Understanding Users' Interaction Behavior with an Intelligent Educational Game: Prime Climb

Alireza Davoodi[1], Samad Kardan[1], Cristina Conati[1]
[1] Computer Science Department, The University of British Columbia,
ICICS/CS Building, 201-2366 Main Mall, Vancouver, B.C., Canada V6T 1Z4,
{davoodi, skardan, conati}@cs.ubc.ca

Abstract. This paper presents work on applying clustering and association rule mining techniques to mine users' behavior in interacting with an intelligent educational game, Prime Climb. Through such behavior discovery, frequent patterns of interaction which characterize different groups of students with similar interaction styles are identified. The relation between the extracted patterns and the average domain knowledge of students in each group is investigated. The results show that the students with significantly higher prior knowledge about the domain behave differently from those with lower prior knowledge as they play the game and that pattern could be identified early during the interactions.

Keywords: Intelligent Educational Games, Behavior Discovery, Association Rule Mining, Open Ended Learning, Scaffolding

1 Introduction

Open-Ended Learning Environments (OELEs) support student-centered learning and allow learners to follow an exploratory interaction behavior to construct their own models of concepts and revise their beliefs subsequent to receiving immediate feedback on their actions [1]. Previous studies have shown that students could not benefit much from an open-ended learning environment if not receiving proper feedback [2]. Among learning environments, educational games are designed to foster motivation and engagement which are shown to be influential in learning [3]. To this end, educational games such as Crystal Island provide exploratory learning environments and encourage autonomous interaction with the game [4]. While such freedom in interaction is required to maintain engagement in the game, it also provides learners with the possibility of showing different interaction patterns. The interaction patterns might be indicative of certain characteristics and understanding such patterns can provide valuable information about the students.

Adaptive OELEs have been designed to answer the need for understanding and intelligently supporting varying learning styles, capabilities, and preferences in individuals in developing their skills. An adaptive educational system maintains a model of student's learning and leverages the student's interactions with the system to provide tailored scaffolding. Many educational systems apply data mining approaches on the logs of students' recorded interactions to extract behavioral patterns and extract high-level information about students [5-7]. Along this line of research, we concentrate on understanding how students interact with Prime Climb (PC), an adaptive educational game (edu-game) and whether there is a connection between behavior patterns of students and their attributes such as prior knowledge. The ultimate goal in an adaptive educational game such as PC is to help a higher number of students learn the desired skills through interacting with the game. Achieving such an objective requires a pedagogical agent which maintains an accurate understanding of individual differences among users and provides more tailored interventions, with the aim of guiding the learners in the right learning direction. For instance, if a pedagogical agent is capable of identifying a group of students with high domain knowledge, it is possible to leverage such information to construct a more accurate user model and intervention mechanism. The user's interaction behaviours can also be provided to developers to improve the design of educational systems [8].

Behavioral discovery has been vastly used in educational systems, but there is limited application in educational games such as Prime Climb, in which educational concepts are embedded and presented in the game scenarios and narratives with minimum explicit technical notation (for instance mathematical notations in PC) to more genuinely support game aspects of the system. In Prime Climb, students do not explicitly practice approaches to number factorization but implicitly follow a self-regulated learning approach [9] to explore and understand the methods and practice them. This paper describes the first step toward leveraging students' behavioral patterns into building a more effective adaptive edu-game. The ultimate goal is devising mechanisms for extracting abstract high-level patterns from raw interaction data and leveraging such understanding for real-time identification of interaction styles to enhance user modeling and intervention mechanism in an edu-game like PC.

Behavior discovery has been recently applied in different educational systems. Kardan et al. [6] leveraged behavior discovery to propose a general framework for distinguishing users' interaction styles in exploratory learning environments. Keshtkar et al. [10] describe an approach to distinguishing players and mentors roles in a multi-chat environment within the epistemic game Urban Science. In another related work, Mccuaig et al. [5] discuss using interaction behaviors to distinguish students who will fail or pass a course in a Learning Management System (LMS). A sequence mining approach has been also used in differentiating behavior patterns in students' interacting with Betty's Brain, a learning-by-teaching environment [7].

Although behavior discovery has been recently applied to many educational systems, there is very limited work on behavior mining in an open ended intelligent educational game like Prime Climb in which learning through playing the game is intended. Additionally, most of the previous works use the entire interaction data to make inferences about the users. In this work, we present the results of behavior mining not only on a big portion of interaction data but also on a truncated data set, which will provide the possibility of constructing an online classifier for early detection of varying patterns of interactions.

2 Prime Climb an Intelligent Edu-game

Prime Climb (PC) is an intelligent educational game for students in grades 5 and 6 to practice number factorization skills. Prime Climb is equipped with an intelligent pedagogical agent which maintains a probabilistic model of the student's knowledge on number factorization skills.

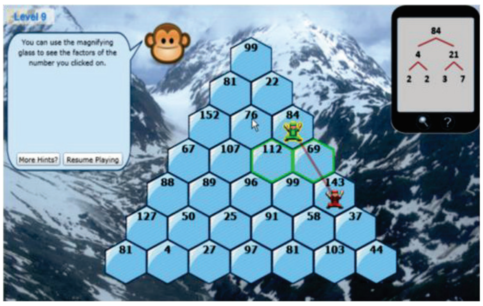

Fig. 1 Prime Climb

The pedagogical agent leverages the probabilistic model to provide an adaptive scaffolding mechanism. If model's assessment about the student's knowledge on a skill falls below a certain threshold, a hint is presented to the player. The hints are given in incremental level of details. In PC, the player and his/her partner climb a series of 11 mountains of numbers by pairing up the numbers which do not share a common factor. There are two main interactions of a player with PC:

***Making Movements*:** A player makes one or more movements at each time, by clicking on numbered hexagons on the mountains. PC provides immediate feedbacks on correctness of movements. If a player makes a wrong movement, s/he falls down.

***Using Magnifying Glass Tool*:** The magnifying glass (MG) tool is always available for the user to benefit from. The MG is used to show the factor tree of a number on the mountains; it is located in the top right corner of the game (Fig. 1).

4 Behavior Discovery in Prime Climb

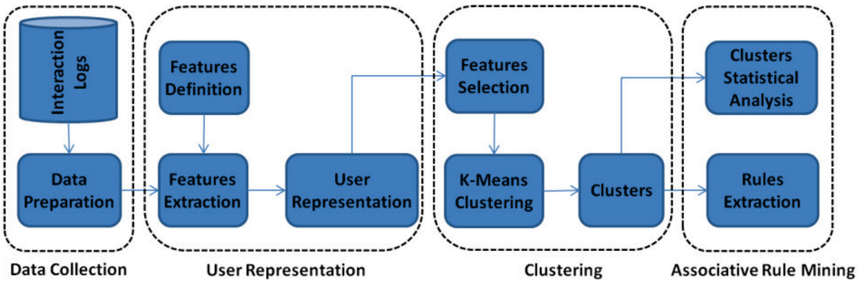

Fig. 2: Behavior discovery methodology in Prime Climb

4.1 Data Collection

Data collection is first component of the behavior discovery methodology in PC shown in Fig. 2. We collected interaction logs of 45 students who played PC voluntarily. Prime Climb consists of 11 levels (mountains), and not all students could manage to reach the last level. Out of the 45 students, 43 completed 9 or more levels. The remaining 2 students who completed fewer levels were excluded from further analyses to ensure that all students in analysis had completed a minimum of 9 levels. For the remaining 43 students, the interaction data for the first 9 mountains was used in the feature extraction process.

4.2 User Representation

Features Definition: Each user is represented by a vector of features. Based on the 2 main groups of interaction previously mentioned (movements and MG), two types of features are defined: (1) Movements features based on statistical measures on movements students made on the mountains and (2) MG features: based on statistical measures on students' usages of the MG tool. Table 1 shows some of these features:

Table 1 Some features used for behavior discovery

Movement Features
[Sum/Mean/STD] of number of [correct/wrong] movements made by a student across mountains
[Sum/Mean/STD] of time on [correct/wrong] movements made by a student across mountains
[Mean/STD] of length of sequences of [correct/wrong] moves made by a student
[Mean/STD] of time spent per sequence of [correct/wrong] moves made by a student
Magnifying Glass (MG) Features
[Sum/Mean/STD] of MG usage
Mean number of [correct/wrong] movements per each MG usage
STD of number of [correct/wrong] movements per each MG usage

Feature Set Definition: Each feature is a measure computed based on user's interactions with one or more mountains. There are two types of feature:
Mountain-Generic Features (m – n), (m >= 1 and n <= 9)**:** Calculated based on the users' interactions with mountains *m* to *n*, inclusively. For instance, the feature, correct-movements (1–9), represents the total number of correct movements made by the user on mountains 1 to 9.
Mountain-Specific Features (k), (1 <= k <= 9)**:** Calculated based on interactions with mountain *k*. For instance, correct-movements (7), represents the total number of correct movements made by the user on mountain 7.
In this paper, we present the behavior discovery results on the two feature sets:
Mountain-Generic Movement(1–9) Set**:** Contains mountain-generic features (1–9) which are related to movement actions the student makes.
Mountains-Generic+Specific-MG+Movement(1–2) Set**:** Contains mountain-generic MG features (1–2), mountain-generic movement features (1–2), mountain-specific MG features (1) and (2), and mountain-specific movements features(1) and (2).

4.3 Clustering

Feature Selection: Prior to performing clustering, feature selection is applied to filter out irrelevant features [11].
Clustering: The optimal number of clusters is determined as the lowest number suggested by C-index, Calinski and Harabasz[12] and Silhouette [13] measures of clustering validity. Once all the students are represented by vectors of selected features, the GA *K*-means (*K*-means for short) clustering algorithm [6], which is a modified version of GA *K*-means [14], is applied to cluster the users into an optimal number of clusters.

4.4 Rule Mining: Higher Prior Knowledge *vs.* Lower Prior Knowledge

Next, the Hotspot algorithm [15] is used to extract the rules for each discovered cluster. Also, we analyzed whether the resulting clusters are significantly different on a measure called *cluster's prior knowledge*, which is defined as follows:

Cluster's Prior knowledge: The cluster's prior knowledge gives the average level of factorization skills of the cluster's members prior to playing the game and is defined as the average of raw pre-test scores of the cluster's members. The following formula is used to calculate the cluster's prior knowledge:

$$\textbf{Cluster's prior knowledge} = \frac{\sum_{\textbf{student}\in\textbf{cluster}} \textbf{pre\_test}(\textbf{student})}{\textbf{Cluster's size}}. \quad (1)$$

where $\text{pre\_test}(\text{student})$ is the student's pre-test score. Before playing the game, a student takes a pre-test on number factorization skills. The maximum score a student can get is 15. The average pre-test score across the 43 students is 11.7, and the standard deviation is 3.29.

Behavior Discovery on Mountain-Generic-Movement(1–9) set: In this feature set, each student is represented by a vector of mountain-generic movement features(1–9). As a result of the features' selection mechanism, 18 features were selected out of the original 30 features. The optimal number of clusters was found to be 2, and the *K*-means method was used to cluster the set of students into 2 groups. The result of a *t*-test showed that there is a statistically significant difference between the prior knowledge of cluster 1 of students (higher prior knowledge (HPK) group) (M = 13.0 , SD = 2.0) and cluster 2 of students (lower prior knowledge (LPK) group) (M = 11.3 , SD = 3.45), p=.03 and Cohen's d= 0.53. Next, the Hotspot association rule mining algorithm was applied on the clusters to extract the associative rules. Table 2 shows the rules extracted for each cluster.

Understanding the Rule Mining Results:

Rules: Each bulleted item in following tables shows an extracted rule. For example, "Mean-Time-on-Movements=Higher" is an extracted rule which applies to at least 25% of the members of cluster 1. (In this study, the threshold of 25% is applied for all rules extracted by the Hotspot algorithm). This rule shows that the values for the feature "Mean-Time-On-Movements(1–9)" across the cluster's members belong to the "Higher" Bin.

Bins: In this study, the Hotspot algorithm considers two bins for values of each feature: (1) Lower bin and (2) Higher bin. Each bin shows a range of values of the features such that the lower bin represents the lower range of values and the higher bin represents the upper range of values for the feature. The cut-off point for splitting a range of values for a feature into two ranges (lower and upper) is calculated specifically for the feature in each extracted rule by the Hotspot algorithm. The lower and higher bins are indicated by the words "Lower" and "Higher" in front of the features in the following tables.

Rule's Support: The other important information is the rule's support shown in square brackets in front of the extracted rules in the following tables. For instance, [6/6=100%] in front of the first rule for the cluster 1 in Table 2 shows that there are in total 6 (in denominator) out of 43 students on which the extracted rule applies and all of these students belong to cluster 1 (6 in the numerator of the fraction). In addition, it can be concluded that this extracted rule applies to 60% (6/10) of the cluster 1 (note that the size of cluster 1 is 10).

Table 2: Extracted Rules for Mountains-Generic-Movement(1–9)

Rules for Cluster 1[HPK]: (Size: 10/43 = 23.26%)
• Mean-Time-on-Movements(1-9) = Higher, [6/6=100%] • Mean-Time-Spent-On-Correct-Movements-On-Mountains(1-9) = Higher, ([5/5=100%])
Rules for Cluster 2[LPK]: (Size: 33/43 = 76.74%)
• Mean-Time-On-Movements(1-9) = Lower, [33/37=89.19%] ○ STD-Time-On-Wrong-Correct-Moves(1-9) = Lower, [33/35=94.29%] • Mean-Time-On-Consecutive-Wrong-Movements(1-9) = Lower, [31/35=88.57%] ○ STD-Time-On-Movements(1-9) = Lower, [31/33=93.94%] ○ STD-Time-On-Correct-Movements(1-9) = Lower, [31/33=93.94]

Discussion and Interpretation: The extracted rules show that the students belonging to the HPK cluster (cluster 1) spent more time on movements and correct movements across 9 mountains. This could indicate that the students with higher prior knowledge were more involved in the game and spent more time before making a movement. Since the time spent on making a correct movement is higher for this group of students, it might mean that a correct move by this group of students is less likely to be due to a lucky guess as compared with the total population. In contrast, the group of students with lower prior knowledge spent less time on making movements as well as making wrong movements. This could be an indication of less involvement in the game by the lower prior knowledge group. It could show that a correct movement by this group of students is more likely due to a guess. In addition, there are some other frequent patterns of interaction for the group of students with lower prior knowledge. These patterns show a lower standard deviation on time spent on making movements and correct movements. This indicates that this group of students showed a consistent pattern of lack of engagement in the game. Therefore, we can conclude that the students with higher prior knowledge showed more engagement in the game than students with lower prior knowledge.

Behavior Discovery on Mountain-Generic+Specific-MG+Movement(1–2) set: This feature set does not employ interaction data from all 9 mountains; instead, only the data from the first 2 mountains is included. Such feature set is mainly valuable for constructing an online classifier to classify students based on their interaction with the game during the game play. The ultimate aim is leveraging such a feature set to step toward building a more accurate individualized student model and intervention mechanism as the student makes progress in the game. For instance, if the classifier can identify a student as a lower/higher knowledgeable student, it could leverage the information for early adjustment of the adaptive intervention mechanism. Similarly *K*-means was applied to cluster the students represented by the Mountains-Generic+Specific-MG+ Movements (1–2) Set. The optimal number of clusters was calculated to be 2, and 25 features out of 51 original ones were selected, as a result of applying the features selection mechanism. The cluster's *prior knowledge* was calculated for each of the discovered clusters and compared using a *t*-test. The result of the *t*-test showed a statistically significant difference between cluster *1's prior knowledge* ($M = 12.45$, $SD = 2.66$) and cluster 2's prior knowledge ($M = 9.22$, $SD = 3.93$), $p = 0.02$, Cohen's $d = 1.08$. Next, association rule mining was applied on the 2 clusters, as shown in Table 3.

***Interpretation and Discussion*:** As shown in Table 3, the results of behavioral discovery on the Mountains-Generic+Specific-MG+Movements(1-2) set is not consistent with the results of behavior discovery on the Mountains-Generic-Movements(1-9) set. Behavior discovery on interaction data from the first two mountains shows that students with higher prior knowledge (M = 12.45 , SD = 2.66) constitute around 79% of the all students and spend less time on making movements. It was previously shown in Table 2 that the students in the HPK cluster constituted approximately 23% of all students and spent more time on making movements when interaction data from all 9 mountains was included. Despite this, we expect that as the students progress in the game, the students with higher prior knowledge would behave differently from the other students and separate themselves from the others. To verify this, we also extracted frequent patterns when more interaction data from upper mountains is included in the clustering and rule mining. When the interaction data from the first 3 mountains is included in patterns mining, 2 clusters are identified which are not significantly different on their prior knowledge. When interaction data from the first four mountains is included, we observe patterns similar to those identified using the interaction data from all 9 mountains as shown in Table 3-right. The result of the t-test shows a statistically significant difference between cluster 1's prior knowledge (M = 13.28, SD = 1.58) and cluster 2's prior knowledge (M = 11.39 , SD = 3.4), p = 0.02, Cohen's d = 0.60. Also, approximately 16% of students belong to the HPK cluster, and 84% belong to the LPK group. This result is very similar to the results when data from all 9 mountains is included. Similar patterns are observed when more interaction data from upper mountains is included in the analysis.

Table 3: Extracted Rules for Mountains-Generic+Specific-MG+ Movements(1-2) [left] and MG+Movements(1-4) [right]

Rules for Cluster 1[HPK] (Size: 33/42=78.57%)	Rules for Cluster 1[HPK] (Size: 7/43=16.28%)
• Mean-Time-On-Movements(1)=Lower, [30/31 =96.77%] • Mean-Time-On-Movements(1-2) = Lower, [29/30 = 96.67%]	• Mean-Time-On-Movements(4) = Higher, [5/5 = 100%] • Mean-Time-On-Correct-Movements(3) = Higher, [3/3 = 100%]
Rules for Cluster 2[LPK] (Size: 9/42=21.43%)	**Rules for Cluster 2[LPK] (Size: 36/43=83.72%)**
• Mean-Time-Spent-On-Mountain(1-2) = Higher, [7/7=100%] • Total-Time-On-Mountain(1) = Higher, [5/5=100%]	• Mean-Time-On-Correct-Movements(1-4) = Lower, [35/35 = 100%] • Mean-Time-On-Movements(1-4)=Lower, [34/34 = 100%]

5 Conclusions and Next Steps

This paper discusses behavior discovery in Prime Climb (PC). To this end, different sets of features were defined. The features were extracted from interaction of students with PC in the form of making movements from one numbered hexagon to another numbered hexagon and usages of the MG tool. To identify frequent patterns of interaction, first, a feature selection mechanism was applied to select more relevant features from the set of all features. Then a K-means clustering was applied to cluster the students into an optimal number of clusters and the Hotspot algorithm of association rule mining was applied on the clusters to extract frequent interaction patterns. Finally, the prior knowledge of the clusters were compared. When interaction data from all 9 mountains was included in behavior discovery, it was found that the students with higher prior knowledge were more engaged in the game and spent more time on making movements. In contrast, the students with lower prior knowledge spent less time on making movements, indicating that they were less involved in the game. Behavior discovery also was conducted on truncated sets of features in which only a fraction of interaction data was included. The results showed that using the interaction data from the first two mountains resulted in groups of students that are statistically different on their prior knowledge.

The scaffolding mechanism in PC relies on the student model so we expect improvements in the model to result in more tailored interventions and guidance. Current PC uses the same student model for all students. Following the results of the presented study, we plan to adjust the model based on the characteristics of each discovered group of students. In addition, an online classifier will be built which identifies frequent patterns of interaction in the students, classifies them into different groups in real time, and leverages such information to build a more personalized user model and adaptive intervention mechanism in PC.

References

1. Papert, S. Mindstorms (2nd ed.), New York: Basic Books inc. (1993) _
2. L. Alfieri, P. Brooks, N. Aldrich, and H. Tenenbaum, "Does Discovery-Based Instruction Enhance Learning," Journal of Education Psychology, vol. 103, no. 1, pp. 1-18, (2011)
3. Lee, J., Luchini, K., Michael, B., Norris, C., Solloway, E., 2004, More than just fun and games: Assessing the value of educational video games in the classroom. Proceedings of ACM SIGCHI 2004, Vienna, Austria, pp. 1375–1378 (2004)
4. Sabourin, J., Rowe, J., Mott, B., & Lester, J. Exploring Affect and Inquiry in Open-Ended Game-based Learning Environments. (2011)
5. Mccuaig J., Baldwin, J.: Identifying Successful Learners from Interaction Behaviour. EDM 2012, Chania, Greece. pp. 160–163 (2012).
6. Kardan, S., and Conati. C.: A Framework for Capturing Distinguishing User Interaction Behaviours in Novel Interfaces. EDM, Eindhoven, the Netherlands, pp. 159–168 (2011)
7. Kinnebrew, J. S., Biswas G.: Identifying Learning Behaviors by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution, EDM 2012, Chania, Greece. pp. 57–64 (2012).
8. Hunt, E., Madhyastha, T.: 2005: Data Mining Patterns of Thought. In: Proceedings of the AAAI Workshop on Educational Data Mining. (2005).
9. Paris, S., Paris, A.: Classroom Applications of Research on Self-Regulated Learning. Educ. Psychol. 36(2), 89–101 (2001)
10. Keshtkar, F., Morgan, B., Graesser, A.: Automated Detection of Mentors and Players in an Educational Game. EDM 2012, Chania, Greece. pp. 212–213 (2012)
11. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182 (2003).
12. Milligan, G.W., Cooper, M.C.: An Examination of Procedures for Determining the Number of Clusters in a Data Set. Psychometrika, 50(2), 159–179 (1985).
13. Rousseeuw, P. J. :Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". *Computational and Applied Mathematics* 20: 53–65. doi:10.1016/0377-0427(87)90125-7, (1987)
14. Kim, K. Ahn, H.: A Recommender System using GA *K*-means Clustering in an Online Shopping Market. Expert Syst. Appl. 34(2), 1200–1209 (2008).
15. Hall, M., Eibi, F., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA Data mining Software: An Update. SIGKDD Explor. 11(1), 10–18 (2009)