Helping Students Manage Personalized Learning Scenarios

Paul Salvador Inventado*, Roberto Legaspi and Masayuki Numao
The Institute of Scientific and Industrial Research, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, Japan, Osaka, 567-0047
{inventado,roberto}@ai.sanken.osaka-u.ac.jp, numao@sanken.osaka-u.ac.jp

ABSTRACT

In personalized learning scenarios, students have control over their learning goals and how they want to learn which is advantageous since they tend to be more motivated and immersed in what they are learning. However, they need to regulate their motivation, affect and activities so they can learn effectively. Our research deals with helping students identify the long-term effects of their learning behavior and identify effective actions that span across learning episodes which are not easily identified without in depth analysis. In this paper, we discuss how we are trying to identify such effective learning behavior and how they can be used to generate feedback that will help students learn in personalized learning scenarios.

Keywords

personalized learning, self-regulated learning, reinforcement learning, user modeling

1. INTRODUCTION

Governments and educational institutions have called for reforms on how students are taught in school to enable them to have more control over their learning [1]. Allowing students to engage in personalized learning grant them skills that prepare them for the needs of the current society and more importantly help shape them into life-long learners.

In personalized learning, students have control over what they learn and how they learn causing them to be more motivated and immersed in what they are learning. Teachers no longer serve as the main sources of information but instead become facilitators of the students' learning process. Although teachers can guide students and give them suggestions about what they are learning, teachers can only assess and provide support for a small number of the challenges that students face. Especially because students learn in situations where teachers are unavailable, students can easily get overwhelmed by challenges and not achieve their aspired learning goals. It is also possible that students would engage in non-learning related activities which might hinder them from learning. Thus, in this kind of learning scenario, self-regulation is essential for students to manage their goals, time, motivation, affective states and hindrances to learning.

Self-regulation is not an easy task because it requires much motivation and effort [5]. There is a high cognitive load when students perform learning tasks while managing it. They would need to continuously monitor the effects of their actions and decide if they should continue doing it or if they should change it. Furthermore, students also keep track of effective learning behavior so they can use them in future learning episodes.

We have been developing a software that helps students monitor their behavior and reflect on what transpired during the learning episode with the help of webcam and desktop snapshots [3]. After each learning episode, students who used the system were asked to review their learning episode then annotate their intentions, their activities and their affective states so they could further understand and analyze their behavior. According to the results, students who used the system discovered behaviors they were initially unaware of and were able to identify ways to improve ineffective learning behavior. We were also able to analyze and process the students' annotated data to have a better understanding of their learning behavior.

Students' reflections from the experiment however, seemed to focus only on immediate effects of their actions and did not consider its long term effects in the learning episode. Also, their reflections did not incorporate their realizations from previous learning episodes. Currently, we are investigating how we can help students identify actions that benefit learning not only in the short-term but also in the long-term. We also want to help students to identify effective learning behavior that span over different learning episodes.

2. STUDENT LEARNING BEHAVIOR

The data we used for this research was gathered from four students engaging in research-related work, which is an example of a personalized learning scenario. One male masteral student and one female doctoral student created a re-

*also affiliated with: Center for Empathic Human-Computer Interactions, College of Computer Studies, De La Salle University, Manila, Philippines

port about their research involving activities such as information search, reading papers, reading books and creating a power point presentation. One male undergraduate student and one female doctoral student wrote a conference paper about their research involving activities such as information search, reading papers, reading books, running programs and simulations to retrieve data from their experiments and paper writing. We gathered two hours of data for five different learning episodes from each student within a span of one week.

Unlike other research, our work dealt with students who freely decided on the time, location and type of activities they did including non-learning related activities. However, they were required to learn in front of a computer running the software we developed for recording and annotating learning behavior.

Although the students worked on different topics and used different applications, all of them processed and performed experiments on previously collected data, searched for related literature and created a report or document about it. Analyzing the data showed that students performed six types of actions – information search (e.g., using a search engine), view information source (e.g., reading a book, viewing a website), write notes, seek help from peers (e.g., talking to a friend), knowledge application (e.g., paper writing, presentation creation, data processing) and off-task (e.g., playing a game).

3. BEHAVIOR EFFECTIVENESS

In a learning episode, students perform many different actions to achieve their goal. Although students can identify the effectiveness of the current action by monitoring its effect, it is more difficult to identify how it will affect or how it has affected their learning in the long run. For example, students spending a long time learning about a topic would seem to be performing well, however they may experience more stress and have a higher chance of making mistakes and getting confused more easily. It would probably be advantageous for the student to also take a rest once in a while. We adapted the concept of *returns* in reinforcement learning [4] to account for this situation wherein the effectiveness of an action was not measured only by its immediate effects but rather its long term effects on the learning episode. Moreover, as the student engaged in more learning episodes, a reinforcement learning algorithm updated the rewards of each action which incorporated the effects of actions from previous learning episodes.

Due to the lack of control in the students' activities while learning, it was not possible to directly gauge the students' learning progress which could have been used to define the rewards of their actions. However, their affective states gave an idea about the events that transpired during the learning episode. D'Mello and Graesser's model of affective dynamics [2] describes the relationship between affective states and events that occur in a learning scenario. For example, confusion indicates instances wherein students need to exert more effort to progress in the current activity. Frustration arises when students are too confused, get stuck and no longer progress in their learning. Too much frustration results in boredom or disengagement from the learning activity. Engagement and delight on the other hand are indicators that a student is moving towards or has achieved the learning goal. Although D'Mello and Graesser's model does not discuss off-task activities in particular, it is logical to consider that they will not directly lead to learning progress. Negative affective states experienced while performing off-task activities might cause a decrease in motivation so it is probably best to avoid them while learning. Based on how each affective state and type of activity affected learning progress, we constructed a reward system (see Table 1) that would be used to update the returns of performing an action.

Table 1: Action-Affect Reward System

Affect	On-task Behavior	Off-task Behavior
Engaged	3	-
Confused	2	-
Frustrated	1	-
Bored	-1	-
Neutral	0	-2
Delighted	3	-2
Surprised	2	-2
Sad	-	-3
Angry	-	-3
Disgusted	-	-3
Afraid	-	-3

Using the reward system we defined, the long-term effectiveness of the actions performed in the learning episode can be discovered using a reinforcement algorithm. Specifically we used Q-learning [4] to discover actions that maximize return when performed in a particular state. In our case, we represented states using the learning context and actions using the activities performed by the student. Specifically, each state was represented using – the current affective state, the amount of time spent in the current state, the previous action performed, the previous affective state experienced, the dominant action previously used and the dominant affective state experienced. States changed when students chose to perform a different activity (e.g., shfting from viewing an information source to seeking help) so this was used to represent an action.

The collected data was manually processed and then converted into state-action pairs. Q-learning was then applied to uncover the returns of performing actions in a particular state. The state-action pairs with their corresponding expected returns were called the student's learning policy.

4. RESULTS

The Q-learning algorithm was applied on each of the student's data separately since we assumed that each student would have a different learning policy. Due to the number of features we used for state representation , there were a lot of states and many of them had high return values. Due to space limitations, we only present some of the notable state-action pairs from one of the student's learning policy in Table 2. Majority of the states with high return values contained state-action pairs that represented transitions inherent to the domain. For example, high returns were given when students applied knowledge after viewing an information source, which happens naturally for example when a

Table 2: Sample State-Action Returns

State	Action	Reward
Engaged while viewing an information source for <5min, Previously engaged while applying knowledge, Mostly felt engaged while applying knowledge	Apply knowledge	6469.20
Confused while applying knowledge for <5min, Previously engaged while viewing information source, Mostly felt engaged while applying knowledge	Apply knowledge	982.80
Engaged while applying knowledge for 5-10min, Previously off-task, Mostly felt engaged while applying knowledge	Off-task	164.30
Neutral while applying knowledge for <5min, Previously engaged while applying knowledge, Mostly felt engaged while applying knowledge	Off-task	-228.60
Delighted while doing off-task behavior for 5-10min, Previously confused while viewing an information source, Mostly felt engaged while viewing an information source	View information source	521.10

student shifts between reading information sources and creates a power point presentation. However, some interesting strategies were discovered such as shifting from an engaged on-task activity to an off-task activity indicating that off task activities may actually have positive long term effects.

Students' answers from surveys and personal interviews regarding their thoughts on a recent learning episode correlated with the reward values produced by the algorithm. For example, students identified the need to continue learning despite encountering challenges (e.g., confusion) and not spending too much time in off-task activities.

5. FUTURE DIRECTION

The next step in the research is helping students find ways to improve their learning behavior. We believe that the behavior identified using the reinforcement learning approach can be used to support students by making them aware of the behavior's long term effects and also informing them of effective learning behavior that have spanned across their learning episodes.

Students' behavior in a learning episode can be evaluated by comparing the actions that a student took in a particular state with the optimal action according to the student's updated learning policy. When a student selects a suboptimal action, the system can inform the student that an ineffective learning strategy might have been used and taking the optimal action could improve their learning effectiveness. Effective learning behavior that span across learning episodes can be identified by keeping track of frequently used state-action pairs that constantly garner high returns in different learning episodes. Students can be informed of such behavior so they will be aware of them and can make sure to apply them in succeeding learning episodes.

We also plan to investigate how students will react to feedback using the policy generated by the reinforcement learning algorithm and observe if it will help students select more effective learning behavior. It will also be interesting to see how differently students will react to feedback when different reward systems are used. Apart from using a students' learning policy we also think that they can benefit from learning about other students' effective learning behaviors taken from other students' learning policies. Moreover, learning behaviors identified by experts which are not exhibited by the student can also be suggested.

Another way to identify more accurate reward values would be to include effectiveness ratings of the actions performed by the students. It might also be good to explore other features for our state representations and see how they affect the resulting learning policy. Lastly, we are also investigating other reward mechanisms that are more flexible so it can handle students' individual differences.

Acknowledgements

This work was supported in part by the Management Expenses Grants for National Universities Corporations from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) and JSPS KAKENHI Grant Number 23300059. We would also like to thank all the students who participated in our data collection.

6. REFERENCES

[1] D. E. Atkins, J. Bennett, J. S. Brown, A. Chopra, C. Dede, B. Fishman, L. Gomez, M. Honey, Y. Kafai, M. Luftglass, R. Pea, J. Pellegrino, D. Rose, C. Thille, and B. Williams. *Transforming American education: Learning powered by technology.* Nov. 2010.

[2] S. D'Mello and A. Graesser. Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2):145–157, Apr. 2012.

[3] P. S. Inventado, R. Legaspi, R. Cabredo, and M. Numao. Student learning behavior in an unsupervised learning environment. In *Proceedings of the 20th International Conference on Computers in Education*, pages 730–737, Dec. 2012.

[4] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning).* The MIT Press, Mar. 1998.

[5] B. J. Zimmerman. Self-regulated learning and academic achievement: An overview. *Educational psychologist*, 25(1), 1990.