

## Run-Time Affect Modeling in a Serious Game with the Generalized Intelligent Framework for Tutoring

Jonathan P. Rowe, Eleni V. Lobene, Jennifer L. Sabourin,  
Bradford W. Mott, and James C. Lester

*Department of Computer Science, North Carolina State University, Raleigh, NC 27695  
{jprowe, eleni.lobene, jlrobiso, bwmott, lester}@ncsu.edu*

**Abstract.** Affective computing holds significant promise for fostering engaging educational interactions that produce significant learning gains. Serious games are particularly well suited to promoting engagement and creating authentic contexts for learning scenarios. This paper describes an ongoing collaborative project between the Army Research Lab (ARL), Teachers College Columbia University, and North Carolina State University to investigate generalized run-time affect detection models in a serious game for tactical combat casualty care, vMedic. These models are being developed and integrated with ARL's Generalized Intelligent Framework for Tutoring (GIFT). Drawing upon our experience with GIFT, we outline opportunities for enhancing GIFT's support for developing and studying run-time affect modeling, including extensions that enhance affective survey administration, leverage mathematical models for formative assessment, and streamline affect data processing and analysis.

**Keywords:** Affect Detection, GIFT, Serious Games.

### 1 Introduction

The past decade has witnessed major advances in research on computational models of affect, endowing software systems with affect-sensitivity and yielding new insights into artificial and human intelligence [1]. Education and training have served as key application areas for computational models of affect, producing intelligent tutoring systems (ITSs) that can model students' affective states [2], model virtual agents' affective states [3], and detect student motivation and engagement [4]. Education-focused work on affective computing has sought to increase the fidelity with which affective and motivational processes are understood and utilized in ITSs in an effort to increase the effectiveness of tutorial interactions and, ultimately, learning.

The rise of affective computing has coincided with growing interest in digital games for learning. Serious games have emerged as an effective vehicle for learning and training experiences [5]. The education community has developed a broad range of serious games that combine pedagogy and interactive problem solving with the salient properties of games (e.g., feedback, challenge, rewards) to foster motivation and engagement [6–8]. Efforts to design serious games for training have also been the subject of increasing interest in the defense community [6, 9].

A notable property of serious games is their potential to serve as virtual laboratories for studying affect in learning and training applications. Serious games are well

suiting to promoting high levels of learner engagement and providing immersive training experiences. These features can have significant impacts on learners' affective trajectories, as well as the relationships between learners' affect and performance. For example, in training tasks that evoke considerable stress or anxiety it is plausible that serious games may foster affective experiences that differ considerably from non-mission-critical domains, significantly impacting learners' abilities to successfully demonstrate their knowledge. Salient features such as these raise questions about how to most effectively study and model learner affect during interactions with serious games, as well as questions about how these methods and models can be generalized to other training environments and domains.

In this paper we describe a collaborative project with Teacher's College Columbia University (TC) and the Army Research Lab (ARL) that uses the Generalized Intelligent Framework for Tutoring (GIFT) to investigate run-time affect modeling in a serious game for tactical combat casualty care. The project draws on recent advances in five areas: minimally-obtrusive and synchronize-able field observations of learner affect [10], empirical studies of serious games [7], educational data mining of affect logs [11–12], hardware sensor-based measurements of affect [13], and generalized intelligent tutoring frameworks [14]. The project's objectives are two fold: 1) create modular intelligent tutor components for run-time affect modeling that generalize across multiple training environments and scale to alternate hardware configurations, and 2) develop tools and procedures to facilitate future research on affective computing in learning technologies. This paper focuses on North Carolina State University's component of the project, which emphasizes sensor-based affect detection, and it outlines recommendations for future enhancements to GIFT in support of run-time affect modeling. Specifically, we outline several opportunities for extending GIFT, which include incorporating support for temporal models of affect such as affect transitions; expanding GIFT's survey tools to serve as a centralized repository of validated instruments with an integrated web-based infrastructure for administering surveys; taking advantage of item response theory techniques to conduct stealth, formative assessment of trainee attitudes during learning interactions; and incorporating features to streamline affect data post-processing.

## **2 Investigating Affect in a Serious Game for Tactical Combat Casualty Care**

The goal of our collaboration with ARL and TC is to model trainee affect in a serious game for combat medic training, vMedic, using GIFT. The research team will utilize machine-learning techniques to induce models for detecting trainee's affective states and levels of engagement during interactions with the vMedic software. Affect and engagement significantly influence learning, and we hypothesize that this will be especially true for the vMedic training environment due to the time-sensitive, life-or-



**Fig. 9.** vMedic serious game for tactical combat casualty care.

death decisions inherent in tactical combat casualty care. In combination with field observations of trainee affect and trace data from the vMedic serious game, the North Carolina State University team will investigate data streams produced by a Microsoft Kinect sensor and Affectiva Q-Sensor to develop and validate affect detection models. The research team seeks to produce models that 1) integrate trace data logs, sensor data, and field observations of trainee emotions; 2) predict emotions accurately and efficiently when hardware sensors are available; and 3) scale gracefully to settings where hardware sensors are unavailable. The models will be developed and utilized to improve trainee engagement and affect when using vMedic, and they will be integrated with interaction-based models devised by colleagues at TC.

The curriculum for the study focuses on a subset of skills for tactical combat casualty care: care under fire, hemorrhage control, and tactical field care. The study materials, including pre-tests, training exercises, and post-tests, are managed entirely by GIFT, which supports inter-module communication through its service-oriented architecture. At the onset of training, learners are presented with direct instruction about tactical combat casualty care in the form of a PowerPoint presentation. After completing the PowerPoint, participants play through a series of scenarios in the vMedic serious game. vMedic presents combat medic scenarios from a first-person perspective (Fig. 1). The learner adopts the role of a combat medic faced with a situation where one (or several) of his fellow soldiers has been seriously injured. The learner is responsible for properly treating and evacuating the casualty. The scenarios include the following elements: a tutorial level for trainees to learn the controls and game mechanics of vMedic; a scenario focusing on a lower leg amputation; a vignette about a patrol that leads to several casualties; and the “Kobayashi Maru” scenario where the trainee cannot save the casualty’s life regardless of her course of medical treatment. vMedic is currently being used at scale by the U.S. Army for combat medic training, and it has been integrated with GIFT by ARL.

The focus of North Carolina State University’s part of the project is leveraging hardware sensor data from a Microsoft Kinect for Windows and Affectiva Q-Sensor to generate affect detection models. Both hardware sensors are integrated with GIFT,

enabling the sensor data to be automatically synchronized with vMedic and PowerPoint interaction logs. This architecture removes the need to directly integrate hardware sensors with individual learning environments. Whenever a new training environment is integrated with GIFT, no additional work is required to use the hardware sensors with the new environment.

The Microsoft Kinect provides four data channels: skeleton tracking, face tracking, RGB (i.e., color), and depth. The first two channels leverage built-in tracking algorithms (which are included with the Microsoft Kinect for Windows SDK) for recognizing a user's skeleton, represented as a graph with vertices as joints, and a user's face, represented as a three-dimensional polygonal mesh. The skeleton and face models can move, rotate, and deform based on the user's head movements and facial expressions. The RGB channel is a 640x480 color image stream comparable to a standard web camera. The depth channel is a 640x480 IR-based image stream depicting distances between objects and the sensor. The latter two channels produce large quantities of uncompressed image data, so configuration options have been added to GIFT to adjust the sample rate (default is 30 Hz), sample resolution, and compression technique. RGB and depth data can be stored in an uncompressed format, in PNG format with zlib compression, or in PNG format with lz4 compression. We intend to utilize data from the Microsoft Kinect to detect user posture, hand gestures, and facial expression. The Affectiva Q-Sensor is a wearable arm bracelet that measures participants' electrodermal activity (i.e., skin conductance), skin temperature, and its orientation through a built-in 3-axis accelerometer. The wireless sensor collects data at 32Hz, and will primarily be used for real-time arousal detection.

Since all technology components in the planned study are managed by GIFT, we have leveraged GIFT's built-in authoring tools to specify the study questionnaires and curriculum tests for assessing trainee knowledge and engagement before and after the learning intervention. We have utilized GIFT's Survey Authoring Tool to rapidly integrate standard presence and intrinsic motivation questionnaires. Additionally, we have used GIFT's sizable repository of reusable content assessment items to create a curriculum test for measuring learning gains across the training sequence.

After specifying the required measures, we used GIFT's Course Authoring Tool to encode the sequence of training and assessment materials that will be presented by GIFT. The Course Authoring Tool includes support for authoring web-based messages that provide instructions to participants, specifying the presentation order of pre- and post-intervention questionnaires and content tests, and specifying the sequence of PowerPoint and vMedic learning activities that occur during the study. When participants take the course, each of these steps is automatically triggered, monitored, and logged by GIFT. It should be noted that authored courses and questionnaires can be easily exported and shared between groups, consistent with GIFT's objective of fostering reusable components.

Currently, our team has established the initial data collection's study procedure, we have tested the integrated hardware sensors, and ensured the reliability of the study's technology setup. In addition to pilot testing field observation tools from TC with GIFT, we are in the process of planning a study at the U.S. Military Academy to investigate cadets' affective experiences during interactions with vMedic.

### 3 Extending GIFT's Capabilities for Run-Time Affect Modeling

GIFT consists of a suite of software tools, standards, and resources for creating, deploying, and studying modular intelligent tutoring systems with re-usable components. GIFT provides three primary functions: authoring capabilities for constructing learning technologies, instruction that integrates tutorial principles and strategies, and support for evaluating educational tools and frameworks. These capabilities provide the foundation for our investigation of generalizable run-time affect models. This section discusses several areas for which extensions to GIFT could support research and development of generalized affect models in ITSs.

#### 3.1 Detecting and Understanding Learner Affect

While considerable work remains to identify the precise cognitive and affective mechanisms underlying learning, significant progress has been made in identifying the emotions that students commonly experience and how these affect the learning process. For instance, both D'Mello *et al.* [15] and Baker *et al.* [16] have shown that students are most likely to remain in the same emotion state over time and that certain emotional transitions are more likely than others. Students who are experiencing *boredom* are much more likely to experience *frustration* immediately following the state of *boredom* than they are to enter into positive learning states such as *flow* [15–16]. In this way affect transition analyses reveal underlying relationships between affect and learning, which occur generally across intelligent tutoring systems.

Since existing research has suggested that affect transition analysis is both a useful and generalizable tool for investigating learner emotions, incorporating affect transition models within GIFT is a natural direction for future research and development. This will likely raise questions about how to effectively, and generally, integrate affect transition modeling capabilities with each tutor module in the GIFT architecture: the sensor module, learner module, pedagogical module, and domain module. When designing these components, one must consider how these components communicate with one another, and how the system should be configured to support cases where affect-sensitive components are missing. For example, physiological sensors are highly beneficial for affect recognition, but may not be available in all cases. Consequently, a learner model relying on output from such a sensor would need to be adapted, or gracefully deactivated, in a manner that minimizes negative impacts on other modules. Similarly, different genres of serious games have distinct capabilities and affordances. For example, serious games with believable virtual agents may present different opportunities for affective feedback than serious games without virtual agents. Pedagogical modules should possess mechanisms for handling cases where alternate learning environments support different types of interventions.

#### 3.2 Advances in Survey Administration using GIFT

In the educational research community there is a persistent need for streamlined instrument access, validation, and administration. GIFT currently provides a rich collection of content test items and questionnaires that can be re-used across studies and training environments. This survey repository could be expanded to serve the broader

research community by systematically adding validated assessment measures and questionnaires used by the education community. GIFT could serve as a searchable, centralized repository of validated instruments, with an integrated web-based infrastructure for administering surveys before and after learning interventions. The instruments could be submitted and listed by category with their important item information such as validity and reliability, links to published papers that describe how the instruments have been used in prior studies, and specific instructions regarding their appropriate use. This type of integrated resource for obtaining, evaluating, and administering questionnaires, content tests, and surveys could help streamline the community's affective computing research efforts, and serve as an entry point for researchers to begin using GIFT. Researchers commonly spend significant effort trying to locate information about study instruments, and GIFT could serve as a tool to facilitate the survey development and selection process. While other domain specific instrument collections are freely available (e.g., [www.IPIPori.org](http://www.IPIPori.org) [17]), they do not include features for integrating instruments into surveys, or administering surveys to users. GIFT could also reduce the time allocated to integrating surveys with systems such as SurveyMonkey or Qualtrics while encouraging the use of high quality validated instruments.

### 3.3 Leveraging Mathematical Models for Formative Assessment in GIFT

Building on the availability of this instrument database, there are opportunities to take advantage of item response theory techniques to conduct stealth, formative assessment during learning interactions. Item response theory (IRT) is a mathematical framework for performing measurement in which the variable is continuous in nature while allowing for an individual person and item to be mapped on the same latent trait continuum [18]. An ideal point response process is an IRT approach based on the idea that an individual only endorses an item if he or she is located close to the item on a latent continuum [19]. In other words, if an item is too extreme in either direction, the individual will respond negatively to the item. It can be used with both dichotomous (e.g., content knowledge test) and polytomous data (e.g., Likert-type attitudes or personality [19]). GIFT is well positioned to integrate ideal point methods within user experiences for stealth and ongoing assessment. To date, little research has investigated embedding adaptive, formative assessment within serious games using intermittent item presentation through ideal point methods with a rich database of instruments from which to select.

GIFT offers the opportunity for assessment of both knowledge and attitudinal (e.g., affective states) variables within immersive training experiences. Using GIFT's capabilities, single items can be "transmitted" as part of the story line within a game experience to the participant. GIFT can run mathematical models in the background to determine the best item to present at the next natural point. Conceptually this approach is similar to computerized adaptive tests designed by major test development companies. For example, if the participant responds negatively to the question "I want to repeat this activity over and over," he or she can be presented with an item lower on the latent trait continuum (e.g., "This activity is interesting for now"). GIFT, having access to all of the item information for each potential question, can strategically present a series of them. By the end of the serious game experience, rich data regard-



ing the individuals' location on a latent trait continuum (e.g., engagement) would be available.

### 3.4 Streamlined Data Processing and Analysis with GIFT

Another opportunity for GIFT to address practical challenges in affective computing research is in post-processing data. Better solutions are needed for merging files and cases and quickly ascertaining basic information from data sets. GIFT could potentially mitigate some of these challenges by introducing standards for data collected during different stages of research; typically data from different stages is encoded in a variety of formats, and a considerable amount of labor is dedicated to data integration after a study has been completed. GIFT could provide a service that automatically links pre, during, and post data for individual participants, thereby reducing labor in data cleaning and transformation steps. GIFT could also be extended to offer quick summary statistics and perform simple operations such as summarizing demographics, computing composite scores for instruments, and providing general summary results. These tools would be especially helpful with affective instruments that often require reverse scoring and other manipulations prior to analysis.

## 4 Conclusion

This paper has described a collaborative project between the Army Research Lab, Teachers College Columbia University, and North Carolina State University that aims to investigate run-time affect modeling in a serious game for combat medic training, vMedic. In addition to describing this project, we have outlined a number of ways to extend GIFT's capabilities to improve affective computing research for educational applications. We anticipate that these opportunities could increase GIFT's future impact and usage as a tool for ITS researchers.

**Acknowledgments.** The authors wish to thank the Army Research Lab for supporting this research. Additionally, we wish to thank our collaborators, Ryan Baker and Jeanine DeFalco at Teacher's College Columbia University, and Bob Sottolare, Keith Brawner, Benjamin Goldberg, and Heather Holden from ARL.

## 5 References

1. Calvo, R. A. & D'Mello, S.K.: Affect Detection: An Interdisciplinary Review of Models, Methods, and their Applications. In: IEEE Transactions on Affective Computing, pp. 18–37. (2010)
2. Conati, C., Maclaren, H.: Empirically Building and Evaluating a Probabilistic Model of User Affect. *User Modeling and User-Adapted Interaction*. 19, 267–303 (2009)
3. Marsella, S.C., Gratch, J.: EMA: A Process Model of Appraisal Dynamics. *Cognitive Systems Research*. 10, 70–90 (2009)

4. Forbes-Riley, K., Litman, D.: When Does Disengagement Correlate with Performance in Spoken Dialog Computer Tutoring? *International Journal of Artificial Intelligence in Education*. (2013)
5. Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van der Spek, E.D.: A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *Journal of Educational Psychology*. (2013)
6. Johnson, W.L.: Serious Use of a Serious Game for Language Learning. *International Journal of Artificial Intelligence in Education*. 20, 175–195 (2010)
7. Rowe, J.P., Shores, L. R., Mott, B. W., & Lester, J.C.: Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments. *International Journal of Artificial Intelligence in Education*. 21, 166–177 (2011)
8. Halpern, D., Millis, K., Graesser, A.: Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity*. 7, 93–100 (2012)
9. Kim, J., Hill, R. W., Durlach, P. J., Lane, H. C., Forbell, E., Core, M., Marsella, S.C., et al.: BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education*. 19, 289–308 (2009)
10. Rodrigo, M.M.T., Baker, R.S.J.d., Agapito, J., Nabos, J., Repalam, M.C., Reyes, S.S., San Pedro, M.C.Z.: The Effects of an Interactive Software Agent on Student Affective Dynamics while Using an Intelligent Tutoring System. In: *IEEE Transactions on Affective Computing*, pp. 224–236. (2012)
11. Pardos, Z.A., Baker, R.S.J.d., San Pedro, M.O.C.Z., Gowda, S.M., Gowda, S.M.: Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In: *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, pp. 117–124. (2013)
12. Sabourin, J., Mott, B.W., Lester, J.C.: Modeling Learner Affect with Theoretically Grounded Dynamic Bayesian Networks. In: *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction*, pp. 286–295. (2011)
13. Grafsgaard, J.F., Boyer, K.E., Lester, J.C.: Predicting Facial Indicators of Confusion with Hidden Markov Models. In: *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction*, pp. 97–106. (2011)
14. Sottolare, R. A., Goldberg, B. S., Brawner, K. W., & Holden, H.K.: A modular framework to support the authoring and assessment of adaptive computer-based tutoring systems (CBTS). In: *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, (2012)
15. D’Mello, S., Taylor, R.S., Graesser, A.C.: Monitoring Affective Trajectories during Complex Learning. In: *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, pp. 203–208. (2007)
16. Baker, R., Rodrigo, M., Xolocotzin, U.: The dynamics of affective transitions in simulation problem-solving environments. In: *Proceedings the 2nd International Conference on Affective Computing and Intelligent Interactions*, pp. 666–677. (2007)
17. Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H.C.: The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*. 40, 84–96 (2006)
18. de Ayala, R.J.: *The Theory and Practice of Item Response Theory*. Guilford Press, New York, NY (2009).
19. Stark, S., Chernyshenko, O.S., Drasgow, F., Williams, B.A.: Examining Assumptions About Item Responding in Personality Assessment: Should Ideal Point Methods Be Considered for Scale Development and Scoring? *Journal of Applied Psychology*. 91, 25–39 (2006)



## Authors

**Jonathan P. Rowe:** Jonathan Rowe is a Research Scientist in the Department of Computer Science at North Carolina State University. He received his Ph.D. in Computer Science from North Carolina State University in 2013. His research is in the areas of artificial intelligence and human-computer interaction for advanced learning technologies, with an emphasis on game-based learning environments. He is particularly interested in intelligent tutoring systems, user modeling, educational data mining, and computational models of interactive narrative. Jonathan has led development efforts on several game-based learning projects, including *Crystal Island: Lost Investigation*, which was nominated for Best Serious Game at the 2012 Unity Awards and the 2012 IITSEC Serious Games Showcase and Challenge. His research has also been recognized with several best paper awards, including best paper at the Seventh International Artificial Intelligence and Interactive Digital Entertainment Conference and best paper at the Second International Conference on Intelligent Technologies for Interactive Entertainment.

**Eleni V. Lobene:** Eleni Lobene is a Research Psychologist in the Center for Educational Informatics at North Carolina State University. She received her Ph.D. in Industrial/Organizational Psychology from North Carolina State University in 2011, where her research focused on K-12 teacher motivations and perceptions. With a background in psychometrics, including classical test theory and item response theory, Dr. Lobene is interested in study design, instrument validation, assessment, and K-16 education. Prior to joining the IntelliMedia group, she worked as a research assistant at the Friday Institute for Educational Innovation, assisting in the assessment of game-based learning environment effectiveness. She has also served as a primary instructor for undergraduate courses in the Department of Psychology at North Carolina State University and provided consulting services to local and global organizations focusing on improving organizational efficiency and implementing data-based change.

**Jennifer L. Sabourin:** Jennifer Sabourin is currently a Ph.D. student at North Carolina State University in the Department of Computer Science. She received her B.S. (2008) and M.S. (2012) in Computer Science from North Carolina State University where she graduated as Valedictorian. She is a recipient of the National Science Foundation Graduate Research Fellowship award. Since 2007 she has been engaged in research examining affective and metacognitive issues associated with intelligent learning technologies. Her research efforts have resulted in over 30 published journal articles, book chapters, and refereed conference proceedings. Her work has been recognized with a Best Student Paper Award at the International Conference on Affective Computing and Intelligent Interaction and several additional nominations. In addition to her research on intelligent technologies for education, Sabourin has played an active role in efforts to broaden participation in computing and STEM fields through informal learning activities. She developed and led a year-long middle school program for introducing young students to computing principles. This program has served over 250 middle school students in the last six years and has been successfully disseminated to other schools and districts.

**Bradford W. Mott:** Bradford Mott is a Senior Research Scientist at the Center for Educational Informatics in the College of Engineering at North Carolina State University. He received his Ph.D. in Computer Science from North Carolina State University in 2006, where his research focused on computational models of interactive narrative. His research interests include intelligent game-based learning environments, computer games, and intelligent tutoring systems. Prior to joining North Carolina State University, he worked in the game industry developing cross-platform middleware solutions for the PlayStation 3, Wii, and Xbox 360. In 2000, he co-founded and was VP of Technology at LiveWire Logic where he led development efforts on the RealDialog™ product suite, an automated natural language customer service solution leveraging corpus-based computational linguistics.

**James C. Lester:** James Lester is Distinguished Professor of Computer Science at North Carolina State University. His research focuses on transforming education with technology-rich learning environments. Utilizing AI, game technologies, and computational linguistics, he designs, develops, fields, and evaluates next-generation learning technologies for K-12 science, literacy, and computer science education. His work on personalized learning ranges from game-based learning environments and intelligent tutoring systems to affective computing, computational models of narrative, and natural language tutorial dialogue. He has served as Program Chair for the International Conference on Intelligent Tutoring Systems, the International Conference on Intelligent User Interfaces, and the International Conference on Foundations of Digital Games, and as Editor-in-Chief of the International Journal of Artificial Intelligence in Education. He has been recognized with a National Science Foundation CAREER Award and several Best Paper Awards.