

Towards Advanced Learner Modeling: Discussions on Quasi Real-time Adaptation with Physiological Data

Emmanuel Blanchard, Pierre Chalfoun, Claude Frasson,
*HERON Lab, Computer Science Department, Université de Montréal,
CP 6128 succ. Centre Ville, Montréal, H3C 3J7, QC, Canada.
{blanchae, chalfoup, frasson}@iro.umontreal.ca*

Abstract

In this paper, we discuss the use of physiological data for quasi real-time adaptation in ITS. We present preliminary results where we analyze learners' reactions while using a game-like virtual learning environment. We also discuss the relevance of adding cerebral data to adaptation by the means of an electroencephalogram.

1. Introduction

The aim of Intelligent Tutoring Systems (ITS) is to properly adapt learning sessions to the user (i.e. the learner). Good student modeling techniques appear to be the key in reaching such an objective. Profiling the learning capabilities of a student have been discussed for a long time now in the ITS field. Recent papers have argued that affective modeling (i.e. determining a learner's emotional or motivational state [2, 3, 6]) should also be considered. Researchers use different techniques to detect learners' affective states.

The easiest method remains the completion and analysis of data collected via questionnaires filled by students. This approach raises many problems. First, asking learners to fill in questionnaire forms interrupts the learning activity, which can disrupt concentration towards the learning content. Frequent and repetitive interruptions (which should be needed to keep a student model up to date) can also affect motivational and emotional states. In other words, the evaluation itself could create *affective noise*. A classic solution to avoid such a situation is to ask learners to answer questionnaires at the end of a learning session and adapt the next learning session accordingly. However, this raises an obvious problem of synchronicity with the affective experience. Second, self reports can become unreliable with people not always fully aware of their affective state [6]. Third, results obtained from

this method are naturally subjective: given individual differences (culture, age, gender, past experiences...), people have personal interpretations of affective vocabulary as well as personal scales to evaluate their affective reactions.

Another often used method to supply affective data to the student model consists of analyzing learners' interactions with the Graphical User Interface (GUI) through the use of sensorial captors (webcam, force feedback devices...). Despite the possibility of quasi real-time performance (i.e. interpretation can be obtained just a short period after the affective experience), problems will arise due to individual differences in affective manifestations. Indeed, body language is culturally sensitive; taciturn or hyperactive learners also represent quite a challenge...

A third method for affective data collection lies in the use of physiological signals. As for conventional sensorial captors cited before, this approach can also be performed in quasi real-time without affecting the dynamic of the learning session. Furthermore, it is resistant to cultural and personal traits differences.

The aims of this paper are two-fold. First, we wish to verify the relevance of using physiological signals for quasi real-time student modeling in ITS learning environments in order to possibly establish a link between learning-related events and signals. Second, we want to inspect the novel use of an electroencephalograph within an ITS environment in regards to its interest, benefits and drawbacks.

This paper is organized as follows: section two presents a brief description of the most meaningful physiological signals related to ITS and Human Computer Interaction (HCI) and the kind of information that can be deduced from their analysis. Section three presents previous related to physiological data. Section four depicts our experience to monitor physiological signals of learners while using a game-like learning environment. Section five presents

preliminary results we obtained. They are discussed in a sixth section that led us to propose the idea of multimodal active learner modeling.

2. Physiological signals in HCI and ITS

The literature regarding the ITS and HCI commonly denotes the use of the following physiological signals: skin temperature (ST), respiration (RESP), mainly heart rate (HR) and blood volume pressure, (BVP) for cardiovascular activity, galvanic skin response (GSR) and surface electromyography (SEMG). Few made use of an electroencephalograph (EEG).

Left-out signals. Two signals were not considered in this experiment. The BVP signal was recorded but not analyzed because of the sensor's high responsiveness to movement thus making the reading unreliable as mentioned in [4]. An electrocardiogram (EKG) is more resistant to movement and would have been preferable to use. However, none was available for this experience. SEMG was not used for a similar reason.

Recorded signals. ST, RESP, GSR and EEG signals were recorded. A description of each follows.

ST changes when the body undergoes a stress response. In an unconscious process, blood then flows away from the extremities such as hands resulting in a decrease in ST, thus colder hands [5].

RESP is accepted as being an automatic response that increases when provoked by physical exercise or by the expression of thoughts resulting in an emotional reaction, namely fear and sadness [9]. Respiration amplitude and frequency are important indicators for the presence of anxiety [12].

GSR is a measure reflecting the electrical resistance of the skin. A high GSR signal is often correlated with a high stress level [4]. GSR has a direct link with arousal [8] and is often used to reflect emotional reactions as well as cognitive activities [5].

EEG signals are a representation of the neural electrical activity present in the brain, called brain waves. Each wave frequency is measured in terms of Hertz (Hz) and microvolts (μV). Frequencies have been broken down in discreet groups in regards to the common cerebral characteristics they represent: *Delta* (1-4 Hz), *Theta* (4-8 Hz), *Alpha* (8-12 Hz), *SMR* (12-16 Hz), *Beta* (16-32 Hz), divided in four groups among which *High Beta* (20-32 Hz) and *Gamma* (38-42 Hz). In this paper, we will particularly focus on High Beta (symptomatic of intensity and anxiety) and Gamma (symptomatic of cognitive processing and learning).

EEG sensors are placed on the head according to the 10-20 international system [7]. We positioned our sensor over the *Fpz* region whose main functions include sustained attention, working memory and initiative [5].

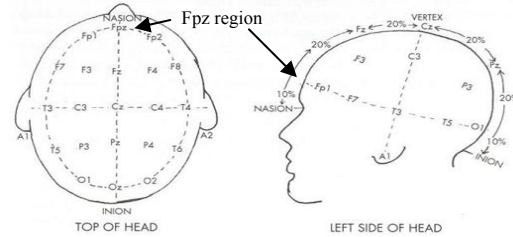


Figure 1: International 10-20 system [6]

3. Previous works on physiological data

We have identified two ways of using physiological data in HCI and ITS. Following are such examples.

On the one hand, feedback of physiological data under different representation is used and integrated to interfaces of collaborative learning environments [10] or chat systems [17] in order to inform peers or the learner of his level of arousal (supposedly related to an emotional state). Cerebral activity has also been used to control a graphical interface of a videogame [14].

On the other, the identification of specific patterns in physiological signals, mainly inspired by Picard [15], is employed for the future goal of proposing adaptive computer environments. Ward and Marsden [19] studied physiological responses when using well or ill designed websites. Amongst other things, they stated that distinguishing the cause of a physiological evolution (affective event, cognitive workload...) wasn't a trivial task. Other researches have used physiological data in the context of emotional detection [3, 12] with an objective of real-time adaptation. Bosma and André [2] have merged two techniques in order to resolve ambiguities in dialogue acts. They used a Bayesian network approach to determine an emotional state and a finite state machine method to study the meaning of dialogue acts according to physiological data.

All those studies did post-hoc analysis of data. Thus, the "*predictive model*" approach of adaptation (adapting according to a model obtained by previously analyzed training data) appears to be the norm.

4. Description of our study

We monitored physiological signals of 18 healthy volunteers while using our system, MOCAS [1]. The sample's mean age was 27 (SD =3.88) and all participants held at least a high school degree.

MOCAS is a game-like learning environment. Learning takes place in a virtual 3D world where the avatar of the learner can interact with avatars of pedagogical agents displaying learning contents (Figure 2).

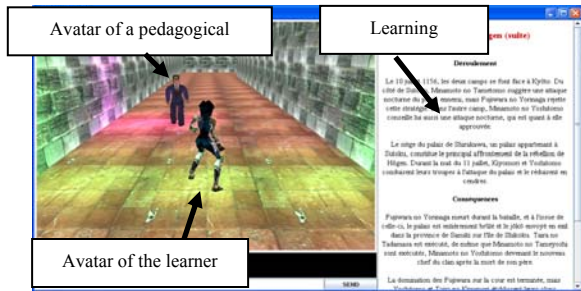


Figure 2: client interface of MOCAS

Physiological sensors are managed by the ProComp Infinity encoder [18]. It has been used by many of the previous works we have mentioned.

Literature warned us about taking great care of the experimental protocol in order to avoid artifacts. Thus:

- the RESP sensor extensible belt was put at diaphragm level,
- the EEG sensor was positioned at Fpz. The resistance between the electrodes was maintained under 5 k Ω while always remaining under 10 k Ω for a clear and useable signal,
- all other sensors were put on the left hand (GSR on 1st and 3rd finger, ST on the 3rd, BVP on the 4th). People were asked to avoid moving as much as possible because sensors (especially BVP) were movement-sensitive,
- two webcams were also used. One monitored the learner's facial activity while the other recorded the learner's interactions on the computer screen.

For each learner, all signals (webcam included) were collected in real-time within a single session using the Biograph Infiniti Software [18]. The organization of the experience is as follows.

A training exercise aiming at getting accustomed to the sensor's presence preceded all monitored activities.

Four different MOCAS 3D environments were created. In each environment, people had to find their way inside a maze. At the beginning and in the middle of the maze, they had to read two different texts but time of reading for the second one was limited to 30 seconds (to induce pressure). After that, learners periodically met avatars asking questions regarding the last text read. Learners responded to the question by following a given direction. A wrong answer led them to a dead-end whereas a good one allowed them to continue in order to reach the next question or text.

There were rests periods before and after each of the four MOCAS environments in order to make people start the activity in the lesser aroused state possible. An oral debriefing period followed the end of each maze.

The experiment took place in a university office. Doors were closed and one researcher was present with the learner to place labels in the recording software when predetermined events occurred.

5. Results

According to previous works, there were clear evidences that many interface or learning activity events frequently triggered physiological reactions. For instance, text reading periods were frequently preceded by a big breath. GSR frequently evolved when the learner met a pedagogical agent avatar. GSR was also very reactive during oral debriefing periods. The only signal that didn't seem short-term reactive to events was ST.

As expected, individual differences in physiological manifestations were noted. Figures 3 and 4 present the GSR and ST for two learners according to the time of the experience in seconds (which could differ for each learner and generally lasted about 50 minutes). Each signal was thoroughly double-checked. These learners were chosen because they were particularly representative of individual variations without being considered exceptions.

In figure 3, the highest change ratio of learner 2 is ~133% whereas it is ~10% for learner 5. It confirms that some people are far more reactive than others [19].

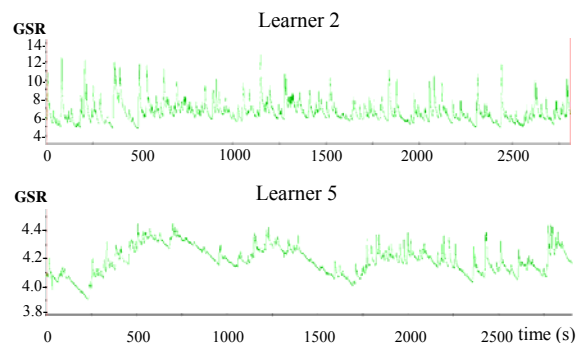


Figure 3: GSR of learners 2 and 5

In figure 4, when comparing mean temperatures, learner 2 is around 33.5 degrees whereas learner 7 is around 24.2 degrees (overall mean ST was between 28 and 32 degrees). The signals also confirm that variations in ST can be significant when observed over an entire session rather than in specific time frames.

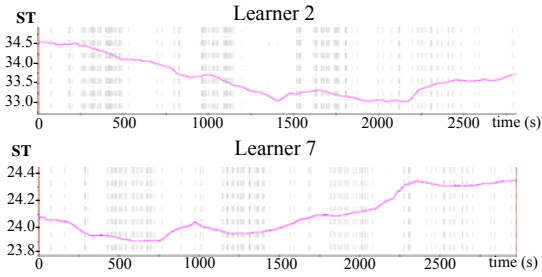


Figure 4: ST of learners 2 and 7

Relevance of the EEG signal. The EEG signal gives precious clues regarding mental activity in a session. These clues inform of a significant mental activity and can be used to help identify the cause of a physiological change.

Figure 5 shows a 30 second time frame during which a learner is reading a text with no time limit imposed.

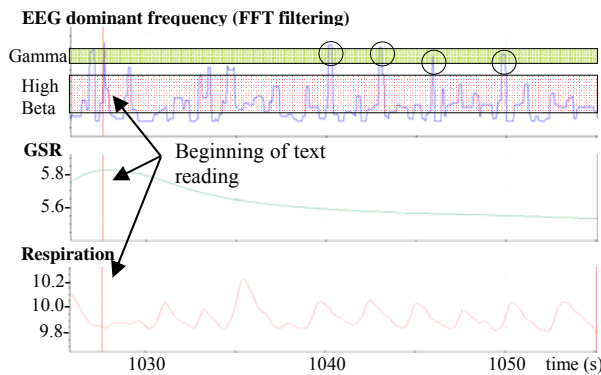


Figure 5: reading text under no pressure

We chose to display a graph of the dominant cerebral frequency (obtained using Fast Fourier Transform methodology: FFT). When reading such graph, people have to concentrate on the value of the dominant frequency at a given time instead of the evolution of the EEG signal. We can clearly see Gamma dominance (circled), indicating that learning is taking place. Such information could not have otherwise been detected.

Figure 6 shows an interesting case where EEG signals provide a crucial clue in opposition to more classical physiological signals. The same learner is now instructed to read a text as quickly as possible within 30 seconds. GSR and RESP show an increase before the beginning of the text but tend to decrease and stabilize as time progresses. This alone could give us the false impression that the learner's arousal levels were high at first but are following a decreasing curve. However, EEG signals demonstrate that High Beta dominance (circled) is clearly more frequent than in figure 5, indicating hyper alertness and anxiety [5].

In the two figures, GSR and RESP provide misleading interpretation rectified by the EEG.

Furthermore, EEG signals have also been used to confirm, and sometimes infirm, the statements given by the learners in the debriefing period following each activity. For example, when asked, the learner said that he had a high stress level when reading the timed text (scored 4 on a 4 points scale). The EEG on figure 6 confirms the high stress/anxiety levels present during this period.

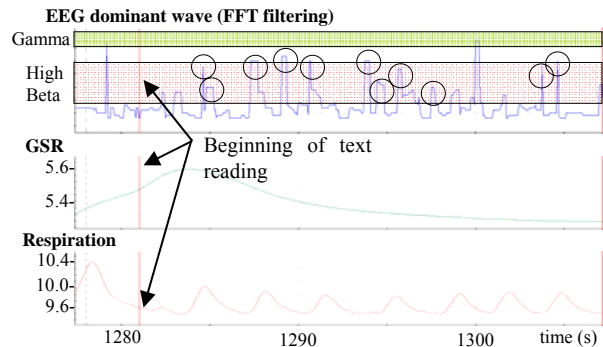


Figure 6: reading text under pressure

6. Discussion of the results

As we said, according to previous findings, we noticed clear evidences of many correlations between activity-related events and evolution of physiological signals (EEG, GSR, RESP). ST appears to be significant for long time observation, making it inappropriate for quasi real-time adaptation.

EEG is of particular interest because it greatly helps in determining mental states and investigating the causation of a physiological change. It also allowed us to confirm that a learner was in a given mental state (for instance concentrated) when we expected him to be. However, the use of EEG sensors is more complicated and intrusive than most of the other sensors, which currently restricts its use to research exploration or industrial usage (military simulators could be a good example for instance).

Our results also lead us to two important points for quasi real-time adaptation.

Necessity of multimodality. In order to be designed as relevant for a learning activity, a particular pattern detected in a given physiological channel *must* be corroborated with other sources in order to confirm or explicit its meaning. Complementary data can come from other physiological signals (GSR decrease in figures 5 and 6 would have been interpreted differently without EEG signals) or from classic channels (the webcam sometimes showed us that a detected Gamma in the EEG signal occurred during a “moving avatar”

period and we thus had great doubts that learning was taking place at such a moment). The idea of multimodality is also evoked in other studies [2, 3, 12].

Doubts about using a predictive model approach. Our results also put forward very high individual differences between the learners. Literature warns us that differences in reactivity can also occur within the same individual [19] due to non detectable elements (such as the level of physical activity before the monitored period). This raises doubts about the relevance of using a predictive model approach for adaptation. Indeed, with such a level of inter and intra individual variability, what could be the significance of deductions obtained from data collected at different times, on different learners, in different conditions when the physiological reference frame is different ?

The solution for such a problem may come from a different learner modeling paradigm called *active learner modeling* [11]. Hence, instead of seeing it as a frequently updated data structure, the learner model should be computed, when there's a necessity for it, using different channels as entry parameters.

Figure 7 shows a process of multimodal active learner modeling. However, further work is necessary to develop student modeling functions that would assess multiple entry channels as our human brain does (giving more or less weight to each channel according to the whole data collected at a given moment).

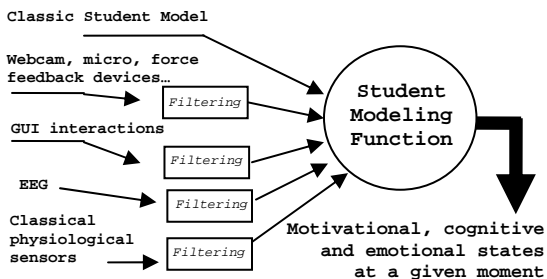


Figure 7: Multimodal active learner model

7. Conclusion

In this paper, we have discussed the use of physiological signals for quasi real-time adaptation in HCI and ITS. The fact that such data can be obtained during the runtime period without interfering with the dynamic of a learning activity makes it very interesting to have. Furthermore, personal traits, culture... don't affect its significance. EEG signals appear to be particularly interesting because they can help distinguish between reactions to affective or cognitive events, amongst other things.

However, in the specific problem of quasi real-time adaptation, oversimplifying the usage of physiological

data may result in great inaccuracy. As many before us, we argue, but also demonstrate, the necessity of multimodality. Signals are also subject to high inter individual variations (and also intra individual supposedly), which makes us raise doubts for the use of a "predictive model approach" for real-time adaptation.

We consider that multimodal active learner modeling is a promising alternative. Further researches have to be done in order to make it possible.

8. References

- [1] Blanchard, E., C. Frasson. Easy Creation of Game-like Learning Environments. Workshop on teaching with robots and agents. In conjunction with ITS2006, Jhongli, Taiwan, 2006.
- [2] Bosma W., E. André: Exploiting emotions to disambiguate dialogue acts. *Intelligent User Interfaces*, 2004, pp. 85-92
- [3] Conati, C., R. Chabbal, H. Maclaren. A Study on using biometric sensors for monitoring user emotions in educational games. *Workshop on Assessing and Adapting to User Attitudes and Affect: Why, When and How?* In conjunction with UM-03, 2003.
- [4] Criswell, E. *Biofeedback and somatics*. Novato, CA : Freeperson Press, 1995.
- [5] Demos, J.N. *Getting started with neurofeedback*. Norton & Compagny, New York, 2005.
- [6] De Vicente, A. *Towards Tutoring Systems that Detect Students' Motivation: an Investigation*. Doctoral Thesis. University of Edimburgh, 2003.
- [7] Jasper, H.H. "The 10-20 Electrode System of the International Federation." *Electroencephalography Clinical Neurophysiology*: 10, 1958, pp. 371-375.
- [8] Lang, P.J. The emotion probe. *American Psychologist*, 50(5), 1995, pp. 372-385.
- [9] Ley, R. *An Introduction to the Psychophysiology of Breathing. Biofeedback and Self-Regulation*, 19(2), 1994.
- [10] Lyons, M.J., D. Klünder, N. Tetsutani. *Enhancing Web-Based Learning by Sharing Affective Experience*, ICALT'04, Joensuu, Finland, 2004.
- [11] McCalla, G., J. Vassileva, J. Greer, S. Bull. *Active Learner Modelling*. *Intelligent Tutoring Systems*, Montreal, Canada, 2000, pp. 53-62.
- [12] Mandryk, R.L., K.M. Inkpen, T.W. Calvert. *Using Psychophysiological Techniques to Measure User Experience with Entertainment Technologies*. *Behaviour and Information Technology (Special Issue on User Experience)*, 25(2), 2006, pp. 141-158.
- [13] Marieb, E.N. *Human anatomy and physiology* (3rd ed.). Redwood City: CA Benjamin/Cummings, 1995.
- [14] Palke, A. *Brainathlon: Enhancing Brainwave Control Through Brain-Controlled Game Play*. Master Thesis. Mills College, 2004.
- [15] Picard, R. *Affective Computing*. The MIT Press, Cambridge, MA, 1997.
- [16] Schwartz, M.S. *Breathing therapies*. In M.S. Schwartz (Ed), *Biofeedback: A practitioner's guide* (2nd ed), New-York: Guilford Press, 1995, pp. 525-559
- [17] Wang, H., H. Prendiger, T. Igarashi. *Communicating emotions in online chat using physiological sensors and animated text*. *Human Factors in Computer Systems (CHI)*, Vienna, Austria, 2004.
- [18] Thought Technology Ltd: <http://www.toughttechnology.com>
- [19] Ward, R.D., P.H. Marsden. *Physiological responses to different WEB page designs*. *International Journal of Human-Computer Studies*, 59, 2003, pp. 199-212.