which our inductive framework monitors situational attributes of the learning environment while the 'wizard' controls empathetic expressions of a companion agent and a learner controls her own character. The inductive framework utilises collected situational data to make predictions of wizard empathetic decisions.

This paper summarises our work to date on modelling affect for the tasks of expression (companion agent empathy) and recognition (student self-efficacy). It is structured as follows. Section 2 discusses the challenges of guided discovery learning and how motivation and self-efficacy can be leveraged in narrative-centred guided

discovery learning environments. Next, Section 3 introduces the Crystal Island learning environment and illustrates its behaviour in a learning scenario. Section 4 summarises the results of affect expression and affect recognition studies conducted with Crystal Island.

2 Affect in narrative inquiry-based learning environments

2.1 Guided discovery learning

It has long been recognised that discovery is a key element of the scientific enterprise, and recent years have seen a growing focus on discovery in education. For more than a decade, science education reform efforts by organisations such as the National Research Council (1996) American Association for the Advancement of Science (1993) have set forth standards promoting a greater emphasis on discovery learning. In discovery learning (or inquiry-based learning), students approach a new topic via learning-by-doing. Instead of being presented problems and solutions in an expository fashion, students are given problems to solve, a rich environment in which to explore the problems, and a set of tools and techniques for constructing solutions. While early accounts of discovery learning focused on concept discovery (Bruner, 1961), contemporary work views discovery learning as scientific investigation. Thus, the process of discovery learning is analogous to the scientific method: students design and perform experiments, collect data, and evaluate hypotheses (de Jong and van Joolingen, 1998). First and foremost, discovery learning is *active* learning. As noted in the National Science Education Standards (National Research Council, 1996), discovery learning is ‘something that students do, not something that is done to them’.

Discovery learning offers several advantages over more didactic approaches. It tends to increase students’ ability to remember what they have learned, to apply their new knowledge, and to transfer it to new tasks more effectively than with more passive approaches such as reading textbooks (Blumenfeld *et al.*, 2000; de Jong and van Joolingen, 1998). In addition to the cognitive benefits of discovery learning, it also offers potential motivational benefits. It enables students to become more active science learners (rather than passive consumers of information), it increases students’ beliefs that scientific theories change as new evidence becomes available (rather than being seen as unchangeable entities), and perhaps most importantly, it makes science more concretely meaningful (rather than seeming too abstract) (White and Fredricksen, 1998).

Despite the potential benefits of discovery learning, in the absence of appropriate scaffolding, discovery learning can be ineffective. Early findings suggested that discovery learning augmented with guidance can be more effective than pure discovery learning in enabling students to apply their knowledge to new problems (Shulman and Keisler, 1996). Furthermore, students may sometimes learn incorrect concepts through discovery learning, and discovery learning may be inefficient (Hammer, 1997). A recent analysis of thirty years of studies of discovery learning suggests that discovery learning accompanied by guidance in the form of feedback and coaching is more effective than unguided discovery learning (Mayer, 2004). Thus, guided discovery appears to be a promising alternative to either didactic instruction or pure discovery learning. In our laboratory we are investigating guided discovery learning in narrative environments for science learning.

2.2 *Motivation in narrative learning environments*

Narrative-centred guided discovery environments for learning may offer motivational benefits. Motivation is critical in learning environments, for it is clear that from a practical perspective, learning environments that fail to engage students will go unused. Game playing experiences and educational experiences that are extrinsically motivating can be distinguished from those that are intrinsically motivating (Malone, 1981). In contrast to extrinsic motivation, intrinsic motivation stems from the desire to undertake activities sheerly for the prospective reward. Narrative-centred discovery learning could provide the four key intrinsic motivators identified in the classic work on motivation in computer games and educational software: challenge, curiosity, control, and fantasy (Malone and Lepper, 1987).

Narrative-centred discovery learning can feature challenging tasks of intermediate levels of difficulty, *i.e.*, tasks that are not too easy and not too difficult, targeting desirable levels of student intrinsic motivation. Dynamically created narratives can feature problem-solving episodes whose level of difficulty is customised for individual students. In discovery approaches, learning is inherently presented as a challenge, a series of problem-solving goals, that once achieved provide a deeper understanding of the domain.

Curiosity requires students to be inquisitive, an essential attribute for successful learning in narrative-centred discovery learning environments. Because discovery learning compels students to obtain knowledge throughout learning episodes, students are likely to question the completeness of their acquired knowledge as they progress and to search for new answers. Narrative-centred discovery learning environments can empower students to take control of their learning experiences; students can choose their own paths, both figuratively (through the solution space) and literally (through the story world), while being afforded significant guidance crafted specifically for them. The narrative structure of guided discovery learning can provide unobtrusive direction by indirectly highlighting a subset of possible goals (*i.e.*, blinking lights in a particular room in the environment, or a character audibly coughing in the student's right audio channel) to inform the student's next action consideration and contribute to the student's perception of control.

Narrative-centred discovery learning is innately fantasy-based. Fantasy refers to a student's identification with characters in the interactive narrative and the imaginative situations created internally and off-screen by the student. All narrative elements ranging from plot and characters to suspense and pacing can contribute to vivid imaginative experiences. The openness of discovery learning affords opportunities to support all levels of student imagination, thereby increasing motivation and engagement. Effective narrative tutorials will employ characters in the story world that either the individual students perceive as possessing some cognitive, emotional, or physical attributes similar to their own, or that the individual student admires, expresses feelings of compassion towards, or for which the student conveys empathetic feelings.

In short, narrative can provide the guidance essential for effective discovery learning and the 'affective scaffolding' for achieving high levels of motivation and engagement.

2.3 Empathetic companion agents

Devising computational models of empathy contributes to the broader enterprise of modelling affective reasoning (Picard, 1997). Beginning with Elliott's implementation (Elliott, 1992) of the OCC model (Ortony *et al.*, 1988), advances in affective reasoning have accelerated in the past few years, including the appearance of a sophisticated theory of appraisal (Gratch and Marsella, 2004) based on the Smith and Lazarus Appraisal Theory (Lazarus, 1991). We have also begun to see probabilistic approaches to assessing users' affective state in educational games (Conati, 2002) and investigations of the role of affect and social factors in pedagogical agents (Baylor, 2005; Burlinson and Picard, 2004; Elliott *et al.*, 1999; Johnson and Rizzo, 2004; Lester *et al.*, 2000; Prendinger and Ishizuka, 2005). Recent work on empathy in synthetic agents has explored their affective responsiveness to biofeedback information and communicative context (Prendinger and Ishizuka, 2005). It has also yielded agents that interact with one another and with the user in a virtual learning environment to elicit empathetic behaviours from its users (Paiva *et al.*, 2005). Empathy has also been investigated in embodied computer agents perceived to care about outcomes of human user experiences in a blackjack game (Brave *et al.*, 2005).

Empathy is a complex socio-psychological construct. Defined as "the cognitive awareness of another person's internal states, that is, his thoughts, feelings, perceptions, and intentions" (Ickes, 1997), empathy enables us to vicariously respond to another via "psychological processes that make a person have feelings that are more congruent with another's situation than with his own situation" (Hoffman, 2000). Social psychologists describe three constituents of empathy. First, the *antecedent* consists of the empathiser's consideration of herself, the target's intent and affective state, and the situation at hand. Second, *assessment* consists of evaluating the antecedent. Third, *empathetic outcomes*, *e.g.*, behaviours expressing concern, are the products of assessment (Davis, 1994) including both affective and non-affective outcomes (*e.g.*, judgement, cognitive awareness). Two types of affective outcomes are possible. In *parallel outcomes*, the empathiser mimics the affective state of the target. For example, the empathiser may become fearful when assessing a target's situation in which the target is afraid. In *reactive outcomes*, empathisers exhibit a higher cognitive awareness of the situation to react with empathetic behaviours that do not necessarily match those of the target's affective state. For example, empathisers may become frustrated when the target does not meet with success in her task, even if the target herself may not be frustrated. Accurately modelling parallel and reactive empathetic reasoning presents significant challenges.

Introducing empathetic companion agents into learning environments may lead to enriched pedagogical agent-student interaction in which agents scaffold student affective experiences in support of learning.

2.4 Self-efficacy in intelligent tutoring systems

To supplement the natural motivational effects of narrative learning environments, we have begun to investigate techniques for modelling student affect. In particular, we are exploring techniques for diagnosing student efficacy. Self-efficacy influences students' reasoning, their level of effort, their persistence, and how they feel; it shapes how they

make choices, how much resilience they exhibit when confronted with failure, and what level of success they are likely to achieve (Bandura, 1995; Schunk and Pajares, 2002; Zimmerman, 2000). While it has not been conclusively demonstrated, many conjecture that given two students of equal abilities, the one with higher self-efficacy is more likely to perform better than the other over time. Highly efficacious students exhibit more control over their future through their actions, thinking, and feelings than inefficacious students (Bandura, 1986). Self-efficacy is intimately related to motivation, which controls the effort and persistence with which a student approaches a task (Lepper *et al.*, 1993). Effort and persistence are themselves influenced by the belief the student has that she will be able to achieve a desired outcome (Bandura, 1997). Students with low self-efficacy perceive tasks to be more challenging than they actually are, often leading to feelings of anxiety, frustration and stress (Bandura, 1986). In contrast, students with high self-efficacy view challenge as a motivator (Bandura, 1986; Malone and Lepper, 1987). Self-efficacy has been studied in many domains with significant work having been done in computer literacy (Delcourt and Kinzie, 1993) and mathematics education (Pajares and Kranzler, 1995). It is widely believed that self-efficacy is domain-specific; whether it crosses domains remains an open question. For instance, students with high self-efficacy in mathematics may be inefficacious in science, or a highly efficacious student in geometry may experience low efficacy in algebra.

A student's self-efficacy is influenced by four types of experiences (Bandura, 1997; Zimmerman, 2000). First, in enactive experiences, she performs actions and experiences outcomes directly. These are typically considered the most influential category. Second, in vicarious experiences, she models her beliefs based on comparisons with others. These can include peers, tutors, and teachers. Third, in verbal persuasion experiences, she experiences an outcome via a persuader's description. For example, she may be encouraged by the persuader, who may praise the student for performing well or comment on the difficulty of a problem. Her interpretation will be affected by the credibility she ascribes to the persuader. Fourth, with physiological and emotional reactions, she responds affectively to situations. These experiences, which often induce stress and anxiety and are physically manifested in physiological responses such as increased heart rate and sweaty palms, call for emotional support and motivational feedback.

Self-efficacy holds great promise for ITSs. Self-efficacy beliefs have a stronger correlation with desired behavioural outcomes than many other motivational constructs (Graham and Weiner, 1996), and it has been recognised in educational settings, that self-efficacy can predict both motivation and learning effectiveness (Zimmerman, 2000). Thus, if it were possible to enable ITSs to accurately model self-efficacy, they may be able to leverage it to increase students' academic performance. Two recent efforts have explored the role of self-efficacy in ITSs. One introduced techniques for incorporating knowledge of self-efficacy in pedagogical decision making (Beal and Lee, 2005). Using a pre-test instrument and knowledge of problem-solving success and failure, instruction is adapted based on changes in motivational and cognitive factors. The second explored the effects of pedagogical agent design on students' traits, which included self-efficacy (Kim, 2005).

3 Crystal Island

In our laboratory we are developing a narrative-centred guided discovery learning environment, Crystal Island (Mott *et al.*, 2006; McQuiggan *et al.*, 2008) (Figure 1). Valve Software's Source™ engine, the 3D game platform developed for Half-Life2, was used to implement the world, interface, and semi-autonomous characters in Crystal Island, which serves as a testbed for investigating affective issues in inquiry-based learning in the domains of microbiology and genetics for middle school students. Crystal Island's narrative takes place at a research outpost situated on a previously unexplored volcanic island. The student is cast as a visitor to the island and the child of the lead scientist. The narrative unfolds as members of the research team begin to fall ill, leaving the student to solve a developing science mystery, thereby saving the expedition. The student progresses by analysing the genetic makeup of chickens responsible for transmitting an unidentified infectious disease through their eggs. In doing so, she is free to explore and interact with the world, its objects, and other characters. Throughout this process, the structure of the developing narrative is characterised by inquiry-based learning activities of question development, hypothesis generation, data collection, and hypothesis testing. To solve the mystery the student must navigate the island, which includes the lead scientist's house, the laboratory, the infirmary, the dining hall, and the men and women's living quarters; manipulate objects, such as eggs, food, books, and mechanical devices; and interact with other research team members to gather relevant information. Ultimately, the student must deduce the species of the chicken that is responsible for the epidemic, and solve the mystery.

Figure 1 Crystal Island Research Station (see online version for colours)



Figure 2 Jin, the camp nurse, with Bryce, the lead research who has fallen ill (see online version or colours)



Consider a ‘typical’ interaction with the Crystal Island environment: a student navigates the world for a period of time, interacting with characters, gradually gathering information about infectious diseases and related topics. As members of the research team fall ill, the accumulated information allows the student to conclude that an infectious disease is the culprit: an illness transmittable between organisms. After learning this concept, an interaction with the nurse suggests that the island’s eggs may be responsible for the spreading illness. If the student can deduce which chickens are responsible for the infected eggs, the mystery may be solved. The student interacts with several of the remaining healthy research team members to learn the relationships between the eggs and chickens and develop the necessary genetics background. The student also utilises an apparatus in the laboratory to perform tests on potentially contaminated eggs. Eventually, the student concludes that white-feathered chickens are responsible for the bad eggs due to a codominant trait, solves the mystery, and reports the finding to the nurse.

4 Experimental results

Effectively modelling student affect requires a representation of situational contexts. Because affect is fundamentally a cognitive process in which a person appraises the relationship between herself and her environment (Gratch and Marsella, 2004; Lazarus,

1991), affect recognition models for learning environments should take into account environmental information used in student appraisals as well as information encoding student responses to their appraisals, such as monitoring physiological changes. To this end we have investigated a rich representation of situational contexts to model empathetic behaviour (McQuiggan and Lester, 2007) and a comparable representation of environmental information extended to include student physiological response data (McQuiggan *et al.*, 2008).

4.1 Empathy modelling for companion agents

A key challenge posed by affective reasoning, particularly for directing synthetic agent behaviour, is devising empirically informed models of empathy that accurately respond in social situations. We have developed Computational Affect Recognition and Expression (CARE), a data-driven affective architecture and methodology for learning models of empathy by observing human-human social interactions (McQuiggan and Lester, 2006). First, in CARE training sessions, a trainer (the student), directs her synthetic agent to perform a sequence of tasks while another trainer manipulates companion agents' affective states to produce empathetic behaviours (spoken language, gesture, and posture). CARE tracks situational data including locational, intentional, and temporal information to induce a model of empathy. At runtime, CARE uses the model of empathy to drive situation-appropriate empathetic behaviours for companion agents. We have conducted two complementary studies investigating the predictive accuracy and perceived accuracy of CARE-induced models of empathy.

4.1.1 Method and procedure

In a formal evaluation, more than 2 h of data were gathered from 31 subjects in an Institutional Review Board (IRB) of North Carolina State University approved user study. The subjects were divided into 25 targets and 6 empathisers. There were 20 male subjects serving as target trainers and 5 female subjects serving as target trainers varying in race, ethnicity, age and marital status who participated as training targets. There were 3 male and 3 female subjects participating as training empathisers. On average, empathisers completed 4 training sessions, each with a unique training target participant.

Each training target participant entered a conference room and was seated in front of a laptop computer with Treasure Hunt, a virtual environment in which a user and an embodied companion agent search for treasures. First, target participants completed the demographic survey at their own rate. Concurrently, empathisers entered a second room and were seated in front of another laptop computer. Training targets were unaware of the empathiser's participation at this point. Empathisers were only aware that a target training participant was in the next room. There was no contact between the participants at any point disabling the empathisers' ability to distinguish any characteristics of the target trainer other than those assumed from the interaction portrayed on their monitor. Empathisers also first completed the same demographic survey as the targets. Next, empathisers completed Davis' IRI questionnaire while targets were given the Half-Life 2 controls reference sheet to read until the practice task was loaded on the laptop in front of the target. Target trainers were able to complete the practice task at their own rate until the task was accomplished. At this point empathisers were given the emotion and empathy reference sheet and instructed to read over the definitions and

empathiser controls. Next, one of the degrees of difficulty was randomly selected and the appropriate Treasure Hunt training environment was loaded on the target machine while the spectator view application was concurrently loaded on the empathiser machine.

Once the training environment was loaded, target trainers had 7 min to explore the environment. Empathisers viewed the interaction and made empathetic behaviour decisions by selecting the appropriate control for the affective state they desired the companion agent to have. When empathetic behaviours were selected by the empathiser, both participants had the opportunity to hear the companion agent's spoken language and see the associated gestural behaviours and posture. Upon completion of the 7 min training session, both training targets and empathisers were given post-session surveys and were interviewed. Finally, target trainers were offered information about the details of the experiment and informed about the presence of the empathiser during the training session.

4.1.2 Summary of results

The following procedural steps were used to generate models of empathy from the training sessions:

- Step 1 Data Construction. Each session log, containing 6000–9000 observation changes, was first translated into a full observational attribute vector. For example, if a treasure box came into view (and all other observable attributes remained constant), then the observational attribute vector would modify the previous vector to account for the noted change.
- Step 2 Data cleansing. After data was converted into the observational attribute vector format the data was ready to be cleaned. This step included generating the dataset containing only records in which the empathiser selected an empathetic emotion.
- Step 3 Naïve Bayes classifier and Decision Tree analysis. Once the dataset was ready it was loaded into the Weka machine learning package (Witten and Frank, 2005), a naïve Bayes classifier and decision tree were learned, and tenfold cross-validation analyses were run on the resulting models. The entire dataset was used to generate models for empathetic assessment (when to be empathetic) and empathetic interpretation (how to be empathetic). Empathetic assessment is determined using the entire dataset, while empathetic interpretation is determined from a transformed dataset containing only empathetic records.

Two sets of models were induced to model the empathisers' decisions. The first model determines when to exhibit an empathetic behaviour. The other model selects which affective state should be expressed by the companion agent. The best performing CARE-induced models have predicted training decisions of 'when' and 'how' to be empathetic with 89% (decision tree model) and 80% (naïve bayes model) accuracy, respectively.¹ These results suggest that the CARE paradigm can provide the basis for effective empathetic behaviour control in embodied companion agents.

The results demonstrate that a decision tree classification approach is sufficient for modelling empathetic assessment and that significantly more training is needed to produce large quantities of empathetic instances for the same approach to have such compelling results for empathetic interpretation. Although the naïve Bayes approach

assumes that all attributes of the observational attribute vector are independent – this assumption is false – it nonetheless induces a model sufficient for controlling empathetic interpretation in companion agents.

4.2 Evaluating empathy models

This section reports on an evaluation of CARE models to determine their perceived accuracy. Perceived accuracy tells us whether the behaviours generated by a model are actually perceived to be socially appropriate in practice. Perceived accuracy is an important aspect of empathetic accuracy because, ultimately, we seek to create models of empathy that will generate behaviours that are deemed to be appropriate for a given social context by human observers.

4.2.1 Method and procedure

In a formal evaluation, 31 undergraduate students, in an IRB of North Carolina State University approved user study, evaluated empathetic responses of the companion agent in video clips from interactions with an interactive learning environment. There were 29 male subjects and 2 female subjects varying in race, ethnicity, and age. Six and a half percent were aged 18–19, 87.0% were aged 20–24, and 6.5% were aged 25–29.

Participants entered a conference room where they were first presented the details of the study and a consent form. They then completed the demographic survey, Davis's IRI questionnaire, and Chapin's Social Insight questionnaire. Next, they read the background on empathy and task directions. Research assistants then fielded any questions from participants regarding empathy and their prescribed task. Participants were then presented, in random order, a series of ten video clips of captured user-interactions in the virtual world. There were four clips of CARE-generated behaviours, three clips of inverse empathetic behaviours, and three clips of human-generated behaviours. After viewing each clip, participants completed the associated response worksheet at their own pace. Following the completion of reviewing and responding to all of the video clips, participants completed the post-experiment survey before the study session concluded. Each subject watched 10 clips, in random order, of empathetic companion agents expressing empathy in a given situation. The clips were derived from three sources including empathetic expressions controlled by human users ($n = 3$), those controlled by a CARE-induced model of empathy ($n = 4$), and crafted controlled expressions (expression of an opposing emotion determined by a CARE model, ($n = 3$)). Subjects rated each clip along three dimensions on a five-point Likert scale (0 to 4): accuracy of the emotional response, accuracy of the timing of the response, and an overall accuracy of the response's appropriateness for the situation.

4.2.2 Summary of results

The perceived accuracy of the displayed emotion when triggered by a CARE-induced model ($M = 2.98$, $SD = 1.02$) was not significantly different from the accuracy of emotions triggered by human users ($M = 2.76$, $SD = 0.97$), $F(1,215) = 2.39$, $p = 0.12$.² However, the perceived accuracy of the timing of empathetic responses driven by a CARE-induced model ($M = 3.06$, $SD = 1.01$) was statistically significant from empathetic responses driven by human users ($M = 2.76$, $SD = 1.16$), $F(1, 215) = 3.97$,

$p = 0.047$. Lastly, the perceived accuracy overall of the empathetic response for the situation depicted in the clip for CARE controlled empathy ($M = 3.03$, $SD = 0.94$) was weakly significantly better than human user driven agent empathetic expression ($M = 2.80$, $SD = 0.98$), $F(1, 215) = 3.26$, $p = 0.073$. These results suggest that CARE-induced empathy models can direct companion agent empathetic expression to the extent that it is perceived to be as accurate as human controlled agent expressions of empathy. In some cases, the CARE model was rated higher, indicating that the consumption of many examples used to induce CARE models of empathy are perceived as more accurate to a wider audience than expressions derived from a single example.

Both human-controlled and CARE-controlled empathetic expressions were perceived to be more accurate than the crafted expressions (designed to use an inappropriate emotion) with convincing significance. For instance, the overall perceived accuracy of the displayed empathetic expression for the depicted situation in the clip for human-controlled expressions ($M = 2.80$, $SD = 0.98$) was significantly better than the crafted expressions ($M = 0.84$, $SD = 1.09$), $F(1, 184) = 165.75$, $p < 0.0001$. Similarly, CARE-controlled expressions ($M = 3.03$, $SD = 0.94$) were also significant from crafted expression ($M = 0.84$, $SD = 1.09$) in the overall rating, $F(1, 215) = 253.99$, $p < 0.0001$. This suggests that humans interacting with empathetic agents can very easily gauge appropriate and accurate empathy from inappropriate empathetic expression.

Participant responses to clips of CARE-generated behaviours cannot be statistically distinguished from the responses to clips of human-generated behaviours from training episodes. This result indicates that CARE models generate empathetic behaviours that are similar to those made by humans and are perceived to be situationally appropriate. The fact that participants were able to distinguish, with statistical significance, inverse empathetic behaviours from both CARE-generated behaviours and human-generated behaviours suggests that both CARE models and human models of empathy differ fundamentally from 'inverse' empathetic models.

4.3 *Modelling student self-efficacy*

We have also begun to investigate inductive approaches to modelling student efficacy in intelligent tutoring systems (McQuiggan *et al.*, 2008). Because self-efficacious students are effective learners, endowing intelligent tutoring systems with the ability to diagnose self-efficacy could lead to improved pedagogy. Accurately modelling self-efficacy requires a representation of the situational context that satisfies two requirements: it must be sufficiently rich to support assessment of changing levels of self-efficacy, and it must be encoded with features that are readily observable at runtime. Because affect is fundamentally a cognitive process in which the user appraises the relationship between herself and her environment (Gratch and Marsella, 2004; Smith and Lazarus, 1990) and because self-efficacy beliefs draw heavily on a student's appraisal of the situation at hand, affect recognition models and models of self-efficacy should take into account both physiological and environmental information. For task-oriented learning environments, self-efficacy models can leverage knowledge of task structure and the state of the student in the learning environment to effectively reason about students' efficacy levels. In particular, self-efficacy models can employ concepts from appraisal theory (Lazarus, 1991) to recognise student efficacy levels generated from their assessment of how their abilities relate to the current learning objective and task. Thus, self-efficacy models can

leverage representations of the information observable in the learning environment – note that this refers to the same information that students may use in their own appraisals – to predict student efficacy levels.

4.3.1 Method and procedure

In a formal evaluation, data was gathered from thirty-three subjects in an IRB of North Carolina State University approved user study. There were 6 female and 27 male participants varying in age, race, and marital status. Approximately 12 (36%) of the participants were Asian, 20 (60%) were Caucasian, and 1 (3%) was Black or African-American. Twenty-seven percent of the participants were married. Participants average age was 26.15 (SD = 5.32).

Each participant entered the experimental environment (a conference room) and was seated in front of the laptop computer. First, participants completed the demographic survey at their own rate. Next, participants read over the online genetics tutorial directions before proceeding to the online tutorial. On average, participants took 17.67 (SD = 2.91) minutes to read through the genetics online tutorial. Following the tutorial, participants were asked to complete the Problem-Solving Self-Efficacy Scale considering their experience with the material encountered in the genetics tutorial. The instrument asked participants to rate their level of confidence in their ability to successfully complete certain percentages of the upcoming problems in the interactive learning environment. Participants did not have any additional information about the type of questions or the domain of the questions contained in forthcoming problems. Participants were then outfitted with biofeedback equipment on their left hand while the interactive learning environment was loaded. Once the system was loaded, participants entered the calibration period in which they read through the problem-solving directions. This allowed the system to obtain initial readings on the temporal attributes being monitored, in effect establishing a baseline for Heart Rate (HR) and Galvanic Skin Response (GSR). Participants used a self-efficacy slider representing the strength of their belief in their answers being correct as they progressed through the environment.

4.3.2 Summary of results

We have explored two families of self-efficacy models: a static self-efficacy model, learned solely from pre-test (non-intrusively collected) data with tutoring environment information, and a dynamic self-efficacy model, learned from both static data as well as runtime physiological data collected with a biofeedback apparatus.

We have utilised naïve Bayes and decision tree modelling techniques for their preliminary predictive power and robustness to large multidimensional data. The highest performing static model is able to predict students' real-time levels of self-efficacy with 85.2% accuracy, while the best physiologically informed dynamic model performed at 86.9% accuracy.³ The experiment revealed two important implications for the design of runtime self-efficacy modelling. First, even without access to physiological data, induced decision-tree models can make reasonably accurate predictions about students' self-efficacy. Sometimes physiological data is unavailable or it would be too intrusive to obtain the data. In these situations, decision-tree models that learn from demographic data and data gathered with a validated self-efficacy instrument administered prior

to problem-solving and learning episodes, can accurately model self-efficacy. Second, if runtime physiological data is available, it can significantly enhance self-efficacy modelling. Given access to HR and GSR, self-efficacy can be predicted more accurately.

5 Conclusions and future work

Recent advances in affective reasoning have demonstrated that emotion plays a central role in human cognition and should therefore play an equally important role in human-computer interaction. To support effective interactions, affect-informed systems must be able to accurately and efficiently recognise user affect from available resources and respond accordingly. A promising approach to constructing models of affect is inducing them rather than manually constructing them. With an inductive approach, machine learning techniques can be leveraged to induce models of affect for intelligent tutoring systems. In this article we have summarised two studies which induced models of self-efficacy and empathy. Incorporated into runtime intelligent tutoring systems, these models offer the potential for increasing motivation and learning effectiveness by alerting tutorial components of changes in student efficacy and recommending empathetic expressions for pedagogical agents.

Several directions for future work are suggested by the results obtained to date. First, it will be important to investigate how to design pedagogical planning components that are informed by models of student affect, such as self-efficacy. Of particular interest here is the problem of dynamically creating pedagogical goals that take into account affective considerations. Second, it will be important to explore how models of self-efficacy can increase the utility of affect recognition models. Combining predictions of student efficacy with detection of student frustration could inform intervention decisions. For example, in potentially frustrating (but pedagogically useful) situations, efficacy may help predict how long a student may be capable of persisting without an intervention. Third, it will be important to integrate models of empathy with tutorial strategies to generate empathetically appropriate pedagogical feedback. Certainly, it is not always the case that pedagogical agents should respond to students with affect – frequently ‘unaffected’ hints or explanations are most appropriate – but it is evident that affect should play an important role in supporting effective learning interactions.

Acknowledgements

The authors would like to thank the members of the IntelliMedia Group at North Carolina State University for their contributions to the implementation of Crystal Island. The authors also wish to thank Valve Software for authorising the use of their Source™ engine and SDK. This research was supported by the National Science Foundation under Grants REC-0632450, IIS-0812291, and DRL-0822200. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- American Association for the Advancement of Science (1993) *Benchmarks for Science Literacy: Project 2061*, New York, NY: Oxford Press.
- André, E. and Mueller, M. (2003) 'Learning affective behavior', *Proceedings of the 10th International Conference on Human-Computer Interaction*, Mahwah, NJ: Lawrence Erlbaum, pp.512–516.
- Baker, R., Rodrigo, M. and Xolocotzin, U. (2007) 'The dynamics of affective transitions in simulation problem-solving environments', *Proceedings the 2nd International Conference on Affective Computing and Intelligent Interactions*, Lisbon, Portugal, pp.666–677.
- Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*, Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1995) 'Exercise of personal and collective efficacy in changing societies', in A. Bandura (Ed.) *Self-Efficacy in Changing Societies*, New York, NY: Cambridge University Press, pp.1–45.
- Bandura, A. (1997) *Self-Efficacy: The Exercise of Control*, New York: Freeman.
- Baylor, A. (2005) 'The impact of pedagogical agent image on affective outcomes', *Workshop on Affective Interactions: Computers in the Loop, in Conjunction with the International Conference on Intelligent User Interfaces*, San Diego, California.
- Beal, C. and Lee, H. (2005) 'Creating a pedagogical model that uses student self reports of motivation and mood to adapt ITS instruction', *Workshop on Motivation and Affect in Educational Software, in Conjunction with the 12th International Conference on Artificial Intelligence in Education*, Amsterdam.
- Blumenfeld, P., Fishman, B., Krajcik, J., Marx, R. and Solloway, E. (2000) 'Creating usable innovations in systemic reform: scaling up technology-embedded project-based science in urban schools', *Educational Psychologist*, Vol. 33, No. 3, pp.149–164.
- Brave, S., Nass, C. and Hutchinson, K. (2005) 'Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent', *International Journal of Human-Computer Studies*, Vol. 62, pp.161–178.
- Bruner, J. (1961) 'The act of discovery', *Harvard Educational Review*, Vol. 31, pp.21–32.
- Burleson, W. (2006) 'Affective learning companions: strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance', PhD thesis, Massachusetts Institute of Technology.
- Burleson, W. and Picard, R. (2004) 'Affective agents: sustaining motivation to learn through failure and a state of stuck', *Workshop on Social and Emotional Intelligence in Learning Environments, in Conjunction with the 7th International Conference on Intelligent Tutoring Systems*, Montreal.
- Conati, C. (2002) 'Probabilistic assessment of user's emotions in educational games', *Applied Artificial Intelligence*, Vol. 16, pp.555–575.
- Conati, C. and McLaren, H. (2005) 'Data-driven refinement of a probabilistic model of user affect', *Proceedings of the 10th International Conference on User Modeling*, New York, NY: Springer-Verlag, pp.40–49.
- Davis, M. (1994) *Empathy: A Social Psychological Approach*, Madison, WI: Brown and Benchmark Publishers.
- De Jong, T. and van Joolingen, W. (1998) 'Scientific discovery learning with computer simulations of conceptual domains', *Review of Educational Research*, Vol. 68, No. 2, pp.179–201.
- De Vicente, A. and Pain, H. (2002) 'Informing the detection of the students' motivational state: an empirical study', *Proceedings of the 6th International Conference on Intelligent Tutoring Systems*, New York, NY: Springer-Verlag, pp.933–943.

- Delcourt, M. and Kinzie, M. (1993) 'Computer technologies in teacher education: the measurement of attitudes and self-efficacy', *Journal of Research and Development in Education*, Vol. 27, No. 1, pp.35–41.
- D'Mello, S., Taylor, R.S. and Graesser, A. (2007) 'Monitoring affective trajectories during complex learning', *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, Austin, Texas, pp.203–208.
- Elliott, C. (1992) 'The affective reasoner: a process model of emotions in a multi-agent system', PhD thesis, Northwestern University, Chicago, Illinois.
- Elliott, C., Rickel, J. and Lester, J. (1999) 'Lifelike pedagogical agents and affective computing: an exploratory synthesis', in M. Wooldridge and M. Veloso (Eds.) *Artificial Intelligence Today, Lecture Notes in Artificial Intelligence*, Sub series of LNCS, Springer-Verlag, Berlin, Special Vol. 1600, pp.195–212.
- Graham, S. and Weiner, B. (1996) 'Principles and theories of motivation', in D. Berliner and R. Calfee (Eds.) *Handbook of Educational Psychology*, New York, NY: MacMillan Publishing, pp.63–84.
- Gratch, J. and Marsella, S. (2004) 'A domain-independent framework for modeling emotion', *Journal of Cognitive Systems Research*, Vol. 5, No. 4, pp.269–306.
- Hammer, D. (1997) 'Discovery learning and discovery teaching', *Cognition and Instruction*, Vol. 15, No. 4, pp.485–529.
- Hoffman, M. (2000) *Empathy and Moral Development: Implications for Caring and Justice*, Cambridge, UK: Cambridge University Press.
- Ickes, W. (1997) *Empathic Accuracy*, New York, NY: Guilford Press.
- Johnson, L. and Rizzo, P. (2004) 'Politeness in tutoring dialogs: 'run the factory, that's what I'd do'', *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, New York, NY: Springer-Verlag, pp.67–76.
- Kim, Y. (2005) 'Empathetic virtual peers enhanced learner interest and self-efficacy', *Workshop on Motivation and Affect in Educational Software, in Conjunction with the 12th International Conference on Artificial Intelligence in Education*, Amsterdam.
- Lazarus, R. (1991) *Emotion and Adaptation*, New York, NY: Oxford University Press.
- Lepper, M., Woolverton, M., Mumme, D. and Gurtner, J. (1993) 'Motivational techniques of expert human tutors: lessons for the design of computer-based tutors', in S. Lajoie and S. Derry (Eds.) *Computers as Cognitive Tools*, Hillsdale, NJ: Erlbaum, pp.75–105.
- Lester, J., Towns, S., Callaway, C., Voerman, J. and FitzGerald, P. (2000) 'Deictic and emotive communication in animated pedagogical agents', in J. Cassell, J. Sullivan, S. Prevost and E. Churchill (Eds.) *Embodied Conversational Agents*, Cambridge, MA: MIT Press, pp.123–154.
- Lester, J., Towns, S. and FitzGerald, P. (1999) 'Achieving affective impact: visual emotive communication in lifelike pedagogical agents', *The International Journal of Artificial Intelligence in Education*, Vol. 10, Nos. 3–4, pp.278–291.
- Maehr, M. and Meyer, H. (1997) 'Understanding motivation and schooling: where we've been, where we are, and where we need to go', *Educational Psychology Review*, Vol. 9, No. 4, pp.371–409.
- Malone, T. (1981) 'Toward a theory of intrinsically motivating instruction', *Cognitive Science*, Vol. 5, No. 4, pp.333–369.
- Malone, T. and Lepper, M. (1987) 'Making learning fun: a taxonomy of intrinsic motivations for learning', in R.E. Snow and M.J. Farr (Eds.) *Aptitude, Learning, and Instruction: Cognitive and Affective Process Analyses*, Hillsdale, NJ: Erlbaum, Vol. 3, pp.223–253.
- Mayer, R. (2004) 'Should there be a three-strikes rule against pure discovery learning?', *American Psychologist*, Vol. 59, No. 1, pp.14–19.

- McQuiggan, S. and Lester, J. (2006) 'Learning empathy: a data-driven framework for modeling empathetic companion agents', *Proceedings of the Fifth International Conference on Autonomous Agents and Multi-Agent Systems*, Hakodate, Japan.
- McQuiggan, S. and Lester, J. (2007) 'Modeling and evaluating empathy in embodied companion agents', *International Journal of Human-Computer Studies*, Vol. 65, No. 4, pp.348–360.
- McQuiggan, S., Mott, B. and Lester, J. (2008) 'Modeling self-efficacy in intelligent tutoring systems: an inductive approach', *User Modeling and User-Adapted Interaction*, Vol. 18, Nos. 1–2, pp.81–123.
- Mott, B., McQuiggan, S., Lee, S., Lee, S.Y. and Lester, J. (2006) 'Narrative-centered environments for guided exploratory learning', *Proceedings of Agent-Based Systems for Human Learning Workshop at 5th International Conference on Autonomous Agents and Multi-Agent Systems*, Hakodate, Japan.
- National Research Council (1996) *National Science Education Standards*, Washington, DC: National Academy Press.
- Ortony, A., Clore, G. and Collins, A. (1988) *The Cognitive Structure of Emotions*, Cambridge, UK: Cambridge University Press.
- Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L. and Zoll, C. (2005) 'Learning by feeling: evoking empathy with synthetic characters', *Applied Artificial Intelligence*, Vol. 19, pp.235–266.
- Pajares, F. and Kranzler, J. (1995) 'Self-Efficacy beliefs and general mental ability in mathematical problem solving', *Contemporary Educational Psychology*, Vol. 20, pp.426–443.
- Picard, R. (1997) *Affective Computing*, Cambridge, MA: MIT Press.
- Porayska-Pomsta, K. and Pain, H. (2004) 'Providing cognitive and affective scaffolding through teaching strategies', *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, New York, NY: Springer-Verlag, pp.77–86.
- Prendinger, H. and Ishizuka, M. (2005) 'The empathic companion: a character-based interface that addresses users' affective states', *Applied Artificial Intelligence*, Vol. 19, pp.267–285.
- Schunk, D. and Pajares, F. (2002) 'The development of academic self-efficacy', in A. Wigfield and J. Eccles (Eds.) *Development of Achievement Motivation*, San Diego, CA: Academic Press, pp.15–31.
- Schunk, D., Pintrich, P. and Meece, J. (2007) *Motivation in Education: Theory, Research, and Applications*, 3rd ed., Upper Saddle River, NJ: Prentice Hall.
- Shulman, L. and Keisler, E. (1996) *Learning by Discovery: A Critical Appraisal*, Chicago, IL: Rand McNally.
- Smith, C. and Lazarus, R. (1990) 'Emotion and adaptation', in L. Pervin (Ed.) *Handbook of Personality: Theory & Research*, New York, NY: Guilford Press, pp.609–637.
- Wang, N., Johnson, W.L., Mayer, R., Rizzo, P., Shaw, E. and Collins, H. (2008) 'The politeness effect: pedagogical agents and learning outcomes', *International Journal of Human-Computer Studies*, Vol. 66, pp.98–112.
- White, B. and Fredricksen, J. (1998) 'Inquiry, modeling, and metacognition: making science accessible to all students', *Cognition and Instruction*, Vol. 16, pp.3–118.
- Witten, I. and Frank, E. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed., San Francisco, CA: Morgan Kaufman.
- Zimmerman, B. (2000) 'Self-efficacy: an essential motive to learn', *Contemporary Educational Psychology*, Vol. 25, pp.82–91.

Notes

- 1 Please see McQuiggan and Lester (2006) and McQuiggan and Lester (2007) for more details.
- 2 Please see McQuiggan and Lester (2007) for more details.
- 3 Please see McQuiggan *et al.* (2008) for more details.