2012 12th IEEE International Conference on Advanced Learning Technologies

Literature Driven Method for Modeling Frustration in an ITS

Ramkumar Rajendran *†‡, Sridhar Iyer †, and Sahana Murthy †

*IITB-Monash Research Academy
†IIT Bombay, Mumbai, India
‡Monash University, Melbourne, Australia
Email:ramkumar.r@iitb.ac.in

***Abstract*—In Intelligent Tutoring Systems, affect-based computing is an important research area. Common approaches to deal with the affective state identification are based on input data from external sensors such as eye-tracker and EEG, as well as methods based on mining of ITS log data. Sensor based methods are viable in laboratory settings but they are tough to implement in real-world scenario which might cater to a large number of students. In our research, we create a mathematical model of frustration based on its theoretical definition. We identify the variables in the model by applying the theoretical definition of frustration to the ITS log data. This approach is different from existing data mining techniques, which use correlation analysis with labeled data. We apply our model to Mindspark, a commercial maths Intelligent Tutoring System, used by several thousand students. We validate our model with human observations of frustration.**

***Keywords*-Affective State Detection; ITS; Modeling Frustration; Student Log Data;**

I. Introduction

An Intelligent Tutoring System (ITS) dynamically adapts the learning content based on the learner's needs and preferences, in order to provide a personalized learning experience to each student. An ITS consists of an adaptation engine, learning content and the student model. The student model typically contains information about students such as their previous knowledge, background and behavior [1]. However, it has been well established that the learning process involves both cognitive and affective processes [2]. Further, the consideration of affective processes has been shown to achieve higher learning outcomes. The importance of the students' motivation and the affective component in learning has led adaptive systems such as ITS to include learners' affective states in their student models. Student models are constructed from log data available in ITS. The log file captures students interaction with the ITS such as response to questions, number of attempts and the time taken for various activities (responding, reading, and so on).

The detection of learners' affective states while interacting with ITS has been the subject of a great deal of research. Several methods have been implemented in ITS to gather data about affective states: human observation, learners' self-reported data of their affective state, machine learning through software log data and more recently, sensing devices such as physiological sensors and face-based emotion recognition systems. While advances in physical sensors and their data analysis techniques are promising in the lab setting, they are as yet difficult to implement at a large scale in a real-world scenario. Other methods rely on features particular only to certain ITS, for example, conversation clues can only be used in an ITS that includes dialogue between the tutor and the student. Due to these limitations, data mining approaches using software log data seems to be a viable alternative for large scale implementation in commercial ITS.

In our research, we consider affective states suggested by R. S. Baker et. al. [3] they are boredom, frustration, confusion, delight, engaged concentration and surprise. Techniques to identify these states using data mining approaches have been reported in previous works [4]–[6]. These approaches rely on correlation analysis between the affective states and the features of the ITS. In this paper, we present a mathematical model for identifying students' frustration when they interact with an ITS. Our model is derived from a well-accepted definition of frustration as stated in psychology literature. We operationalize the theoretical definition of frustration in terms of common features available in commercial ITS. We apply the model to a commercial mathematics ITS, Mindspark. We validate our model by comparing its identification of students' frustration with human observation of students interacting with Mindspark. In section 2, we discuss related work which identifies students' affective states. Our proposed research approach is described in section 3. We discuss the application of our proposed model to Mindspark, and its validation in section 4. In section 5, we discuss results of our model. We summarize and discuss future directions of research in section 6.

II. Related work

To identify affective states, R. S. Baker et. al. [3] suggest three different methodologies: human observation, using hardware sensors and machine learning techniques using data from student log files. There are various research studies that identify learners' affective states using data from physiological signals [7] such as electrocardiogram (ECG), facial electromyograph (EMG) and galvanic skin response

978-0-7695-4702-2/12 $26.00 © 2012 IEEE
DOI 10.1109/ICALT.2012.167

(GSR), from various sensors such as blue eye camera [8], posture analysis seat, pressure mouse and skin conductance bracelet [4], and from conversation clues [6]. A recent review paper on affective states detection [9] summarizes that identification of affective states using sensor signals, facial expression, text analysis and voice analysis are widely researched compared to data mining methods.

In our research, we focus on systems which identify students' affective states from log files, where the data are collected from students' interaction with the system. An example of such a system is Autotutor [6], in which students' frustration is identified from log data such as subtopic number, response time and turn number. Emotions like joy/distress are identified from students' goals, interaction patterns and actions while playing the Prime-Climb math game [5]. In yet another example, students' average frustration in computer programming exercises across different labs is detected based on information from compiler data, such as the average time between compilations and consecutive pairs of same error etc [10]. In all systems mentioned above, the features used to detect students' affective state are specific to the ITS or the gaming environment. The Prime-Climb system [5] has goals such as 'beat one's partner' and 'avoid falling', which are not common in all games or ITS. Features in Autotutor [6] are specific to dialogue-based tutoring systems.

Mathematical models used in data mining techniques are created by identifying which features from student log data are important. This is done by correlation analysis to select all features which are significantly correlated with affective states and then applying dimensionality reduction techniques like principle component analysis (PCA) [6]. PCA is used to to reduce the number of features (dimensions) by removing the correlation between features and to produce a smaller, uncorrelated set of features. The systems used in existing approaches are rich in the data they log. In our approach we use the commercial ITS which records relatively less data.

III. Proposed Research Approach

Our method differs from the above approaches in the manner we extract features. We base our approach on the definition of frustration from literature. We begin with a well-accepted definition of the frustration, and the factors that could lead to the experience of the particular emotion by a human being. We then create a generic linear regression model using these factors from the fundamental definition. To apply the model, we identify features in the student log data of the system (such as the ITS under study) that could give rise to the factors leading to the affective state. We use features in our model that are commonly found in most ITS, such as students' response to questions, or time taken to answer.

The classic definition of the frustration from psychology [11] is "Frustration refers to the blocking of behavior directed towards a goal". From this definition, we see that we will first need to identify features in an ITS that correspond to students' goals, and then use those features to model "blocking of behavior directed towards a goal". We restrict the scope of our method to identify frustration that occurs due to students' goal blockage while interacting with ITS. At this stage, we do not include frustration that might have occurred due to factors external to students' interaction with the ITS.

Generic Linear Regression Model for Frustration

To create a mathematical model for frustration we identify the goals of the student with respect to their interaction with the ITS, and select the top n goals. Based on information from the student log, a blocking factor bf for each of the n goals is identified, for example $goalj_{bf}$ represents the blocking factor for the $goalj$. We formulate a linear function F_i, as the frustration index at i^{th} question based on the blocking behaviors of student goals. The linear regression formulation of frustration is given below:

$$F_i = \alpha[w_0 + w_1 * goal1_{bf} + w_2 * goal2_{bf} + ... \\ +w_n * goaln_{bf}] + (1-\alpha)[F_{i-1}] \quad (1)$$

Here w_0, w_1, w_2, w_n and α are weights. The frustration index of the student at the previous question answered, F_{i-1}, is added to account for the cumulative effect of frustration building up over consecutive questions [12]. The value of α determines the contribution of frustration at $(i-1)^{th}$ question to frustration at i^{th} question, α ranges from 0 to 1. We assume that the student is not frustrated at the beginning of their interaction with the ITS, and hence choose $F_i = 0$ for $i = 1, 2$.

To calculate the frustration index F_i for a student interacting with a particular ITS, we have to create the frustration model in the context of that ITS, that is, we have to identify the specific goals and blocking factors of each goals. Considering the data logged in the ITS, operationalize the blocking factors of the goals. In next section we show how to create frustration model by applying our prosed method to a Math ITS, Mindspark.

IV. Creation and Validation of Frustration Model

In this section we explain the system, Mindspark, in which we created and validated the frustration model, based on the discussion in the above section.

A. Context: Mindspark ITS

We use Mindspark, a commercial mathematics ITS developed by Educational Initiatives India (EI-India), to implement and test our proposed model. Mindspark is being used as a part of the school curriculum for different age groups (grade 3 to 8) of students [13]. Mindspark is currently being implemented in sixty schools and being used by 30,000 students. Mindspark teaches mathematics by asking questions to the students and provides the a detailed feedback and explanation after receiving the answer from the student. Mindspark adapts the next questions based student's answer to current question and performance in the topic which allows student to move at her own pace. If the student's performance at the current topic is below 75%, she will be moved to previous level in the same topic which is basic level compared to current level.

In Mindspark, if a student answers three consecutive questions correctly, she receives a Sparkie (extra motivational points). If a student answers five consecutive questions correctly, then he/she receives a Challenge Question, which is tougher than normal questions. If the student answers the Challenge Question correctly, he/she receives five Sparkies. Every week, the highest Sparkie collectors (Sparkie Champ) are identified and their names are published in the Mindspark website[1]

B. Creation of Model for Mindspark

Students' Goals: We identified the three most common goals of students while interacting with Mindspark. To identify these goals, we conducted interviews with the staff of EI-India. Mindspark staffs interacted with students and recorded the students' goal while interacting with Mindspark. The goals ($goal$) and the blocking factor ($goal_{bf}$) to achieve the goals are given Table I.

Table I
STUDENT GOALS AND BLOCKING FACTORS

Student Goal	Blocking factor
$goal1$: To get the current question correct	$goal1_{bf}$: The answer to the current question is wrong
$goal2$: To get a Sparkie	$goal2a_{bf}$: If answers to last two questions are correct and to current question is wrong $goal2b_{bf}$: If an answer to last question is correct and to current question is wrong
$goal3$: To reach the challenge question	$goal3a_{bf}$: If answers to last four questions are correct and to current question is wrong $goal3b_{bf}$: If answers to last three questions are correct and to current question is wrong
$goal4$: To get the challenge question correct	$goal4_{bf}$: The answer to challenge question is wrong

To model each $goal_{bf}$ we consider students' response to a question, a feature captured in Mindspark student log file.

For $goal1$ of '*to get the current question correct*' the blocking factor is getting the answer to current question is wrong. We use a_i to represent the answer to current current question, $a_i = 1$ if correct, $a_i = 0$ if incorrect. The blocking factor of $goal1$ is captured using

$$goal1_{bf} = (1 - a_i) \tag{2}$$

[1]http://www.mindspark.in/login/

For $goal2$, the student will get a Sparkie if he/she answers three consecutive questions correctly. This goal can be blocked if the student gets either one question wrong, or two questions wrong. Thus, $goal2$ can be blocked in two ways. The blocking factor of $goal2$ has two components:

$$goal2_{bf} = goal2a_{bf} + goal2b_{bf} \tag{3}$$

One way to block the $goal2$ is, if student answers first two questions correctly in a sequence of three questions and the third question is wrong. This is captured by blocking factor $goal2a_{bf}$

$$goal2a_{bf} = (a_{i-2} * a_{i-1} * (1 - a_i))$$

The second way to block the $goal2$ is, if the student answers only the first of a sequence of three questions correctly and second question is wrong. This is captured by blocking factor $goal2b_{bf}$

$$goal2b_{bf} = a_{i-1} * (1 - a_i)$$

Similar to $goal2$, blocking factor for $goal3$ also split into two components:

$$goal3_{bf} = goal3a_{bf} + goal3b_{bf} \tag{4}$$

The blocking factor $goal3a_{bf}$ is captured by

$$goal3a_{bf} = (a_{i-4} * a_{i-3} * a_{i-2} * a_{i-1} * (1 - a_i))$$

and the blocking factor $goal3b_{bf}$ is captured by

$$gaol3b_{bf} = (a_{i-3} * a_{i-2} * a_{i-1} * (1 - a_i))$$

For $goal4$: '*To get the challenge question correct*' the blocking factor is getting the answer to challenge question wrong. The blocking factor of $goal4$ is captured using

$$goal4_{bf} = I * (1 - a_i) \tag{5}$$

where, I is the indicator for whether the current question challenge question or not. $I = 0$ for normal question and $I = 1$ for challenge question. The mathematical model to predict frustration for Mindspark data is given in eq (6).

$$F_i = \alpha[w_0 + w_1 * goal1_{bf} + w_2 * goal2_{bf} + w_3 * goal3_{bf} + w_4 * goal4_{bf}] + (1 - \alpha)[F_{i-1}] \tag{6}$$

C. Validation of Model

To validate the above model and to determine the weights, we need to identify students' affective state using an independent method. In this study, we used human observation as the independent method. We recorded students' facial expressions in one Mindspark session of eleven students. Each Mindspark session is 25-30 minutes long and a student answers 30 to 40 questions per session. The student's facial expression during the interaction with Mindspark is recorded using a web camera. The student's interaction with Mindspark is recorded using Camstudio open source[2], free streaming video software.

The student's facial expressions are coded after every answer the student submits, to capture his expression when the goal is blocked. Based on guidelines given in [8] and [14] the student's facial expressions such as outer brow raise, inner brow raise, pulling at her hair, statements like "what", "this is annoying", "arey" (hindi word for disappointment) and so on are considered as frustration. We captured 423 observations in all. Among those, 72 observations were classified as frustration and remaining as non-frustration.

We represent the values obtained from human observation as B_i at the i^{th} instance, $B_i = 0$ for non-frustration and $B_i = 1$ for frustration. The predicted frustration P_i is identified by applying a threshold to the frustration index F_i to determine if the state is frustration (Frus) / non-frustration (non-Frus). The threshold value to determine the frustration is the average of values used to represent Frus and non-Frus for training the model. In this experiment we use 0 and 1 to represent non-Frus and Frus respectively, so the threshold used is 0.5. Our goal is to minimize the error $(P_i - B_i)$ by varying the weight values w_0, w_1, w_2, w_3, w_4.

$$min(P_i - B_i)^2$$

$$by\ varying\ w_0, w_1, w_2, w_3, w_4$$

To solve this linear regression problem, we used GNU Octave[3], a open source numerical computation software. We selected six students' data (corresponding to $2/3$ of the data) and used it as training set to determine the weights, and used the trained model to predict the frustration for rest of the students' data ($1/3$ of the total data). Our method leads to a converged set of weights after 10000 iterations. We manually vary α, which represents the proportion of frustration arising from previous questions, from 0.1 to 1 in steps of 0.1, and check our results to determine the α value.

To evaluate our model we use the metrics precision and accuracy, calculated from the contingency table for our model, as shown in Table II.

Table II
CONTINGENCY TABLE

	Actual Frus	Actual Non-Frus
Predicted Frus	True Positive (TP)	False Positive (FP)
Predicted Non-Frus	False Negative (FN)	True Negative (TN)

and

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

[2]www.camstudio.org

[3]http://www.gnu.org/software/octave/

As mentioned earlier it is not feasible to predict the frustration experienced by a student due to all sources, especially those extrinsic in nature. So in our research work we are more interested in reducing FP (Precision) compared to FN (Recall).

V. RESULTS

For our analysis, we randomly selected six students' data for training the model and used the trained model to predict remaining students' affective state. To avoid the bias on training and testing data we did a three-fold cross validation analysis. The results of our study is given in Table III. The value of $\alpha = 0.8$, and the threshold to classify frustration versus non-frustration is 0.5 in Table III.

Table III
PERFORMANCE OF LINEAR FRUSTRATION MODEL IN MINDSPARK

Method	*Precision in %*	*Accuracy in %*
Method 1: 6 students' data training and predicting 4 students' frustration	85	94.97
Method 2: Three fold cross validation	75.38	90.78

In first method we used 7 students data (264 observation) to determine the values for weights. We applied the identified weight values in eq (6) and used remaining 4 students data (159 observations) to test the model. Contingency table for method 1 shown in Table IV is used to calculate accuracy and precision given in III. Similar contingency table were used for method 2 to calculate the results. In method 2 we used three fold cross validation method. We combined the results of three fold (423 observations) to produce a single estimation.

Table IV
CONTINGENCY TABLE FOR METHOD 1

	Actual Frus	Actual Non-Frus
Predicted Frus	TP = 17	FP = 3
Predicted Non-Frus	FN = 5	TN = 134

One parameter in the model we need to vary and choose is α. In the model for the Frustration Index of a student (1) at a given question, the value of α appears as a weight to the goal blockage at the current question, and $1 - \alpha$ appears as a weight to the frustration index at the previous question. Thus α represents the contribution of previously accumulated frustration, capturing the idea that frustration is cumulative. In our analysis, we applied different values for α and analyzed the precision and accuracy. We found that $\alpha = 0.8$ gave the best results of precision as well as accuracy. This means that the frustration level at i^{th} is question is 80% depends on goal blockage at i^{th} question and 20% on previously accumulated frustration, if any.

VI. Summary and Future Work

In this paper, we proposed a mathematical model to identify the frustration of a student while interacting with an ITS. The model is developed from a fundamental definition of the psychological emotion of frustration, and operationalizing it using data from student log. Based on the results, the mathematical model can predict the frustration as well as in previous research methods [6], [10].

Since our method uses features from log data which are available in most commercial ITS, it would be relatively simple for any ITS to adapt our technique. Each ITS would first need to identify the important goals, and map how the blocking of goals would manifest in the features available in student log data. Careful thought is required to Operationalize the blocking factors of goals from log data. The results of the theory driven model is depend on how well the goals are captured and blocking factors of the goals are operationalized. Once the affective state of a student is identified, an ITS can include it in its adaptation logic.

A limitation of our current study is that our results are based on a relatively small number of students. In order to generalize our results, we need to validate our model on a larger sample. Secondly, the proposed frustration model is a linear combination of goals. However, the model may perform better for the second or higher order polynomial. To explore this issue, we plan to compare different models. Once we test our model with a larger sample size and different mathematical functions, we plan to use a similar technique to model other affective states such as boredom and confusion. Another possible avenue of future work is to include the identification of affective states in Mindspark's adaptation logic.

Acknowledgment

The authors would like to thank Prof. Campbell Willson and Prof. Judy Sheard for their guidance and comments on this paper. Also we thank all students who accepted to participate in our study, staff and Managing Director of Mindspark for their support to conduct our research.

References

[1] Peter Brusilovsky and Eva Millán. User models for adaptive hypermedia and adaptive educational systems. In *The Adaptive Web, Methods and Strategies of Web Personalization*, pages 3–53. 2007.

[2] Sidney K. D'Mello, Scotty D. Craig, Jeremiah Sullins, and Arthur C. Graesser. Predicting affective states expressed through an emote-aloud procedure from autotutor's mixed-initiative dialogue. *Int. J. Artif. Intell. Ed.*, 16:3–28, January 2006.

[3] Ryan S. J. d. Baker, Sidney K. D'Mello, Ma. Mercedes T. Rodrigo, and Arthur C. Graesser. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *Int. J. Hum.-Comput. Stud.*, 68:223–241, April 2010.

[4] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper, and Rosalind Picard. Affect-aware tutors: Recognizing and responding to student affect. *Int. J. Learn. Technol.*, 4:129–164, October 2009.

[5] Cristina Conati and Heather Maclaren. Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, 19:267–303, August 2009.

[6] Sidney K. D'Mello, Scotty D. Craig, Amy Witherspoon, Bethany Mcdaniel, and Arthur Graesser. Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction*, 18:45–80, February 2008.

[7] M. S. Hussain, Omar AlZoubi, Rafael A. Calvo, and Sidney K. D'Mello. Affect detection from multichannel physiology during learning sessions with autotutor. In *International Conference on Artificial Intelligence in Education*, pages 131–138, 2011.

[8] Sidney K. D'mello, Scotty D. Craig, Barry Gholson, and Stan Franklin. Integrating affect sensors in an intelligent tutoring system. In *In Affective Interactions: The Computer in the Affective Loop Workshop at Intl. Conf. on Intelligent User Interfaces*, pages 7–13, 2005.

[9] Rafael A. Calvo and Sidney D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affect. Comput.*, 1:18–37, January 2010.

[10] Ma. Mercedes T. Rodrigo and Ryan S.J.d. Baker. Coarse-grained detection of student frustration in an introductory programming course. In *International workshop on Computing Education Research*, pages 75–80, 2009.

[11] Clifford T. Morgan, Richard A. King, John R. Weisz, and John Schopler. *Introduction to Psychology*. McGraw-Hill Book Company, seventh edition edition, 1986.

[12] J Klein, Y Moon, and R. W. Picard. This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, 14:119–140, January 2002.

[13] Suchismita Srinivas, Muntaquim Bagadia, and Anupriya Gupta. Mining information from tutor data to improve pedagogical content knowledge. In *International Conference on Educational Data Mining*, pages 275–276, 2010.

[14] Ma. Mercedes T. Rodrigo, Ryan S. J. d. Baker, Maria C. V. Lagud, Sheryl A. L. Lim, Alexis F. Macapanpan, Sheila A. M. S. Pascua, Jerry Q. Santillano, Leima R. S. Sevilla, Jessica O. Sugay, Sinath Tep, and Norma J. B. Viehland. Affect and usage choices in simulation problem-solving environments. In *International Conference on Artificial Intelligence in Education*, pages 145–152, 2007.