

Student 1

3 : (strong accept)

Paper [20]: Filtered Time Series Analyses of Student Problem-Solving Behaviors in Game-based Learning

Summary

In this paper, the authors present a general framework for analyzing students' problem-solving behaviors in game-based learning environments by first filtering them into a gameplay trajectory using a dimensionality reduction filter followed by further analysis with an expert solution and learning metrics. This framework was tested with data from 68 college students in a game-based learning environment for microbiology students, Crystal Island. Using the temporal representation, the authors calculated both the slope and the distance (in comparison to an expert solution) of the problem solving trajectories. Their results show that both of these metrics are correlated with the normalized learning gain and problem solving performance of the students.

Main strengths

- The authors present a framework that targets time-series analysis of student behaviors in a relatively open-ended environment which I think is a much needed framework because: 1) student behaviors in open-ended environments are challenging to analyse and 2) time-series analysis is usually more challenging but much more insightful for interventions purposes in real-time compared to summary statistics.
- They calculated both baseline and temporal distances and the comparison later on highlights the advantages of using a temporal analysis that can incorporate additional information regarding the problem-solving behavior path.
- The authors are thorough in measuring the distance metrics. For example, since the distance measurement is sensitive to misalignment in time, in order to account for similar segments of student trajectories out of place within Crystal Island, the distance over each gameplay phase is calculated.
- I liked the thoroughness of the related work and how the authors compared their work with each set of related work elaborating how it's different or similar to the previous work.
- The discussion on the various outcomes are quite thorough which give the readers richer insights into the current results and limitations and what can be done to mitigate or investigate them further.

Main weaknesses/points of discussion

- Can there only be one expert path? Wouldn't it be more robust to capture trajectories of various experts to generate an expert path? This has been later mentioned in the discussion section though so I assume the authors plan to investigate it in the future work.
- Since the sequences are of unequal length, why not use a distance method more suited to unequal length sequences such as dtw distance instead of lock-step euclidean distance with padding?
- In this work, it was shown that the problem-solving trajectory distance to an expert solution are

related to learning and in-game student engagement; however, I am not sure if scaffolding the students to be as close to an expert solution in the first trial would necessarily lead to the highest learning gains because it could be that failing more or guess-and-check behavior initially can be beneficial to remove misconceptions in the long-run such as shown by the work on Productive Failure. An alternative could be to try to converge to expert paths over a series of trials rather than in the initial trial.

Connection to the bigger picture

In this paper, the authors present a framework that targets time-series analysis of student behaviors in a relatively open-ended environment which I think is a much needed and a relevant framework because: 1) student behaviors in open-ended environments are challenging to analyse and 2) time-series analysis is usually more challenging but more insightful for interventions purposes in real-time compared to summary statistics. Hence, this paper would be a valuable contribution to the EDM community.

Student 2

3: (strong accept)

Main contributions

- * Analyzing student behavior in a game-based learning environment using time series.
- * Showing that there is a relationship between game trajectories and learning gains. This could indicate that there are trajectories that lead to higher learning gains and that by using hints, it could be possible to guide students in the “expert” direction to maximize their learning gains in the game.

Strengths

- + Addressing an open-ended problem.
- + Using trajectories instead of transition matrix and using the full gameplay sequence rather than fragments
- + Participants demographics mentioned
- + Assessing learning gains with pre-test and post-test scores.

Weakness

- It would have been interesting to see how the results from 68 students in a laboratory study compare to the results of a larger population and in different environments.
- On the one hand, the motivation for reducing dimensionality is that a single value reduces the noise in distance measurements. On the other hand, this "noise reduction" also brings "information reduction" because the single point is obtained with the first component that only explains 37% of the variance of the data. It would be interesting to see how the results varied when using other dimensionality reduction methods different from PCA (Example: T-SNE) or by constructing multidimensional trajectories with 2 components or more.
- As mentioned in the paper, there are multiple solution paths and probably different types of “gamers” or playing/learning strategies. I wonder if there would be any value in clustering the trajectories to find different playing strategies and customize the scaffolding depending on the cluster’s specific needs.
- As mentioned by the authors, as time increases there is more heteroskedasticity in the trajectories. The increased variance decreases the confidence of the results.
- If students who followed a predetermined path achieved significantly higher normalized learning gains than students who had the freedom of control, I wonder if the students should just be given hints to follow the expert’s path. However, could it be possible that there are some learning gains that are not being measured like creativity or curiosity?

Connection to the bigger picture

- * Understanding student's behavior in game-based learning environments allows a personalized gaming experience that helps students learn more and be more engaged. The time-series (trajectories) analysis shown in the paper can be extended to other open-ended problems with multiple solutions to gain insight into student's behavior.

Student 3

2: (accept)

Summary

The authors present a framework to examine the problem solving behaviors of students in game based learning environments through a filtered time series analysis. They track cumulative student actions through gameplay and use principal component analysis to turn these vectors into scalar values. The trajectory of the student then becomes a list of these values measured periodically as they play the game. The paper shows that the slope of this trajectory is negatively correlated with learning as measured by normalized learning gain. It is also shown that the similarity with an expert's trajectory is correlated with learning as measured by the same metrics. By comparing to just using final summaries of gameplay sessions, the authors show that the temporal nature of trajectories contain valuable information about a student's learning progress.

Strengths

Being able to quantify and predict a user's learning as they interact with a game based environment can be useful in designing new learning environments and helping users get the most out of their experiences. The authors present a useful framework that can be applied in such environments to extract trajectories for students and show several ways of inferring knowledge from these trajectories. It is promising that the trajectories have a relationship with both the normalized learning gain (as measured by pre and post-tests) and the final game score (as measured by game-specific criteria).

The framework is potentially applicable to many learning environments, as the only information that needs to be collected is the sequences of actions. An expert trajectory is harder to obtain in some cases but still usually feasible.

The inclusion of a baseline metric that only uses the final summary information of students is a thoughtful addition and successfully shows the value of using similarity measures that take into account full gameplay trajectories of users.

The authors correctly point out a potential problem that could arise from segment misalignments and propose a matching procedure to avoid it.

Weaknesses

Some of the results reported in the paper are not statistically significant. They are either reported as marginally significant or approaching significance, but this takes away some of the impact of the results. Perhaps a replication study with more students could remove this doubt about the results.

Reducing whole action counts into a single scalar by just taking a single principal component might be sacrificing too much information that could have been useful. The reported variance explained of the chosen component is rather low.

Padding the trajectories with the final value might distort results, although a good effort to keep segments aligned is made by the authors.

The framework's reliance on an expert trajectory should be investigated more deeply. While it is not necessarily wrong, it would be useful to have more information on the expert. Especially in open-ended environments, there could be many different ways of optimizing learning and there might be even better indicators of success than similarity with this single expert's trajectory. It would be interesting to have multiple expert trajectories, and if a learning environment is designed with a scaffolding that guides students toward expert trajectories, one should be careful not to bias the environment to favor certain students.

Overall Evaluation

While there are weaknesses to be addressed, the presented framework seems to be useful and potentially applicable in many environments. After more research and investigation, it can hopefully be used to improve learning environment designs and help us understand student learning processes better. I vote for acceptance.

Student 4

1: (weak accept)

Overview:

This paper presents a general framework for analyzing students' problem-solving behavior in game-based learning environments. To do so, the authors analyze the temporal sequence of student problem-solving behaviors in comparison to expert solution paths, while encoding the student's actions into time-series trajectories. The authors evaluated their method by comparing student trajectories with expert ones and also by measuring the learning and engagement in the game-based learning environment. The paper concludes that there is a significant promise for assessing the temporal nature of student problem solving during game-based learning.

Major Strengths:

- The problem of adaptive student guidance to achieve effective problem-solving student behavior is a pragmatic one, and this paper addresses a part of this issue. By incorporating the sequential nature of student actions within the educational game, it provides a supporting tool to design the learning guidance and creating an effective game-based learning environment.
- The proposed methodology paves the way for a high interpretable and understandable use of student data in the open-world learning environments. One of the strongest points of this approach is the fact that it can be generalized to any game-based open-world learning environment by formulating the user input as structured student trajectories. Furthermore, the designer of the system can leverage the results of both the exploratory and the correlation analysis, without having extensive domain expertise.
- Authors present a powerful tool that can be used for the design of adaptive learning environments. In the era of online inference and reinforcement techniques in ITSs, a statistical and light-weighted technique for analyzing the data and draw meaningful conclusions is essential.

Major Weaknesses:

- The main concern regarding the proposed framework is around the assessment of the slope as an important insight, as well as its correlation with the student's overall learning gain. As authors explain in their work, the slope can be viewed as the pace of problem-solving actions, and since some actions can contribute to a different amount based on their weight, the slope depends on the number and the type of actions a student did. This approach is highly dependent on the design of the actions and the overall flexibility of the user inside the gameplay scenario. Thus, the slope of the trajectory can be misleading or unuseful in the wrong weight assignment or bad game design.
- In continuation with the previous point, as the authors also confirm in their work, there is a need for multiple optimal paths (expert paths) since there is no one path for all students' different learning behaviors and paces. Thus, this approach does not provide personalized features for individual student needs. Overall, it is not clear how the slope and trajectory experiments can reflect the actual learning, and it is not clear if the provided results are the product of correlation or actual causation. There needs to be a further investigation of the characteristics of the participated students.

- There should be further analysis regarding the scalability of this methodology with more complicated paths and a larger amount of possible actions. An interesting approach would be to focus on the specific actions and behaviors that are associated with high performance, rather than just the overall trajectory.

Student 5

1: (weak accept)
CS-702: Week 9

Review of Sawyer et al.

Summary

In this paper, Sawyer et al. examine the dynamics of student problem-solving behaviour within a game-based learning environment. Specifically, they build on findings and data from previous studies related to the microbiology education game *Crystal Island*. They first present a framework for filtering problem-solving behaviours by using principal component analysis to conduct a dimensionality reduction from students' action sequences to a time series. They then test their framework using data generated from a lab study comprising 63 students (42 female), for which students had interacted with *Crystal Island* and conducted pre-tests and post-tests to assess their learning gains. The authors analyse the slope of the students' trajectories as well as their distances from an example of expert gameplay and calculate the correlation of these metrics with normalised learning gains. Their analysis shows that (i) the slope of the trajectory was negatively correlated with learning gains, signifying that a slower pace indicated positive learning outcomes; (ii) the distance from the expert trajectory was negatively correlated with learning gains, signifying that trajectories closer to the expert path were indicative of higher learning gains; (iii) the temporal distance used, which was the average Euclidean distance over time, performed better than using a baseline of just the distance at the end of the sequence; (iv) the magnitudes of the correlations observed between with slope and expert solution distance is similar to the correlation observed between a hand-crafted score and learning gains, providing a baseline for the utility of the automatically-generated measures. Nevertheless, it is important to note that the significance for most of these correlations are marginal and the effect size is not particularly high.

Strengths

This paper has several strengths. First, the motivation for the study is clear and the study is timely, as the use of game-based learning scenarios becomes more widespread. I was particularly encouraged to see a game catered to university-level students, as I have seen fewer games used for tertiary education. Second, the authors provide a well-structured background section, with each paragraph describing a facet of their approach and ending with how related work has been incorporated and/or how their work differs from past approaches. Third, the authors do a really good job building on previous findings. This is evident from how they coalesce the outcomes of several past papers related to *Crystal Island*, including the expert gameplay, the student engagement indicator, the correlation between information gathering and improved problem-solving efficiency, and the result of higher learning gains for students following predetermined paths. Finally, the use of real data, with external indicators such as pre-tests and post-tests, to validate their approach bolsters the applicability of their findings.

Weaknesses

There are several weaknesses in this paper that are worth underlining. First, there is a minor gender imbalance in their data, which poses the question of the effects of demographics. It would've been

interesting to see if analysing subsets of their data led to interesting findings. Second, their method of projecting the cumulative count vector onto one dimension is not thoroughly motivated, in my opinion. Given that it is one of the core contributions of their work, I would have appreciated a more intuitive explanation of why this approach is the best fit for this process. Third, the fact that they pad shorter sequences to facilitate the distance comparison between trajectories seems a bit arbitrary to me and complicates the intuition behind using the temporal distance. Although they seem to suggest that this can be mitigated by using the trajectory distance per interval, I could not see whether this was applied to their main analysis. Fourth, I'm not sure that their method of calculating distance is well motivated. Although they do provide a comparison between calculating the average distance over time, they compare it to just calculating the final time-step. Is this a useful comparison? It seems arbitrary to provide a baseline of just the last distance between trajectories. Fifth, some parts of the paper seem to introduce assumptions that are not properly backed. For example, they mention that the "slow place of play... reflects the... deliberate and efficient on-task problem-solving path", which in turn "demonstrates positive problem-solving strategies, such as reading texts thoroughly and planning the next action". These statements without conditionals are possibly misleading. Finally, the p values are somewhat weak and therefore the impact of the approach should be taken with a pinch of salt.

Connection to the Bigger Picture

The fact that the authors are analysing game-based learning at the university level is particularly encouraging, as this type of gamified approaches to digital education has proven useful. Seeing an immersive game dedicated to a complex subject such as microbiology is quite exciting. I believe that motivating the use of games in the classroom with a good understanding of how students interact with them and how this correlates with learning gains is especially helpful in broadening the use of gamified software in educational contexts.