

COGNITIVE STATE ESTIMATION FOR ADAPTIVE LEARNING SYSTEMS USING WEARABLE PHYSIOLOGICAL SENSORS

Aniket A. Vartak¹, Cali M. Fidopiastis², Denise M. Nicholson², Wasfy B. Mikhael¹

¹*Department of Electrical Engineering and Computer Science, ²Institute of Simulation & Training
University of Central Florida, Orlando, FL, U.S.A.*

Dylan D. Schmorow

Office of Naval Research, U.S.A.

Keywords: Intelligent tutoring, psychophysiological metrics, augmented cognition, signal processing, wearable sensors.

Abstract: This paper presents a historical overview of intelligent tutoring systems and describes an adaptive instructional architecture based upon current instructional and adaptive design theories. The goal of such an endeavor is to create a training system that can dynamically change training content and presentation based upon an individual's real-time measure of cognitive state changes. An array of physiological sensors is used to estimate the cognitive state of the learner. This estimate then drives the adaptive mitigation strategy, which is used as a feed-back and changes how the learning information is presented. The underlying assumptions are that real-time monitoring of the learners cognitive state and the subsequent adaptation of the system will maintain the learner in an overall state of optimal learning. The main issues concerning this approach are constructing cognitive state estimators from a multimodal array of physiological sensors and assessing initial baseline values, as well as changes in baseline. We discuss these issues in a data processing block wise structure, where the blocks include synchronization of different data streams, feature extraction, and forming a cognitive state metric by classification/clustering of the features. Initial results show our current capabilities of combining several data streams and determining baseline values. Given that this work is in its initial staged the work points to our ongoing research and future directions.

1 INTRODUCTION

The design of metrics to determine cognitive state changes in real-time of persons performing tasks in their work environment is an emerging field of research. For example, Human Factors and Augmented Cognition research endeavors suggest the use of psychophysiological measures to determine best practices when developing trainers for military (Nicholson et al., 2006) and medical (Scerbo, 2005) occupations in an effort to optimize the learning state of the user. Further, a valid and reliable metric of cognitive state has far reaching utility in the field of intelligent tutoring, which has further implications for cognitive rehabilitation and assistive brain-computer interfaces.

This type of research is not possible without portable, unobtrusive psychophysiological sensing devices. However, utilizing physiological metrics

such as electroencephalography (EEG) is difficult due to the many factors that influence cognitive processes intra and interpersonally. Some such factors include external demands (e.g., loud noises), trait characteristics (e.g., personality), and physical states (e.g., levels of fatigue). More importantly, the neurobiology underlying constructs defining cognitive states (e.g. working memory) are not fully elucidated (Cabeza & Nyberg, 2003), thus operationally defining "cognitive state" is difficult as is identifying a theoretical approach for studying it. Thus, the most straight forward approach to developing these metrics is by establishing a convergent methodology that is multimodal in nature (Karamouzis, 2006).

In this paper, we discuss the historical aspects of developing adaptive intelligent tutoring using psychophysiological metrics. Additionally, we describe our Adaptive Instructional Architecture,

which features multimodal sensors. We discuss challenges in developing a convergent methodology for using multimodal sensors. Finally, we present initial work on data fusion techniques necessary for driving the adaptive tutoring system.

2 ADAPTIVE TUTORING SYSTEMS

In 1958, Skinner challenged educators to become more efficient and effective in their teaching strategies by using “teaching machines”. These machines would not only deliver learning content, but also allow the learner to interact with the system in a manner appropriate for learning to occur. The strength of this approach was the potential for customized instruction in an anytime anywhere format. However, teaching machines from this era neglected the knowledge base of the learner and focused more on “contingencies of reinforcement” or the presentation of learning material (Wenger, 1987).

“Intelligent Tutoring Systems” (ITS) was first coined by Sleeman & Brown (1982); however, it was Wenger (1987) who advocated for cross-talk among education, cognitive, and artificial intelligence researchers to shape the future of ITS design. This collaborative approach shifted emphasis from purely computational solutions to those that integrated Cognitive Psychology constructs (e.g., working memory) and new research in Education Psychology (e.g., experiential learning). The improved flexibility of these designs supported the successful transition of some adaptive systems into classrooms and workplaces (Anderson, et. al., 1995; Parasuraman et. al., 1992).

While previous ITS theories emphasized the knowledge state of the learner, current instructional design methods consider the learner’s cognitive state, (i.e., cognitive load state) as more predictive of learning outcomes (Paas et. al., 2004). Cognitive load theorists contend that learning complex tasks (e.g., performing surgery) is optimal when the learning environment matches the cognitive architecture of the learner (Sweller, 1999). Thus, the learning environment should account for individual differences in the unique ways that individuals cognitively process data.

Physiological metrics of cognitive load such as pupil dilation and heart rate may map a learner’s cognitive state to the learning task (Paas et al, 2003, p. 66). Another suggested use of psychometric data is to drive the adaptive response in the ITS

(Karamouzis, 2006; Scerbo, 2006). In previous work, we have proposed an Adaptive Instructional Architecture (AIA) that merges the constructs of experiential learning, cognitive load, and adaptive trainers into a testbed simulation capable of measuring multimodal psychophysical responses (Nicholson et al., 2007). In the next sections we provide an overview of the AIA and give a description of the sensors used within the training environment. In addition, we provide pilot data from current studies which use multiple sensors to determine the learner’s cognitive states. These studies are discussed in the context of data fusion strategies and point to future work in the field

3 OVERVIEW AIA, SENSORS, FEATURE EXTRACTION & DATA FUSION

3.1 Adaptive Instructional Architecture Overview

Figure 1 provides an overview of the Adaptive Instructional Architecture (AIA) within a simulator testbed. As shown the learner interacts with context based stimuli that follow the continuum from real world to simulated real world multi-sensory content. The psychophysiological sensors (e.g., heart rate) attached to the learner collect information about the learner’s cognitive state. The sensor data streams are sent through a signal processing block (Figure 3) where data fusion techniques determine such constructs as learner engagement, arousal, and workload.

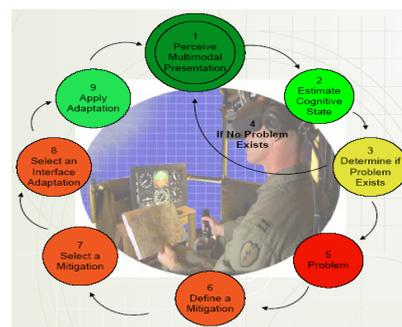


Figure 1: Adaptive Instructional Architecture Overview.

If the learner is experiencing higher than baseline values of these state references, the system chooses an appropriate mitigation strategy from a database of options. The system interface is then

adapted to adjust to the learner and the training scenario continues. This decision tree cycle is continued until the training session ends.

The novel features of the AIA are the potential to assess the cognitive state changes of the learner in real-time, change the learning scenario as the learner transitions in knowledge states, and assess performance outcomes concomitantly with the cognitive state assessment. Two main design issues faced are: 1) defining metrics derived from the multimodal data streams that reliably predict the learner's cognitive state and 2) determining the relationship of the metric and that of mitigation selection. Our current focus is on deriving meaningful metrics from the multimodal data stream. In the next sections, we introduce the psychophysical sensors and measures that we are currently exploring.

3.2 Physiological Sensors and Cognitive State Estimation

Various proposed cognitive states such as arousal, and workload are quantified in terms of physiological parameters. For example, heart rate variability (HRV) can provide a measure of arousal (Hoover & Muth, 2005). Eye position tracking may indicate visual attention and stress. The EEG can provide brain based measures of psychological constructs such as cognitive workload. Thus, a multi-modal data acquisition strategy may be necessary for accurate cognitive state estimation (Erdogmus et al., 2005; Cerutti et al., 2006). However, synchronizing and determining relevant meaning of the multiple data streams is an ongoing issue.

Figure 2 represents examples of state-of-the-art psychophysiological sensing devices within our lab. The ASL 6000 eye tracker (www.a-s-l.com) shown in Figure 2 utilizes a head tracker with pan tilt capabilities to track the corneal reflection of the user. The B-Alert EEG (www.b-alert.com) provides classifications for engagement, mental workload, distraction and drowsiness (Berka et al., 2005). The Wearable Arousal Meter (WAM, www.ufiservingscience.com) also measures arousal however does so by utilizing inter-heartbeat interval (IBI) changes associated with task performance. Changes in IBI reflect the Respiratory Sinus Arrhythmia (RSA), which correlates with autonomic nervous system states (Hoover & Muth, 2004). Also shown are the respiratory, temperature, and GSR sensors of Thought Technologies InfitiPro wireless system (www.thoughttechnologies.com). Overall,

the sensors provide a portable solution for capturing real-time neural and behavioral responses training in a naturalistic environment.



Figure 2: Sensor suite examples.

3.3 Block-wise Multimodal Signal Processing/ feature Extraction

The data generated from various sensors over time is enormous. To draw meaningful conclusions and to classify cognitive state in real-time, while also providing the feedback to the learner, the data may be effectively handled in a block processing procedure. Figure 3 provides a general overview of block processing as it applies to multimodal signal processing.

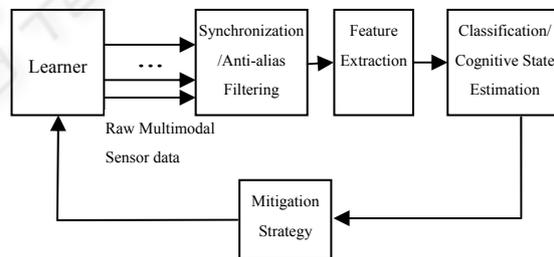


Figure 3: Multi-modal signal processing block.

The first block of the system synchronizes the data from various sensors. Multi-rate Digital Signal Processing (DSP) techniques such as decimation/interpolation are used to match the sampling frequency of various sensors. The data also needs to be time-synchronized to a unique clock-time, so that there is no error interpreting the data in further blocks.

The next block of feature extraction is a very important step in processing the data emanating from the sensor suite. The physiological measure will dictate what type of feature is to be extracted and the level to which this feature will provide

meaningfulness to the derived metric. In the following sub-section we will give an overview of typical features used from various sensors in the literature.

3.3.1 Heart Rate Features

The most popular feature used from the ECG data is the power spectral density (PSD) of the IBI. The PSD analysis provides a means to evaluate various autonomic nervous system influences on the heart efficiently. Most of the recent research focuses on quantifying the change in RSA as a measure of vagal tone activity influencing the heart (Hoover & Muth, 2005; Keenan & Grossman, 2006; Aysin & Aysin, 2006).

3.3.2 Blood Pressure Features

Blood pressure also affects heart rate modulation through the baroreceptor reflexes (Sleight & Casadei, 1995). The main challenge is to obtain a continuous measure of arterial blood pressure (ABP). The photoplethysmogram (PPG) signal is much more accessible and easily acquired in continuous manner as compared to direct measurement of the ABP signal. Recent work by Shaltis et al, (2005) discusses the calibration of the PPG signal to ABP signal.

3.3.3 Eye Tracking Features

The ASL 6000 eye tracker uses an IR camera to capture images of the eye. An image processing algorithm detects the dark pupil area in the eye and the glint of light coming off of the eye. Using these two measures, the learner's point of gaze (POG) is calculated. After proper calibration, the learner's POG can be transformed into a point on the screen correspond to where he or she is looking.

Various features could be extracted from the horizontal and vertical co-ordinate data, such as fixation intervals, speed of eye movement, and direction of eye movement. Marshall (2007) used these features as inputs to a neural network to classify cognitive states such as *relaxed/engaged*, *focused/distracted*, and *alert/fatigued*. The authors also state that as the data captured at the rate of 60-250 Hz, the states could be predicted in real time.

3.4 Data Fusion, Cognitive State Estimation

Once appropriate features psychometric data are extracted, a strategy is needed for defining the

mathematical relationship between the feature the state change. For example, Marshall (2007) used features extracted from the eyetracker (e.g., eye blinks, eye movement, pupil size, and divergence) to classify cognitive activity into 'low' and 'high' activity measures. The authors used discriminant function analysis to create a linear classification model. A feed-forward neural network architecture was trained with backpropagation learning scheme to create a non-linear classification using the eyetracker features.

We are in the process of creating multidimensional classifiers based upon feature analysis across multiple psychophysiological metrics. These classifiers will eventually index levels of cognitive state, which in turn will drive the mitigation selection process of the AIA. The pilot work presented in the next section highlights current results.

4 PRELIMINARY RESULTS

4.1 Sensor Sensitivity in Cognitive State Estimation

We are currently investigating the sensitivity of the multimodal sensors to define cognitive state changes dynamically. For example, Figure 4 shows eyetracker data merged with the instantaneous arousal level of the observer, as the observer passively views a series of varying visual stimuli. The arousal metric is calculated from the heart rate data and was obtained using the WAM (Hoover & Muth, 2005).

In Figure 4(d), the ellipse represents the current viewing location of the observer. When the observer moves his or her eyes in a vertical direction, the major axis of the ellipse appears as vertical. A diagonal movement of the eyes will produce a circle as shown in Figure 4(a) and 4(c). Fixations are illustrated in 4(c). As the observer fixates onto a point of interest, the ellipse becomes a dot. The fixation time can be presented along with the fixation point in real-time or in an after action review format.

The arousal levels are mapped to the ellipse via colors ranging from red for high, yellow for medium, and green for low. The scale used to change the color will be verified experimentally using a variation of the International Affective Picture Sort (Lang et al., 2005). These transformed features may further be used to develop

multidimensional metrics with which to predict visual attention and arousal states of the learner.

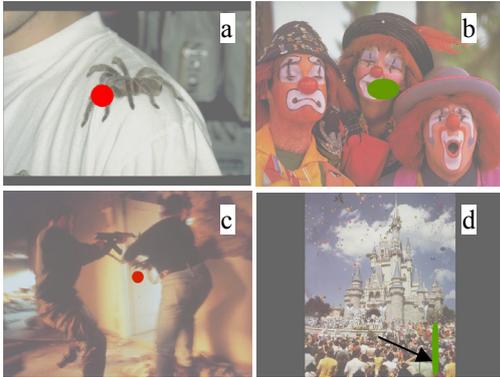


Figure 4: Four screen captures from our system, showing the observer's current gazing location along with the arousal (Images: Lang et al., 2005).

4.2 Identifying Baseline Values

Understanding how multimodal psychometric data combine to predict cognitive states is only one part of the problem in AIA design. Another issue is identifying initial baseline values that will set the system indices and determine the appropriate classification of the learner's cognitive state. Not only will these baseline values vary based upon individual difference, they may also vary during the training session.

In a recent study, we monitored the arousal state of persons placed in a mixed reality scenario representing an every day social experience. The social interaction was classified as friendly (e.g., mutual regard) or rude (e.g., confrontational). Figure 5 shows the percent high engagement as measured by the EEG and the mean skin conductance for a single participant. We used a multiple baseline approach to identify points in the scenario that may indicate a new baseline score.

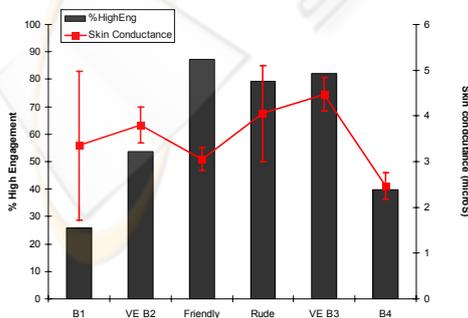


Figure 5: Skin conductance mean amplitude with 95% Confidence Interval and % High Engagement as measured by the EEG.

As shown, high engagement alone would not capture the change in state of the participant accurately. Regardless of variability, the sustained arousal carried over from experiencing the rude interaction may indicate a change in baseline that must be account for in order to appropriately select the next mitigation. Multimodal data is necessary to construct an appropriate metric to capture this type of sustained effect.

5 CONCLUSIONS

In this paper we reviewed the historical aspects of ITS design and discussed a new direction in combining current learning theory with adaptive system theory. The resulting AIA represents a step forward in providing on-demand training in a complex and contextually relevant training environment. The addition of physiological measures to estimate the cognitive state of the learner is not a novel; however, the data fusion techniques and the use of the multimodal data drive mitigation selection may present a worthwhile contribution to the field.

ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research and an In-House Grant provided by UCF's Institute for Simulation and Training.

REFERENCES

- Nicholson, D., Stanney, K., Fiore, S., Davis, L., Fidopiastis, C., Finkelstein, N., & Arnold, R. 2006, 'An adaptive system for improving and augmenting human performance', In D.D. Schmorow, L.M. Reeves, and K.M. Stanney (eds.): *Foundations of Augmented Cognition* 2nd Edition, Arlington, VA: Strategic Analysis, Inc., pp. 215-222.
- Scerbo, M. 2005 'Biocybematic systems: Information processing challenges that lie ahead', *Proceedings of the 11th International Conference on Human-Computer Interaction*.
- Cabeza, R., Nyberg, L. 2000, 'Imaging cognition II: An empirical review of 275 PET and fMRI studies', *Journal of Cognitive Neuroscience*, vol. 12, pp. 1-47.
- Karamouzis, S. 2006, 'Artificial Intelligence Applications and Innovations', *IFIP International Federation for Information Processing*, Vol. 204, Springer, Boston, pp. 417-424.

- Sleeman, Brown, J. 1982, *Intelligent tutoring systems*, New York: Academic Press.
- Nicholson, D., Fidopiastis, C., Davis, L., Schmorrow, D. & Stanney, K. 2007, 'An adaptive instructional architecture for training and education', *Proceedings of HCI International (in press)*.
- Skinner, B. 1958, *Teaching Machines*, Science 128, pp. 969-77.
- Wenger, E. 1987, *Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge*, Morgan Kaufmann Publishers, Inc., Los Altos, CA.
- Anderson, J., Corbett, A., Koedinger, Pelletier, K. 1995, 'Cognitive Tutors: Lessons Learned', *Journal of the Learning Sciences*, vol. 4, no. 2, pp. 167-207.
- Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. 1992, *Theory and design of adaptive automation in adaptive systems (Progress Report No. NAWCADWAR-92033-60)*. Warminster, PA: Naval Air Warfare Center, Aircraft Division.
- Paas, F., Renkl, A., & Sweller, J. 2004, 'Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture', *Instructional Science*, vol. 32, pp. 1-8.
- Sweller, J. 1999, *Instructional Design in Technical Areas*. Australian Council for Educational Research Press, Camberwell, Australia.
- Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P.W. M. 2003 'Cognitive load measurement as a means to advance cognitive load theory', *Educational Psychologist*, vol. 38, pp. 63-71.
- Hoover, A., Muth, E. 2005, 'A real-time index of vagal activity', *International Journal of Human-Computer Interaction*, vol. 17 no. 2, pp. 197-209.
- Erdogmus, D., Adami, A., Pavel, M., Lan, T., Mathan, S., Whitlow, S., Dorneich, M. 2005, 'Cognitive state estimation based in EEG for augmented cognition', *Proceedings of the 2nd International IEEE EMBS Conference in Neural E engineering*, Arlington, Virginia, March 16-19.
- Downs, J., Downs, T., Robinson, W., Nishimura, E., Stautzenberger, J. 2005, 'A new approach to fNIR: The optical tomographic imaging spectrometer', *Proceedings of the 1st International Conference on Augmented Cognition*, Las Vegas, NV, 22-27 July 2005.
- Berka, C., Levendowski, D., Cvetinovic, M., Davis, G., Lumicao, M., Zickovic, V., Popovic, M., Olmstead, R. 2005, 'Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset', *International Journal of Human-Computer Interaction*, vol. 17 no. 2, pp. 151-170.
- Takahashi, M., Kubo, O., Kitamura, M., Yoshikawa H. 1994, 'Neural network for human cognitive state estimation', *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems '94*.
- Cerutti, S., Bianchi, A., Reiter, H. 2006, 'Analysis of sleep and stress profiles from biomedical signal processing in wearable devices', *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, Aug 30-Sept 3.
- Crosby, M., Ikehara C. 2005, 'Using physiological measures to identify individual differences in response to task attributes', *In D.D. Schmorrow, L.M. Reeves, and K.M. Stanney (eds.): Foundations of Augmented Cognition 2nd Edition*, Arlington, VA: Strategic Analysis, Inc., pp. 162-168.
- Keenan, D., Gorssman, P. 2005, 'Adaptive filtering of heart rate signals for an improved measure of cardiac autonomic control', *International Journal of Signal Processing*, vol. 2, no. 1, pp. 52-8.
- Aysin, B., Aysin, E. 2006, 'Effect of respiration in heart rate variability (HRV) analysis', *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York City, USA, Aug 30-Sept 3.
- Sleight, P., Casadei, B. 1995, 'Relationships between Heart rate, respiration and blood pressure variabilities', *Heart Rate Variability*, Futura Publishing Company, Armonk, NY.
- Shaltis, P., Reisner, A., Asada, H. 2005, 'Calibration of the photoplethysmogram to arterial blood pressure: capabilities and limitations for continuous pressure monitoring', *Proceedings of the 27th IEEE EMBS Annual International Conference*, Shanghai, China, Sept 1-4.
- Salvucci, D., Anderson, J. 1998, 'Tracing eye movement protocols with cognitive process models', *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*, pp. 923-8.
- Marshall, S., 2007, 'Identifying cognitive state from eye metrics', *Aviation, Space and Environmental Medicine*, Vol. 78, no. 5, pp. 165-75.
- Lang, P., Bradley, M., Cuthbert B. 2005, 'International affective picture system (IAPS): Affective ratings of pictures and instructional manual', *technical report A-6*, University of Florida, Gainesville, FL.