Process Mining of Interactions During Computer-Based Testing for Detecting and Modelling Guessing Behavior

Zacharoula Papamitsiou(✉) and Anastasios A. Economides

IPPS in Information Systems, University of Macedonia, Thessaloniki, Greece
papamits@uom.edu.gr, economid@uom.gr

Abstract. Detecting, recognizing and modelling patterns of observed examinee behaviors during assessment is a topic of great interest for the educational research community. In this paper we investigate the perspectives of *process-centric inference of guessing behavior patterns*. The underlying idea is to extract knowledge from real processes (i.e., not assumed nor truncated), logged automatically by the assessment environment. We applied a three-step process mining methodology on logged interaction traces from a case study with 259 undergraduate university students. The analysis revealed sequences of interactions in which low goal-orientation students answered quickly and correctly on difficult items, without reviewing them, while they submitted wrong answers on easier items. We assumed that this implies guessing behavior. From the conformance checking and performance analysis we found that the fitness of our process model is almost 85 %. Hence, initial results are encouraging towards modelling guessing behavior. Potential implications and future work plans are also discussed.

Keywords: Assessment analytics · Educational data mining · Guessing behavior · Pattern recognition · Process mining · Student interaction analysis

1 Introduction

The rise of educational data mining and learning analytics promises to deconstruct the deeper learning processes into more simple, distinct mechanisms, for understanding and supporting human learning accordingly [1]. This is envisaged to be achieved by tracking *every type of interactions* within any type of information system supporting learning or education (formal, informal, ubiquitous, mobile, virtual or real-world), converting them into explorable educational datasets, subjecting them into mining, and coding them into interpretable and useful schemas. *Interaction analysis* covers a number of methods for empirically exploring the space of humans' activities with each other and with objects in their environment via the use of artefacts and technology, for identification of practices and problems, and the origins for their solution [2]. In the educational context, technology-mediated learning supports learners' ability to interact with other learners, tutors, content interfaces, features and digital environments, and provides a great opportunity for recording, filtering and processing logged interaction trace data regarding systems' usage and *user activity indicators* [3].

P. Zaphiris and A. Ioannou (Eds.): LCT 2016, LNCS 9753, pp. 437–449, 2016.
DOI: 10.1007/978-3-319-39483-1_40

Recently, researchers attempt to extract indicators of students' behavior from multiple and diverse logged data sources. The exploration of the underlying relationships between these indicators and learning outcomes, its forthcoming analysis and the endeavor to identify patterns and model students' performance based on their actual interactions, has attracted increased attention (e.g. [4–6]). In these cases, advanced data mining and machine learning techniques have been utilized, beyond traditional statistical analysis methods, in order to investigate the abovementioned relationships.

1.1 Problem Statement - Motivation of the Research and Research Questions

As apparent, *recognizing patterns of students' behaviors during assessment* – which is closely related to measuring performance – is crucial for the research community. When it comes to computer-based testing procedures – which is a typical, popular and widespread method of online assessment – one of the *observed unwanted examinee behaviors* that needs to be detected and appropriately managed and that critically affects the assessment result (e.g., score) is *guessing* of the correct answer on testing items.

The prevalent methods for modelling guessing behavior include Item Response Theory (IRT)-based techniques and Bayesian Knowledge Tracing (BKT), both of which adopt a probabilistic approach for hypothesizing and investigating students' behavior either within testing environments [7, 8] or within Intelligent Tutoring Systems [9]. In these approaches, researchers defined thresholds for discriminating non-effortful guessing responses from solution behavior upon test speededness (i.e., amounts of times to answer the question) [10, 11], explored different combinations of IRT-parameters (e.g. difficulty-based guessing models based on test-taking motivation, the corresponding effort expenditure, the correctness of the answer and the estimated examinee ability) [12], contextualized the estimation of the probability that a student has guessed or slipped [13], and enhanced previous results with skill difficulty driven by the estimation of knowledge acquisition during each step of the problem solution procedure [14].

However, the abovementioned methodologies (a) follow an *outcome-centric probabilistic consideration*, in terms of employing the student's ability/performance estimation, and (b) *do not "dive" into the causation and origins* of the occurring interactions.

In order to overcome these shortcomings, the novelty of the present approach resides in the following facts: (a) we investigate the perspectives of *process-centric (rather than outcome-centric) inference* of guessing patterns, (b) we explore *full-fledged process models with concurrency patterns,* unlike most of the traditional educational data mining techniques which focus on data or simple sequential structures [15], and (c) we associate *student's goal-orientation* to exhibiting guessing behavior during assessment. Thus, the research question (*RQ*) is:

> *"Can we discover behavioral patterns (sequence/repetition/alternation/frequency/duration) within event logs that can be associated with guessing during testing?"*

The underlying idea in the proposed approach is to employ process mining in order to extract knowledge from event logs tracked automatically by the testing environment [16, 17]. In particular, we suggest a three-step process mining methodology on logged trace data: (a) an initial control flow perspective, (b) next, the identification of sequences of events, (c) and finally, the classification of these sequences of events based on *students' time-spent* on correctly and wrongly answered questions, *student's goal-orientation* and *questions' difficulty*, and their mapping to respective behavior schemas.

In this paper we present the results from a study that we conducted in order to explore the capabilities of the proposed methodology on recognizing meaningful patterns that imply guessing behavior. 259 undergraduate students from a Greek University participated in an assessment procedure designed for the study. We employed the LAERS assessment environment [6] to collect the data (i.e., track students' interactions logs) during testing. For the mining purposes we used the ProM process mining tool [18] –a generic open-source framework for implementing process mining tools in a standard environment. The analysis revealed patterns of interactions in which low goal-orientation students frequently answered quickly and correctly on difficult items, without reviewing or altering them, while they submitted wrong answers on easier items. We classified this as guessing behavior. In order to measure the model's ability to re-reproduce all execution sequences that are in the log, we performed conformance checking and performance analysis. The fitness of our process model was almost 85 %. In essence, we suggest that process mining of temporal traces, taking into consideration each student's goal-orientation can be used for modelling guessing behavior during testing.

The rest of the paper is organized as follows: in Sect. 2, we provide an overview of process mining applied in the educational domain. In Sect. 3, we present the experiment methodology, the data collection procedure and the analysis methods that we applied, while in Sect. 4, we analyze the results from the case study. Finally, in Sect. 5, we discuss on the major findings, possible implications and future work plans.

2 Process Mining: An Overview

Process mining is a relatively new technology which emerged from the business community, and at the same time, a field of research situated at the intersection of data mining and business process management. The main objective of this technology is to allow for process-related knowledge extraction from event logs, automatically recorded by Information Systems [16]. The target is "to discover, monitor and improve real processes" [19, p. 34]. In other words, the purpose of process mining is to identify, confirm or extend process models based on actual data.

The core component of all process mining tasks is an *event log*. An event log is a set of finite *event sequences*, whereas each event sequence corresponds to a particular *process instance* (i.e., a *case*) of an activity, and can have a timestamp and an actuator executing or initiating the activity [19]. The sequence of events executed for a case is called a *trace*. Thus, within an event log, multiple cases may have the same trace.

The most prominent process mining technique is process model discovery (i.e., structures that model behavior), which includes the complete process model production from event-based data, without using any a-priori information. The constructed process model reflects the behavior observed in the original log and is able to reproduce it.

In the educational domain, typical examples of event logs may include learner's activity logs in e-learning environments (e.g. learning management systems, intelligent tutoring systems, etc.), use of pedagogical/educational resources, examination traces, participation and engagement in collaborative activities, etc. Examples of cases where process mining has been the central methodology, include the discovery of processes followed by the learners in different contexts, such as in self-regulated learning [20], in collaborative learning [21, 22], in collaborative writing [23, 24], in multiple-choice questions tests [25], and the discovery of learning habits based on MOOC data [26].

More precisely, in [25], process model discovery and analysis techniques (such as Petri nets, Heuristic and Fuzzy miner) were used to analyze assessment data (e.g. correctness of the answer, certitude, grade, time-spent for answering the question, etc.) from online multiple choice tests and to investigate the students' behavior during online examinations. In the collaborative learning context, the authors explored regulatory processes [21], and analyzed collaborative writing processes and how these correlate to the quality of the produced documents [23, 24]. In addition, the analysis of behavioral learner data (i.e., related to modeling and prototyping activities during a group project and the respective scores) with process mining techniques – targeting a complex problem solving process – shed light on the cognitive aspects of problem-solving behavior of novices in the area of domain modeling, specifically regarding process-oriented feedback [27]. Yet, in [26] the objective was to provide insights regarding students and their learning behavior (watching videos in a recommended sequence) as it relates to their performance. Finally, in the context of enhancing self-regulated learning [20], the authors analyzed the temporal order of spontaneous individual regulation activities.

In these examples from the educational domain, the prevailing process model discovery techniques were control-flow mining algorithms, which allow the discovery of educational processes and learning paths based on the dependency relations that can be inferred from event logs. The results of mining educational datasets with process mining provided useful insight regarding the improvement of understanding of the underlying educational processes, allowing for early detection of anomalies [23, 24]. These results were used for generating recommendations and advice to students [26], to provide feedback [27] to either students, teachers or/and researchers, to help students with specific learning disabilities, to improve management of learning objects [20], etc.

In our approach, we applied a three-step process mining methodology during a testing procedure, and explored its capabilities on recognizing meaningful patterns of guessing behavior during examination. We elaborate on this methodology in Sect. 3.

3 Methodology

3.1 Research Participants and Data Collection

In this study, data were collected from a total of 259 undergraduate students (108 males [41.7 %] and 151 females [58.3 %], aged 20-27 years old (M = 22.6, SD = 1.933, N = 259) from the Department of Economics at University of Macedonia, Thessaloniki, Greece. 12 groups of 20 to 25 students attended the midterm exams of the Computers II course (related to introduction to databases, information systems and e-commerce). For the purposes of the examination, we used 34 multiple choice quiz items. Each item had two to four possible answers, but only one was the correct. Finally, the participation to the midterm exams procedure was optional. As an external motivation to increase the students' effort, we set that their score would participate up to 30 % to their final grade.

In our study, we used the LAERS assessment environment [6], which is a Computer-Based Assessment system that we are developing. At the first phase of its implementation, we configured a testing unit and a tracker that logs the students' interaction data. The testing unit displays the multiple choice quiz items delivered to students separately and one-by-one. Within the duration of the test, the students can temporarily save their answers on the items, before submitting the quiz, can skip or re-view them and/or alter their initial choice by selecting the item to re-view from the list underneath. They submit the quiz answers only once, whenever they estimate that they are ready to do so.

The second component of the system records the students' interaction data during testing. In a log file we tracked students' time-spent on handling the testing items, distinguishing it between the time on correctly and wrongly answered items. In the same log file, we also logged the times the students reviewed each item and the times they changed their answers, and the respective time-spent during these interactions. In a separate file we also calculated the effort expenditure on each item and estimated the item's difficulty level [28]. Finally, we embedded into the system a pre-test questionnaire in order to measure each student's goal expectancy (GE) (a measure of student goal-orientation and perception of preparation [29]) in a separate log file. The final collected dataset includes the features illustrated in Table 1.

Table 1. Features from the raw log files

1. The student's ID	*9. The total time the student spends on viewing the tasks and submitting the correct answers*
2. The task the student works on (question Id)	*10. The total time the student spends on viewing the tasks and submitting the wrong answers*
3. The answer the student submits (answer Id)	*11. How many times the student views each task*
4. The correctness of the submitted answer	*12. How many times the student changes the submitted answer for each task*
5. The timestamp the student starts viewing a task	*13. The student's effort required*
6. The timestamp the student chooses to leave the task (saves an answer)	*14. The item's difficulty*
7. The idle time the student spends viewing each task	*15. The student's goal-expectancy (GE)*
8. The student's total time on task	*16. The student's actual performance (final score).*

In this study we applied a three-step process mining methodology on logged trace data and explored its capabilities on recognizing meaningful patterns of interactions that imply guessing behavior: (a) initially we adopted a control flow perspective (Petri-Nets [30]), (b) next, we identified sequences of events (*traces*), and finally, (c) we classified these sequences of events based on the *students' time-spent* on correctly and wrongly answered questions, the *student's goal-orientation* and the *question's difficulty*, and mapped them to respective behavior schemas.

3.2 Data Pre-processing and Construction of the Petri Net

Data pre-processing allows the transformation of original data into a suitable shape to be used by process mining algorithms. During this process, and within the dataset, we identified abstract behaviors of students regarding the testing items (i.e. students' actions), which we then coded into *tasks*. In our study we define as *task* T = {View(v), Answer Correctly(ac), Answer Wrongly(aw), Review(r), Change to Correct(chc), Change to Wrong(chw)}, the simplest learner's action. In addition, and since students' time-spent on each task is a continuous variable, that is difficult to subject into mining, we classified the students' temporal behavior in 4 clusters by applying the k-means algorithm (with k = 4). We experimented and executed the k-means algorithm for a number of iterations with different values of k (k = 3, k = 4, k = 5, k = 10). We computed the sum of squared error (SSE) for these values of k and plotted k against the SSE. According to the "Elbow" method [31], we finally selected k = 4. For simplicity reasons we call *Cluster* C = {medium-slow(ms), quick(q), medium-quick(mq), slow (s)}. Table 2 shows a sample of the consolidated event log with each row representing one event.

Table 2. Features after the data pre-processing

CaseId	Resource	timestamp	questid	answid	spenttime	Cor/ness	task	Cluster
170	*16610*	*13:36:38*	*59*	*231*	*54.487*	*0*	*aw*	*ms*
172	*30314*	*13:51:07*	*60*	*235*	*24.759*	*0*	*aw*	*mq*
174	*30314*	*12:33:30*	*61*	*239*	*65.909*	*0*	*aw*	*ms*
175	*22514*	*13:36:44*	*56*	*219*	*10.498*	*1*	*chc*	*q*
177	*16610*	*13:37:12*	*60*	*232*	*31.133*	*1*	*ac*	*ms*
179	*22514*	*13:36:59*	*57*	*221*	*11.391*	*1*	*r*	*q*
180	*22514*	*13:37:26*	*58*	*225*	*27.091*	*1*	*r*	*mq*
181	*18805*	*13:37:41*	*58*	*undef*	*32.591*	2	*v*	*ms*
182	*25711*	*13:37:44*	*56*	*219*	*30.332*	*1*	*ac*	*ms*
184	*16610*	*13:37:40*	*61*	*239*	*26.254*	*0*	*aw*	*mq*
185	*22514*	*13:37:51*	*59*	*231*	*24.562*	*0*	*chw*	*mq*
187	*18805*	*13:39:55*	*59*	*undef*	*110.063*	2	*v*	*s*
188	*30314*	*12:34:03*	*61*	*239*	*5.167*	*0*	*r*	*q*

Our analysis of the logged data explores the temporal behavior of students. Hence, the final clustered tasks considered in this study included 24 *events classes* E = {quick view(qv), quick review(qr), quick correct(qc), quick wrong(qw),..}.

Next, we performed a dotted chart analysis in order to gain some insight in the underlying processes and the respective performance. Figure 1 illustrates the results of this analysis. All the instances (one per student) are sorted by the duration of the computer-based assessment.

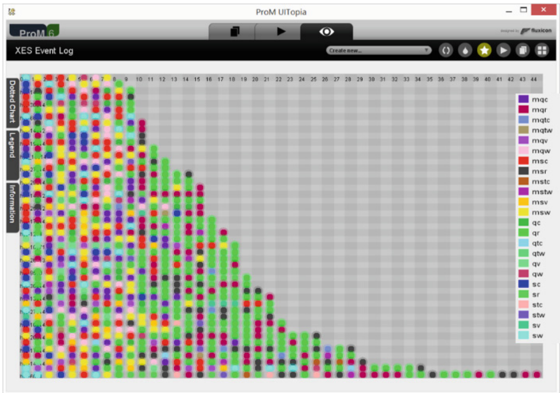

Fig. 1. Dotted Chart Analysis (Color figure online)

The basic idea of the dotted chart is to plot a dot for each event in the log according to the time. Thus, it enables visually examining the complete set of data in an event log and highlighting possible interesting patterns within the log. The dotted chart has *three orthogonal dimensions*: time and component types. The time is measured along the horizontal axis of the chart. The component types (e.g., originator, task, event type, etc.) are shown along the vertical axis. Note that the first component considered is shown along the vertical axis, in boxes, while the second component of the event is given by the color of the dot. Let us also note that in a dotted chart, common patterns among different cases are not clearly visible.

For detecting common patterns between the behavioral "traces" (response strategies) of the students during testing, in our case study, we mined the event log for Petri Nets using Integer Linear Programming (ILP). The ILP Miner is known for the fact that it always returns a Petri Net that perfectly fits a given event log [32]. Figure 2 illustrates the generated Petri Net which describes the generic pattern of answering questions, allowing for answer-reviews and changes. In this figure, the states (i.e., the events) and the transitions between them, including sequences, braches and loops between events are summarized for the whole sample, modeling the testing behavior of the participants. Every question can be answered correctly or wrongly and the student can spent a lot or less of time on answering the question. Further, a question can be viewed or reviewed and the student may change the submitted answer. The latter decision is modeled by an internal transition (painted in black) that goes to the final place of the net.

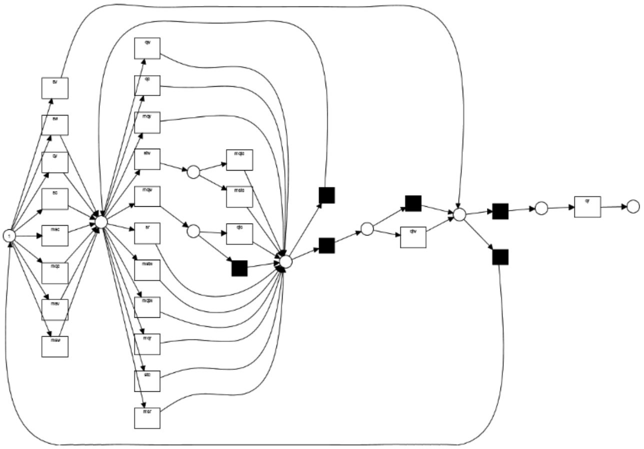

Fig. 2. The Petri Net that models the handling of testing items in our study

3.3 Identification of Traces - Conformance Checking and Performance Analysis

A process that specifies which event need to be executed and in what order is a workflow process model. In our study, we identified 47 sequences of events, i.e. *traces*. We define every unique sequence of events as *trace* $TR_i = \{TR_1 = \{qv,qw,qr,mqv, mqc, \ldots\}, TR_2 = \{sv,qw,msv,mqc,qr,msr,\ldots\}, \ldots\}$. Figure 3 shows all the paths detected within the event log, corresponding to the solution strategies the students follow during testing. All the 47 traces are illustrated in this figure. The numbers on the arrows indicate how many cases follow the specific trace. Sequences of events, branches and loops are also illustrated in this figure.

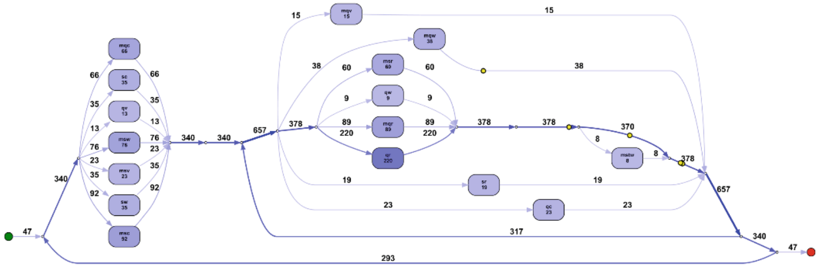

Fig. 3. Paths and traces detected in the event log of student's interactions with the testing-items

Before performing the conformance checking and performance analysis, we enhanced the process mining technique with trace alignment. In fact, trace alignment prepares the event logs in a way that can be explored easily and it complements existing process mining techniques focusing on discovery and conformance checking.

Trace alignment allows for similarity detection between traces, inducing interesting patterns of testing-item manipulation by the students during assessment. Given the great heterogeneity in traces, only few of the produced clusters delivered a good trace alignment.

Next, we performed conformance checking and performance analysis This analysis may be used to detect deviations, to locate and explain these deviations, and to measure their severity. We found that the fitness of our process model (i.e. whether the log traces comply with the description in the model) is almost 85 % (40 out of the 47 traces were re-produced correctly). This is particularly useful for finding out whether (or how often) the students exhibit guessing behavior. Then, we classified these sequences of events based on the students' time-spent on correctly and wrongly answered questions, their GE and the question's difficulty, and mapped them to respective behavior schemas.

4 Results

4.1 Recognition of "Guessing Behavior" Pattern

Figure 4(a) and (b) are samples of the traces followed by students who answered correctly and wrongly to the most of the questions respectively.

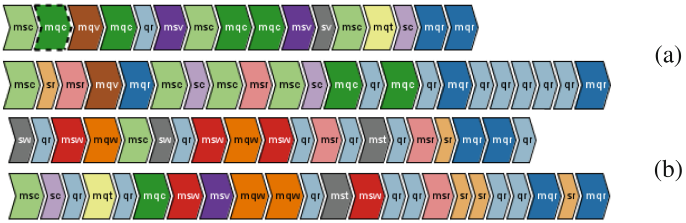
msc mqc mqv mqc qr msv msc mqc mqc msv sv msc mqt sc mqr mqr
msc sr msr mqv mqr msc sc msc msr msc sc mqc qr mqc qr mqr qr qr qr qr qr mqr
(a)
sw qr msw mqw msc sw qr msw mqw msw qr msr qr mst qr msr sr mqr mqr qr
msc sc qr mqt qr mqc msw msv mqw mqw qr mst msw qr qr msr sr sr qr qr mqr sr mqr
(b)

Fig. 4. Traces of (a) high achieving students, (b) low achieving students

As seen from Fig. 4(a), high achieving students answer correctly on the items, review the items and in those that they initially submitted a wrong answer, they revise it, think about it and submit a new, correct answer. Similarly, from Fig. 4(b) one can tell that low achieving students answer wrongly on the questions, they will not revise them and will not change their answers. Note that, high achieving students also denote high goal-expectancy (GE). In [6] it was found that there is a positive effect of GE on their time to answer correctly, while there is a negative effect of GE on time to answer wrongly, indicating that poorly prepared students will spend less time on questions.

However, the major category of the students are those who achieve an intermediate score. In this case, two major behaviors have been identified: those who will try their best, but not answer all items correctly and may have slipped some answers, and those who may have guessed some of the answers. The traces of these two categories are illustrated in Fig. 5.

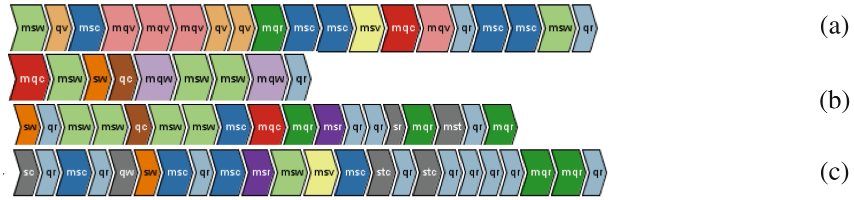
msw qv msc mqv mqv mqv qv qv mqr msc msc msv mqc mqv qr msc msc msw qr
(a)
mqc msw sw qc mqw msw msw mqw qr
sw qr msw msw qc msw msw msc mqc mqr msr qr qr sr mqr mst qr mqr
(b)
sc qr msc qr qw sw msc qr msc msr msw msv msc stc qr stc qr qr qr qr mqr mqr qr
(c)

Fig. 5. Traces of students exhibiting (a) solution, (b) guessing and (c) slipping behavior

As seen from Fig. 5 (a), the students view and review the items and try to “solve” the questions and submit the correct answers. They spent a considerable amount of time on dealing with the question and in some case, they might change their answers. On the contrary, Fig. 5 (b) corresponds to a trace that implies guessing behaviour. That is because in both traces, the students have answered fast and correctly on an item that has been found to be a difficult one, while they have submitted false answers on less difficult items. Furthermore, in both cases, the students do not revise the “suspicious” item. In an analogous way, in Fig. 5 (c), the students have slipped an easy item, by submitting fast a wrong answer, while answering correctly on the most difficult items.

5 Discussion and Conclusions

The issue of detecting and appropriately managing the observed examinee guessing behavior during testing is a central topic for the educational research community. In general, guessing behavior is expressed as rapidly occurring random responses, implying either that the students did not exhibit effort exertion or that they did not fully consider the testing item. Previous methods from related work, follow a probabilistic consideration on the identification of guessing behavior, that is outcome-centric and do not “dive” into the causation of the interactions that take place. In order to overcome these shortcomings, the novelty of the present approach resides in the following facts: (a) we investigate the perspectives of process-centric (rather than outcome-centric) inference of guessing patterns, (b) we explore full-fledged process models with concurrency patterns, and (c) we associate student’s goal-expectancy to exhibiting guessing behavior during assessment. In essence, the core research question of this study concerned the discovery of behavioral patterns (sequence/repetition/alternation/frequency/duration) within event logs that can be associated with guessing during testing.

In the suggested approach, the underlying idea was to extract knowledge from event logs (i.e., real processes) tracked automatically by the testing environment. Hence, in order to address the research question, we applied a three-step process mining methodology on logged trace data. In our approach, we initially employed Petri Nets from a control flow perspective, next we identified sequences of tasks (traces), and finally, we classified these traces based on the students’ time-spent on correctly and wrongly answered questions, their goal-expectancy and the question’s difficulty, and mapped them to respective behavior schemas. We conducted a study with 259 undergraduate university students who participated in an assessment procedure appropriately designed. We employed the LAERS assessment environment to track examinees' interactions logs during testing, and the ProM process mining tool for mining the logs. We discovered 47 behavioral traces (patterns) in total. The analysis revealed patterns of interactions in which low goal-orientation students frequently answered quickly and correctly on difficult items, without reviewing or altering them, while they submitted wrong answers on easier items (Fig. 5). We classified this as guessing behavior. This is partially in agreement with previous research results [8, 10], according to which the response time of guesses is usually very short compared to the amount of time required for the items.

In order to measure the model's ability to re-reproduce all execution sequences that are in the log, we performed conformance checking and performance analysis. The fitness of our process model was almost 85 %, with 40 out of the 47 traces to conform to the description in the model, and be correctly re-produced. Initial results are encouraging, indicating that process mining of temporal traces, taking into consideration each student' s goal expectancy can provide reliable modelling of guessing behavior.

However, it is important to note that an event log contains only example behavior, i.e., we cannot assume that all possible traces have been observed. In fact, an event log often contains only a fraction of the possible behavior [16]. Moreover, and in agreement with [20], although the proposed methodology may be useful for gaining insight into the students' interactions with the learning and assessment items, however, one identified disadvantage of process mining and descriptive modelling is that they are not directly suitable for statistical testing (e.g., significance testing).

According to [33], guessed answers increase the variance error of test scores and lower the test reliability. The accurate modelling of guessing behavior could lead to using the frequency of this behavior as an indicator of students' disengagement with the test. Identification of guessing behavior patterns could also assist in assessing the quality of multiple-choice items, re-designing the testing items, and change those that have caused guessing behavior too frequently. Furthermore, and since process mining is a promising methodology for behavioral pattern recognition within educational logged data, one possible research direction would be to explore the optimum size of the test (number of items) as well as the position of the items within the test, and associate these with fatigue and lack of focus, that could hinder guessing behavior. In [25] the authors employed process mining on assessment data and found that 35 % percent of the students answered the first question right and had high confidence. It would be interesting to measure the correct answers on this item if it was delivered as the last item of the assessment process and considering the students' confidence [34].

In the educational context, the application of process mining to learner's interaction trace data can become valuable assets for discovering, monitoring and improving real processes by extracting knowledge from learning-oriented event logs. We believe that analysis of behavioral learner data with process mining can add value in addition to the currently available learning analytics tools and techniques.

References

1. Papamitsiou, Z., Economides, A.A.: Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence. Educ. Technol. Soc. **17**(4), 49–64 (2014)
2. Jordan, B., Henderson, A.: Interaction analysis: foundations and practice. J. Learn. Sci. **4**(1), 39–103 (1995)
3. Donnelly, R.: Interaction analysis in a "Learning by Doing" problem-based professional development context. Comput. Educ. **55**, 1357–1366 (2010)
4. Huang, S., Fang, N.: Predicting student academic performance in an engineering dynamics course: a comparison of four types of predictive mathematical models. Comput. Educ. **61**, 133–145 (2013)
5. Macfadyen, L.P., Dawson, S.: Mining LMS data to develop an "early warning system" for educators: a proof of concept. Comput. Educ. **54**(2), 588–599 (2010)
6. Papamitsiou, Z., Terzis, V., Economides, A.A.: Temporal learning analytics during computer based testing. In: 4th International Conference on Learning Analytics and Knowledge, pp. 31−35. ACM, New York (2014)
7. Lee, Y.-H., Jia, Y.: Using response time to investigate students' test-taking behaviors in a NAEP computer-based study. Large-scale Assessments Educ. **2**(1), 1–24 (2014)
8. Kong, X.J., Wise, S.L., Bhola, D.S.: Setting the response time threshold parameter to differentiate solution behavior from rapid-guessing behaviour. Educ. Psychol. Meas. **67**(4), 606–619 (2007)
9. Baker, R.S., Corbett, A.T., Gowda, S.M., Wagner, A.Z., MacLaren, B.A., Kauffman, L.R., Mitchell, A.P., Giguere, S.: Contextual slip and prediction of student performance after use of an intelligent tutor. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 52–63. Springer, Heidelberg (2010)
10. Schnipke, D.L., Scrams, D.J.: Modeling item response times with a two-state mixture model: a new method of measuring speededness. J. Educ. Meas. **34**, 213–232 (1997)
11. Wise, S.L., Kong, X.: Response time effort: a new measure of examinee motivation in computer-based tests. Appl. Meas. Educ. **16**, 163–183 (2005)
12. Cao, J., Stokes, S.L.: Bayesian IRT guessing models for partial guessing behaviors. Psychometrika **73**(2), 209–230 (2008)
13. Baker, R.S.J.d., Corbett, A.T., Aleven, V.: Improving contextual models of guessing and slipping with a truncated training set. In: Baker, R.S.J.d., Barnes, T., Beck, J.E. (eds.) 1st International Conference on Educational Data Mining, pp. 67−76 (2008)
14. Gowda, S.M., Rowe, J.P., Baker, R.S.J.d., Chi, M., Koedinger, K.R.: Improving models of slipping, guessing and moment-by-moment learning with estimates of skill difficulty. In: Pechenizkiy, M., Calders, T., Conati, C., Ventura, S., Romero, C., and Stamper, J. (eds.) 4th International Conference on Educational Data Mining, pp. 199−208 (2011)
15. Trčka, N., Pechenizkiy, M., van der Aalst, W.P.M.: Process mining from educational data (Chapter 9). In: Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S.J.d. (eds.) Handbook of Educational Data Mining. pp. 123−142. CRC Press, London (2010)
16. van der Aalst, W.P.M.: Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer, Berlin (2011)
17. van der Aalst, W.P.M.: Process mining: making knowledge discovery process centric. SIGKDD Explor. Newsl. **13**(2), 45–49 (2012)
18. van Dongen, B.F., de Medeiros, A.K.A., Verbeek, H., Weijters, A., van der Aalst, W.M.: The ProM framework: a new era in process mining tool support. In: Ciardo, G., Darondeau, P. (eds.) ICATPN 2005. LNCS, vol. 3536, pp. 444–454. Springer, Heidelberg (2005)

19. van der Aalst, W.M.: Process mining in the large: a tutorial. In: Zimányi, E. (ed.) eBISS 2013. LNBIP, vol. 172, pp. 33–76. Springer, Heidelberg (2014)
20. Bannert, M., Reimann, P., Sonnenberg, C.: Process mining techniques for analysing patterns and strategies in students' self-regulated learning. Metacogn. Learn. **9**(2), 161–185 (2014)
21. Schoor, C., Bannert, M.: Exploring regulatory processes during a CSCL task using process mining. Comput. Hum. Behav. **28**(4), 1321–1331 (2012)
22. Bergenthum, R., Desel, J., Harrer, A., Mauser, S.: Modeling and mining of learnflows. In: Jensen, K., Donatelli, S., Kleijn, J. (eds.) ToPNoC V. LNCS, vol. 6900, pp. 22–50. Springer, Heidelberg (2012)
23. Southavilay, V., Yacef, K., Calvo, R.A.: Process mining to support students' collaborative writing. In: Baker, R.S.J.d., Merceron, A., Pavlik, P.I.Jr. (eds.) 3rd International Conference on Educatinal Data Mining, pp. 257−266 (2010)
24. Calvo, R.A., O'Rourke, S.T., Jones, J., Yacef, K., Reimann, P.: Collaborative writing support tools on the cloud. IEEE Trans. Learn. Technol. **4**(1), 88–97 (2011)
25. Pechenizkiy, M., Trčka, N., Vasilyeva, E., van der Aalst, W.M.P., de Bra, P.: Process mining online assessment data. In: Barnes, T., Desmarais, M.C., Romero, C. and Ventura, S. (eds.) 2nd International Conference on Educational Data Mining, pp. 279–288 (2009)
26. Mukala, P., Buijs, J.C.A.M., Leemans, M., van der Aalst, W.M.P.: Learning analytics on Coursera event data: a process mining approach. In: Ceravolo, P., Rinderle-Ma, S. (eds.) 5th SIMPDA, pp. 18−32 (2015)
27. Sedrakyan, G., De Weerdt, J., Snoeck, M.: Process-mining enabled feedback: "tell me what I did wrong" vs. "tell me how to do it right". Comput. Hum. Behav. **57**, 352–376 (2016)
28. Papamitsiou, Z., Economides, A.A.: A temporal estimation of students' on-task mental effort and its effect on students' performance during computer based testing. In: 18th International Conference on Interactive Collaborative Learning, pp. 1136–1144. IEEE (2015)
29. Terzis, V., Economides, A.A.: The acceptance and use of computer based assessment. Comput. Educ. **56**(4), 1032–1044 (2011)
30. van der Aalst, W.M.P.: The application of petri nets to workflow management. J. Circuit. Syst. Comput. **8**(1), 21–66 (1998)
31. Sugar, C.A., James, G.M.: Finding the number of clusters in a dataset: an information-theoretic approach. J. Am. Stat. Assoc. **98**, 750–763 (2003)
32. van der Werf, J.M.E.M., Dongen, B.F., van Hurkens, C.A.J., Serebrenik, A.: Process discovery using integer linear programming. Fundam. Inform. **94**(3–4), 387–412 (2009)
33. Bereby-Meyer, Y., Meyer, J., Flascher, O.M.: Prospect theory analysis of guessing in multiple choice tests. J. Behav. Decis. Making **15**(4), 313–327 (2002)
34. Lamboudis, D., Economides, A.A.: Adaptive exploration of user knowledge in computer based testing. WSEAS Trans. Commun. **3**(1), 322–327 (2004)