Process Mining Online Assessment Data

Mykola Pechenizkiy, Nikola Trčka, Ekaterina Vasilyeva, Wil van der Aalst, Paul De Bra
{m.pechenizkiy, e.vasilyeva, n.trcka, w.m.p.v.d.aalst}@tue.nl, debra@win.tue.nl
Department of Computer Science, Eindhoven University of Technology, the Netherlands

Abstract. Traditional data mining techniques have been extensively applied to find interesting patterns, build descriptive and predictive models from large volumes of data accumulated through the use of different information systems. The results of data mining can be used for getting a better understanding of the underlying educational processes, for generating recommendations and advice to students, for improving management of learning objects, etc. However, most of the traditional data mining techniques focus on data dependencies or simple patterns and do not provide a visual representation of the complete educational (assessment) process ready to be analyzed. To allow for these types of analysis (in which the process plays the central role), a new line of data-mining research, called *process mining*, has been initiated. Process mining focuses on the development of a set of intelligent tools and techniques aimed at extracting process-related knowledge from event logs recorded by an information system. In this paper we demonstrate the applicability of process mining, and the ProM framework in particular, to educational data mining context. We analyze assessment data from recently organized online multiple choice tests and demonstrate the use of process discovery, conformance checking and performance analysis techniques.

1 Introduction

Online assessment becomes an important component of modern education. It is used not only in e-learning, but also within blended learning, as part of the learning process. Online assessment is utilized both for self-evaluation and for "real" exams as it tends to complement or in some cases even replace traditional methods for evaluating the performance of students.

Intelligent analysis of assessment data assists in achieving a better understanding of student performance, the quality of the test and individual questions, etc. Besides, there are still a number of open issues related to authoring and organization of different assessment procedures. In Multiple-Choice Questions (MCQ) testing it might be important to consider how students are supposed to navigate from one question to another, i.e. should the students be able to go back and forward and also change their answers (if they like) before they commit the whole test, or should the order be fixed so that students have to answer the questions one after another? Is it not necessarily a trivial question since either of two options may allow or disallow the use of certain pedagogical strategies. Especially in the context of personalized adaptive assessment it is not immediately clear whether an implied strict order of navigation results in certain advantages or inconveniences for the students. In general, the navigation of students in e-Learning systems has been actively studied in recent years. Here, researchers try to discover individual navigational styles of the students in order to reduce cognitive load of the students, to improve usability and learning efficiency of e-Learning systems and support personalization of navigation [2]. Some recent empirical studies demonstrated the

feasibility and benefits of feedback personalization during online assessment, i.e. the type of immediately presented feedback and the way of its presentation may significantly influence the general performance of the students [9][10]. However, some students may prefer to have less personalization and more flexibility of navigation if there is such a trade-off. Overall, there seem to be no "best" approach applicable for every situation and educators need to decide whether current practices are effective.

Traditional data mining techniques including classification, association analysis and clustering have been successfully applied to different types of educational data [4], also including assessment data, e.g. from intelligent tutoring systems or learning management systems (LMS) [3]. Data mining can help to identify group of (cor)related questions, subgroups (e.g. subsets of students performing similarly of a subset of questions), emerging patterns (e.g. discovering a set of patterns describing how the performance in a test of one group of students, i.e. following a particular study program, differs from the performance of another group), estimate the predictive or discriminative power of questions in the test, etc. However, most of the traditional data mining techniques do not focus on the process perspective and therefore do not tell much about the assessment process as a whole. *Process mining on the contrary focuses on the development of a set of intelligent tools and techniques aimed at extracting process-related knowledge from event logs recorded by an information system.*

In this paper we briefly introduce process mining [7] and our ProM tool [8] for the EDM community and demonstrate the use of a few ProM plug-ins for the analysis of assessment data coming from two recent studies. In one of the studies the students had to answer to the tests' questions in a strict order and had a possibility to request immediate feedback (knowledge of correct response and elaborated feedback) after each question. During the second tests student had a possibility to answer the questions in a flexible order, to revisit and earlier answers and revise them as well.

The remainder of the paper is organized as follows. In Section 2 we explain the basic process mining concepts and present the ProM framework. In Section 3 we consider the use of ProM plug-ins on real assessment data, establishing some useful results. Finaly, Section 4 is for discussions.

2   Process Mining Framework

Process mining has emerged from the field of Business Process Management (BPM). It focuses on extracting process-related knowledge from event logs[1] recorded by an information system. It aims particularly at discovering or analyzing the complete (business, or in our case educational) process and is supported by powerful tools that allow getting a clear visual representation of the whole process. The three major types of process mining applications are (Figure 1):

1) *conformance checking* - reflecting on the observed reality, i.e. checking whether the

---

[1] Typical examples of event logs may include resource usage and activity logs in an e-learning environment, an intelligent tutoring system, an educational adaptive hypermedia system.

modeled behavior matches the observed behavior;

2) *process model discovery* - constructing complete and compact process models able to reproduce the observed behavior, and

3) *process model extension* - projection of information extracted from the logs onto the model, to make the tacit knowledge explicit and facilitate better understanding of the process model.

Process mining is supported by the powerful open-source framework ProM. This framework includes a vast number of different techniques for process discovery, conformance analysis and model extension, as well as many other tools like convertors, visualizers, etc. The ProM tool is frequently used in process mining projects in industry. Moreover, some of the ideas and algorithms have been incorporated in commercial BPM tools like BPM|one (Pallas Athena), Futura Reflect (Futura Process Intelligence), ARIS PPM (IDS Scheer), etc.
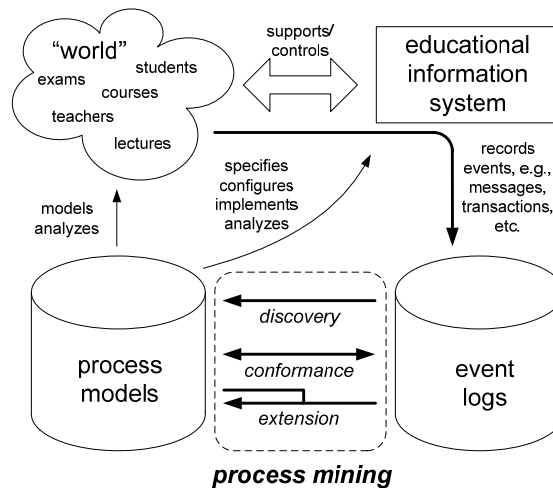

Figure 1. **The process mining spectrum supported by ProM**

3   Case Studies

We studied different issues related to authoring and personalization of online assessment procedures within the series of the MCQ tests organized during the mid-term exams at Eindhoven University of Technology using Moodle[2] (Quize module tools) and Sakai[3] (Mneme testing component) open source LMSs.

To demonstrate the applicability of process mining we use data collected during two exams: one for the Data Modeling and Databases (DB) course and one for the Human-Computer Interaction (HCI) course. In the first (DB) test students (30 in total) answered to the MCQs (15 in total) in a strict order, in which questions appeared one by one. Students after answering each question were able proceed directly to the next question

---

[2] http://www.moodle.org
[3] http://www.sakai.org

(clicking "Go to the next question"), or first get knowledge of correct response (clicking the "Check the answer") and after that either go the next question ("Go to the next question") or, before that, request a detailed explanation about their response ("Get Explanations"). In the second (HCI) test students (65 in total) had the possibility to answer the MCQs (10 in total) in a flexible order, to revisit (and revise if necessary) the earlier questions and answers. Flexible navigation was facilitated by a menu page for quick jumps from one question to any other question, as well as by "next" and "previous" buttons.

In the MCQ tests we asked students to also include the confidence level of each answer. Our studies demonstrated that knowledge of the response certitude (specifying the student's certainty or confidence of the correctness of the answer) together with response correctness helps in understanding the learning behavior and allows for determining what kind of feedback is more preferable and more effective for the students thus facilitating personalization in assessment [3].

For every student and for each question in the test we collected all the possible information, including correctness, certitude, grade (determined by correctness and certitude), time spent for answering the question, and for the DB test whether an answer was checked for correctness or not, whether detailed explanation was requested on not, and how much time was spent reading it, and for the HCI test whether a question was skipped, revisited, whether answer was revised or the certitude changed.[4]

In the remainder of this section we demonstrate how various ProM plug-ins supporting dotted chart analysis, process discovery (Heuristic Miner and Fuzzy Miner), conformance checking, and performance analysis [1][6] allow to get a significant better understanding of the assessment processes.

3.1 Dotted Chart Analysis

The dotted chart is a chart similar to a Gantt chart. It shows the spread of events over time by plotting a dot for each event in the log thus allowing to gain some insight in the complete set of data. The chart has three (orthogonal) dimensions: one showing the time of the event, and the other two showing (possibly different) components (such as instance ID, originator or task ID) of the event. Time is measured along the horizontal axis. The first component considered is shown along the vertical axis, in boxes. The second component of the event is given by the color of the dot.

Figure 2 illustrates the output of the dot chart analysis of the flexible-order online assessment. All the instances (one per student) are sorted by the duration of the online assessment (reading and answering the question and navigation to the list of questions). In the figure on the left, points in the ochre and green/red color denote the start and the

---

[4] Further details regarding the organization of the test (including an illustrative example of the questions and the EF) and the data collection, preprocessing and transformation from LMS databases to ProM MXML format are beyond the scope of this paper, but interested readers can find this information in an online appendix at http://www.win.tue.nl/~mpechen/research/edu.html.

end (passed/failed) of the test. Triangles denote the moment when the student submits an answer or just navigates to another question. Green triangles denote correct responses with low (LCCR – light green) and high (HCCR – dark green) certainty, red triangles correspondingly – wrong responses (light red – LCWR, dark red – HCWR), white triangles – the cases when the student navigated to the next question without providing any response. The blue squares show the moments when the students navigated from the list of the questions (menu) to a question of the quiz (or just submitted the whole test).
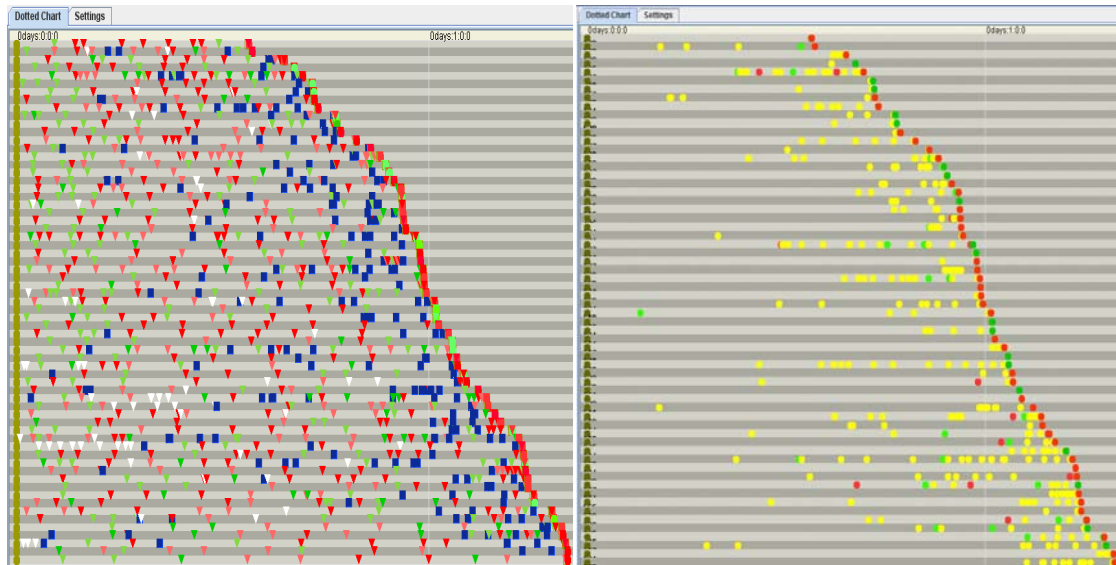

Figure 2. Two dotted charts extracted from the test with flexible order navigation; (1) the overall navigation and answering of questions (left chart), and (2) the effects of changes (right chart)

We can clearly see from the figure that most of the students answered the questions one by one, and provided more correct answers for the first questions of the test than for the last questions. They used the possibility to flexibly navigate mainly at the end of the test: students navigating to the list of the questions and then to the different questions from the list. It can be also clearly seen that only few students read and skipped some questions, not providing their answers first, and then returning to those questions back to provide an answer.

In the figure on the right, we can see the when students revisited the questions. Points in yellow correspond to the situations when correctness of the answers did not change, and points in red and green correspond accordingly to changes to wrong and correct answers. We can see that in a very few cases the correctness was changed, most changes do not affect correctness (e.g., a wrong answer was changed to another wrong answer). Moreover, changes from right to wrong or from wrong to write had similar frequencies, thus not significantly changing the end results.

3.2   *Process discovery*

In some cases, given a usage log we may have limited knowledge about the exact procedure of the assessment but want to discover it based on the data from the log. There exist several algorithms that can automatically construct a depiction of a process. This

process representation typically comes in form of a (formal) mathematical model supporting concurrency, sequential and alternative behavior (like, e.g., the model of Petri nets, Heuristic or Fuzzy miner).

Figure 3 illustrates for the DB test a part (for the first 3 questions) of the discovered process (left) as a Heuristic net, and animation of the same part after conversion to the Fuzzy model (middle), and for the HCI test the complete Heuristic net (right), abstracted from the type of the answer, but from which it is clear which jumps between the questions were popular. From the visualization of the DB test process we can see what possibilities students had, and what the main "*flows*" were globally or at a particular time.
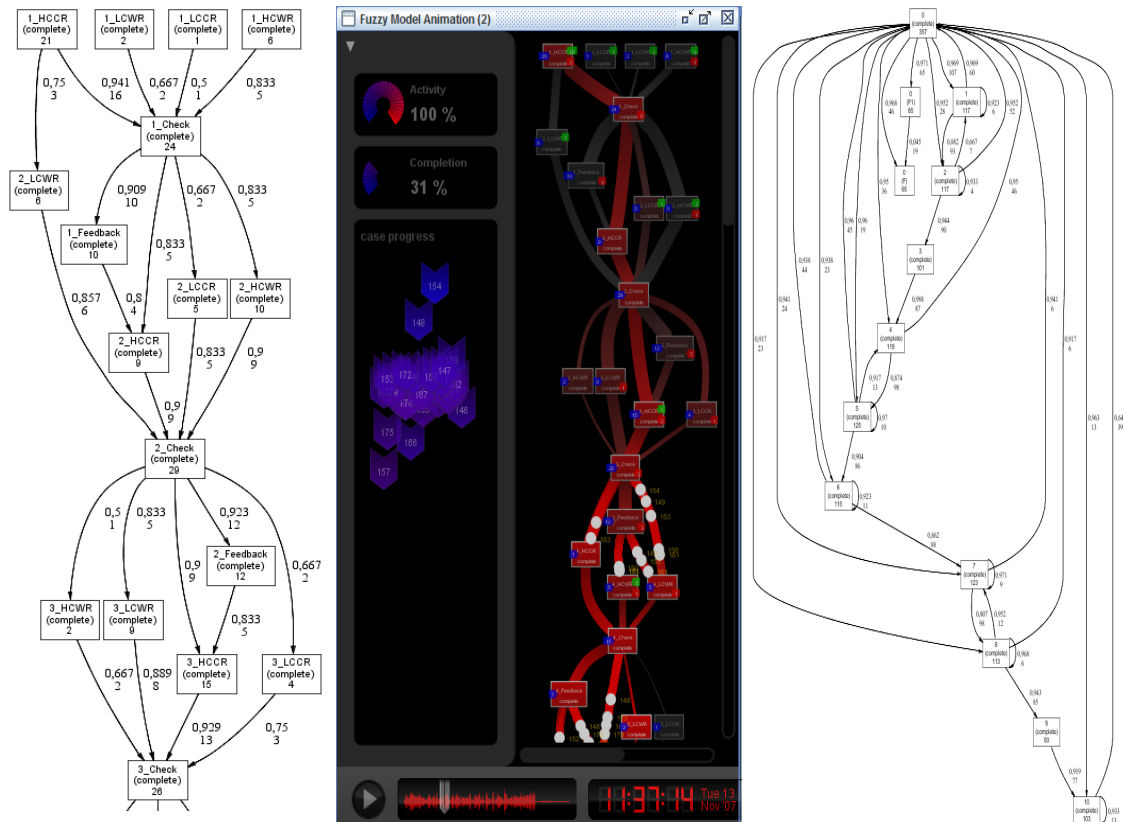

Figure 3. Heuristic nets of strict order (left) and flexible order tests (right)

3.3   *Process analysis*

In some cases, the goal is not to discover the real learning process but to analyze some normative or descriptive model that is given a-priori. For example, the Petri net shown in Figure 4 (formally) describes the generic pattern of answering questions in the DB test allowing for answer-checks and feedbacks. Now it is interesting to see whether this model conforms to reality (and vice versa) and augment it with additional information learned from the event logs. The advantage of having the answering pattern represented as a Petri net is that this allows for many different analysis techniques. ProM offers various plug-ins to analyze Petri nets (verification, performance analysis, conformance, etc.). Models like the one in Figure 4 can be discovered or made by hand. It is also possible to first discover a model and then refine it using the tool Yasper (incorporated

into ProM). Figure 4 was constructed using Yasper and this was a one-time task for this test-type and in principle an authoring tool can be developed to facilitate an automatic translation of the multiple-choice tests with varying properties to Petri nets.

As every question can be answered correctly or wrongly, and with either high or low confidence, there are four possibilities for the first step in the net from Figure 4. The transition HCCR, for example, denotes that the answer is given with high confidence and that it was correct; the other three starting transitions are similar. After answering the question the student can check his answer or just go the next question. The latter decision is modeled by an internal transition (painted in black) that goes to the final place of the net. In case the student has decided to check the answer, he can also ask for some feedback afterwards.
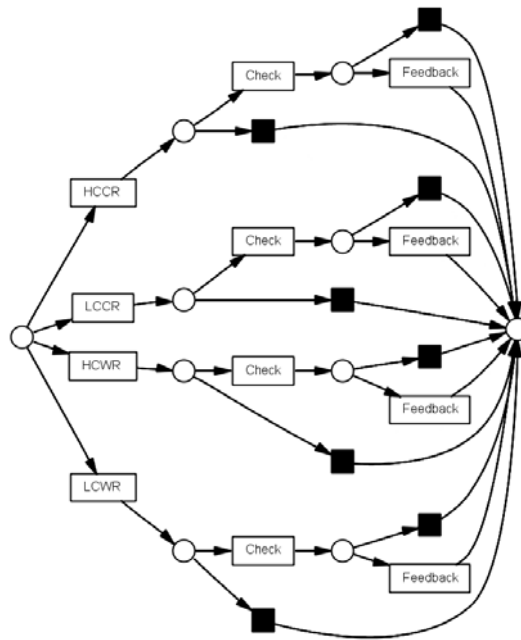

Figure 4. **A Petri net representing the question pattern**

To illustrate the many analysis possibilities of ProM, we show some results obtained using the *Conformance checker* and the *Performance Analysis with Petri net* plugin.

The purpose of conformance analysis is to find out whether the information in the log is as specified. This analysis may be used to detect deviations, to locate and explain these deviations, and to measure the severity of these deviations. We are mostly interested in the notion of *fitness* which is concerned with the investigation whether a process model is able to reproduce all execution sequences that are in the log, or, viewed from another angle, whether the log traces comply with the description in the model (the fitness is 100% if every trace in the log corresponds to a possible execution of the model). This notion is particularly useful for finding out whether (or how often) the students respected the specified order for answering questions (to discover frauds, for example).

Figure 5 shows the result of conformance checking when applied on our log and the Petri net from Figure 4. In this, so-called *log perspective* of the result, each trace from the log

has all its mismatched events colored in orange. In our case, however, there are no orange events, therefore there are no mismatches between the specified answering pattern and the actual exam data.
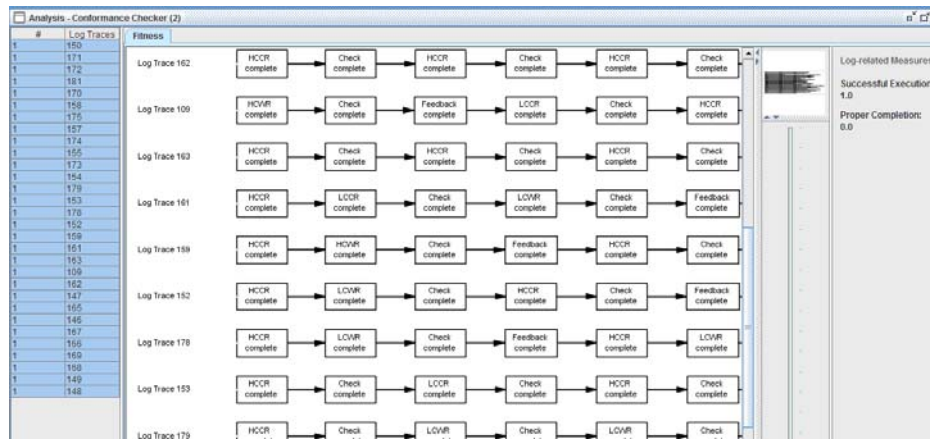

Figure 5. Result of conformance checking showing a 100% fitness

Our next analysis is of a different kind. Instead of checking for the correctnes of the exam behavior, we provide a means to assess the performance of the answering process. The *Performance analysis with Petri net* plugin can extract the Key Performance Indicators from the log, summarizing them in an intuitive way, and graphically present them on a Petri net describing the process under consideration. For our purpose we apply the plugin with the exam data log and the answering pattern from Figure 6 (only for the first question of the test).
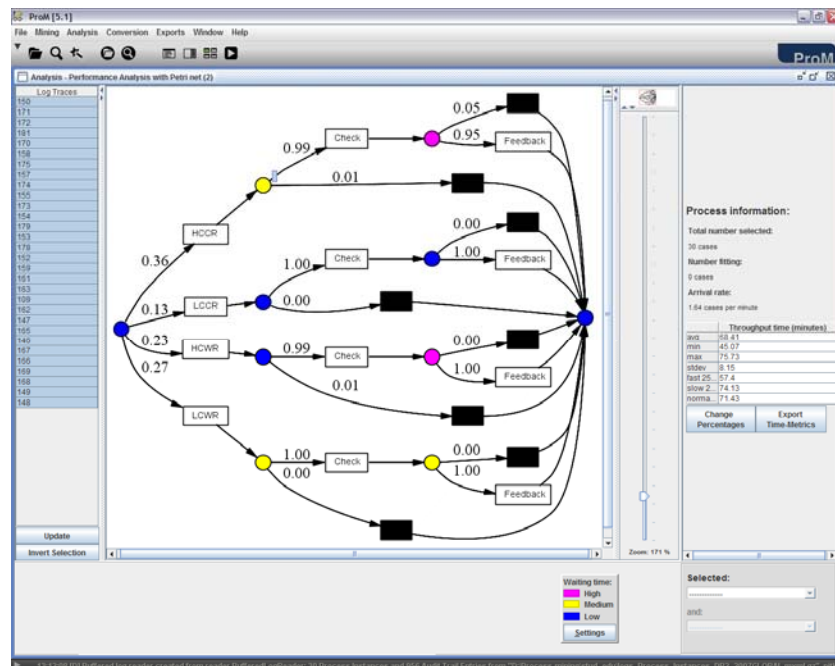

Figure 6. **Results of applying the *Performance analysis with Petri net* plug-in**

The result of the analysis is shown in Figure 6. In the right panel different throughput-type metrics are displayed; from there we, e.g., see that the average duration of the test was 64.41 minute. The central panel shows the answering pattern, colored and annotated with performance information. The numbers on the arcs represent probabilities. As shown, 35% percent of the students answered the first question right and had high confidence. We could also see that almost all students checked their answers and asked for feedback afterwards. Places are colored with respect to their soujourn time, i.e., with respect to the time the process spends in this place. From the picture we can thus see that the answering time was short (the first question was easy), and that the students who answered with high confidence spent more time on the feedback (regardless on the correctness of the answer).

4 Conclusions and Future Work

Data mining techniques have been successfully applied to different types of educational data and have helped to address many issues by using traditional classification, clustering and association analysis techniques. Although the process perspective in educational domains has received some attention, most of the traditional intelligent data analysis approaches applied in the context of educational data mining do not consider the process as a whole (i.e., the focus is no data or simple sequential structures rather than full-fledged process models).

In this paper, we illustrated some of the potential of process mining techniques applied to online assessment data where students in one of the tests were able to receive tailored immediate EF after answering each of the questions in the test one by one in a strict order, and in the other test – to receive no feedback but to answer question in a flexible order. This data was of a sequential nature, i.e. it did not include concurrency. However, other educational processes have lots of concurrency and this can be discovered by ProM. Applying process mining techniques for other types of assessment data, e.g. grades for traditional examinations is therefore an interesting possibility.

ProM 5.0 provides a plugable environment for process mining offering a wide variety of plug-ins for process discovery, conformance checking, model extension, model transformation. Our further work includes the development of EDM tailored ProM plug-ins. On the one hand, this would help bringing process mining tools closer to the domain experts (i.e. educational specialists and researchers), who not necessarily have all the technical background. On the other hand, this will help to better address some of the EDM specific challenges related to data preprocessing and mining. Besides this, the development of the authoring tools for assessment modules with specialized ProM plug-ins would allow to significantly simplify some of the processes for conformance analysis as e.g. a Petri net representing certain assessment procedure can be generated completely automatically.

Acknowledgements

References

[1] Günther, C.W., van der Aalst, W.M.P. Fuzzy Mining: Adaptive Process Simplification Based on Multi-perspective Metrics. In: G. Alonso et al. (eds), *Proc. of Int. Conf. on Business Process Management*, LNCS 4714, p. 328-343. Springer-Verlag, 2007.

[2] Makany, T., Engelbrecht, P.C., Meadmore, K., Dudley, R., Redhead, E.S., & Dror, I.E.: Giving the learners control of navigation: Cognitive gains and losses. In L. Gomez et al. (Eds.), *Proceedings of INTED'07*, 2007.

[3] Romero, C., Ventura, S., García, E. Data Mining in Course Management Systems: MOODLE Case Study and Tutorial. *Computers and Education*, 51. pp. 368-384, 2007.

[4] Romero, C., Ventura, S. Educational Data Mining: A Survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), p. 135-146, 2007.

[5] Rozinat, A., van der Aalst, W.M.P. Conformance Checking of Processes Based on Monitoring Real Behavior. *Information Systems* 33(1), p. 64-95.

[6] Song, M., van der Aalst, W.M.P. Supporting Process Mining by Showing Events at a Glance. In K. Chari, A. Kumar (eds), *7th Annual Workshop on Information Technologies and Systems* (WITS'07), p. 139–145, 2007.

[7] van der Aalst, W.M.P., Weijters, A.J.M.M., Maruster, L. Workflow Mining: Discovering Process Models from Event Logs. *IEEE Transactions on Knowledge and Data. Engineering* 16(9), p. 1128–1142, 2004.

[8] van Dongen, B.F., de Medeiros, A.K.A., Verbeek, H.M.W., Weijters, A.J.M.M., van der Aalst, W.M.P. The ProM framework: A New Era in Process Mining Tool Support. In: Ciardo, G., Darondeau, P. (eds.) *Application and Theory of Petri Nets*, LNCS 3536, p. 444–454. Springer, Heidelberg, 2005.

[9] Vasilyeva, E., De Bra, P., Pechenizkiy, M., Puuronen, S. Tailoring feedback in online assessment: influence of learning styles on the feedback preferences and elaborated feedback effectiveness. In: *Proc. of 8th Int. Conf. on Advance Learning Technologies* (ICALT 2008), IEEE CS Press, p. 834-838, 2008.

[10] Vasilyeva, E., Pechenizkiy, M., and De Bra, P.: Adaptation of Elaborated Feedback in e-Learning, In: W. Nejdl et al. (Eds.), *Proc. of Int. Conf. on Adaptive Hypermedia* (AH'08), LNCS 5149, Springer-Verlag, Berlin, Heidelberg, p. 235-244 (2008)