Modeling Task-Based vs. Affect-based Feedback Behavior in Pedagogical Agents: An Inductive Approach

Jennifer L. ROBISON[a,1], Scott W. MCQUIGGAN[b] and James C. LESTER[a]
[a] *Department of Computer Science, North Carolina State University, Raleigh NC*
[b] *SAS Institute, Cary NC*

Abstract. Affect has been the subject of increasing attention in cognitive accounts of learning. Many intelligent tutoring systems now seek to adapt pedagogy to student affective and motivational processes in an effort to increase the effectiveness of tutorial interaction and improve learning outcomes. However, the majority of research on tutorial feedback has focused on pedagogical content, often at the expense of the affective component of the learning process. It is unclear under which circumstances it is more appropriate to focus directly on student affect and when support is best offered through task-related feedback. This paper proposes an inductive framework for modeling task-based and affect-based feedback to inform the behavior of pedagogical agents within a narrative-centered learning environment.

Keywords. Empathy, Affect Modeling, Task-Based Feedback, Pedagogical Agents.

1. Introduction

Affect has begun to play an increasingly important role in intelligent tutoring systems. The AI in Education community has seen the emergence of work on affective student modeling [1], detecting frustration and stress [2, 3], modeling agents' emotional states [4, 5], devising affectively informed models of social interaction [6, 7], detecting student motivation [8], and diagnosing and adapting to student self-efficacy [9]. All of this work seeks to increase the fidelity with which affective and motivational processes are understood and utilized in intelligent tutoring systems in an effort to increase the effectiveness of tutorial interactions and, ultimately, learning.

This level of emphasis on affect is not surprising given the effects it has been shown to have on learning. Student affective states impact problem-solving strategies, the level of engagement the student feels with the environment and how motivated she is to continue with the learning process [10, 11, 12, 13]. All of these factors have the potential to impact both how a student learns immediately and her learning behaviors in the future. In this work, we investigate techniques for keeping students in an affective state that is conducive to learning through the use of pedagogical agent feedback in a narrative-centered learning environment, CRYSTAL ISLAND.

[1] Corresponding Author: Jennifer L. Robison, Department of Computer Science, North Carolina State University, Raleigh NC 27695; E-mail: jlrobiso@ncsu.edu

A challenging problem is determining when it is most useful to focus on the students' affective state, or when students prefer assistance with the learning task at hand. Some research into this area suggests that cognitive feedback is superior to motivational support, which often has little or negative effects [14, 15]. However, these results are found in one-to-one tutorial dialog sessions on well-defined problem sets, and the type of affective response is limited to concepts such as "praise" or "encouragement." In narrative-centered learning environments, which are game-based learning environments populated with a rich cast of interactive characters, a much more flexible approach to affective feedback is permitted. Here, the notion of a pedagogical agent and a companion agent may be effectively blended.

This paper reports on a study that is a preliminary exploration of the factors that may impact the utility of task-based and affect-based feedback styles in pedagogical agents in narrative-centered learning environments. It presents an empirically-based model that can predict situations in which students prefer either task-based or affect-based feedback. While it is likely that effective feedback strategies may include both aspects of affect and task-based feedback, these strategies are considered separately for the purposes of this investigation. The proposed model has been induced using training sessions during which the system monitors student situation data (actions, visited locations, and intentions) and affective states while a training user directs her virtual agent to perform a sequence of tasks in a narrative-centered learning environment. Meanwhile, pedagogical characters respond to student situations with task-based hints and suggestions or affect-based empathetic responses. During interactions with pedagogical agents, students are able to evaluate the helpfulness of these responses. These ratings are used to induce a model of appropriate feedback behavior which is then used at runtime to provide students with the most appropriate feedback.

2. Background

2.1 Affect in Interactive Learning Environments

Understanding and categorizing student affective experiences has been the subject of a large body of research conducted with a broad range of learning environments. These range from AutoTutor [17], a dialog-based problem-solving environment, the Incredible Machine, a well-structured problem-solving game [18], and CRYSTAL ISLAND, an exploratory narrative-based learning environment [19]. Even with the differences these environments exhibit, striking similarities in affective experience have been found, such as the dominance of the flow state [17, 18, 19]. There are also important differences categorizing these experiences, such as the increased prevalence of narrative focused emotions such as excitement and anger [19].

2.2 Character Feedback

The characters in the CRYSTAL ISLAND environment serve both narrative and pedagogical roles. This permits a wide range of possible behavior, including empathetic behaviors, which can form the basis for the delivery of an affective feedback strategy. Empathy is defined as an awareness of another's affective state that generates emotions in the empathizer that reflect more than her own situation in attempt to foster a feeling of understanding or to motivate a more positive affective state [20]. Empathetic approaches to student affect have been shown to alter the affective state of the student as well as other qualities such as motivation [17, 19]. Recent work has yielded models of when empathetic response is appropriate [21], how it ought to be delivered and when parallel or reactive empathy is preferable [22]. These behaviors have also been shown to have an impact on the affective experiences of students [19]. Collectively, the results of these studies provide the foundation for the affective feedback used in the study reported here.

The following strategies were used to deliver task-based feedback in response to student affect. If students report a positive affective state, it is assumed that they are proceeding well through the environment, or are at least confident that they will be able to achieve the goal on their own. Therefore, to avoid interrupting this state, we provide feedback in the form of a summary of their current progress. This serves to ground their current state within the environment and provide reinforcement of previously learned material [23]. However, if a student reports a negative state, we intervene to try to assist them in overcoming the source of this problem. Because of the exploratory nature of the CRYSTAL ISLAND environment, sources of negative emotions are difficult to distinguish. A student may be confused over some aspect of pedagogical content or over the next problem-solving action she is supposed to perform in the narrative. Rather than attempt to guess at the source of negative emotion and risk an inappropriate response, we direct students towards information that will help them complete the goal. This may mean reading a book to acquire additional domain knowledge or talking to a character who can provide further direction. In addition to these hints, we also provide a short summary of learned information similar to the response to positive affect to again provide reinforcement and grounding. All responses are kept brief (2-3 sentences) to avoid overloading the student or leading the student to avoid reading the material all together [15].

3. The CRYSTAL ISLAND Learning Environment

The study was conducted in a narrative-centered inquiry-based learning environment, CRYSTAL ISLAND (Figure 1). This environment is being created in the domain of microbiology for middle school students. It features a science mystery set on a recently discovered volcanic island where a research station has been established to study the

Figure 1. The user, Alex, with Jin, the camp nurse, on CRYSTAL ISLAND.

unique flora and fauna. The student plays the protagonist, Alex, who is attempting to discover the source of an unidentified infectious disease at the research station. The story opens by introducing the student to the island and the members of the research team for which her father serves as the lead scientist. As members of the research team fall ill, it is her task to discover the cause and the specific source of the outbreak. She is free to explore the world and interact with other characters while forming questions, generating hypotheses, collecting data, and testing hypotheses. Throughout the mystery, she can walk around the island and visit the various locations. She can pick up and manipulate objects, and she can talk with characters to gather clues about the source of the disease. In the course of her adventure she must gather enough evidence to correctly diagnose the members of the research team. For more details on the CRYSTAL ISLAND learning environment, please see [22].

4. Inductive Framework

Accurately modeling feedback strategies requires a representation of the situational context that satisfies two requirements: it must be sufficiently rich to support assessment of affective and problem-based features, and it must be encoded with features that are readily observable at runtime. Affect can be viewed as a cognitive process in which the student appraises the relationship between herself and her environment [5, 24]. Similarly, affect-based empathy draws heavily on appraisal of the situation at hand in addition to student affect. Therefore, feedback models should take into account environmental information in addition to student affective states. Task-based feedback strategies should leverage knowledge of problem and learning task structures as well as the state of the student in the environment to effectively assess if possible hints will be useful or needed. The specific goal that the student is currently attempting to achieve and their progress toward achieving that goal may determine the effectiveness of task-based feedback. Additionally, individual characteristics of the student may impact their appraisal of various feedback strategies. Traits such as goal orientation and empathetic reactivity may strongly affect student preferences. Therefore, attributes relating to student situation (such as the student's current actions, locations, and goals, as well as artifacts manipulated, previous locations visited, and the characters with which the student has interacted), student characteristics (demographics, personality traits, goal orientation, etc.) and student affect (obtained directly from students via pedagogical agent inquiries) were collected for use in modeling.

5. Method

To empirically investigate preferred feedback strategies, a study was conducted with subjects interacting with pedagogical agents. The subjects of the study consisted of 41 college students ranging in age from 19 to 38 (M = 24.0, SD = 3.96) including 12 females and 29 males. Among these students, approximately 73% were Caucasian (n = 30), 17% were Asian (n = 7), and 10% were Other (n = 4).

Participants were given an overview of the experiment agenda, and they completed the pre-experiment questionnaires including a demographics survey, a goal orientation survey [25], a personality questionnaire [26] and the interpersonal reactivity index survey [27], which measures subjects' empathetic nature. They were given 35 minutes to solve the mystery. Solving the mystery consisted of completing 15 goals including learning about various diseases, compiling the symptoms of the sickened researchers, testing a variety of possible sources, and reporting the solution (cause and source) back to the camp nurse.

When subjects chose to interact with one of the six virtual agents, the following schema was used to direct subject-character interactions and virtual character feedback:

1. The agent queries the subject for a self-reported affective state. The subject is presented with a dialog box posing the question, "Hi Alex, how are you feeling?" The subject may respond by selecting one of the 9 available emotions (*anger, anxiety, boredom, confusion, curiosity, delight, excitement, flow,* and *frustration*).
2. The agent then responds to the subject's reported affective state with a randomized response of either affect-based empathy or task-based hints or summary.
3. A follow-up dialog box is then presented to the subject asking her to respond with the prompt, "… and you respond." The subject is able to choose from 4 Likert-scaled responses designed to evaluate the appropriateness and effectiveness of the virtual character's empathetic response. Subjects can issue responses ranging from (1) "That does not help me at all." to (4) "Thanks, that helped a lot!"
4. The agent responds with a one-word quip (e.g., "Thanks," or "Great!") directed towards the subject's evaluation response (Step 3).
5. At the conclusion of the interaction, the agent again asks the subject how she feels. The subject is presented a dialog box similar to the one described in Step 1 without the character greeting. Here, the character prompts the subject with, "How are you feeling now?"

6. Results

In Step 3 of the user-agent interaction schema presented above, subjects evaluated agent responses as part of the "conversation" with the character. The distribution of responses (affect-based empathy and task-based suggestions) and the associated ratings are detailed in Table 1. The first two rows show the percentage of responses that were

Table 1. Distribution of empathetic responses and associated subject evaluation ratings (4 = high to 1 = low).

		Character Responses		
		Task-based	**Empathy**	**Total**
Evaluative Rating	**4**	19%	12%	31%
	3	13%	8%	21%
	2	10%	14%	24%
	1	10%	13%	24%
	Total	53%	47%	100%

Table 2. Model results by data type used for learning (left column bold items) and dataset used. Highest-Rated refers to the dataset containing responses rated a 4. Favorably-Rated refers to the dataset containing responses rated either a 3 or 4.

	Dataset	
Induced Models	**Highest-Rated**	**Favorably-Rated**
Baseline (Task-based)	0.62	0.61
Situation Attributes Only		
Naïve Bayes	0.65	0.66
Decision Tree	0.71	0.70
Situation Attributes + User Affect		
Naïve Bayes	0.72	0.70
Decision Tree	0.83	0.82
Situation Attributes + Affect + User Characteristics		
Naïve Bayes	0.76	0.72
Decision Tree	0.95	0.96

found to be appropriate by subjects (i.e., the instances of empathy and task-based responses that were given an evaluative rating of 3 or 4), while the next two rows indicate responses that were found to be inappropriate by subjects (i.e., the instances of empathy and task-based empathetic behaviors that were given an evaluative rating of 1 or 2). Overall, task-based responses received a higher rating (M = 2.77, SD = 0.02) than empathy responses (M = 2.39, SD = 0.02), $t(342) = 13.2$, $p < 0.0001$.

Naïve Bayes and decision tree models were induced from data collected in the training sessions using the WEKA machine learning toolkit [28]. All models were constructed using a tenfold cross-validation scheme for producing training and testing datasets, a widely used method to obtain an acceptable estimate of error [28].

Two distinct datasets were used. The first dataset was comprised only of responses receiving high ratings of 4 (n = 1006) from subjects during conversations with virtual characters. The second dataset was comprised of responses receiving a rating of either a 3 or 4 (n = 1750). Within each dataset, three versions of models were learned from the various types of observational attributes (situation data and student characteristics and affect reports). Table 2 provides results of all induced models and baselines.

A baseline measure determines the most frequent class label (in this case task-based response) and predicts all responses to call for task-based responses. For the dataset containing only responses rated level 4 (*Highest-Rated*), task-based responses accounted for 62% of the instances. Task-based responses occurred in 61% of the instances in the dataset containing responses rated as level 3 or 4 by subjects (*Favorably-Rated*).

All induced models outperformed baseline models. The improvement of induced models over baselines is statistically significant. For instance, the least accurate induced model from the Highest-Rated dataset is the naïve Bayes model (65% accuracy), which was learned from situation attributes only. The increase in accuracy of this naïve Bayes model over baseline accuracy was statistically significant, $\chi^2(1, N = 1006) = 4.46$, $p = 0.04$. The best performing model was a decision tree induced from the favorably-rated situation data, student affect and student characteristics. This model is able to accurately predict whether to use a task-based or empathetic response 96% of the time.

7. Conclusion

Pedagogical agents can utilize both task-based and affective feedback to mediate student affect in conjunction with context-appropriate selection of feedback strategies. A promising approach for acquiring multi-strategy feedback models is by learning them from training sessions. In a study of multi-strategy feedback, it was found that task-based feedback was more frequently rated as helpful behavior, and that in certain situational and affective conditions, it appears that some students benefit more from affect-based empathetic responses than from hints and suggestions. Additionally, the increased performance of models including affect over those monitoring situational data alone highlight the importance of incorporating affective factors when devising feedback strategies. By integrating models such as the ones reported here into pedagogical agents, it may be possible to devise agents that can simultaneously optimize for both learning gains and motivation.

This study provides a first step towards modeling multiple types of feedback, and there are many promising directions for future work. Perhaps first among these is analyzing the impact of feedback strategies on student affect. While student ratings offer one indicator of the benefits of each strategy, they do not tell the whole story. To devise an optimal feedback model, features such as impact on affect, motivation and learning gains must be taken into account. Another promising line of work is incorporating automated affect detection into pedagogical agents. In this way, we will no longer be reliant on student reports and may also gain the ability to intervene with agent behavior at the onset of negative emotions rather than waiting for potentially damaging interactions to occur. Finally, given the encouraging results of this first attempt, the study underscores the need for further exploration of task-based feedback strategies and their integration with empathetic behaviors.

Acknowledgments

The authors wish to thank the members of the IntelliMedia research lab for their assistance in implementing CRYSTAL ISLAND, Omer Sturlovich and Pavel Turzo for use of their 3D model libraries, and Valve Software for access to the Source™ engine and SDK. This research was supported by the National Science Foundation under REC-0632450, IIS-0757535, DRL-0822200 and IIS-0812291. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. 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

[1] C. Conati and H. Mclaren, Data-driven refinement of a probabilistic model of user affect, in *Proceedings of the 10th International Conference on User Modeling* (2005), 40-49.

[2] W. Burleson, 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, 2006.

[3] S. McQuiggan, S. Lee and J. Lester, Early prediction of student frustration, in *Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction*, (2007).

[4] E. André and M. Mueller, Learning affective behavior, in *Proceedings of the 10th International Conference on Human-Computer Interaction* (2003), 512-516.

[5] J. Gratch and S. Marsella, A domain-independent framework for modeling emotion, *Journal of Cognitive Systems Research*, **5** (2004), 269-306.
[6] L. Johnson and P. Rizzo, Politeness in tutoring dialogs: “Run the factory, that’s what I’d do”, in *Proceedings of the 7th International Conference on Intelligent Tutoring Systems* (2004), 67-76.
[7] A. Paiva, J. Dias, D. Sobral, R. Aylett, S. Woods, L. Hall and C. Zoll, Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, **19** (2005), 235-266.
[8] A. de Vicente and H. Pain, Informing the detection of the students’ motivational state: an empirical study, in *Proceedings of the 6th International Conference on Intelligent Tutoring Systems* (2002), 933-943.
[9] C. Beal and H. Lee, Creating a pedagogical model that uses student self reports of motivation and mood to adapt ITS instruction, in *Workshop on Motivation and Affect in Educational Software, in conjunction with the 12th International Conference on Artificial Intelligence in Education* (2005).
[10] B. Kort, R. Reilly and R. Picard, An affective model of interplay between emotions and learning: Reengineering educational pedagogy—building a learning companion, in *Proceedings of the IEEE International Conference on Advanced Learning Technology: Issues, Achievements and Challenges* (2001), 43-48.
[11] D. Goleman, *Emotional Intelligence.* Bantam Books, New York, 1995.
[12] R. Picard, S. Papert, W. Bender, B. Blumberg, C. Breazeal, D. Cavallo, T. Machover, M. Resnick, D. Roy and C. Strohecker, Affective learning – a manifesto, *BT Technology Journal*, **22** (2004).
[13] N. Schwarz, Emotion, cognition, and decision making, *Journal of Cognition and Emotion*, **14** (2000), 443-440.
[14] K. Boyer, R. Phillips, M. Wallis, M. Vouk and J. Lester, Balancing cognitive and motivational scaffolding in tutorial dialog, in *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (2008), 239-249.
[15] V. Shute, *Focus on Formative Feedback,* ETS, Princeton, NJ, 2007.
[16] S. D’Mello, R. Taylor and A. Graesser, Monitoring affective trajectories during complex learning, in *Proceedings of the 29th annual meeting of the cognitive science society* (2007), 203-208.
[17] S. D’Mello, T. Jackson, S. Craig, B. Morgan, P. Chipman, H. White, N. Person, B. Kort, R. el Kaliouby, R. Picard and A. Graesser, AutoTutor Detects and Responds to Learners Affective and Cognitive States, in *Workshop on Emotional and Cognitive issues in ITS in conjunction with the 9th International Conference on Intelligent Tutoring Systems* (2008), 31-43.
[18] R. Baker, M. Rodrigo and U. Xolocotzin, The dynamics of affective transitions in simulation problem-solving environments, in *Proceedings on the 2nd Annual International Conference of Affective Computing and Intelligent Interactions* (2007), 666-677.
[19] S. McQuiggan, J. Robison and J. Lester, Affective transitions in narrative-centered learning environments, in *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (2008).
[20] M. Davis, *Empathy: A Social Psychological Approach*, Brown and Benchmark Publishers, Madison, WI, 1995.
[21] S. McQuiggan and J. Lester, Learning empathy: a data-driven framework for modeling empathetic companion agents, in *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multi-Agent Systems* (2006), 961-968.
[22] S. McQuiggan, J. Robison, R. Phillips and J. Lester, Modeling parallel and reactive empathy in virtual agents: an inductive approach, in *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems* (2008).
[23] A. Graesser, N. Person and J. Magliano, Collaborative dialogue patterns in naturalistic one-to-one tutoring, *Journal of Applied Cognitive Psychology* **9** (1995), 495-522.
[24] A. Kapoor and R. Picard, Multimodal affect recognition in learning environments, in *Proceedings of the 13th annual ACM International Conference on Multimedia* (2005), 677-682.
[25] A. Elliot and H. McGregor, A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, **80** (2001), 501-519.
[26] R. McCrae and P. Costa, *Personality in adulthood: A five-factor theory perspective* (2nd ed.), New York: Guilford Press, 1993.
[27] M. Davis, Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, **44** (1983), 113-126.
[28] I. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.), Morgan Kaufman, San Francisco, CA, 2005.