Methods to Improve the Field of Intelligent Tutoring Systems using Emotion-based Agents

Michael Roth

Georgia Institute of Technology, Atlanta, GA

themichaelroth@gatech.edu

Abstract. The aim of this paper is to review select current methods used in the field of Intelligent Tutoring Systems (ITS) with respect to the use of emotion-based agents and how those systems interact with the learner to capture critical data, store the data, and effectively process the data to produce valuable feedback. From this data collected, proposed methods are presented on how to improve existing ITS systems and how to make new ITS's more effective.

Keywords: Intelligent Tutoring Systems, emotion-based agents, depth camera, data deduplication.

1 Introduction

The field of intelligent tutoring systems (ITS) and emotionally responsive agents is an ongoing field of study that involves research and understanding on the affective state of a learner with the use of learning technologies. "An Intelligent Tutoring System is computer software designed to simulate a human tutor's behavior and guidance. It can assist students studying a variety of subjects by posing questions, parsing responses, and offering customized instruction and feedback." (1). Early in my research into ITS's, I was drawn to the idea of a system that would act as a personal mentor for students and customize the learning experience for each student to effectively help the student achieve one or more specific learning goals, such as language learning or one particular subject matter within a STEM field.

The primary principle behind ITS's is to help students learn. Think of it this way, take thousands of tutors or subject matter experts and compress them into a computer system so they can help train you and answer your questions. Voilà, you now have an ITS that is capable of teaching students in one or more specific subjects. Open that system to the Internet for students to use as a learning resource and you now have the potential to impact an unlimited number of users and how they learn. That is exciting educational technology!

2 Methods

After performing initial research into the ITS field, I heard mention of the idea behind ITS's using emotional agents to increase the efficiency of those systems and my thoughts about automated, mentor-based solutions drastically changed and redirected my curiosity towards what happens when some level of artificial intelligence is integrated into an ITS and the effect it could have on the learner. At the Artificial Intelligence in Education (AI-ED) conference researchers stated "…we believe that accurately identifying a learner's cognitive-emotive state is a critical indicator that will determine how to assist the learner in appreciating an understanding of the efficiency and pleasure of the learning process." (2), which supports the idea that ITS's emotional agents and the capability of reading body-language is the next logical step. This statement helps pinpoint that the approach of using emotional agents in ITS's would prove an effective tool to better assist the learner.

This approach seems relevant even today in the classroom. I have had discussions with teachers regarding an ITS and emotional agents to help recognize the affective state of students and we came to the conclusion, from our own personal perspective, that teachers in the classroom continue to deal with the same affective state pattern recognition problems that researchers are most likely facing with ITS's. The ability to properly assess a student's current emotional state will most likely be an evolutionary process that continues to be refined and improved, just as teachers experience new student emotions and have to learn to adapt to that scenario for each student. I think that's one of the most exciting concepts behind solutions like ITS's, where there is not just one answer to the problem and we never work on it again, but instead it's a never-ending process that continues to find advancements to improve the process just like we do as humans.

2.1 Emotion-based Agents

The one unique element of ITS's that is still a challenge to replicate is the social interaction that you can only receive, up to this point, from a human. The ability to understand body language, recognize the mood of the learner, and alter the learning process based on these events. These are capabilities largely beyond the functionality of ITS's. Consider pairing the issue with ITS's and its ability to detect human emotion with measuring the effectiveness of the system, then you have an entirely new metric to account for now. We're now looking at the performance of the student and how well they're learning while incorporating their emotions into the mix to really build a rich data set that can be used to customize and tailor the training for that particular user.

Researchers at the Artificial Intelligence in Education conference stated, "We also assume that computers will, much sooner than later, be more capable of recognizing human behaviors that afford strong inferences about affective state." (2). That brings up one of the key points and issues with ITS's, in that any added emotional agents that record and process learner affective state data will be expected to return an equitable increase in effectiveness considering the increased computational resources and ability needed to accomplish this task. I tend to highly agree with the researcher's assumption about computers being capable of handling the computational increases required to accurately and quickly assess the affective state of the learner. Moore's Law is one of the first thoughts that comes to mind when considering solutions for accurately capturing a learner's affective state, such that as continual advances are made in computational capacity, in turn this should increase the overall efficiency of a system's ability to process the emotional state data of a learner. Solutions like using graphics cards to handle complex computations could be of benefit as well given their computational capabilities over traditional processors.

One of the more commonly referenced papers in the field of ITS's, "Evaluation Methodologies for Intelligent Tutoring Systems", published in 1993 has been referenced as having just as much relevance today as it did when it was published. The journal article researchers revisited this paper in 2015 with some interesting observations in the ITS field including using techniques such as data mining to handle the influx of data collected. "Many educational systems gather fine-grained click-stream data documenting every learner action. Similarly, video logs, eye-tracking data, LMS log files and MOOC platforms accumulate huge amounts of data. Attempting to make sense of very fine-grained log data poses a significant technical challenge for data mining." (3).

Gathering the resources for ITS's can easily be cost-prohibitive, especially considering the amount of time it takes to develop the systems to support these types of platforms. Some ITS's require up to 200 hours of development to every 1 hour of instructional content created (4). That's an unbelievable metric. Given these factors, developers and investors want to know that the model they're going to develop or have developed will be an effective ITS tool. One issue with ITS's are the ability to adequately measure the overall effectiveness of the system and how well the student is learning. In order to gain interest with investors and users is to demonstrate the effectiveness of the system. If it works well, then the ITS is more likely to garner attention and gain user volume.

2.2 Data Analysis

With these challenges continuing today, the question whether verbal processing is more feasible and less computationally taxing than video-based emotional agents is evident. Video processing, compared to audio processing, requires larger storage space to store the data and more CPU and memory resources to process and analyze the data. One of my concerns with only using audio-based emotional agents, though, is knowing how effective it can be over using multiple agents together at the same time, like voice audio detection and eye movement tracking. Think about a phone call you've had where you heard someone say something but you misinterpret what they said because you couldn't read their body language through the phone. It would be beneficial for the person on the phone to be able to analyze the body-language of the person on the other end of the phone and then compare that to the voice portion of the conversation to make a better decision on how to interpret what was said. I think situations like this would be prevalent if only audio-based methods were used and could benefit from other body behavior tracking information, such as eye or head movement. Solutions like cloud computing services could be used to handle the data storage and computational requirements needed to store the data collected from sensors like cameras and microphones. Cloud computing solutions offer more efficient "availability, elasticity and the adaptability of services, where the user will be able to access his programs remotely and he can have a larger capacity of storage and processing available, without the necessity of owning more expensive equipment" (5).

2.3 Existing Solutions

AttentiveLearner.

Researchers are continually looking at new solutions that can be integrated into ITS's for monitoring people's eye movement, body expressions, and other signatures like heat and heart rate. One such solution, AttentiveLearner, is a mobile learning system that makes use of the embedded camera in mobile devices, like smartphones, to track the user's heart rate and infer the learner's attention. The primary focus with AttentiveLearner is to track the attention of learner's that are enrolled in MOOC's. "In a 24-participant study, we found heart rates extracted from noisy image frames via mobile cameras can be used to predict both learners' 'mind wandering' events in MOOC sessions and their performance in follow-up quizzes. The prediction performance of AttentiveLearner (accuracy = 71.22%, kappa = 0.22) is comparable with existing research using dedicated sensors. AttentiveLearner has the potential to improve mobile learning by reducing the sensing equipment required by many state-of-the-art intelligent tutoring algorithms." (6)

Affectiva.

Affectiva was developed at MIT and provides a solution that monitors a learner's facial expressions in real-time to provide feedback regarding their current emotional affective state. Building from video data collected over the Internet of people viewing online media (such as advertisements from the Super Bowl like the Doritos, Google, and Volkswagen commercials), researchers collected data of their facial expressions to build datasets to use for analyzing their responses to the videos at certain points and correlate that to their current facial expression. There was "considerable effort required in coding" and some of the videos were hand labeled to be made available for public release. (7) Considering the video data was only recorded at a 320x240 resolution, there assumptions that can be made about the challenges that potentially prevent the system from accurately detecting the current affective state of the learner.

Collaborative Authoring for Intelligent Tutoring.

Considerations for how to author new ITS's have been proposed to assist with defining standards for their development through the Generalized Intelligent Framework for Tutoring (GIFT). "As an intelligent tutoring framework, GIFT is unique in that it is open source, domain independent, includes a sensor framework, and is designed to integrate with external training applications." (8) Using a cloud architecture is considered ideal for reasons such as "supporting concurrent access by multiple users, cloud infrastructure typically includes support for several key elements of a collaborative system such as authentication, storage, versioning, and scalability." (8) Again, the challenge with storing data is presented as a common theme with ITS's when used with emotion collecting sensors like audio and visual devices. Using frameworks like GIFT, new ITS's can be designed to build around issues with scalability and already include a sensor framework that can integrate with audio and video devices for collecting data from the learner.

3 Proposed Solutions

As discussed throughout this paper, ITS's are presented with many technical challenges when it comes to integrating sensors used for collecting information about the learner and their emotional affective state. These challenges can be overcome with the following proposed solutions to help make existing solutions more effective or integrated with newly developed ITS.

3.1 Video Sensors

Use of RGB-D (red, green, blue, and depth) sensors, like that in the Microsoft Kinect sensor, can be used to "provide high-resolution depth maps in real time at a very low cost." (9) The Microsoft Kinect is a popular add-on for the Xbox game console and used as an interaction device for the gamer to use their body to control the game or other console applications with hand gestures, body motions, or voice-controlled actions.

Another solution using the RGB-D sensor technology is Intel, which might be a more reasonable device to use and connect to modern personal computers via USB. The Intel RealSense 3D Camera "sensor consists of an IR and color camera, as well as an IR coded light projection and has a maximum depth map resolution of 640 x 480 pixels, where only 320 x 240 are used (both RGB and depth channels) for increased performance." (10) The addition of the depth camera sensor is what helps accurately capture facial expressions, which in turn can be used with ITS's for analyzing the learner's expressions and making inferences regarding their current emotional affective state.

3.2 Data Storage

One of the primary issues with using emotional agents in an ITS can be storing the data that is captured from the sensors monitoring the learner. Audio data is relatively small compared to video. Video data is much larger and typically requires significant more disk space to store. In order to keep emotional agents cost-effective, solutions like Microsoft's Data Deduplication could greatly benefit ITS projects as it stores data in "chunks" to reduce the amount of storage space required to store the same bytes of information and uses compression as well. (11) In other words, if there are many files stored on a storage device where, at the block-level, data is identical, then Data Deduplication works by removing references, or pointers, to all of the identical block-level data and keeps one as the main reference point for that "chunk" of data. Therefore, as referenced in the name of the solution "Data De-duplication," the data is de-duplicated or un-duplicated, as to remove all duplicate references.

In a more technical sense, data deduplication "works by splitting files into multiple chunks using a content-aware chunking algorithm like Rabin fingerprinting and using SHA-1 hash signatures for each chunk to determine whether two chunks contain identical data." (12)

Using Data Deduplication can be used as a form of disk space and cost savings, especially when used with other large data sets. Here are two examples of tests using multimedia files stored with Data Deduplication running on Windows Server 2012 R2. Figure 1 shows results with a smaller subset of data with up to 9% disk space savings with video files, saving 5+ GB in total disk space. Figure 2 shows results with a much larger subset of data with up to 59% when combining video and photos, saving 17.8 TB in disk space.

Fig. 1.

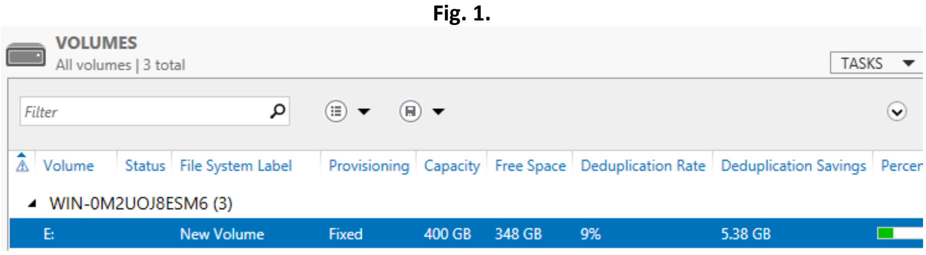
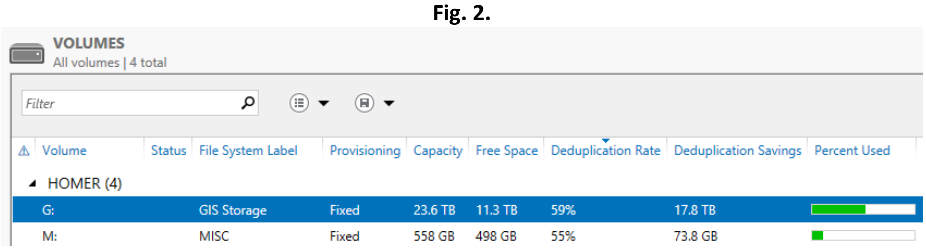


Fig. 2.



3.3 Computational Processing

Cloud computing services are an appealing option to pair with ITS's because the continual advancements that are made in the field demand newer technology to collect, store, and compute data. With cloud computing, "users can access database resources via the Internet from anywhere, for as long as they need, without worrying about any maintenance or management of actual resources. It advantages to mention but a few include scalability, resilience, flexibility, efficiency and outsourcing non-core activities." (13) One of the key issues related with the use of technology in the field of ITS is the cost of implementing your own hardware and software solutions to process the data. Systems can be under-utilized, whereas with cloud computing the costs are typically based on a per-use case. Plus the costs associated with maintenance, renewals, and refresh cycles are of no concern with using cloud-based technologies. (14)

Solutions like Amazon Web Services, Microsoft Azure and Google Cloud are options to consider with integrating existing ITS's or, especially, for newly developed ITS's so that the entire framework is designed in the cloud using the available resources and tools included with the selected cloud service. Researchers that are authoring or designing new ITS's can benefit from cloud computing services by sharing one instance of the program or software code and simultaneously making changes and modifications to improve the overall time required to develop the solution. (15)

4 Conclusion

If ITS's can play the part of providing a social interaction with the learner, then we're achieving so much more than just a student learning a new foreign language or how to solve a math problem. We're looking at designing ITS's that are capable of interacting with students on a whole new level to help give them the confidence that a human tutor would most likely do and to excel in that particular field as well. The primary goal in this research paper was to help advance the field of study in ITS's with emotion-based agents by looking at existing models as examples and find ways to either improve those models or create new methods that are more efficient and effective at collecting data with sensor technology like the Microsoft Kinect and Intel RealSense cameras, storing data using storage reducing technologies like Microsoft Data Deduplication, computing and analyzing data in the cloud to make better use of new technology to process the data and provide for scalability, all in an effort to help with determining the affective state of the learner so that they can progress forward.

References

1. **Educause Learning Initiative (ELI).** *7 Things You Should Know About Intelligent Tutoring Systems.* Louisville : Educause, 2011.

2. *External Representation of Learning Process and Domain Knowledge: Affective State as a Determinate of its Structure and Function.* **Barry Kort, Rob Reilly, Rosalind W. Picard.** San Antonio : M.I.T. Media Laboratory, 2001. Artificial Intelligence in Education.

3. *Evaluation Methods for Intelligent Tutoring Systems Revisited.* **Jim E Greer, Mary Mark.** 2015, International Journal of Artificial Intelligence in Education (IJAIED), pp. 1, 5.

4. *Cognitive Tutors: Lessons Learned.* **John R. Anderson, Albert T. Corbett, Kenneth R. Koedinger and Ray Pelletier.** 2, s.l. : Taylor & Francis, Ltd., 1995, Vol. 4.

5. *U-ALS: A Ubiquitous Learning Environment.* **Piovesan, Sandra Dutra, Passerino, Liliana Maria and Medina, Roseclea Duarte.** Madrid, Spain : International Association for Development of the Information Society, 2012.

6. *AttentiveLearner: Improving Mobile MOOC Learning via Implicit Heart Rate Tracking.* **Pham, Phuong and Wang, Jingtao.** Madrid, Spain : Cham Switzerland, 2015.

7. *Affectiva-MIT Facial Expression Dataset (AM-FED): Naturalistic and Spontaneous Facial Expressions Collected "In-the-Wild".* **McDuff, Daniel Jonathan, et al.** Portland : Institute of Electrical and Electronics Engineers (IEEE), 2013. 9780769549903.

8. *From desktop to cloud: collaborative authoring for intelligent tutoring.* **Ragusa, C.** Los Angeles : Cham Switzerland, 2015.

9. *An Improved Colorimetric Invariants and RGB-Depth-Based Codebook Model for Background Subtraction Using Kinect.* **Julian Murgia, Cyril Meurie, Yassine Ruichek.** Tuxtla Gutiérrez : Springer International Publishing, 2014. 978-3-319-13647-9.

10. *Facial Performance Sensing Head-mounted Display.* **Li, Hao and Trutoiu, Laura and Olszewski, Kyle and Wei, Lingyu and Trutna, Tristan and Hsieh, Pei-Lun and Nicholls, Aaron and Ma, Chongyang.** New York : ACM, 2015.

11. *ChunkStash: Speeding up Inline Storage Deduplication using Flash Memory.* **Li, Biplob Debnath and Sudipta Sengupta and Jin.** Boston : USENIX, 2010.

12. *FlashStore: High Throughput Persistent Key-Value Store.* **Li, Biplob Debnath and Sudipta Sengupta and Jin.** Singapore : Very Large Data Bases Endowment Inc., 2010.

13. *Cloud Computing and the Libraries.* **Nazir, Tawfeeq.** New Dehli : National Law University Delhi Press, 2015. 978-93-84272-03-6.

14. *An Overview of E-Learning in Cloud Computing.* **Fernandez, A and Peralta, D and Herrera, F and Benitez, JM.** Berlin : Springer, 2012.

15. *Unwrapping GIFT: A Primer on Authoring Tools for the Generalized Intelligent Framework for Tutoring.* **Hoffman, Michael and Ragusa, Charles.** Pittsburgh : s.n., 2015.