Determining the Precursors of Boredom

Michelle Marie C. Dagami[1]
rhyzz_craig_08@yahoo.com

Bao Guo[1]
guobao7@hotmail.com

Mark Berlon Pating[1]
markpating@gmail.com

Thea Faye G. Guia[1]
theafayeguia@yahoo.com

William Bernard S. Leonor[1]
william_leonor12@yahoo.com

Ma. Marcedes T. Rodrigo[1]
mrodrigo@ateneo.edu

ABSTRACT

This paper presents analysis of user action log files and human observation data gathered from students using Aplusix. The objective of the study is to determine the precursors of student boredom, an affective state that is negatively correlated with learning. The researchers matched the action logs with the human-coded observation by using the time window synchronization technique. The researchers segregated the data into bored and non-bored time windows. The researchers then compared the features of these sets, e.g. frequency of different actions (i.e. hit of a key, special keys and arrow keys) and average duration between actions. The analysis identified three probable precursors of boredom: frequent use of the *ask for solution*, *ask for verification* and *ask for score,* rare use of the Number keys and rare use of the Special buttons.

Keywords

Aplusix, precursors of boredom, log files, observation, time windows.

1. INTRODUCTION

Fisher defines boredom as “an unpleasant, transient affective state in which the individual feels a pervasive lack of interest in and difficulty concentrating on the current activity” [1]. Studies that have explored the relationships between student behavior and affective states and student learning have found that boredom is negatively correlated with student achievement [2, 3, 4].

Students begin to experience boredom when their interest in using the system decreases, leading to off-task behavior that in turn lead to less learning [3]. Rodrigo et al. in [3] found that boredom tends to co-occur and precede gaming the system. Gaming is defined by Baker as a behavior aimed at obtaining correct answers and advancing within the tutoring curriculum by systematically taking advantage of regularities in the software’s feedback and help [5].

Furthermore, boredom tends to persist. Baker et al. [6] found that once a student was bored, he tended to stay bored. These results suggest that boredom is the primary cognitive-affective state which interactive learning environments should focus on detecting and quickly responding to.

The aim of this research was to determine precursors of boredom of students using Aplusix, an Intelligent Tutoring System (ITS) for Algebra. A “precursor” is the one that precedes and indicates the approach of another [7]. The study focused on the probable features that might indicate that a student is becoming bored, an undesirable state that has a detrimental effect on learning [3]. Determining these precursors might enable us to anticipate when a student is starting to become bored. Early detection might then enable us to create interventions that prevent boredom from taking place.

2. LITERATURE REVIEW

Past studies have successfully built detectors of student affective states. Frustration, for example, was an area of interest. McQuiggan et al. attempted to build an early detector of student frustration. They used an inductive approach to generating affect recognition models, wherein models are induced from observations of students interacting within a learning environment in which student actions, locations, goals, and temporal information are monitored [8]. Their induced affect recognition models were both accurate and efficient.

Rodrigo and Baker in [9] attempted to automatically detect student frustration at a coarse-grained level using measures distilled from student behavior within an interactive development environment for Java. They gathered logs of student compilations as well as human observations of novice programmers as they worked on laboratory problems. Their analysis resulted in a model that could predict novice programmer frustration based on the number of consecutive compilations with the same edit location, the number of consecutive pairs with the same error, the average time between compilations and the total number of errors.

There has been a research that attempted to model the motivational state of a learner in an ITS. Rebolledo-Mendez et al. aimed to measure the motivational states of the learners and scaffold the motivational strategies to adjust the system’s reactions according to individual pupils through analysis of relationship of the concept of the ‘Zone of Proximal Development’ (ZPD) and motivational variables [10]. They proposed to further develop Ecolab, a Vygotskyan inspired ITS, to consider motivational issues.

Further investigation on the role of emotion on student performance was conducted by Lee et al [11]. They aimed to build a model that can be used for detecting boredom and confusion through the compile logs from BlueJ environments. The results of their study were poor and the models were not very reliable.

3. METHODS

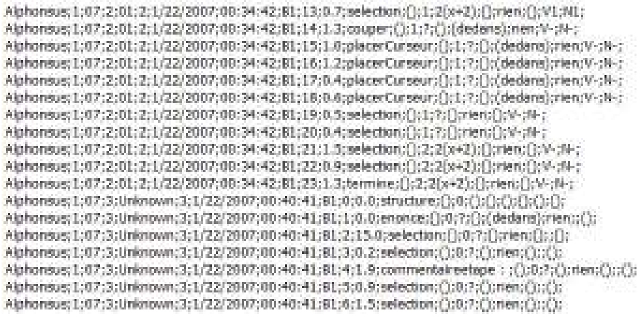
The data used for this study were the user log files and human coded observation files. These data were from the study made by Maria Carmina V. Lagud and Ma. Mercedes T. Rodrigo. Students are 140 first year and second year High School students from 4 private schools within Metro Manila and a school from Cavite. The students have an average age of 13.5 [2].

3.1 Aplusix

Aplusix II: Algebra Learning Assistant [12] is an ITS for algebra. Its content is divided into six topics: numerical calculation, expansion and simplification, factorization, solving equations, solving inequalities and solving systems of equations or inequalities. These topics are broken down into four to nine levels of difficulty. After a student chooses a problem, the system provides a problem to be solved by the student. Step-by-step calculations will be made by the students, using an advanced editor of algebraic expressions. After each step, the system indicates whether the prior step and current step were mathematically equivalent. A student can end the exercise, ask for a hint, or ask for the final answer at any time.

3.2 Log Files

Aplusix automatically logs all user interactions, and each exercise is recorded in one log. These logs were refined into a more comprehensible single log file (Figure 1).



```
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;13;0.7;selection;();1;2(x+2);();rien;();V1;N1;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;14;1.3;couper;();1;?;();(dedans);rien;V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;15;1.0;placerCurseur;();1;?;();(dedans);rien;V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;16;1.2;placerCurseur;();1;?;();(dedans);rien;V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;17;0.4;placerCurseur;();1;?;();(dedans);rien;V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;18;0.6;placerCurseur;();1;?;();(dedans);rien;V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;19;0.5;selection;();1;?;();rien;();V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;20;0.4;selection;();1;?;();rien;();V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;21;1.5;selection;();2;2(x+2);();rien;();V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;22;0.9;selection;();2;2(x+2);();rien;();V-;N-;
Alphonsus;1;07;2;01;2;1/22/2007;00:34:42;B1;23;1.3;termine;();2;2(x+2);();rien;();V-;N-;
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;0;0.0;structure;();0;();();();();();();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;1;0.0;enonce;();0;?;();(dedans);rien;;();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;2;15.0;selection;();0;?;();rien;();;();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;3;0.2;selection;();0;?;();rien;();;();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;4;1.9;commentaireetape : ;();0;?;();rien;();;();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;5;0.9;selection;();0;?;();rien;();;();
Alphonsus;1;07;3;Unknown;3;1/22/2007;00:40:41;B1;6;1.5;selection;();0;?;();rien;();;();
```

Fig. 1. A part of distilled user action log file

From the log files, information can be distilled. Lagud [2] describes the contents of the log file as follows:

1. School – the name of the participating school
2. Run – the student's run or batch number. Three to four batches of 10 students each were observed per school. The run number ranges from 1 to 3.
3. Student No. – the identification number of the student within the run taking the exercise. The student number ranges from 1 to 10.
4. Set No. – the set number of the current exercise. A set is composed of a group of items under a specific exercise category
5. Problem No. within the Set – the item number within the set number chosen.
6. Absolute Problem No. – the item number relative to all the problems answered by the student.
7. Date – the date when the exercise was done.
8. Time Started – the time when the student started with the specific problem.
9. Level – the degree of difficulty of the topic.
10. Step No. – the step number which is the number of the current step.
11. Duration – the number of seconds describing how long each step was done.
12. Action – the action performed by the student. Terms used are expressed in French (Fr.).
13. Error – the error committed by the student while solving the problem
14. Etape (Fr.) – the stage or phase of the solution.
15. Expression – the state of the mathematical expression.
16. Etat (Fr.) – the current state of the solution.
17. Cursor – location of the cursor.
18. Selection – selected values in the solution.
19. Equivalence – indicates whether the equation is correct or not.
20. Resolution – indicates whether the problem has been solved or not.

The log file was again improved for more clarity using the Log File Analyzer, as shown in Figure 2. The researchers focused only on information 1 to 12 from the log file.

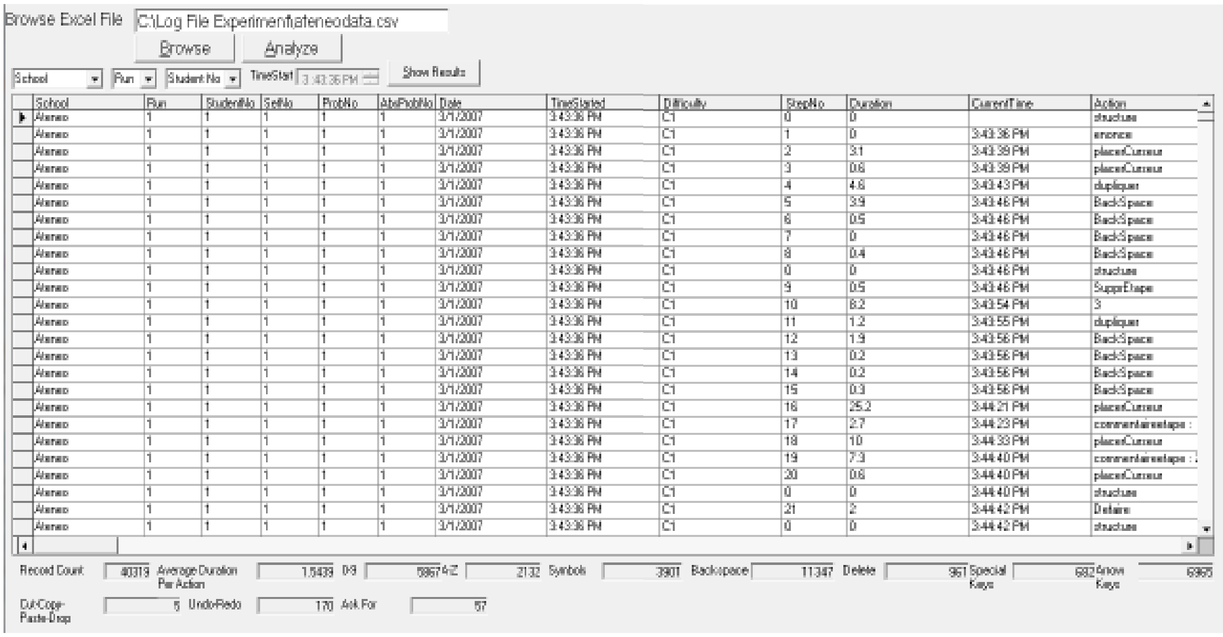
3.3 Observation

Affective profiles of students were collected using the observation method discussed in [2]. There were six coders who carried out the observation. Each observation has a duration of twenty seconds. The affective categories used were: boredom, confusion, delight, surprise, flow, frustration and neutral. Since the inter-rater reliability in [2] was acceptably high and the single consistent observer is Jess, the researchers used her labels only for identifying the affective profile of each student.

At the point of actual data collection, the human observations and the user action logs were not synchronized automatically. There was no automated system or common clock that synchronized the two data streams. The synchronization took place after collection.

3.4 Dataset Creation

A Visual Basic application called Log File Analyzer (See Figure 2) was created to filter and convert the distilled interaction log file into a more readable format.



Browse Excel File: C:\Log File Experiment\ateneodata.csv

School | Run | Student No | TimeStart 3:43:36 PM

School	Run	StudentNo	SetNo	ProbNo	AbsProbNo	Date	TimeStarted	Difficulty	StepNo	Duration	CurrentTime	Action
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	0	0		structure
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	1	0	3:43:36 PM	enonce
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	2	3.1	3:43:39 PM	placerCurseur
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	3	0.6	3:43:39 PM	placerCurseur
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	4	4.6	3:43:43 PM	dupliquer
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	5	3.9	3:43:46 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	6	0.5	3:43:46 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	7	0	3:43:46 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	8	0.4	3:43:46 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	0	0	3:43:46 PM	structure
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	9	0.5	3:43:46 PM	SupprEtape
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	10	8.2	3:43:54 PM	3
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	11	1.2	3:43:55 PM	dupliquer
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	12	1.9	3:43:56 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	13	0.2	3:43:56 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	14	0.2	3:43:56 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	15	0.3	3:43:56 PM	BackSpace
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	16	25.2	3:44:21 PM	placerCurseur
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	17	2.7	3:44:23 PM	commentaireetape :
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	18	10	3:44:33 PM	placerCurseur
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	19	7.3	3:44:40 PM	commentaireetape :
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	20	0.6	3:44:40 PM	placerCurseur
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	0	0	3:44:40 PM	structure
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	21	2	3:44:42 PM	Defaire
Ateneo	1	1	1	1	1	3/1/2007	3:43:36 PM	C1	0	0	3:44:42 PM	structure

Record Count: 40319 | Average Duration Per Action: 1.5439 | 0-9: 5867 | A-Z: 2132 | Symbols: 3901 | Backspace: 11347 | Delete: 961 | Special Keys: 692 | Arrow Keys: 6965

Cut-Copy-Paste-Drop: 5 | Undo-Redo: 170 | Ask For: 57

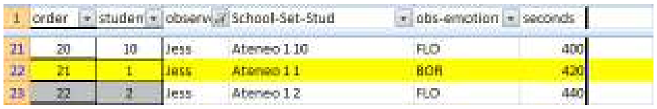
Fig. 2. A screenshot of Log File Analyzer

To create the datasets, we followed the time window synchronization technique of Walonoski and Heffernan [13]. A time window is defined as a dilation of time around a recorded observation time [13]. For example, given an observation made at time t and a time window of size 2 minutes, the user actions associated with that observation will include all actions made between t-1 minute before and t+1 minute after. In Walonoski and Heffernan's exploration of gaming behavior within the Assistments system the researchers used time windows 2, 2.5 and 3 minutes.

In an earlier study, Rodrigo et al. [12] also used time-window to synchronize human observations and interaction logs using 2-, 2.5- and 3-minute time windows to create the feature data sets for Aplusix data. They attempted to determine whether it was possible to meaningfully identify distinct groups of students based on interaction logs alone. Two categories of students were identified in their study: Cluster 0 and Cluster 1.

In this study, the researchers created the dataset by first adding a new column called *CurrentTime*. *CurrentTime* is the time when the student performs the current action; it is the basis for the computation of time windows. In obtaining the *CurrentTime* of an action, what the researchers did was to get the sum of the previous action *CurrentTime* and the current action *Duration*.

We synchronized the action logs with the affective state by matching the observation time and action current time. Since the affect might not be definitely observed at the exact observation time, for this study, the researchers extended the action *Current Time* to 1 minute after. From there all actions made 2, 2.5 and 3 minutes prior to the each observed affective state were collected and included in the dataset. The time windows were then labeled "Bored" if the observer noted boredom at this time and "Not bored" on all other affective states that are not "Bored" (e.g. flow, neutral, delight). As seen in Figure 3, "Bored" observations are noted distinctly in this paper.



1	order	studen	observ	School-Set-Stud	obs-emotion	seconds
21	20	10	Jess	Ateneo 1 10	FLO	400
22	21	1	Jess	Ateneo 1 1	BOR	420
23	22	2	Jess	Ateneo 1 2	FLO	440

Fig. 3. Sample observation of a student as "Bored" (in yellow highlight)

To determine the precursors of boredom, statistical computations of frequency analysis and mean were conducted. The features of interest are as listed:

1. Frequency of actions:
   1.1. Hit of a key
      1.1.1. Numbers (*0-9*)
      1.1.2. Letters (*A-Z*)
      1.1.3. Symbols (+, -, *, /, ^, =, <, >, (,), ., ,)
      1.1.4. *BackSpace*
      1.1.5. *delete*
   1.2. Special buttons or keys (<=, >=, <>, *{[£]/[?]} fraction* – selection to numerator, *{[?]/[£]} fraction* – selection to denominator, *(?)* – two parentheses, *{@^[2]}* – square, *{@^[?]}* – power, *and, or*)
   1.3. Arrows (*left, right, home, end, up, down*)
   1.4. *cut, copy, paste, drop*
   1.5. *undo, redo*
   1.6. *ask for verification, ask for solution, ask for score*
2. Average duration between actions (in seconds)

For each time window, frequency counts of different actions and average duration between actions for the bored and non-bored time windows were obtained from the datasets. Another Visual Basic application called Consolidated Results was used to separate the bored time windows from the non-bored time windows and export them as separate worksheet to Microsoft Excel. After the bored and non-bored data were separated, the researchers consolidated each feature within each time window.

The final task was to compute and look for significant results using t-test.

4. RESULTS

To be able to determine if there were significant differences between means of frequency counts and average durations of “Bored” observations and those of “Not bored” observations, the researchers used Two-tailed T-test with equal variances.

We found that the 3-minute time window resulted in the most number of significant differences (See Tables 1, 2, 3).

Table 1. Average frequency of actions and average duration between actions in 2-minute time window; statistically significant differences are shaded in dark gray and marginally significant differences in light gray.

Feature Set	Bored	Not bored
No. of hits of Number keys	6.2692	8.4306
No. of hits of Alphabetical keys	2.4615	3.0935
No. of hits of Symbol keys	4.4231	5.4706
No. of hits of *Backspace*	15.1795	18.7981
No. of hits of *delete*	0.2536	0.7671
No. of hits of Special buttons	0.5385	0.9679
No. of hits of Arrow keys	4.5256	6.3782
No. of *cut, copy, paste* and *drop*	0.0128	0.0166
No. of *undo* and *redo*	0.4487	0.2511
No. of *ask for verification*, *ask for solution* and *ask for score*	0.1410	0.0887
Average duration between actions	2.9428	3.5869

Table 2. Average frequency of actions and average duration between actions in 2.5-minute time window; statistically significant differences are shaded in dark gray and marginally significant differences in light gray.

Feature Set	Bored	Not bored
No. of hits of Number keys	8.0641	10.5310
No. of hits of Alphabetical keys	3.0385	3.8275
No. of hits of Symbol keys	5.5769	6.8190
No. of hits of *Backspace*	18.8205	23.5134
No. of hits of *delete*	0.2821	0.9434
No. of hits of Special buttons	0.6923	1.1806
No. of hits of Arrow keys	5.3462	7.9786
No. of *cut, copy, paste* and *drop*	0.0128	0.0224
No. of *undo* and *redo*	0.4615	0.3082
No. of *ask for verification*, *ask for solution* and *ask for score*	0.2308	0.1154
Average duration between actions	2.8255	3.5777

Table 3. Average frequency of actions and average duration between actions in 3-minute time window; statistically significant differences are shaded in dark gray.

Feature Set	Bored	Not bored
No. of hits of Number keys	9.6282	12.3082
No. of hits of Alphabetical keys	3.5769	4.5523
No. of hits of Symbol keys	6.5385	8.0321
No. of hits of *Backspace*	23.5513	27.9551
No. of hits of *delete*	0.3205	1.0983
No. of hits of Special buttons	0.7949	1.3745
No. of hits of Arrow keys	6.9231	9.2628
No. of *cut, copy, paste* and *drop*	0.0128	0.0283
No. of *undo* and *redo*	0.4872	0.3457
No. of *ask for verification*, *ask for solution* and *ask for score*	0.3077	0.1410
Average duration between actions	2.8684	3.4226

When analyzing the data using 2-minute time windows (Table 1), one statistically significant and one marginally significant difference emerged. The non-bored students used the Number keys statistically significant more than bored students (t(1948)=-2.22; p=0.03). Also, non-bored students hit Special buttons marginally significantly more than bored students (t(1948)=-1.85; p=0.06).

When analyzing the data using 2.5-minute time windows (Table 2), we found two statistically significant and one marginally significant difference emerged. The non-bored students used the Number keys statistically significant more than bored students (t(1948)=-2.18; p=0.03). Also, non-bored students hit Special buttons marginally significantly more than bored students (t(1948)=-1.93; p=0.05). Bored students used *ask for verification*, *ask for solution* and *ask for score* buttons statistically significant more than non-bored students (t(1948)=2.16; p=0.03).

When analyzing the data using 3-minute time windows (Table 3), three statistically significant differences emerged. The non-bored students used the Number keys and Special buttons statistically significant more than bored students (t(1948)=-2.1, p=0.04 and t(1948)=-2.11, p=0.03 respectively). On the contrary, Bored students used *ask for verification*, *ask for solution* and *ask for score* buttons statistically significant more than non-bored students (t(1948)=2.69; p=0.01).

Across three sizes of time-windows, the average number of hits of Number keys was consistently seen as having statistically significant difference between bored and not-bored students.

While the number of hits of Special buttons and the number of hits of *ask for verification*, *ask for solution*, and *ask for score* buttons appeared as having either marginal significant or statistically significant difference between the two affective states.

Since 3-minute time window resulted in the most number of statistically significant differences, the researchers used the findings in this time window size in identifying the probable precursors of boredom.

5. CONCLUSION

Based on the results presented here, the researchers found three precursors of boredom: (1) frequent usage of *ask for verification*, *ask for solution*, and *ask for score* buttons, (2) less hitting of Number keys and (3) less hitting of Special buttons. This finding suggests that the frequent usage of *ask for verification*, *ask for solution*, and *ask for score* buttons and infrequent hitting of Number keys and Special buttons will most probably lead students to the negative affective state of boredom soon while using Aplusix. The researchers deem the enumerated features might be indicative of boredom since the findings seem to be reasonable that students who are about to get bored tend to lack the interest of advancing in the learning environment with exerted effort.

These precursors of boredom will be helpful to the future development of intelligent tutoring systems. ITSs can add a feature that automatically predicts whether a student is about to experience boredom and it then can apply immediate, appropriate intervention or motivation.

6. ACKNOWLEDGMENTS

The researchers would like to thank the Department of Science and Technology for the grant entitled "Design of Affect-Sensitive Interfaces".

7. REFERENCES

[1] Fisher, C. D., "Boredom at work: A neglected concept," *Human Relations*, vol. 46, no. 3, p. 395-417, (1993).

[2] Lagud, M. C. V. and Rodrigo, M. M. T., "The Affective and Learning Profiles of Students when using an Intelligent Tutoring System for Algebra," *2010 Intelligent Tutoring Systems Conference* (2010).

[3] Rodrigo, M. M. T., Baker, R. S., and Nabos, J. Q., "The Relationships between Sequences of Affective States and Learner Achievement," 18th *International Conference for Computers in Education*, (2010).

[4] Craig, S. D., Graesser, A. C., Sullins, J., and Gholson, B., "Affect and Learning: an Exploratory Look into the Role of Affect in Learning with AutoTutor," *Journal of Educational Media*, vol. 29, no. 3, p. 241-250, (2004).

[5] Baker, R. S., Corbett, A. T., Koedinger, K. R., Wagner, A. Z., "Off-Task Behavior in the Cognitive Tutor Classroom: When Students 'Game the System'," *Proceedings of ACM CHI 2004: Computer-Human Interaction*, p. 383-390, (2004).

[6] Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., and Graesser, A. C., "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," *International Journal of Human-Computer Studies*, vol. 68, no. 4, p. 223-241, (2010).

[7] precursor. http://www.merriam-webster.com/dictionary/precursor, 2010.

[8] McQuiggan, S. W., Lee, S., and Lester, J. C., "Early Prediction of Student Frustration" *Proc. 2nd Intl. Conf. Affective Computing and Intelligent Interaction*, Vol. 4738, p. 698-709, (2007).

[9] Rodrigo, M. M. T. and Baker, R. S., "Coarse-Grained Detection of Student Frustration in an Introductory Programming Course," *International Computer Education Research Workshop ICER 2009*, (2009).

[10] Rebolledo-Mendez, G., "Modeling the motivational state of the learner in a Vygotskyan inspired ITS," *Proceedings of the 6th HCT postgraduate workshop, University of Sussex*, (2002).

[11] Lee, D. M. C., Ramos, C. C., Sy, C. W. C., Rodrigo, M. M. T., Baker, R. S., "Detecting Boredom and Confusion among Novice Programmers using BlueJ Compile Logs," *2010 Philippine Computing Society Congress, Ateneo de Davao*, (2010)

[12] Rodrigo, M. M. T., Anglo, E. A., Sugay, J. O., and Baker, R. S., "Use of Unsupervised Clustering to Characterize Learner Behaviors and Affective States while using an Intelligent Tutoring System," P*roceedings of International Conference on Computers in Education*, (2008).

[13] Walonoski, J. A., and Heffernan, N. T., "Detection and Analysis of Off-Task Gaming Behavior in Intelligent Tutoring Systems," *ITS 2006*, LNCS 4035, p. 382-391, (2006).

8. AUTHORS' INSTITUTIONAL AFFILIATION

[1]Department of Information Systems and Computer Science, Ateneo de Manila University, Loyola Heights, Quezon City, Philippines, +63 (2) 426-6071