Lifelike Pedagogical Agents and Affective Computing: An Exploratory Synthesis

Clark Elliott[1], Jeff Rickel[2], and James Lester[3]

[1] Institute for Applied Artificial Intelligence,
School of Computer Science, Telecommunications, and Information Systems,
DePaul University, 243 South Wabash Ave., Chicago, IL 60604,
elliott@ils.nwu.edu,
WWW home page: http://condor.depaul.edu/~elliott/
[2] Information Sciences Institute and Computer Science Department,
University of Southern California,
rickel@isi.edu,
WWW home page: http://www.isi.edu/isd/VET/vet.html
[3] Multimedia Laboratory,
Department of Computer Science,
North Carolina State University,
lester@eos.ncsu.edu,
WWW home page: http://multimedia.ncsu.edu/imedia

1 Introduction

Lifelike pedagogical agents have been the subject of increasing attention in the agents and knowledge-based learning environment communities [2, 17, 19–21]. In parallel developments, recent years have witnessed great strides in work on cognitive models of emotion and affective reasoning [4, 18, 22]. As a result, the time is now ripe for exploring how affective reasoning can be incorporated into pedagogical agents to improve students' learning experiences.

This chapter investigates how these two converging research efforts may yield a new form of pedagogical agent that is sensitive to students' emotive state and can reason about affective aspects of problem-solving contexts. Initial forays have been taken into pedagogical emotion generation [1] and reasoning about learners' emotions [3], indicating the potential richness offered by affective learner-system interactions. These efforts suggest important new functionalities for learning environments. Rather than speculating in the abstract about how these new functionalities may come about, we explore them with the particulars of a specific computational model of emotion and specific lifelike pedagogical agents. We discuss preliminary conceptual work on integrating the Affective Reasoner, being developed at DePaul University, with two extant learning environments: the Soar Training Expert for Virtual Environments (Steve) being developed at the Information Sciences Institute (Figure 1), and the Design-A-Plant (Herman the Bug) system being developed at North Carolina State University (Figure 2). These types of integration are undertaken in an effort to create pedagogical agents with rich models of personality, context-sensitive emotional responsiveness, and affective user modeling.

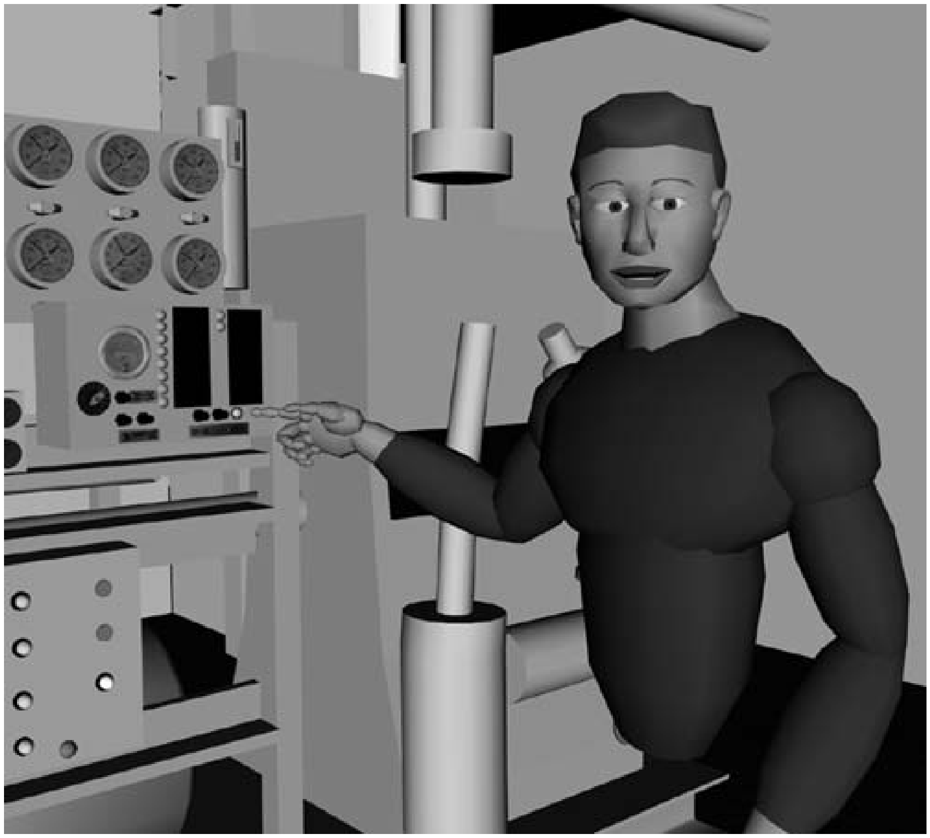

Fig. 1: The Steve Agent in the VET Learning Environment

In the following section we describe the motivations for our work, arguing that affective reasoning will make pedagogical agents better teachers. The remainder of the paper outlines the basic elements of the Affective Reasoning framework and how it might apply to Steve and Herman. We use Steve as a platform for discussing application of personality and emotional responsiveness in pedagogical agents, whereas we use Herman as a platform for discussing affective user modeling. However, this is only for convenience of exposition; each kind of emotion reasoning applies equally well to both systems.

2 Objectives

Good teachers are often good motivators. Motivation is a key ingredient in learning, and emotions play an important role in motivation. We therefore believe that pedagogical agents will be more effective teachers if they display and understand emotions. This could facilitate learning in several ways:

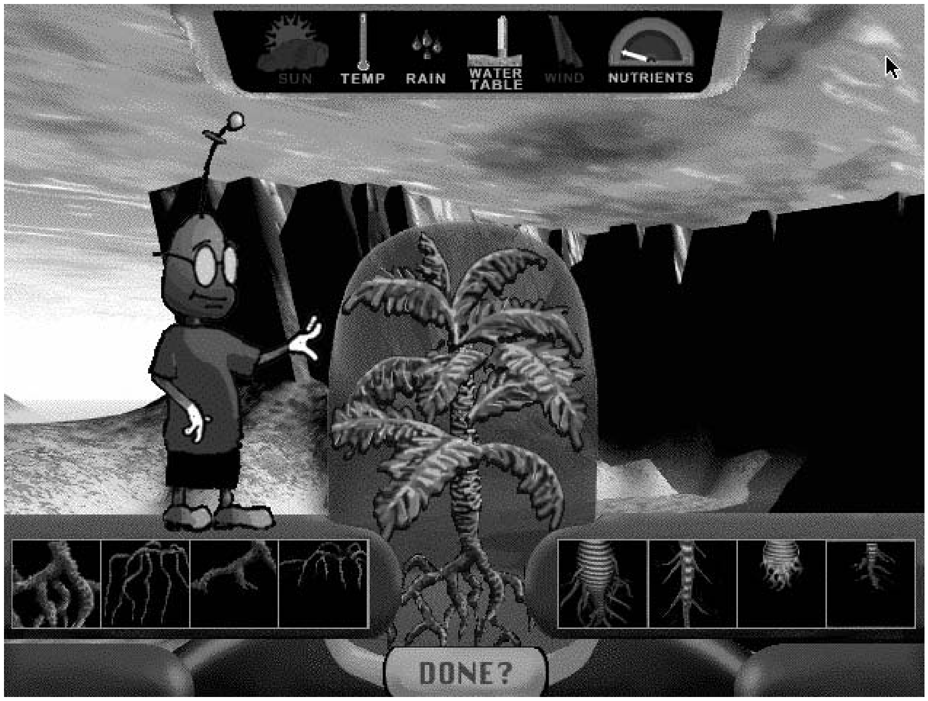

Fig. 2: The Herman Agent in the Design-A-Plant Learning Environment

1. A pedagogical agent should appear to care about students and their progress. This can foster in a student a feeling that she and the agent are "in things together," and can encourage the student to care about her own progress, and the agent's opinion of her.
2. A pedagogical agent should be sensitive to the student's emotions. For example, the agent must recognize a student's frustration so as to intervene with assistance and encouragement before the student loses interest.
3. A pedagogical agent should convey enthusiasm for the subject matter, in order to foster similar enthusiasm in the student. To achieve a credible appearance of enthusiasm in an agent, it is useful to model the emotions that underlie it.
4. A pedagogical agent with a rich and interesting personality may simply make learning more fun. A student that enjoys interacting with a pedagogical agent will have a more positive perception of the whole learning experience. A student that enjoys a learning environment will undoubtedly spend more time there, which is likely to increase learning.

We cannot, at this point, claim to elevate these much beyond the level of intuition, but they are highly commonsensical, and are also testable hypotheses (and c.f. [13]). With respect to the effectiveness of lifelike characters, a recent large-scale study on animated pedagogical agents has demonstrated the *persona* *effect*, which is that the presence of a lifelike character in an interactive learning environment can have a strong positive effect on students' perception of their learning experience [11]. Other studies have revealed similar findings [2, 9]. By directly comparing the effectiveness of pedagogical agents with and without various types of emotional capabilities, we can better understand the role of emotion in learning.

3 The Affective Reasoning Platform

What we refer to as "emotions" in this paper arise naturally in many human social situations as a byproduct of goal-driven behavior, principled (or unprincipled) behavior, simple preferences, and relationships with other agents. This includes many situations not normally thought of as emotional (e.g., becoming annoyed at someone, a mild form of anger in our theory), but explicitly excludes representation of any physical (i.e., bodily) properties of emotions.

At DePaul University, in our current research on the Affective Reasoner (AR), embodied in a large-scale Lisp-based implementation, we build agents capable of responding "emotionally" to other agents and interactive users, as a function of their concerns. Agents are given unique pseudo-personalities modeled as both a set of *appraisal frames* representing their individual goals (with respect to events that arise), *principles* (with respect to perceived intentional actions of agents), *preferences* (with respect to objects), *moods* (temporary changes to the appraisal mechanism), and as a set of about 440 differentially activated *channels* for the expression of emotions [4, 7]. Situations that arise in the agents' world may map to twenty-six different emotion types (e.g., *pride*, as approving of one's own intentional action), twenty-two of which were originally theoretically specified by Ortony, *et al.* [16]. Qualities, and intensity, of emotion instances in each category are partially determined by some subset of roughly twenty-two different *emotion intensity variables* [8].

To communicate with users the AR agents use various multimedia modes including facial expressions, speech, and even music. Agents have about 70 line-drawn facial expressions, which are morphed in real time, yielding about 3,000 different morphs. (A central assumption of our work is that social interaction based on emotion states *must* run in real time.) This is extensible since the morphs are efficiently computed on the client computer each time they are displayed. Agents can select facial expressions, speed of morph, size of the display, and color of the foreground and background. Agents, whose mouths move when they speak, communicate with users through minimally inflected text-to-speech software, which allows us to dynamically construct spoken sentences at run time.

With regard to music, to add qualitatively to their expression of emotions, agents have access to a large database of MIDI files, any portion which they can retrieve in less than a second, and in which they can index down to 1/1000th of a second. Each of these (mostly music) files are real performances (that is, are creations of human performers, not of computers playing sounds from scores). Speech recognition software, used with some AR applications, has allowed children as young as two years old to interact with AR application agents. In all cases, the agents respond in real time to input from the world around them: when spoken to, they speak back [6]. These particular agents have been shown quite effective at communicating a wide range of emotion, comparing favorably with a human actor [5].

4 Creating Virtual Affective States in Steve

The Soar Training Expert for Virtual Environments (Steve) [19] is a pedagogical agent for virtual environments (Figure 1). Steve's objective is to help students learn to perform physical, procedural tasks, such as operating or repairing complex equipment. Students are immersed in a 3D computer simulation of their work environment where they can improve their skills through practice on realistic tasks. Steve cohabits the virtual environment with them, continually monitoring the state of the environment and periodically controlling it through virtual motor actions. Steve appears in the environment as a full upper body, and he communicates to students via text-to-speech software. Steve helps students in various ways: he can demonstrate tasks, answer questions about the rationale behind task steps, and monitor students while they practice tasks, providing help when requested. He functions as part of a larger Virtual Environments for Training (VET) system being developed jointly by the USC Information Sciences Institute, the USC Behavioral Technology Laboratory, and Lockheed Martin.

The VET system includes two main components in addition to Steve. The first of these, the virtual reality software, handles the interface between students and the virtual world, updating the view on their head-mounted display as they move around and detecting their interactions with virtual objects. The second of these, a world simulator, maintains the state of the virtual world as agents (such as Steve) or students interact with it. When the student interacts with the virtual world, Steve gets a message from the virtual reality software describing the interaction, and he gets messages from the simulator describing the resulting changes in the world (in terms of attribute-value pairs). In this way, Steve is aware of both the student's action (e.g., pulling on the dipstick), and the result of that action (e.g., the dipstick is pulled out).

How can affective reasoning be integrated into Steve? We provide a tentative but concrete answer to this question by outlining some potential roles for affective reasoning in Steve and show how these map to elements of the existing Affective Reasoning framework. Throughout, we use examples from Steve's current domain, that of operating a high pressure air compressor. Although Steve is designed as a domain independent tutor, capable of providing instruction in a variety of different domains given the appropriate knowledge, this domain has served as a testbed for his early development.

4.1 Emotion Antecedents in Steve

As discussed above, in the AR scheme, emotions arise as a result of agents' appraisals of their world. Goal-based emotions, such as joy and distress, are generated when there is a match between some event, and the goal-based concerns of the agent. Similarly, principle-based emotions, such as admiration and reproach, are generated when there is a match between what is seen as an accountable act of some agent and the beliefs of an observing agent, with respect to right and wrong. Lastly, preference-based emotions, such as liking and disliking, are generated when there is a match between appealing or unappealing objects that the observing agent attends to, and the preferences of that agent.

In this architecture, one part of the pseudo-personality of an agent—that which stems from their emotional predisposition—is based on the way they see the world, and this in turn is partially determined by the goals, values, and likes they hold. In any particular context (such as that in which Steve operates), it is thus necessary to describe the personality of the automated agents in terms of the situations that arise, characterized as sets of goals, principles, and preferences in the content domain.

In Steve's virtual world such situations arise, and are formally represented, thereby allowing us to test for possible appraisals with respect to his personality. An "emotionally intelligent" Steve might, or might not, have emotion responses to these situations, depending upon his concerns. Similarly, Steve might have emotions based on the presumed state of the student, and about future and past events. Here we examine some basic mappings from situations to emotions.

Sample Goals for Steve Steve has a set of goals which help define his personality. These goals may match situations that arise in the simulation. In some cases, these matches represent achieved goals; in others, they represent thwarted goals. In this way, we model Steve's *desires* so that when the world takes on a new state, events represented by that state may combine with Steve's goals to determine how he "feels." Some goals are *serendipity* goals which can only be matched when achieved, some are *preservation* goals which can only be matched when thwarted, and still others are *bi-valenced* which can be matched either way. The following are simple examples of such goals for Steve.

- **Goal 1: I will give explanations about the subject domain and discuss interesting details about the domain with the student.** For example, we want Steve to be happy when he has a chance to give the student useful, albeit not necessarily critical, information about the underlying domain. By exhibiting happiness when he has a chance to share information about the subject domain, Steve manifests a rudimentary form of enthusiasm about the subject.

  Example situation: (a) Student is learning to start the compressor, (b) Student goes through pre-startup routine but neglects to open the cut-out valve, (c) Student attempts to start compressor, (d) Steve intervenes and says, "Before starting compressor, you must always open the cut-out valve. It is actually pretty interesting what can happen if you do not," and gives the explanation.

- **Goal 2: I will engage the student.** For example, we want Steve to be distressed, or anxious, when the student appears bored with the tasks at hand.
  Example situation: (a) Steve is giving an explanation about why the cut-out valve must be opened when starting the compressor, (b) Steve and the cut-out valve are not in the student's field of vision, i.e., the student is looking elsewhere. Steve frowns and in an irrirated voice says, "Look over here."
- **Goal 3: The student will retain task knowledge.** For example, we want Steve to be upset with himself if the student does not retain critical knowledge; we want Steve to be happy if the student remembers important points. By caring about how well the student performs, Steve shows the student that he is interested in the student's progress.
  Example situation: (a)–(d) (from Goal 1), then later the student repeats the startup procedure, and either (e) student attempts to start the compressor without opening the cut-out valve, or (f) the student has not opened the cut-out valve and asks Steve if there is anything else to be done before starting the compressor, or (g) the student remembers to open the cut-out valve. Should (e) or (f) obtain, Steve's goal would be thwarted, and he would feel distressed; should (g) obtain, he would feel joy.
- **Goal 4: The student will be cautious.** For example, we want Steve to have an "Oh no!" distress, or fear, response to the situation when the student is careless and exposes both Steve, and herself, to danger.
  Example situation: (a)–(e) (from Goal 3), and (f) failing to open the cut-out valve is potentially dangerous.

Such goals alone might not fully determine Steve's response. Steve's emotions may also be affected by his *relationship* with the student, and the perceived appraisal of the situation by the student. Relationships are independent of the goals of the agent, but can combine with the goals to determine fortunes-of-others emotions that may arise. For example, suppose that Steve believes a student to be distressed about forgetting a particular task. If Steve is in a *friendship* relationship with the student he might feel *sorry-for* her. If he is in an *animosity* relationship with her (which could be useful for modeling competition), he might *gloat* over her misfortunes. In both cases the argument can be made that Steve is exhibiting not only care about what the student does, but also a rudimentary form of care about the student herself.

Fear and *hope* over prospective future events—these are emotions which are also based on an agent's goals—must be handled somewhat differently. Since these emotions may later be resolved (i.e., with satisfaction, fears-confirmed, relief, or disappointment), the internal representation of the matched situation must remain active. With a cautious Steve, for example, there might be both a recurring hope, and a recurring fear with respect to an accident happening to a student, subsequent to training. In this case each instance of student caution would raise the threshold which controls whether or not a *rumination* occurrence (implemented as a cyclic, self-generating, event) actually leads to an instance of *fear* on Steve's part. Similarly each instance of sloppiness, leading to a simulation accident (or potential accident), would lead to a reduced threshold for activation of *fear*. With identical structure, but in a contrasting manner, ruminations leading to *hope* that an accident will not occur would take the same form: the more the student exhibits caution, the more *hope* will tend to arise because of a lowered threshold for activation; the more the student exhibits sloppiness the less *hope* will tend to arise because of a raised threshold.

Example Principles for Steve. Like his goals, Steve's principles would also help to determine the make-up of his personality. Situations that arise in the simulated world can be appraised by Steve in terms of their effects as events (e.g., looking only at *what happened*) relevant to his goals, but can also, in some cases, be appraised as perceived accountable actions of some observed agent (e.g., by making a determination about who is to take credit, or blame, for the situation coming about). In still other cases one situation can be appraised as being simultaneously relevant to both an agent's goals and the agent's principles. The following examples illustrate this difference between principles and goals:

- **Principle 1: Student should attend to me when I am talking with them.** For example, we want Steve to show annoyance when the student is not paying attention.
  Example situation: (a) Steve is giving an explanation about why the cut-out valve must be opened when starting the compressor, and (b) Steve, and the cut-out valve, are not in the student's field of vision, i.e., the student is looking elsewhere]. Note that these constraints are the same as with the parallel goal above, but here Steve's concerns have a different focus. That Steve could be wrong about the student paying attention (e.g., when the student is visualizing the problem internally) is not something we worry about because people are often wrong in their perceptions as well; wrong or right, it should still be clear to the student that Steve *cares*.
- **Principle 2: Student should be cautious when it is appropriate to do so.** For example, we want Steve to express admiration for the quality of being careful with the expensive equipment, or reproach when a student is not cautious.
  Example situation: (a)–(d) (from goal 1), then later the student repeats the compressor start-up sequence and then either (e) fails to open the cut-out valve before attempting to start the compressor, or (f) remembers to open the cut-off valve. In contrast to the event-based emotion detailed above, here Steve's focus is on the blameworthiness of the student forgetting to open the valve, or the praiseworthiness of remembering.
- **Principle 3: I should be patient with the user.** For example, we want Steve to get angry with a student (serving to get her attention), but then feel shame or remorse about losing his temper with her (serving to assuage bad feeling).
  Example situation: As a result of Goal 4, and Principle 2 above, Steve gets angry at the student (where *anger* is the result of the student's blameworthy

action thwarting one of Steve's goals), yielding, (1) Steve gets angry, and (2) the anger is directed at the student.

- **Principle 4: I should tutor my students well enough that they make adequate progress on the domain tasks.** For example, we want Steve to feel shame, or remorse, when his students fail to show good progress.

  Example situation: (a) a student has established a benchmark time for successfully completing the pre-start procedure for the compressor although it is below competency level, (b) the student fails to improve on this time in three successive attempts.

Emotion Generation in Steve. Although it is not possible to cover the many details that will go into a rich, plausible, emotion model for Steve, the following touch on a few salient points not covered in the preceding examples. Because Steve's emotion states are based on antecedents, it is possible for him to give rich explanations for why he feels the way he does. These in themselves can be useful pedagogical tools: One emotion state may be the result of many different goals, principles, and a large collection of intensity values. The student might, for example, be motivated to ask Steve what makes him happy, or angry. The explanation can serve to inform the user that Steve cares about the student's progress, wants the user to learn safety rules, and so forth.

Agents such as Steve can have multiple, and even conflicting, goals and principles, just as people do (e.g., he might be simultaneously experiencing *hope* that the user will not have an accident subsequent to training, based on instances of caution, and *fear* that the user will have an accident, based on instances of sloppiness). In designing solutions to tutoring-interaction problems it is possible to come up with several, possibly conflicting, approaches. For example, Steve might take on different personalities for different students giving him such personality characteristics as sanguine, melancholy, supportive, competitive, warm, formal, etc., yielding a wide array of different types to use for instruction, engagement, humor, and the like. Moreover, by design, the emotion model supports having more than one agent at a time so it is possible to have more than one Steve agent simultaneously interacting with the student, yielding, e.g., "good cop / bad cop," a chorus of approval, or agents with interests in different, albeit equally important, aspects of the domain.

Although not likely to be stressed in the first pass of this work, interesting, yet cohesive, dynamic changes in the ways Steve appraises the world can be made, by using emotion theory. For example, a negative emotion might lead Steve to have "self-directed attributions of negative worth." This in turn can be expressed as a change in the thresholds for variables used in the match process, putting him in a bad mood, or making it easier for him to be saddened, and harder for him to feel happy. Such a technique would allow us to support a pedagogical model that includes Steve becoming "depressed" over a student's continual poor attention, which in turn is consistent with fostering the belief that Steve and the student are both in the learning process together.

4.2 Emotion Intensity in Steve

In generating emotions, it is critical to assign them intensities that is commensurate with the situation in which they will be expressed. The following examples illustrate the two most important (from a full set of twenty-two) variables used in AR emotion intensity calculations. (For a full exposition see [8].)

In general, there are two major variables that contribute to emotion intensity. There are correlates for both principle-based emotions (e.g., *pride*), and goal-based emotions (e.g., satisfaction). The first of these, *simulation-event* variables, measure those factors which are external to any particular agent and might simultaneously, and differentially, affect more than one agent at the same time. An example of this might be the *number of parts damaged* when a compressor overheats. For example, consider that Steve has the following goals:

- **High Level Goal: Do not allow students to damage the (virtual) equipment.**
- **(Steve) Goal 1: Teach the student to check the oil level before starting the compressor (inherits from High Level Goal).**

Example situation: (a) compressor overheats and parts are damaged, and (b) history does not show an instance of the student checking the oil level.

In this case, the number of compressor parts which are damaged might affect the modeled experience of failure on Steve's part. The greater the number of damaged parts, the greater the failure. Now consider that the student has the following goals and principles (see Affective User Modeling section, below):

- **(Student) Goal 1: Do not fail on any tasks that the tutor gives me to do.**
- **Principle 1: It is wrong to break the virtual equipment.**
- **Principle 2: Steve should teach me well enough that I do not end up breaking the equipment.**

Simultaneously, the number of compressor parts which are damaged might also affect a student, Sarah, with respect to the above goal: she might be disappointed that she damaged the compressor; she might be ashamed of having damaged the compressor; she might be angry at Steve for failing to teach her how to avoid damaging the compressor. In each case, the extent of damage to the compressor is likely to be directly proportional to the degree of intensity in the negatively valenced emotions.

By contrast, the *stable disposition* variables are those which determine the *importance* of a particular goal, or principle, to an agent. These values are internal, and changes in them do not affect any other agents. For example, one Steve might be very concerned about safety, and damage to the equipment, whereas another Steve might be more concerned with exposing the student to explanations. For Safety-Steve, the *importance* of the equipment being damaged might be quite high, whereas for Explanation-Steve the importance might be quite low (or even help him to achieve a goal through affording him a chance to give an explanation). Since these are internal variables, and help to give the agents their own dispositional personalities, changes in the importance values for one agent will not affect another.

5 Affective User Modeling

So far, we have considered how affective reasoning can be used to generate emotive communication in pedagogical agents. The complementary functionality required of agents is the ability to model students' emotive states. We now turn to affective student modeling and illustrate its operation with examples drawn from a second pedagogical agent, Herman the Bug [10], and the learning environment he inhabits, Design-A-Plant (Figure 2), developed at North Carolina State University's Multimedia Laboratory.

Design-A-Plant is a knowledge-based learning environment project to investigate interactive problem-solving with animated pedagogical agents within the design-centered learning paradigm. With Design-A-Plant, students learn about botanical anatomy and physiology by graphically assembling customized plants that can thrive in specified environmental conditions. The Design-A-Plant work focuses on several intertwined aspects of introducing lifelike pedagogical agents into learning environments: dynamically sequencing the explanatory behaviors of animated pedagogical agents [20], user modeling and artifact-based task modeling [12], focusing learners' problem-solving activities [14], and increasing the believability of animated pedagogical agents [13]. Design-A-Plant's agent, Herman, performs a variety of lifelike, entertaining actions as it supplies advice to students when they solve design problems.

5.1 Overview of Affective User modeling Architecture

The general idea behind the model we are investigating is that AR agents have relatively reusable structures for appraising the world. The same structures that give them their own dispositions can be built and maintained for other agents as well. The vehicle for attempting to model some rudimentary form of the affective state of users is based on the following insights:

- AR agents have a dispositional component which determines how they appraise the world. This frame-based structure allows them to interpret situations that arise in ways that may give rise to emotion responses.
- Because agents have emotions about the fortunes of other agents, it is necessary for them to also maintain similar structures for these other agents. For example, if an agent's team wins he will be happy for himself, but might gloat over an agent rooting for the other team. To effect this the agent's own appraisal structure must result in an appraisal of an achieved goal from the situation, but the agent's own structure of the *presumed* goals of the second agent must result in an appraisal of a blocked goal from that same situation.
- Agents, who already keep these concerns-of-others structures, can maintain them for users as well.

A perfect structure of each individual user's goals, principles, and preferences, (e.g., a perfect affective user model, albeit begging the question of updating it correctly) would allow a great many correct inferences to be made about their emotion responses to the situations that arise while using the system. Since such a structure is not possible, it is necessary for us to use multiple types of inference in an attempt to approximate it using the following mechanisms:

1. *Inquiry:* Ask the user. In work with the AR, it appears to be true that users are *motivated* to express themselves to a computer agent who appears to have some understanding of how they feel.
2. *Stereotypes:* Use other known information to make assumptions about user *types*. Some users like to win, some like to have fun, some prefer to follow the rules, some are impatient. These qualities will tend to remain constant across tasks and domains.
3. *Context:* Use context information. For example, a user who has just repeatedly failed is likely to feel bad, whereas one who has been successful is likely to feel good.
4. *Affective Stereotypes:* Infer how *most users* would feel. The more user models extant, the stronger a prototype we have for a typical user.
5. *Self-Inspection:* If all else fails, infer what the agent would feel if it happened to him. Agents have affective "lives" too. One can always ask how they themselves would feel, and make the assumption that the user would feel that way too, i.e., the agent would filter the situation through its own appraisal mechanism and examine the resulting emotions which do, or do not, arise.

In general, our hypothesis is that we can be at least minimally effective at filling in missing information when working from a structure that specifies (a) what is *likely to be true* of the antecedents of user emotion, and (b) gives us a high-level understanding of different plausible affective user models, in a relatively comprehensive (albeit purely descriptive) model of human emotion. In other words, having a rigorously defined model of what user affect we are looking for helps us to map system events, and responses from the user to direct queries, effectively.

5.2 Examples of Student Goals and Principles

In this section we present a few brief examples of what we might want to model on behalf of the student, based on observations of seventh graders interacting with the Design-A-Plant software.

Example Student Goals

- **Goal 1: I want to do well on each task in the learning environment.** This is reflected in wanting always to get the right answer by selecting the correct component (roots, stems, and leaves) for the current environment. Sample situation: (a) student, or tutoring sequence, selects component (e.g., [stem, root, leaves]). (b) *correct-value* is bound to the correct value for chosen

component (e.g., type of stem). (c) *selected-value* is bound to the student-selected value for the chosen component. (d) components match, or (e) components do not match.
Sample intensity variables:

1. *Importance to student:* This is *high* by default, but this can be modified to fit individual personalities.
2. *Effort*: This increases as the student spends more time trying to solve this particular set of tasks.
3. *Anxiety-Invincibility:* If the student has had a pattern of recent success, then invincibility is increased. If a student has had a pattern of recent failures, then anxiety is increased.
4. *Arousal:* The default is *normal*. This can be modified to fit individual student profiles, and context. Context heuristics can include rules like the following: (1) If the student is performing normally, then, all other things being equal, the student is aroused at the *normal* level. If the student is either having an above-average (for the student) success rate, or a below average success rate then *arousal* is increased proportionally. (2) If student responses show little wall-clock dead time then student engagement is assumed to be higher and higher engagement affects arousal similarly.

- **Goal 2: I want to be entertained.**
  The developers of Design-A-Plant believe there to be a very strong "immediate gratification" element in student interaction with the learning environments. In general there are good mechanisms for providing what is likely to be entertaining (based on the long history of computer games, and entertainment systems of varied types), and general ad hoc heuristics for measuring entertainment levels (e.g., by market share). What we do *not* have is a method for agents to assess this dynamically as the interaction is taking place.
  Affective user modeling can address this in two ways: (1) It may be provably true that interesting interactive agents are more entertaining than static, or impoverished, agents are, or systems that operate without interactive agents. Agents with rich affective lives of their own can be extremely interesting to interact with. (2) Through observation we have some clues as to what is entertaining. Having an agent that makes inferences about student state by tracking situations believed to be entertaining may help the overall timing of the system. For example, we might believe that events we have tagged as, *funny*, *visual*, *with audio appeal*, *exhibiting cartoon effects* may be entertaining. A learning environment may call on resources to effect its limited arsenal of entertaining actions at more appropriate times (e.g., to cheer up a student perceived to be distressed).
- **Goal 3: I want to learn the material.**
  Through built-in testing procedures it is possible to effect an *estimate* of the *student's* assessment of how well they learned the material.

Example Student Principles. Principles are less well defined at this stage, but here are some that are based on observation with students of Design-A-Plant. These are *sample* principles, and it is clear that different students will hold different subsets of these, and others, in practice.

- **Principle 1: It is good to work cautiously and avoid making any errors.**
  Students who hold this principle will feel (some form of) *shame* when they make a mistake, and will be *proud* of themselves if they complete a set of tasks relatively error-free. This can be important, because for students who hold this to be true, it is inappropriate to praise them for, e.g., getting seven out of ten correct. In other words, it might be better to agree that it was sloppy work, and then suggest that if the student slows down, it might be possible to get them all right—it is sometimes easier to help a student achieve their own standards than it is to get them to change those standards.
  The principle might be adopted, for example, for students who respond in a particular way to simple questions (e.g., Herman: "If someone works fast, and gets seven out of ten right, is this good or bad?"; Student: "Bad"; Herman, "O.K. – I'll remember that. Thanks!")
  The general idea is that while most everyone would agree that it is important to address the situation when a student is frustrated, this is not always easy to assess: seven out of ten for one student means frustration, while for another it means success.
- **Principle 2: It is right to have fun with the system.**
  Most students would fall in this category, rather than the one above. They would not like to see themselves get too serious about the system.
- **Principle 3: It is right to be successful solving problems.**
  Not only might students be happy over achieving the goals of problem solving, they might also be *proud* of themselves for performing an admirable action in doing so.
- **Principle 4: Long explanations (e.g., of photosynthesis) at inappropriate times (and possibly all times are inappropriate) are wrong and should not be inflicted on me.**
  Herman might benefit from knowing that giving long explanations might make the student angry, even though he might still give the explanation for other reasons. It might also be possible for Herman to discriminate between a typical student in a typical state, and one that is particularly unreceptive to a long explanation: it would not be wise to give a long explanation to a student already perceived to be in an angry or frustrated state.

Inference with Intensity Variables. Simulation events in the AR are a frame representation of the salient points of situations that arise in the course of interaction with the user. In learning environments these would take a theoretically equivalent form, regardless of the actual implementation. Agents maintain internal representations of what they believe to be true of the appraisal mechanisms (e.g.,the *dispositions*) of the students, and call on these for interpreting the supposed effect of simulation events on a student. For example, if an agent believes that student Sarah has a strong desire (e.g., the stable-disposition variable *importance* is high) to succeed on Task A, but that she does not care much about Task B, then the agent might feel *pity* for Sarah if she fails on Task A, or *happy-for* her if she succeeds, but would have no *fortunes-of-other* response for Sarah's relative success, or lack thereof, with Task B.

As part of the internal models that agents keep of other agents, including the student, they may update *mood variables* dynamically, which tend to affect the thresholds at which emotions arise. Therefore, if an agent believed Sarah to be feeling particularly *anxious*, he might, after all, feel *pity* for Sarah's failure on Task B, because failure on even a relatively unimportant (to her) task such as Task B might be perceived as surpassing the lowered threshold for activation of the distress emotion. Similarly, if the agent believed Sarah to be feeling particularly *invincible* (e.g., after a string of grand successes), he might not believe Sarah to be distressed about failure on the important (to her) Task A, and hence might not feel *pity* for her.

6 Concluding Remarks

This paper has provided a brief overview of how an affective reasoning framework can be introduced into lifelike pedagogical agents. Our objective has been to sketch a model of how an an agent might express emotions and evaluate the effects of these emotions on student learning. The required affective reasoning naturally decomposes into two areas: emotional responsiveness of the tutor itself, and affective user modeling. We have illustrated some of the issues involved in these two areas using two extant agents, Steve and Herman the Bug.

The next steps in integrating affective reasoning into lifelike pedagogical agents are suggested by the following research agenda. First, techniques must be created to perform comparisons of an agent's goals and principles against simulation events and student actions to identify appropriate triggering conditions for emotions. A promising starting place for this work is example goals and principles described in this paper. Second, mechanisms need to be created for converting these triggering conditions into appropriate emotional states. Conversion processes are likely to be complex and, therefore, non-trivial to design. Finally, mechanisms need to be created for expressing the internal emotive states of agents to user. We have seen some progress in the latter area [15] and will be pursuing the others activities in our future work.

7 Acknowledgments

We would like to thank Lewis Johnson and Gary Stelling for their valuable comments throughout the course of this work. The authors also gratefully acknowledge Charles Callaway for his assistance in preparing in the manuscript and Bradford Mott for comments on an earlier draft of this paper. Support for this work was provided by the Office of Naval Research under grant N00014-95-C-0179; an internal research and development grant from the USC Information Sciences Institute; the National Science Foundation under grants CDA-9720395 (Learning and Intelligent Systems Initiative) and IRI-9701503 (CAREER Award Program); the North Carolina State University IntelliMedia Initiative; and an industrial gift from Novell.

References

1. Sassine Abou-Jaoude and Claude Frasson. Emotion computing in competitve learning environments. In *Working Notes of the ITS '98 Workshop on Pedagogical Agents*, pages 33–39, San Antonio, Texas, 1998.
2. Elisabeth André, Thomas Rist, and Jochen Müller. Integrating reactive and scripted behaviors in a life-like presentation agent. In *Proceedings of the Second International Conference on Autonomous Agents*, pages 261–268. Minneapolis, 1998.
3. Angel de Vicente and Helen Pain. Motivation diagnosis in intelligent tutoring systems. In *Proceedings of the Fourth International Conference on Intelligent Tutoring Systems*, pages 86–95, San Antonio, Texas, 1998.
4. Clark Elliott. *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System.* PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.
5. Clark Elliott. I picked up catapia and other stories: A multimodal approach to expressivity for "emotionally intelligent" agents. In *Proceedings of the First International Conference on Autonomous Agents*, pages 451–457, 1997.
6. Clark Elliott and Jack Brzezinski. Autonomous agents as synthetic characters. *AI Magazine*, 19(2), 1998.
7. Clark Elliott and Andrew Ortony. Point of view: Reasoning about the concerns of others. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, pages 809–814, Bloomington, IN, August 1992. Cognitive Science Society.
8. Clark Elliott and Greg Siegle. Variables influencing the intensity of simulated affective states. In *AAAI technical report SS-93-05 for the Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, pages 58–67. American Association for Artificial Intelligence, 1993. Stanford University, March 23-25, Palo Alto, CA.
9. Pentti Hietala and Timo Niemirepo. The competence of learning companion agents. *International Journal of Artificial Intelligence in Education*, 1998.
10. James Lester, Brian Stone, and Gary Stelling. Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments. *User Modeling and User-Adapted Interaction*, 1999. In press.
11. James C. Lester, Sharolyn A. Converse, Susan E. Kahler, S. Todd Barlow, Brian A. Stone, and Ravinder Bhogal. The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of CHI'97 (Human Factors in Computing Systems)*, pages 359–366, Atlanta, 1997.
12. James C. Lester, Patrick J. FitzGerald, and Brian A. Stone. The pedagogical design studio: Exploiting artifact-based task models for constructivist learning. In *Proceedings of the Third International Conference on Intelligent User Interfaces*, pages 155–162, Orlando, Florida, 1997.

13. James C. Lester and Brian A. Stone. Increasing believability in animated pedagogical agents. In *Proceedings of the First International Conference on Autonomous Agents*, pages 16–21, Marina del Rey, CA, 1997. ACM Press.
14. James C. Lester, Brian A. Stone, Michael A. O'Leary, and Robert B. Stevenson. Focusing problem solving in design-centered learning environments. In *Proceedings of the Third International Conference on Intelligent Tutoring Systems*, pages 475–483, Montreal, 1996.
15. James C. Lester, Stuart G. Towns, and Patrick J. FitzGerald. Achieving affective impact: Visual emotive communication in lifelike pedagogical agents. *International Journal of Artificial Intelligence in Education*, 1999. In press.
16. Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.
17. Ana Paiva and Isabelle Machado. Vincent, an autonomous pedagogical agent for on-the-job training. In *Proceedings of the Fourth International Conference on Intelligent Tutoring Systems*, pages 584–593, San Antonio, Texas, 1998.
18. Rosalind W. Picard. *Affective Computing*. MIT Press, 1997.
19. Jeff Rickel and W. Lewis Johnson. Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, 1998. Forthcoming.
20. Brian A. Stone and James C. Lester. Dynamically sequencing an animated pedagogical agent. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, pages 424–431, Portland, Oregon, 1996.
21. Stuart Towns, Charles Callaway, Jennifer Voerman, and James Lester. Coherent gestures, locomotion, and speech in life-like pedagogical agents. In *Proceedings of the Fourth International Conference on Intelligent User Interfaces*, pages 13–20, San Francisco, 1998.
22. Juan D. Velasquez. Modeling emotions and other motivations in synthetic agents. In *AAAI-97: Proceedings of the Fourteenth National Conference on Artificial Intelligence*, pages 10–15, 1997.