DISCUSSION OF THE BENEFIT POTENTIALS OF PROCESS MINING FOR E-LEARNING PROCESSES

Marianne Holzhüter, Dirk Frosch-Wilke
University of Applied Sciences Kiel, Sokratesplatz 2, 24149 Kiel, Germany

Salvador Sánchez-Alonso
Universidad de Alcalá, Carretera Madrid-Barcelona, C.P. 28871, Alcalá de Henares, Madrid, Spain

Holzhüter M., Frosch-Wilke D. and Sánchez-Alonso S. (2010).
DISCUSSION OF THE BENEFIT POTENTIALS OF PROCESS MINING FOR E-LEARNING PROCESSES.
In *Proceedings of the 2nd International Conference on Computer Supported Education*, pages 407-411
DOI: 10.5220/0002770104070411

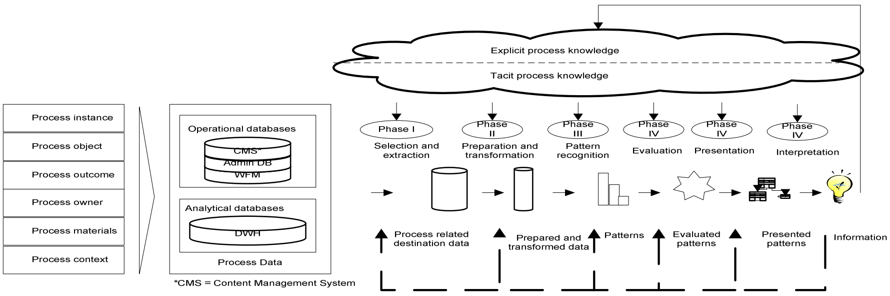

Figure 1: Process mining phases control (Bensberg & Coners, 2008, p. 271).

2 PROCESS MINING

Process mining – inter alia – aims at building a process model in order to describe the behavior contained in event logs of information systems (de Medeiros et al., 2005). The event logs are produced by process-aware information systems (e.g. Workflow Management Systems). Typically, these event logs contain information about the start/completion of process steps together with related context data (e.g. actors and resources). Furthermore, process mining is a very broad area both in terms of applications and techniques.

The goal of process mining is to extract information (e.g., process models) from event logs. Process mining requires the data categories represented in figure 1, which should briefly be explained. First of all process instances, i.e. processes which have been executed, need to be identified. It might be useful to enrich these instances by process objects. In case the process instance is some „learning process" e.g. it would possess the process object „learner". By means of process outcomes as well as relating to the process goals, the process instances can possibly be evaluated. Furthermore the so-called process owner (e.g. the teacher), the process materials (e.g. literature) as well as the process context (e.g. start and end, level of external control) are of interest. Metadata can be additionally considered for the documentation of the process execution.

Explicit process knowledge can be generated from different available process data sources as mentioned above. In particular, decision tree induction methods permit the generation of descriptive rule sets which are able to predict process quality. These rule sets can be used as operational knowledge base to ensure effectivity and efficiency of process executions. (Grob et al., 2008).

3 FRAMEWORK FOR PROCESS MINING OF E-LEARNING PROCESSES

In this section we introduce a novel framework for Process Mining of e-learning-processes. The main motivation for this approach consists in the process improvement by integration of rules as control basis into appropriate e-learning processes. First of all, it may be of strong interest whether rule-based control of e-Learning processes makes sense at all. Amidst several individual learning methods it is possible to identify processes with a more or less uniform structure. Analyzing these processes by means of process mining, conclusions may be drawn for the optimization of learning modules and online learning sessions.

3.1 Data Preparation and Application of the Method

In order to assess and generate rules it is necessary to analyze the correlations between the process characteristics, e.g.

- process object = learner
- process object characteristic = level of knowledge

and the extent to which the process goal has been reached, e.g. increase of knowledge level.

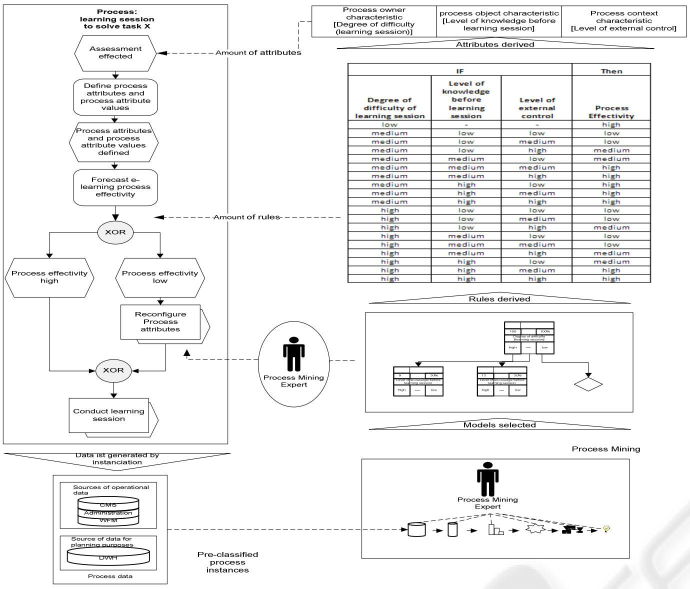

Figure 2: Integration approach for a rule-based control of e-learning processes.

To effect that the analyzed correlations are made accessible as rules, rule based data mining methods may be useful. These methods let the results obtain the structure of IF-THEN rules. Decision trees qualify as an appropriate method since the training period remains short and the results produced dispose of high comprehensibility. Decision trees generate classification models out of the set of pre-classified objects in order to describe the classes as well as to forecast new objects (Grob et al., 2008, pp. 270).

Effective and ineffective process instances are screened. The represented method contributes to process improvement by deriving a generalized action statement from a created rule (see figure 2).

For the purpose of improving processes by integrating rules as a control basis into appropriate e-learning processes, descriptive process knowledge is transformed into a normative action statement. An assessment instruction to analyze the extent, to which the rule conditions have been met, facilitates the use of the rule as control element. The purpose of the assessment is to avoid an ineffective learning process execution. In case of the necessity to modify the process flow a learning designer may reconfigure the manipulable process attributes.

After having identified an effective influence of variables on the process flow via statistical methods and having analyzed the correlation between characteristics and the dependent variables, the variables need to be approved by a specialist regarding the respective processes.

In order to use the classification model in a production rule system within an LMS the prognostic validity needs to be ensured. This may be affected by estimating the cross-validity. Ideally, these kind of approaches are complemented by the inspection of a process expert. A further important criteria for prognostic validity is the data quality. Hereto the different data categories act as toehold. The following data quality criteria need to be considered:

- disposability, integrity and consistency of the respective database
- integration ability in terms of logic (data can be unified in an overall relational schema)
- appropriate timely reach.

3.2 Selection Model for Appropriate e-Learning Processes

This section deals with the identification of e-learning process characteristics for which our process mining framework can be used. Grob et al. (2008) have built a model selecting business process characteristics which facilitate process mining procedures. This model can be applied to e-learning processes as well (see figure 3).

The more structured a process the better it is suited for process mining. Structured Processes can mainly be found where student performance requirements to solve a task can clearly be defined (such as mathematical problems) in contrast to less definable requirements regarding tasks to which

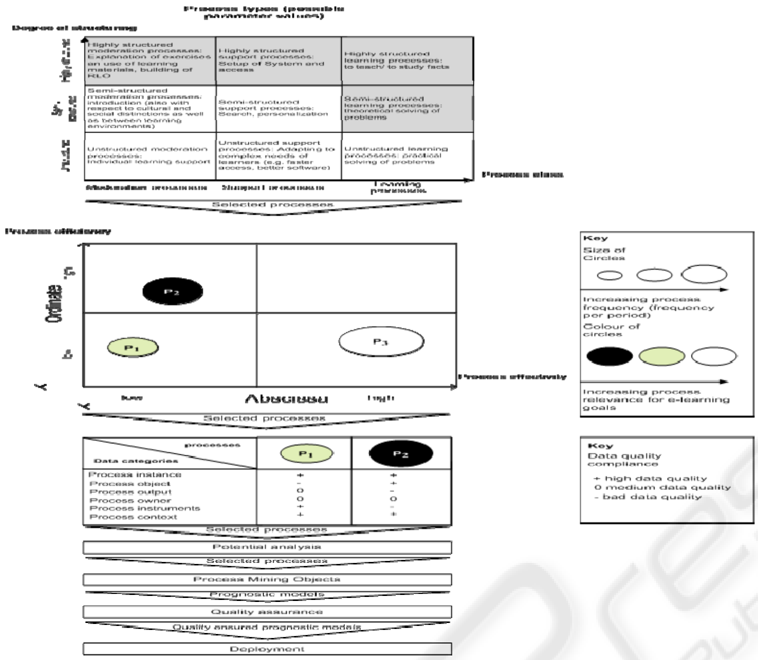

Figure 3: Selection model for choice of e-learning processes for rule based-control.

many possible solutions exist (such as an essay on a broadly defined topic). Processes with possibly high contribution to the e-learning goals (e.g. to gain a higher knowledge level or the reduction of learning time) as well as high execution frequency are of special interest regarding process mining. From these kind of processes, such processes will be selected, which do not sufficiently reach their effectiveness (e.g. gain of higher knowledge level) and efficiency (e.g. reduction of learning time).

Process controlling provides means to identify improvement potentials for such deficient processes, which can be used to reconfigure the respective learning processes.

But, the disadvantage of this method as an a-posterio method is that the process has already been executed before evaluation and reaction. Furthermore certain weaknesses may not be discovered this way. These two aspects may militate for process mining.

4 IMPLEMENTATION

In this section we introduce an implementation model of our process mining framework (see figure 4), which integrates the rules defined by the process mining system into a Learning Management System. Cesarini et al. (2004) provide a web-based approach which we combine with the process mining aspect.

In this implementation model the process mining system hands over its models via Predictive Model Markup Language (PMML)-interface to the rule management system. The rule management system needs to be able to integrate and interpret these models as well as to access the necessary data. The models are leveraged at the point of time they are evaluated in the course of the process runtime by the process leading systems. These systems dispose of interfaces towards the respectively valid process model and facilitate the determination of actual parameters of the attributes relevant to the forecast of the process instance to be executed. The e-learning application produces and makes use of the operational data sources. The application is controlled by the rules generated from the models.

The web server enables dynamic interactions in the web, whereas the users – the teacher as well as the student – profit from the possibility to change the particular data in the database via HTML-pages.

In the course of creating learning objects and tutoring, teachers profit from information how to improve the processes involved. A typical procedure illustrating the benefit of the approach could be the following:

The effectiveness of e-learning sessions with different difficulty levels is anticipated by analyzing historical and actual data. According to the results of the prognosis, the respective learning objects, knowledge levels and kinds of external control are combined in a process model by the teacher. The time for learning and studying learning objects built according to this process model can be reduced by ensuring a higher learning effecivity in contrast to a scenario which does not employ the process model.

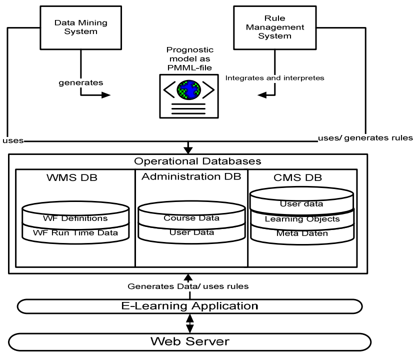

Figure 4: Implementation Model.

5 CONCLUSIONS

The presented approach partly responds to the need for efficient technical systems as well as innovative didactical methods to support the knowledge exchange by enabling the control, analysis and improvement of already executed e-learning processes and can even support not yet executed e-learning processes.

The approach reveals high potentials for use in companies and educational institutions, e.g. gain of valuable time and effort eliminating ineffective learning procedures and accelerating the learning process of complex topics. Nevertheless, there is still a need for evaluations in practice. Further studies could deal with more concrete implementation scenarios going further into technical detail or a substantiated cost-benefit-ratio.

To complete the analysis of the potentials of the approach there is a need to render the application of process mining in e-learning marketable and ready for implementation. Furthermore the method can be analyzed in order to identify possibilities to fill the gap of missing learning process models and lack of learning process adaptability.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge comments received on earlier versions of this paper from Michael Doktor, Sathya Laufer and Matthias Barz.

REFERENCES

Cesarini, M., Monga, M., & Tedesco, R. (2004). *Carrying on the e-learning process with a workflow management engine.* In *Proceedings of the 2004 ACM symposium on Applied computing* (S. 940 - 945). Nicosia, Cyprus : Association for Computing Machinery.

Datta, A. (1998). *Automating the Discovery of As-Is Business Process Models: Probabilistic and Algorithmic Approaches.* Information Systems Research, 9 (3), 275-301.

de Medeiros, A. A., Weijters, T. A., & van der Aalst, W. (September 2005). *Genetic Process Mining: A Basic Approach and its Challenges.* In *Proceedings of the Workshop on Business Process Intelligence*. Nancy, France.

Federal Ministry of Education and Research (Germany) (2009). bmbf.de – *Lifelong Learning.* Retrieved 09/09/2009 from bmbf.de: http://www.bmbf.de/en/411.php.

Franconi, E., & Sattler, U. (1999). *A Data Warehouse Conceptual Data Model for Multidimensional Aggregation*. In *Proceedings of the International Workshop on Design and Management of Data Warehouses.*

Frosch-Wilke, D.; Sànchez-Alonso, S.; Dodero, J.M. (2008*). Evolutionary Design of Collaborative Learning Processes through Reflective Petri Nets*. In *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies*, Santander (Spain), pp. 423-427

Grob, H. L., Bensberg, F. & Coners, A. (April 2008). *Regelbasierte Steuerung von Geschäftsprozessen - Konzeption eines Ansatzes auf Basis des Process Mining*. Die WIRTSCHAFTSINFORMATIK, pp. 268-281.

Helic, D. (2006). *Technology-Supported Management of Collaborative Learning Processes*. International Journal of Learning and Change, Vol. 1, No. 3, pp. 285-298.

Ingvaldsen, J.E., & J.A. Gulla. (2006). *Model Based Business Process Mining*. Journal of Information Systems Management, Vol. 23, No. 1, Special Issue on Business Intelligence, Auerbach Publications

Palmer, C.-E. (2003). *E-Learning - die Revolution des Lernens gewinnbringend einsetzen*. In U. Scheffer, & M. Charlier (Eds.), *E-Learning* (pp. 13). Stuttgart, Germany : Klett-Cotta Verlag.

Schönwälder, S. (1997). *Portfoliomanagement für betriebliche Informationssysteme*: *computergestützter Ansatz zur partizipativen Einführung und Gestaltung.* Wiesbaden, Germany: Gabler.