Muñoz, K., Noguez, J., Neri, L., Kevitt, P. M., & Lunney, T. (2016). A Computational Model of Learners Achievement Emotions Using Control-Value Theory. *Educational Technology & Society, 19* (2), 42–56.

A Computational Model of Learners Achievement Emotions Using Control-Value Theory

Karla Muñoz[1*], Julieta Noguez[2], Luis Neri[2], Paul Mc Kevitt[1] and Tom Lunney[1]
[1]Ulster University, Magee campus, North Ireland // [2]Tecnologico de Monterrey, Mexico City campus, Mexico // munoz_esquivel-k@email.ulster.ac.uk // jnoguez@itesm.mx // neri@itesm.mx // p.mckevitt@ulster.ac.uk // tf.lunney@ulster.ac.uk
[*]Corresponding author

ABSTRACT

Game-based Learning (GBL) environments make instruction flexible and interactive. Positive experiences depend on personalization. Student modelling has focused on affect. Three methods are used: (1) recognizing the physiological effects of emotion, (2) reasoning about emotion from its origin and (3) an approach combining 1 and 2. These have proven successful only in labs, or use theories of emotion not associated with an educational setting. The Control-value theory of *achievement emotions* holds that appraisals of control and value are most meaningful when determining emotion. This paper focuses on the design and evaluation of an emotional student model of Control-value theory applied to online GBL environments using Approach 2. This model is implemented using a dynamic sequence of Bayesian Networks (BNs). *PlayPhysics* - an emotional GBL environment for teaching Physics - was designed, implemented and evaluated with 118 students at ITESM-CCM. To evaluate our model, we employed cross-validation and Cohen's Kappa. Our model achieved a fair to moderate accuracy of classification, but the results are promising. Future work will focus on identifying other variables that can improve classification.

Keywords

Affective student model, Game-based learning, Control-value theory, Dynamic sequence of Bayesian Networks

Introduction

Formal instruction is transforming into a more flexible and interactive process, focusing on student preferences for learning and engagement (Moore, Dickson-Deane, & Galyen, 2011). As a result, Virtual Learning (VL) and Game-Based Learning (GBL) environments have gained popularity and acceptance. GBL environments typically comprise features, such as storytelling, sound effects and feedback, which facilitate an emotional connection with the learner (Sykes, 2013). The key to attaining positive and successful experiences in VL and GBL environments is to achieve personalization (Janssen, van den Broek, & Westerink, 2011).

Emotion has shown to be important in many contexts including Evolution and Neuroscience. Here we focus on Educational Psychology and Computing contexts, where student modeling has recently focused on affect, because it has been shown to influence student understanding, performance and motivation. However, to date, the methods employed to reason about emotion in ITSs have shown highly promising in laboratories, but not in the classroom (Arroyo et al., 2009). For reasoning about emotion, the majority of the models use theories that have not originated from an educational setting (Conati & Maclaren, 2009; Jaques, Vicari, Pesty, & Martin, 2011; Landowska, 2013). Therefore, it may be possible that the targeted emotions do not actually occur during the teaching/learning experience. Also, the classification accuracy of these models is presented mainly using only percentages, so it is unclear that the effects are not random. On the other hand, the Control-value theory of achievement emotions by Pekrun, Frenzel, Goetz and Perry (2007), assumes that control and value appraisals are the most essential to determine emotion in an educational context. Achievement emotions are derived from performing activities and the pursuit of goals. Performance and achievement are judged against previously defined standards of quality. It was observed that this theory has not previously been utilized to create a computational model of student emotion. Therefore, here we focus on this objective.

For reasoning about student emotion, we employ a Cognitive-based Affective User Modeling approach, which allows applying what is known, in this case in the psychological educational field, to predict emotion (Martinho, Machado, & Paiva, 2000). Our model predominantly uses answers during in game dialogues and contextual variables, e.g., mouse location and the number of times help is asked, because it is mainly targeted at on-line GBL environments. Our model is implemented using a dynamic sequence of Bayesian Networks (BNs), since they can effectively manage the uncertainty of the domain and appropriately represent the temporal interdependencies. A preliminary version of our model was discussed in Muñoz, Mc Kevitt, Lunney, Noguez and Neri (2013), the model presented here is the final result of performing further tests and employing more formal tools to conduct our analysis.

Related work

GBL environments, i.e., Edutainment, enhance learning by providing immediate feedback to student actions in simulated contexts. They have proven to attain student attention and engagement more easily than VL environments (Muñoz et al., 2009). Their success depends on the composition of diverse elements that gives them an emotional character (Sykes, 2013), such as penalizing errors and rewarding learning, e.g., through sounds, colors, narrative, and scoring. These elements combine to create a unique game-experience known as gameplay. Lazzaro (2004) argues that this emotional experience is the source of the appeal of playing games. GBL environments can also be combined with ITSs to achieve adaptable and personalized instruction (Conati & Maclaren, 2009).

The new generation of ITSs aims to recognize and respond appropriately to student affect (Alexander, Sarrafzadeh & Hill, 2008; Conati & Maclaren, 2009; D'Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008; D'Mello, Olneyc, Williams, & Hays, 2012; Jaques et al., 2011; Landowska, 2013; Porayska-Pomsta, Mavrikis, & Pain, 2008; Sabourin, Mott, & Lester, 2011). The main motivation for modeling emotions and moods arises from the field of Affective Computing (Picard, 1995). It has been noted that GUIs that do not consider student affect may impede and limit performance (Brave & Nass, 2008).Three main approaches are employed by the new generation of ITSs for recognizing or reasoning about student affect: (1) identifying physical and physiological effects of emotion, (2) reasoning about observable behavior from its origin, i.e., Cognitive-Based Affective User Modeling and (3) a hybrid approach combining both. Identifying the physical and physiological effects involves acquiring data related to student behavior using hardware, e.g., cameras, sensors and microphones. This data is processed and relevant features are selected and mapped to emotional states using opinions of judges or self-reports (D'Mello et al., 2008; Landowska, 2013; Sarrafzadeh, Alexander, Dadgostar, Fan , & Bigdeli, 2008). Processing this kind of data requires high bandwidth, which may deteriorate performance. The facial coding system by Ekman and Friesen (1978) is often used as a reference to map facial gestures to emotional states.

Reasoning about emotion from its origin, i.e., Cognitive-Based Affective User Modeling (Martinho et al., 2000), involves using cognitive psychology theories as a reference to reason about the elements that determine emotion. The most common theory employed using this approach is the OCC model (Ortony, Clore, & Collins, 1990). This the theory has been adapted to be applied to the learning experience, because it was originally created to explain emotion in personal diaries. Therefore, it is possible that some of the emotions do not happen or do not occur in the described manner in educational settings. This approach can employ contextual variables related to student behavior, which are considered low bandwidth variables; as a result, it can be applied to diagnosis of emotion during on-line learning. This approach has shown promising (Jaques & Vicari, 2007; Sabourin et al., 2011), but has not been as successful as the previous approach. A hybrid approach that combines both approaches is expected to be more successful than its constituent parts. However, it also inherits the weakness of its composite approaches (Conati & Maclaren, 2009). The hybrid approach involves acquiring data related to the student interaction (context), physical changes and physiological signals, and then it uses all this information in conjunction to a cognitive theory to determine student emotion.

Contextual variables have been successfully employed to determine student motivation (Del Soldato, 1993), self-efficacy (McQuiggan, Mott, & Lester, 2008) and goals and attitudes (Arroyo & Woolf, 2005). After reviewing Control-value theory (Pekrun et al., 2007), we noticed that these variables are also related to determining student emotion. Del Soldato (1993) uses variables, such as the number times the student asked for help, performance and the number of times the student quit, to determine student motivation. McQuiggan et al. (2008) employs intentional (e.g., number of problems solved), locational (e.g., current learning goal), physiological (e.g., heart rate) and temporal (e.g., time in current location) variables to classify student self-efficacy in the GBL environment Crystal Island. Arroyo and Woolf (2005) employed variables like the time invested per problem and the average number of hints given per problem to effectively determine student goals and attitudes. Also, it is observed that student gaze can be used to infer student attention (D'Mello et al., 2008; D'Mello et al., 2012). Rust (2010) show that the mouse position is also an alternative to infer a person's concentration or attention.

Theoretical framework

Dynamic sequence of Bayesian networks for Affect modeling

For defining the structure of Bayesian models, it is necessary to know the conditional independent or dependent relations (CIDRs), which can be defined with assistance of a domain expert or obtained through statistical tests of historical domain data using a learning algorithm such as Peter-Clark (PC) or Necessary Path Condition (NPC). The chosen algorithm depends on the available amount of data to derive, train and evaluate the model. For defining the parameters of the Bayesian model, a learning algorithm such as Expectation Maximization (EM) can be applied to discrete chance nodes from observed data (Jensen & Nielsen, 2007). The selection of the evaluation method also depends on the quantity of data available. With a large quantity of data available for training and testing, a hold-out procedure (Bouckaert et al., 2012) can be employed. However, when data is scarce, a cross-validation approach is employed, i.e., the dataset is divided into ***n*** sub-samples and one of these is held for testing the model and the ***n-1*** sub-samples are employed for training. Probabilistic Relational Models (PRMs) can be used also to facilitate the derivation of Bayesian models. They are object representations of the domain (Sucar & Noguez, 2008). As a result, the domain is characterized as series of entities with properties and relationships between them (Koller, 1999).

Pekrun theory

Pekrun et al. (2007) proposes the Control-value theory to explain how emotion arises in educational settings. Control-value theory focuses on *achievement emotions*, which arise from activities and outcomes that are judged against standards of quality. This theory focuses on understanding when students feel in and out of control of relevant activities and outcomes. Control and value appraisals are the key cognitive elements employed to define achievement emotions. Control refers to student beliefs about their abilities, e.g., skills and strategies, to perform an activity and attain its goal. Value relates to the assigned value of the activity or the outcome from the student perspective, which can be focused on achieving success or avoiding failure, where success and failure have positive and negative connotations respectively.

Pekrun et al. (2007) argues that if one of the appraisals is lacking, there is no emotion. There are three kinds of achievement emotions: *prospective-outcome*, *activity* and *retrospective-outcome emotions*. Two dimensions are considered to define the type of emotion that a person is feeling: the object in focus (activity/outcome) and the time frame (during an activity or before/after an outcome). Table 1 shows the definition corresponding to activity emotions in terms of control and value appraisals. It was observed that this theory has not previously been used to create a computational model of student emotion.

Table 1. Activity outcome emotions

Object Focus	Value	Control	Emotion
Activity (during)	Positive/Negative	High	Enjoyment
	Positive/Negative	High	Anger
	None	Low	Frustration
		High/Low	Boredom

Experimental design

Goal

In this work, the proposed computational model of student achievement emotions considers the Control-value theory (Pekrun et al., 2007) as a reference.

Hypothesis

The hypothesis of this work is that an emotional student model, based in Control-value theory and using answers to questions in game-dialogues and contextual variables, will reason about student emotion non-randomly and accurately.

We decided to focus on diagnosing student emotion in on-line GBL environments, since we would have access to a larger student population. For reasoning about student emotion, we employ a Cognitive-Based Affective User Modeling approach, since it employs low bandwidth variables.

Recognition variables employed for reasoning about emotion

To select the recognition variables, we examined the Achievement Emotions Questionnaire (AEQ) by Pekrun, Goetz and Perry (2005) corresponding to emotions that arise before/during/after a lecture, which comprises motivational, cognitive, affective and physiological factors. After identifying the factors employed by Pekrun et al. (2005), which we summarized in Table 2, we decided to focus on the cognitive and motivational factors while diagnosing emotion in on-line GBL environments, because these can be inferred from the interaction and the context of the learning activity. The affective factors signaled by Control-value theory are considered as student self-report of emotion during game interaction.

Table 2. Summary of cognitive and motivational factors

Before	During	After
• Attitude towards subject/activity • Confidence beliefs towards probable outcome (self-efficacy) • Attitude towards investing effort • Prospective level of difficulty (subject/activity) • Internal/ external motivation to perform & achieve an activity	• Current attitude towards subject/activity • Current level of confidence • Current effort invested • Perceived level of difficulty (subject/activity) • Student Level of concentration • Status of progress on fulfilling the activity goals • Avoiding requesting or asking for help	• Past outcome/outcomes • Willingness to keep performing/mastering the activity (investing effort) • Eagerness to make the outcome public • Resultant attitudes towards subject/activity • Internal/external attribution of the obtained outcome • Resultant confidence on own capacity/skills

For the prospective outcome emotions, corresponding to the time frame before performing the learning activity, we observed that student attitudes and beliefs related to the future performance, i.e., outcomes are key to determining these emotions. Therefore, we decided to enquire about this by using game-dialogues, while introducing the task and the game story. For the activity emotions, regarding the time frame while the student is interacting with the GBL environment, we decided to employ contextual variables that have proven to be significantly related to variables such as confidence, effort and self-efficacy in related work to diagnose student motivation (Del Soldato, 1993), self-efficacy (McQuiggan et al., 2008) and goals and attitudes (Arroyo & Woolf, 2005). We also decided to use as a basis the classification of variables, e.g., temporal, intentional, locational and physiological, proposed by McQuiggan et al. (2008) to define our contextual variables. However, our locational variables correspond to where student attention resides.

To diagnose retrospective-outcome emotions, we use the latest state of the variables presented in Table 3, in specific the latest outcome, independence (attribution of the final result), and the type of outcome (the willingness to keep interacting). We also include a new variable, publishing outcome, that is a variable related to the student intention to make the outcome public.

Table 3. Contextual variables for recognizing activity emotions

Type of variable	Variable	Description	Associated factors to control or value				
			Effort	Confidence	Perceived level of difficulty	Attitude towards the activity	Concentration
Temporal	Interval of Interaction	The total time that the student has interacted, since the game challenge is started	✓				
	Time to achieve learning goal(s)	The time that the student invested in achieving the learning goal the first time	✓	✓	✓		
Intentional	Outcome	The result that is most likely to be achieved and directly associated to student performance		✓	✓	✓	
	Times asked help	The number of times that the student asked for help	✓	✓			
	Attempts alone	The number of attempts by the student to solve the challenge alone (without help)	✓	✓			
	Estimated independence	Results from the difference between the number of attempts alone and the number of times that the student asked for help	✓	✓			
	Overall attempts	The total number of student attempts with and without help	✓	✓		✓	
	Average quality of tutoring feedback	The average value calculated from the student qualitative evaluation related to how useful, he/she finds the help or instruction provided	✓		✓	✓	
	Type of outcome	Indicates whether the student obtained a successful outcome, committed a misconception or quit the game challenge		✓	✓	✓	
Locational	Focus coarse value	The average value of the mouse position on the screen associated to student location	✓			✓	✓

Proposal for representing students achievement emotions

Through examining control-value theory (Pekrun et al., 2007), it was observed that there is mutual causation between antecedents and effects of achievement emotions over time. Also, control and value are defined as categorical variables in the Control-value theory. BNs and Binary and Multinomial Logistic Regression (BLR/MLR) can handle categorical variables appropriately. As a result, we decided to implement a dynamic sequence of BNs to represent student achievement emotions.

To define the dynamic sequence of BNs, we employed the methodology shown in Figure 1. We focused mainly on defining the BNs structure and learning their parameters. Once the design of game challenges is known and their elements are described in a class diagram, we employed Probabilistic Relational Models (PRMs) to define a preliminary and generic structure of our emotional student model.

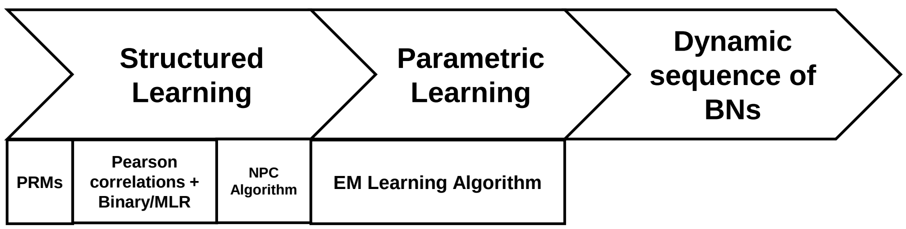

Figure 1. Methodology to define dynamic sequence of BNs

Figure 2 shows the generic PRM derived and corresponding to control-value theory. However, in this structure relationships between causes and effects are not completely defined, see the activity emotions PRM in Figure 3, where all the relations are indicated as dotted lines, meaning they are uncertain.

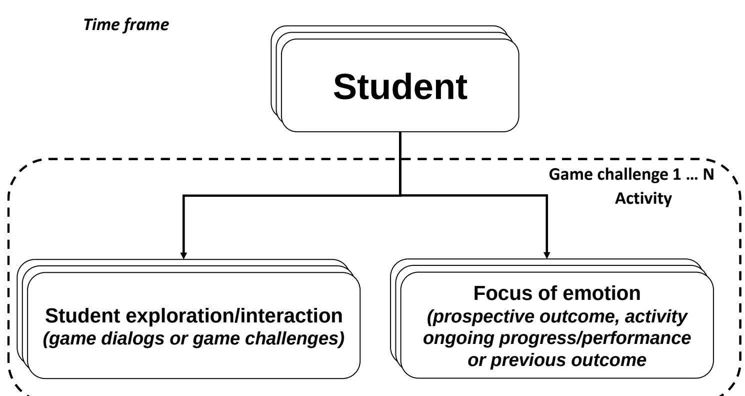

Figure 2. PRM of Control-value theory

Therefore, we will have to acquire data corresponding to the student interaction with the GBL environment, which we can analyze by employing Pearson correlations and BLR/MLR (gaining more insight about the variables that enhance the classification). Then, all the information acquired through Pearson correlations and BLR/MLR is used to solve uncertain relations and complete the definition of BN structures applying the necessary Path Condition (NPC) algorithm. Finally to define the BN parameters, we apply the Expectation Maximization (EM) learning algorithm using the collected data.

Figure 4 illustrates the concept of time frame ***t-1*** in the context of our emotional student model. The interaction with the GBL environment may be visualized as a film tape, but its execution is not necessarily in sequence, since students' actions define the order in which the elements of the GBL environment are accessed. Concurrently, each achievement emotions network serves a purpose in time. For example, the prospective-outcome emotions network is employed for reasoning about emotion in *PlayPhysics'* game dialogues. The activity emotions network is employed for reasoning about emotions while students interact with the challenges in the GBL environment. Finally, the retrospective-outcome emotions network is used to reason about emotion in the instant of time that the outcome of the game challenge is presented to the student.

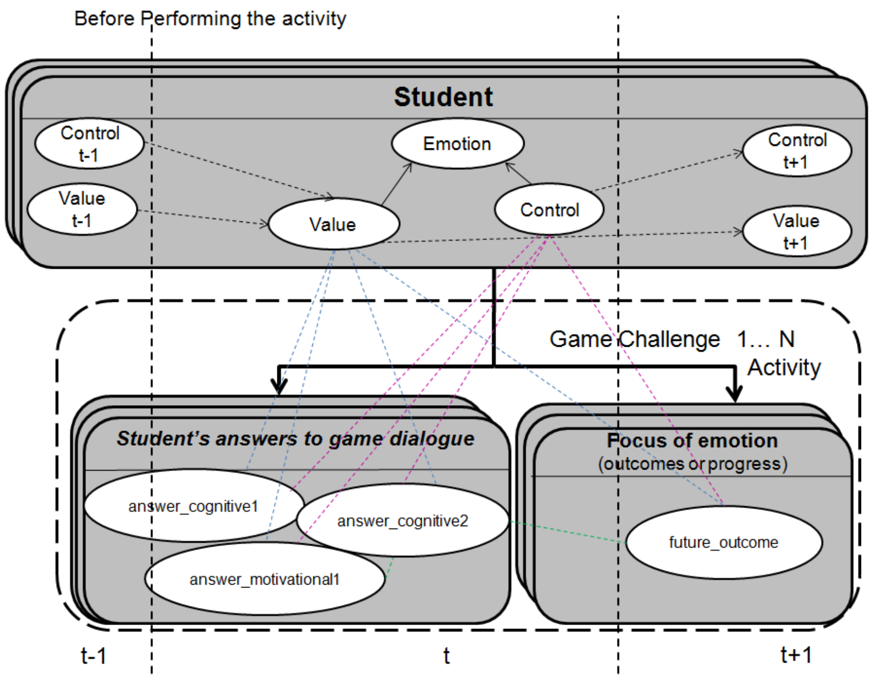

Figure 3. Activity emotions PRM

At the beginning, the student is presented with a game dialogue that introduces the game's plot, the next game challenge and enquires about student beliefs/attitudes. At the end of each game dialogue students self-report their emotions and the prospective outcome emotions network can be employed at this moment to reason about emotion (see Figure 4(1)). Then students may proceed to interact with the game challenge, if students self report their emotion before completing the challenge or evaluate the feedback provided by the learning companion, the activity emotions network can be used at that instant of time for reasoning about emotion (Figure 4(3)). If the latest entry corresponds to the ongoing interaction with a game challenge immediately after the game dialogue, value *t*-1 and control *t*-1, the state of the contextual variables is evaluated using the activity-outcome emotions network (Figure 4(2)). It is also possible that the latest interaction corresponds to the event of notifying students of their outcome. However, since students can retry game challenges as many times as desired after receiving their result, value *t*-1 and control *t*-1 can may also come from the retrospective-outcome emotions network (Figure 4(4)). Finally, when students have been presented with their game outcome, they must self-report their emotion towards it and may decide to proceed with another challenge, thereby starting another game dialogue. In this case, value *t*-1 and control *t*-1 come from the retrospective-outcome emotions network (Figure 4(5)).

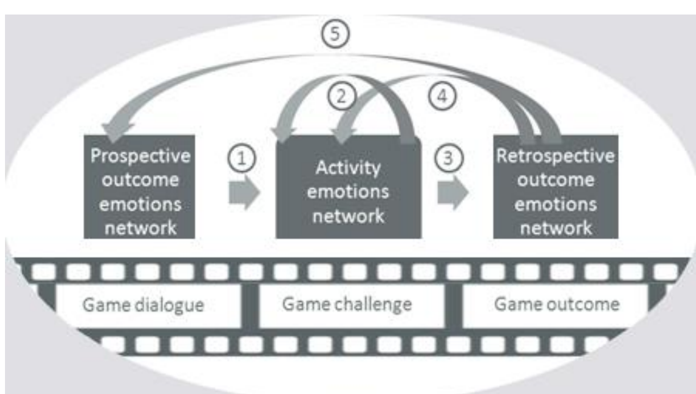

Figure 4. Time *t*-1 in our student model

Experiment design

PlayPhysics

To acquire data related to student interaction with a GBL environment and enable students to communicate their emotion over time, we created *PlayPhysics*, an emotional GBL environment for teaching Physics. *PlayPhysics* includes our emotional student model and will enable testing of the hypothesis of this work. However, *PlayPhysics* also has to assist students in learning Physics. This section discusses key aspects of *PlayPhysics'* design. For looking at more detail see Muñoz et al. (2013).

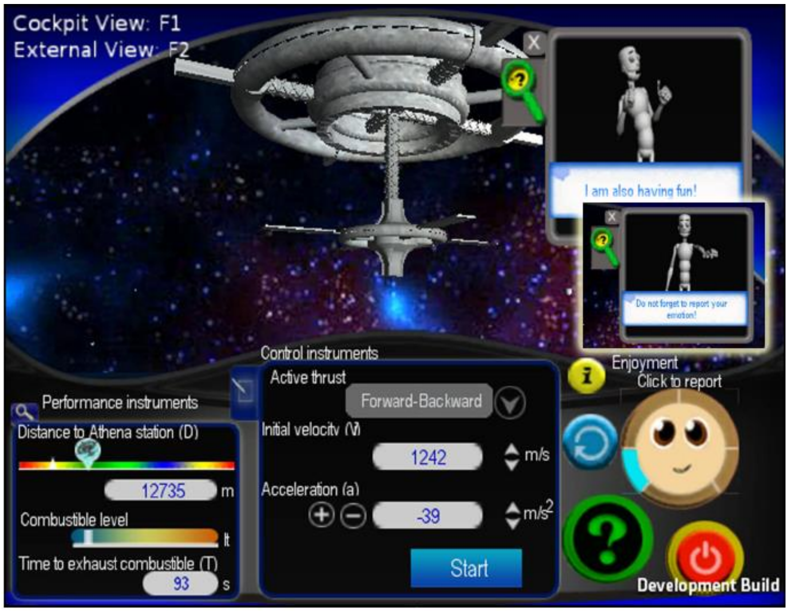

Figure 5. Game challenge and GUI of *PlayPhysics*

PlayPhysics functional and non-functional requirements were defined by conducting an on-line-survey of a course in introductory physics from the Tecnologico de Monterrey, Mexico City campus (ITESM-CCM) and Trinity College, Dublin. As a result, *PlayPhysics* is focused on teaching the topics of Newton's laws for particles and rigid bodies, Dynamics and Kinematics and vectors in 3D, which were judged as the most challenging topics.

PlayPhysics is a Role-Playing Game (RPG) and space adventure comprising challenges that must be overcome using knowledge of physics. The student is an astronaut with the mission of saving Captain Foster, who is trapped in space station Athena. The mission begins when the student is going to be launched from Earth to travel to Athena, which is located between Mars and Jupiter, and which is rotating with a constant acceleration. Each challenge is associated with one game level. Here, we focus on the first challenge.

An expert in Astrophysics from the ITESM-CCM assisted us in defining the game- scenarios. The first game-challenge comprises the Alpha Centauri spaceship, which has been launched from the Earth, and the Athena station. Alpha Centauri is heading at constant speed towards Athena, see Figure 5. The main goal is that the student sets suitable values for physics variables of Alpha Centauri to stop along Athena's rotational axis. However, to make this goal challenging, the student has to fulfill conditions such as defining a position that facilitates docking and entering to Athena before the fuel is exhausted.

Subjects

For our investigation, we invited students enrolled in a related Engineering undergraduate degree at ITESM-CCM, and in an age range between 18 and 23 years old. We acquired the data from 118 participants that interacted with *PlayPhysics*.

Objects

The first challenge is related to the topic of one-dimensional rectilinear motion. To achieve a successful outcome, constant deceleration has to be applied. Two constraint variables in this challenge are: (1) the initial distance ($\boldsymbol{D}$) from Alpha Centauri to Athena and (2) the time remaining until fuel is exhausted ($\boldsymbol{T}$). As a result, both variables are assigned randomly within specific value ranges: $\boldsymbol{D} \in [17, 50]$ m and $\boldsymbol{T} \in [80, 120]$ s, to make the solution of the challenge non- trivial. Students must concentrate on setting the values of the exploration variables.

The factors that must be considered to solve *PlayPhysics*' first challenge appropriately are:
- Choosing the correct direction for the acceleration a of Alpha Centauri spaceship.
- Setting the magnitude of the acceleration of Alpha Centauri considering that humans black-out if $\boldsymbol{a} > 4\boldsymbol{g}$, where $\boldsymbol{g}$ is the gravity acceleration at sea level ($\boldsymbol{g} = 9.81$ m/s^2)
- Not going beyond the fuel exhausting time, $\boldsymbol{t_s} \leq \boldsymbol{T}$, and achieving the lowest relative error, $\boldsymbol{e_d} \leq 2\%$, in the breaking distance ($\boldsymbol{d_s}$).

$$e_d = \frac{d_s - D}{D} \times 100$$

- Defining the lowest value for the breaking time (t_s).

These factors were implemented in *PlayPhysics* as rules to diagnose student knowledge. The simulation model is concerned with the representation of the physics domain.

Instrumentation

Students solved a pre-test, and afterwards interacted with the first challenge of *PlayPhysics*, and finally solved a post-test and qualitative questionnaire. Students self-reported their emotional state before, during and after performing the game activity.

During the interaction with the game challenge, the student's emotion can be reported at any time, using the *EmoReport* wheel (Figure 6 (a)). The emotion relating to the outcome at the end of the challenge is always enquired (Figure 6 (b)), whether the challenge finishes due to an error or misunderstanding or due to a successful end. Learning companion M8- robot provides an emotional response every time the student reports his or her emotional state (See Figure 6).

Figure 6. (a) *EmoReport* wheel and (b) Learning companion M8- robot

Data collection, cleaning and analysis

We collected the data from 118 students at ITESM-CCM from the Faculty of Computing and Engineering, who interacted freely with *PlayPhysics* during one week. Through applying to the collected data NPC - in combination with the information obtained from applying BLR/MLR and Pearson correlations using SPSS - and EM algorithms using Hugin Lite, we obtained the dynamic sequence of BNs, comprised of the *prospective-outcome*, *activity* and *retrospective-outcome* networks. The resultant activity emotions network is shown in Figure 7. We used 708 cases related to student game interaction to derive this BN, where 136, 122, 262 and 188 cases corresponded to anger, boredom, enjoyment and frustration. WEKA was employed to perform stratified random sampling in order to obtain 499 cases from the original 708 cases that we had, since we used the free version of Hugin Lite, which is limited to 50 states and 500 cases. Also, we employed WEKA to convert continuous to categorical variables using equal frequency binning to divide the variables into two or three categories, bearing in mind that control and value are also divided into three and two categories respectively.

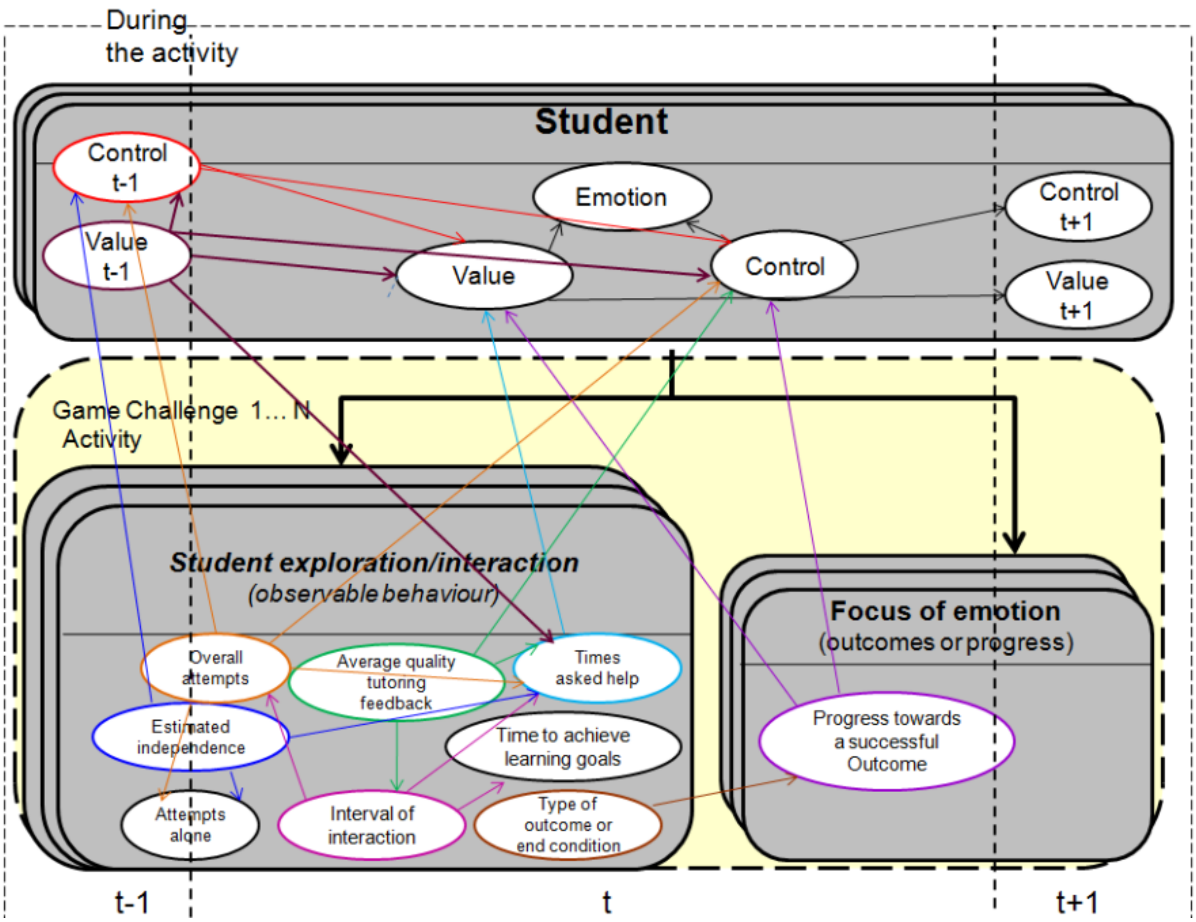

Figure 7. Activity emotions network

We employed 10-fold cross-validation using the available data to determine the performance of each network over fresh data. The objective was to compare the performance of each network, and we obtained sensitivity, specificity, precision and accuracy measures of the networks (Han & Kamber, 2006).

True positives (t_{pos}) are positive tuples that were correctly labelled by the classifier. True negatives (t_{neg}) are negative tuples that were labelled by the classifier. False positives (f_{pos}) are negative tuples that were negatively labelled by the classifier. False negatives (f_{neg}) are positive tuples that were incorrectly labelled by the classifier.

Considering these definitions; it is possible to define sensitivity, specificity and precision from them. Sensitivity (ss) is the true positive **(**t_{pos}) recognition rate. Specificity (sp) is the true negative (t_{neg}) rate. Precision (prec.) is the percentage of tuples that actually belong to each labelled category. Accuracy (acc.) is a function of sensitivity and specificity. Results corresponding to the classification of the different types of achievement emotions are presented in Table 4-6.

From the prospective- outcome emotions, anxiety and hope are classified with 80% and 67.5% accuracy (see under *ss*). This is owed to classifying more appropriately control in the “Medium” category rather than the “High” category using answers to questions in game-dialogues.

Table 4. Performance of the prospective-outcome emotions network

Prospective outcome emotions								
Observed	Predicted							
	Anticipatory joy	Anticipatory relief	Anxiety	Hope	sp	ss	prec.	acc.
Anticipatory joy	6	3	1	10	0.857	0.300	0.375	0.567
Anticipatory relief	2	10	2	6	0.914	0.500	0.625	
Anxiety	0	1	8	1	0.925	0.800	0.571	
Hope	8	2	3	27	0.660	0.675	0.613	

Note. sp = specificity; ss = sensitivity; prec. = precision; acc. = accuracy.

From the activity emotions (Table 5), enjoyment and frustration are classified with accuracies (see under *ss*) of 67.8% and 60.0%, respectively. However, anger and boredom are classified with accuracies of 48% and 20%. This is due to being unable to recognize value appropriately in its category “None,” also the precision of frustration, is not very high, as a result, there is a high probability of classifying the other emotions as frustration.

Table 5. Performance of the activity emotions network

Activity emotions								
Observed	Predicted							
	Anger	Boredom	Enjoyment	Frustration	sp	ss	prec.	acc.
Anger	48	2	27	23	0.848	0.480	0.440	0.532
Boredom	15	18	31	26	0.954	0.200	0.486	
Enjoyment	30	6	122	22	0.740	0.678	0.595	
Frustration	16	11	25	78	0.808	0.600	0.523	

Note. sp = specificity; ss = sensitivity; prec. = precision; acc. = accuracy.

Table 6. Performance of the retrospective-outcome emotions network

Retrospective outcome emotions										
Observed	Predicted									
	Anger	Gratitude	Joy	Pride	Sadness	Shame	sp	ss	prec.	acc.
Anger	77	3	1	1	10	8	0.613	0.770	0.554	0.504
Gratitude	5	0	0	1	3	1	0.932	0.000	0.000	
Joy	11	1	2	3	3	0	0.992	0.100	0.500	
Pride	11	8	1	9	0	1	0.974	0.300	0.600	
Sadness	22	3	0	0	27	8	0.880	0.450	0.529	
Shame	13	2	0	1	8	16	0.918	0.400	0.471	

Note. sp = specificity; ss = sensitivity; prec. = precision; acc. = accuracy.

On the other hand, from the retrospective-outcome emotions (Table 6), anger is classified with an accuracy of 77%. However, its precision it is not very high. Gratitude is not classified accurately at all.

Table 7. Cohen’s Kappa for the achievement emotions networks

Dependent variable	Prospective-outcome emotions BN		Activity emotions BN		Retrospective outcome-emotions BN	
	κ	Significance	κ	Significance	κ	Significance
Emotion	0.369	7.732E-9	0.348	5.108E-39	0.310	4.377E-21
Value	0.550	1.811E-7	0.381	2.726E-29	0.437	1.881E-13
Control	0.311	0.003	0.429	2.433E-22	0.306	1.127E-12

Cohen’s Kappa (κ), an intra-class correlation coefficient, is employed as a measure of agreement that adjusts the observed proportional agreement by considering the amount of agreement expected by chance. Kappa can take values in a range [−1, 1], but only values in a range [0, 1] are meaningful, where the value of zero corresponds to random classification. For hypothesis testing, if Kappa lies on the range: $0.2 < \kappa \leq 0.4$, it corresponds to a fair agreement between the observed and the predicted values. If Kappa lies on the range: $0.4 < \kappa \leq 0.6$, it corresponds to a moderate agreement. If Kappa lies on the range: $0.6 < \kappa \leq 0.8$, it corresponds to a substantial agreement. The values of Cohen's Kappa calculated for the achievement emotions networks are presented in Table 7. Control, value and emotion achieve fair-moderate classification accuracy and results are not random. This gives enough evidence to accept our alternative hypothesis.

Evaluation and discussion

PlayPhysics teaches physics and was created with the intention of providing instruction to students in an introductory course of physics. Therefore, *PlayPhysics* targets students in the last year of High school and first years of undergraduate education. In this work, we focused principally on assessing a student model of emotion for the target population using a Cognitive-Based approach. In this case, the Control-value theory by Pekrun et al. (2007) is the cognitive psychological theory used to derive the model.

Conati and Maclaren (2009) used the cognitive psychological theory of the OCC model (Ortony et al., 1990) as a reference for their model. However, this theory was not originally created to explain emotion in an educational context, but instead was created for reasoning about emotion in personal diaries. So, it is not clear if the emotions chosen are relevant to, or will arise in the same manner during the teaching-learning experience. Conati and Maclaren (2009) employed an Embodied Pedagogical Agent (EPA) to remind students to self-report their emotion. In a similar manner, *PlayPhysics* employs the learning companion M8- robot. Students can also use a pop-up window to report their emotion in Prime-Climb, and in similar manner, students using *PlayPhysics* can employ the *EmoReport* wheel. However, this is always present in *PlayPhysics'* game challenges screen. Joy, distress, admiration and reproach are the emotions identified by PrimeClimb. Conati and Maclaren (2009), as other researchers, present the results corresponding to their emotional model using percentages of agreement between student self-reports and the predictions of the emotional model (69.59%, 62.30%, 67.42%, and 38.66% accuracy for joy, distress, admiration and reproach respectively), which makes it difficult to appreciate its reliability.

Sabourin et al. (2011) also focuses on recognising student achievement emotions using CRYSTAL ISLAND as does *PlayPhysics*. But, CRYSTAL ISLAND uses the appraisal based theory of learning emotions by Elliot and Pekrun (2007) as a reference. It differs from Control-value theory in that it relates the attainment of performance or mastery of goals and its valence with the experience of achievement emotions. In similar way, Sabourin et al. (2011) do not consider the category of "no-emotion" in their model, as in our investigation, since it is not defined by either theory. Their results are also reported as percentages of agreement, so it cannot be known whether the agreement is or is not random. Sabourin et al. (2011) focused on identifying student confusion, curiosity, excitement, focus, anxiety, boredom and frustration. The latter two were identified with accuracies of 18% and 28% respectively, whilst *PlayPhysics* identifies these two emotions using Control-value theory with accuracies of 20% and 60% respectively employing only contextual variables. Our emotional student model is the first and only model to date that was implemented using Control-value theory.

Another theory, adapted and employed to identify emotions in education using facial expressions is the theory by Ekman (1999), which has been successfully employed by Autotutor (D'Mello et al., 2008) in laboratories. However, this approach has still not proven effective in classrooms or on-line environments. Autotutor uses artificial neural networks to classify features of emotion. As a result, the emotional model is more like a black box and does not result in an intelligible model of emotion, i.e., does not provide further information about the participants or the affective domain. *PlayPhysics*' emotional student model is intelligible and assists us in identifying factors that are considered actual predictors of control and value and the manner in which these are associated. The model also assists us in achieving an enhanced understanding of the student population.

We employed PRMs to achieve an enhanced understanding of the variables that may be considered while creating our emotional student model. Therefore, they facilitate defining Bayesian student models. This approach has been employed previously by Sucar and Noguez (2008), but for the purpose of defining a student model capable of identifying the level of a student knowledge or understanding. The application of the NPC algorithm for structural learning has been successfully employed in the area of telecommunications (Bashar, Parr, McClean, Scotney, &

Nauck, 2010) when scarce data is available. Here, we employ the same approach in combination with information acquired through applying BLR/MLR and Pearson correlations to solve uncertain relations. Pearson Correlations have been successfully employed as criteria for defining the structure of a Bayesian student model of attitudes (Arroyo & Woolf, 2005). We use the results of applying BLR/MLR as criteria for creating the network structure, since Bayesian models are a kind of Logistic Regression (Roos, Wettig, Grünwald, Myllymäki, & Tirri, 2005) and we can know the contribution of each selected variable to the prediction. We are not aware of any other research that employs BLR or MLR for this same purpose.

Conclusion and future work

We presented here an investigation about whether the creation of a computational model of student emotions using Control-value theory (Pekrun et al., 2007) can achieve a reasonable accuracy recognising student emotions in online GBL environments. *PlayPhysics* was implemented to test whether our emotional student model can be applied to GBL environments. Results showed that our model attains fair-moderate accuracy with results that are not random using answers in game dialogues and contextual variables. But, the resulting model is not highly accurate (Values of Cohen's Kappa where $\kappa \geq 0.75$). Therefore, future work will focus on utilising other observable variables such as facial expressions, sentiment and speech to identify other features to enhance the classification of control and value. Also, the approach that we employed to derive the dynamic sequence of BBNs proved effective in creating an intelligible emotional student model and may be employed to derive other dynamic and intelligible data models to attain an enhanced understanding in areas other than education, e.g., e-Commerce and Genetics, in addition to the prospective areas of Affective Student Modelling and Adaptable Computer Tutoring.

References

Alexander, S., Sarrafzadeh, A., & Hill, S. (2008). Foundation of an affective tutoring system: Learning how human tutors adapt to student emotion. *International Journal of Intelligent Systems Technologies and Applications, 4*(3/4), 355-367.

Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, C., & Christopherson, R. (2009). Emotion sensors go to school. In *Artificial Intelligence in Education (AIED 2009), Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (Vol. 200, pp. 17-24). Amsterdam, The Netherlands: IOS Press.

Arroyo, I., & Woolf, B. P. (2005). Inferring learning and attitudes from a Bayesian network of log file data. In C. K. Looi, G. Mc Calla, B. Bredeweg, & J. Breuker (Eds.), *Proceedings of the 12th International Conference on Artificial Intelligence in Education, Frontiers in Artificial Intelligence and Applications* (Vol. 125, pp. 33-40). Amsterdam, The Netherlands: IOS Press.

Bashar, A., Parr, G., McClean, S., Scotney, B., & Nauck, D. (2010). Learning-based call admission control framework for QoS management in heterogeneous networks. In F. Zavoral, J. Yaghob, P. Pichappan & E. El-Qawasmeh (Eds.), *Proceedings of the Second International Conference Networked Digital Technologies (NDT 2010), Part II* (pp. 99-111). doi:10.1007/978-3-642-14306-9_11

Bouckaert, R., R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., & Scuse, D. (2012). *WEKA Manual for version 3-7-6*. Hamilton, New Zealand: University of Waikato.

Brave, S., & Nass, C. (2008). Emotion in human-computer interaction. In A. Sears & J. A. Jacko (Eds.), *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2th ed., pp. 77-92). New York, NY: Lawrence Earlbaum Associates, Taylor & Francis Group.

Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction, 19*(3), 267-303.

D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B. T., & Graesser, A. C. (2008). Automatic Detection of Learner's Affect from Conversational Cues. *User Modeling and User-Adapted Interaction, 8*(1-2), 45-80

D'Mello, S. K., Olneyc, A., Williams, C., & Hays, P. (2012). Gaze tutor: A Gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies, 70*(5), 377-398.

Del Soldato, T. (1993). *Motivation in tutoring systems* (Unpublished doctoral dissertation). The University of Sussex, Brighton, England, UK.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & T. Power (Eds.), *The Handbook of Cognition and Emotion* (pp. 45-60). Sussex, UK: John Wiley and Sons, Ltd.

Ekman, P., & Friesen, W. V. (1978). *Facial coding system: A Technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Elliot, A., & Pekrun, R. (2007). Achievement emotion in the hierarchical model of approach avoidance achievement motivation. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in Education* (pp. 57-74). London, UK: Elsevier.

Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques* (2nd ed.). San Francisco, CA: Elsevier.

Janssen, J. H., van den Broek, E. L., & Westerink, J. H. D. M. (2011). Tune in to your emotions: A Robust personalized affective music player. *User Modeling and User-Adapted Interaction, 22*(3), 255-279. doi:10.1007/s11257-011-9107-7

Jaques, P. A., & Vicari, R. M. (2007). A BDI approach to infer student's emotions in an intelligent learning environment. *Journal of Computers & Education, 49*(2), 360-384.

Jaques, P. A., Vicari, R. M., Pesty, S., & Martin, J.-C. (2011). Evaluating a cognitive-based affective student model. In S. D'Mello, A. Graesser, B. Schuller & J. C. Martin (Eds.), *Proceedings of the 4th International Conference of Affective Computing and Intelligent Interaction (ACII 2011) Part I* (pp. 599-608). doi:10.1007/978-3-642-24600-5_63

Jensen, F. V., & Nielsen, T. D. (2007). *Bayesian networks and decision graphs* (2nd ed.). Berlin, Germany: Springer.

Koller, D. (1999). Probabilistic relational models. In S. Džeroski & P. Flach (Eds.), *Proceedings of the 9th International Workshop of Inductive Logic Programming (ILP-99)* (Vol. 1634, pp. 3-13). Pittsburgh, USA: Springer.

Landowska, A. (2013). *Affect-awareness framework for intelligent tutoring systems. In Proceedings of 6th International Conference on Human System Interaction (HSI 2013)* (pp. 540-547). doi:10.1109/HSI.2013.6577878

Lazzaro, N. (2004). *Why we play games: Four keys to more emotion without story XEO design*. Oakland, CA: XEODesign Inc. Retrieved from http://www.xeodesign.com/xeodesign_whyweplaygames.pdf

Martinho, C., Machado, I., & Paiva, A. (2000). A Cognitive approach to affective user modelling. In A. Paiva (Eds), *Affect Interactions* (pp. 64-75). doi:10.1007/10720296_6

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2008). Modeling self-efficacy in intelligent tutoring systems: An Inductive approach. *User Modeling and User-Adapted Interaction, 18*(1 - 2), 81-123.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning and distance learning environments: Are they the same? *The Internet and Higher Education, 14*(2), 129-135.

Muñoz, K., Mc Kevitt, P., Lunney, T., Noguez, J., & Neri, L. (2013). An Emotional student model for game-based learning. In D. Griol, Z. Callejas, & R. López-Cózar (Eds.), *Technologies for Inclusive Education: Beyond Traditional Integration Approaches* (pp. 175-197). Hershey, PA: IGI Global.

Muñoz, K., Noguez, J., Mc Kevitt, P., Neri, L., Robledo-Rella, V., & Lunney, T. (2009). Adding features of educational games for teaching Physics. In *Proceeding of the 39th IEEE International Conference Frontiers in Education* (pp. 1-6). doi:10.1109/FIE.2009.5350630

Ortony, A., Clore, G. L., & Collins, A. (1990). *The Cognitive structure of emotions*. New York, NY: University Press.

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The Control value theory of achievement emotions. An Integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in Education* (pp. 13-36). London, UK: Elsevier.

Pekrun, R., Goetz, T., & Perry, R. P. (2005). *Achievement Emotions Questionnaire (AEQ).User's manual* (Unpublished Manuscript). University of Munich, Germany.

Picard, R. W. (1995). *Affective computing. Vision and modeling* (Technical report No. 21). Cambridge, Massachusetts MA: Institute of Technology (MIT).

Porayska-Pomsta, K., Mavrikis, M., & Pain, H. (2008). Diagnosing and acting on student affect: The Tutor's perspective. *User Modeling and User-Adapted Interaction, 18*, 125-173.

Roos, T., Wettig, H., Grünwald, P., Myllymäki, P., & Tirri, H. (2005). On Discriminative Bayesian network classifiers and logistic regression. *Machine Learning, 59*(3), 267-296.

Rust, A. (2010). *Google nabs patent to monitor your cursor movements* [Blog]. Retrieved from http://news.techeye.net/internet/google-nabs-patent-to-monitor-your-cursor-movements

Sabourin, J., Mott, B. W., & Lester, J. C. (2011). Modelling learner affect with theoretical grounded dynamic Bayesian networks. In S. D'Mello, A. Graesser, B. Schuller & J. C. Martin (Eds.), *Proceedings of the 4th International Conference of Affective Computing and Intelligent Interaction (ACII 11)* (pp. 286-295). doi:10.1007/978-3-642-24600-5_32

Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., & Bigdeli, A. (2008). How do you know that I don't understand? A look at the future of intelligent tutoring systems. *Computers in Human Behavior, 24*(4), 1342-1363.

Sucar, L. E., & Noguez, J. (2008). Student modeling. In O. Pourret, P. Naïm, & B. Marcot (Eds.), *Bayesian Networks: A Practical Guide to Applications* (pp. 173-185). West Sussex, UK: J. Wiley & Sons.

Sykes, J. M. (2013, October). Technology—"Just" playing games? A Look at the use of digital games for language learning. *The Language Educator* (pp. 32-35). Retrieved from https://www.actfl.org/sites/default/files/pdfs/TLE_pdf/TLE_Oct13_Article.pdf