

A PROCESS MODEL FOR WORKFLOW MINING

G. N. Singh¹ & Sandeep Aggarwal²

It is common practice that even data is logged by most of information systems about executed tasks. Explicit process models drive modern workflow management. Creating a workflow design is a complex time-consuming process and normally there are Discrepancies between the actual workflow processes and the processes as apparent by the management.

Therefore, we have suggested technique for (re)discovering workflow models. Starting point for such techniques are workflow logs containing all the information about the workflow process as it is actually being executed.

Keywords: Workflow Mining, Workflow Management, Data Mining, Petri Nets, Workflow Model

1. INTRODUCTION

Numerous type of information systems record event data about executed tasks. Workflow mining is concerned with the origin of a graphical workflow model out of this data. Workflow mining is a highly interactive process. The mining expert repeatedly approaches the result by changing the parameter of the mining tool and verifying the mined models. The modeling of business process in a computerized manner either fully or partially is called as workflow. Workflow is used for representing a repeatable sequence of operations of a business process. It is a activity pattern used for systematic organization of resources, investigation, operational implementation to represent real work for further assessment. Workflow is related to various fields like operations research, artificial intelligence or in some special type of industries like printing, automation of events or where processes are executed by particular rules for completing the business process. A workflow management system is a computer software system that manages and specifies a series of tasks to generate a final outcome within an organization. Workflow Management System define various workflows for different types of processes. workflow management concepts and technologies have been functional in many enterprise information systems.

Work Flow Management is a fast growing technology which is gradually more being exploited by businesses in various industries. Its major feature is the computerization of various processes which involve combinations of human behavior and machine-based actions, especially those involving communication with IT applications and tools.

Many software systems ERP (Enterprise Resource Planning) systems such as SAP, PeopleSoft, and Oracle, CRM (Customer Relationship Management) software, etc. have adopted workflow technology. Even with its promise,

many problems are encountered while applying workflow knowledge. One of the problems is that these systems require a workflow design, i.e., a designer need to construct a detailed model accurately unfolding the routing of work. Modeling a workflow is far from trivial: It requires understanding of the workflow language and detailed discussions with the workers and management involved. By creation of the graphical process definitions, i.e., models describing the life-cycle of a classic case (workflow instance) in segregation, one can classify these systems to maintain business processes.

2. WORKFLOW: WHERE AND HOW

- It supports mainly the analysis and the design of the new system. It helps to understand how people work with the new system.
- It reduces the analysis time and cost because in most of information system, log data is readily available without any additional cost.
- It provides objective information about any system under consideration because workflow logs are an impartial reflection of executed activities.
- It allows formal verification of workflow properties by simulating the process before its execution.
- It supports multiple views of the same process by synthesis algorithms and model conversions.
- It allows the involuntary derivation of special paths from acyclic normal flow of the process whenever required.
- It enacts repeated model changes on the running instance of a workflow, shielding them from undesired outcomes.

3. PROPOSED MODEL



Fig. 1:

¹Assistant Professor, Sudarshan College, Lalgaoon (M.P.) India

²Lecturer, D.A.V. College, Abohar (Pb.) India

E-mail: ²sandeepaggarwal10@gmail.com

Modeling: Workflow Modeling denote the formation of a workflow type, which is a formal depiction of various aspects of a workflow, such as various activities to be performed, recognition of processing entities performing the behavior and dependencies that exist among the activities i.e. data flow etc. for a workflow type an arbitrary number of workflows instances can be created. The set of workflow types defined at a certain point in time forms the workflow model. Process knowledge is written and compiled into one process. A process description can be textual as well as graphical. For reproduction and validation purposes, sometimes a formal process model is constructed. The process modeling may include the capability to model processes in the perspective of an organization structure where an organization model is integrated. In the planning phase, a process definition comprises manual as well as workflow activities and is typically on a high level of abstraction. The final yield of modeling phase of the model is a complete definition of process which can be defined and parsed at runtime by workflow stations. Modern products of workflow modeling are in the form suitable to a workflow management system that was under consideration.

Definition & Implementation: The definition normally consists of a number of distinct activity steps, which are related with computerization or automation of business process with the help of various activity steps. The definition of process should be represented in an easily understandable form that corresponds to some official language representation so that developers can easily understand it. Also some times dynamic changes in process environments may be required from actual working environments. A part of this high level process is more polished to a workflow definition and implemented into a process aware information system. Within a workflow management system, the workflow is implemented in form of an explicit graphical process model. Using conventional implementation techniques, the workflow is reflected within the regular code.

Ratification: The workflow ratification interprets the process definition and controls the complete definition of a particular instance of systematic processing to achieve the goals of a desired business process. It can be seen as realization of mechanisms through one or more cooperating processes which simulates the working of particular instance. The workflow ratification service maintains the execution of same plan repeatedly. Here the control data for various processes can store internal state information that manages flow speed and flow volumes of a particular instance in consideration. Also it can include check points in a particular process and recovery information of the instance of a process to be used in case of a failure. All this information is stored either centrally or in a distributed manner among engines of various processes. Then, the

process aware information system is put into operational use. Various Researches show that the effect of a business process definition on work practice is that of a standard. Work practice is not determined by this standard but influenced. The new work practices emerge with the ratification of the new system and become visible within the actions of the process participants executed to achieve the business process goals.

Verification: The purpose of verification is to check whether the expected properties are satisfied by the operational description or not. Apart from Workflow allied data, the Execution log may also contain industrial data recorded during the Workflow ratification phase. Those data are used in order to check the competence of the Business process model to the goals of the organization. If this is not the case, a revision of the Business process model becomes compulsory. The Business process reengineering phase leads to the Business process model analysis, where the Business process model is revised. The updated model is the starting point of a new iteration of the processes of stabilization and debugging.

Analysis: In process or workflow analysis the resulting work practice is analyzed. This analysis may include, for example, the determination of the frequency of activity occurrences or the application of workflow mining. Once a stable version of the Business Process model is available as a Workflow design, a full-featured, operational version is generated by converting modeling concepts into an executable illustration. As a result of the execution or ratification of the Workflow a log is created which collects information about the process execution e.g., tasks executed, starting and ending of tasks execution, resources used, etc. that can be analyzed. Also to detect anomalous situations like as deadlocks, bottlenecks, etc. derived from a bad Workflow design; once an irregularity is detected, a revision of the Workflow model must be performed, i.e. a new stabilization process must be carried-out.

4. IMPLEMENTATION

By creating graphical process definition, i.e., models relating to the life-cycle of a classic case (workflow instance) in separation, one can organize these systems to maintain business processes. In addition to pure workflow management systems many other software systems have adopted workflow expertise. Instead of starting with a workflow design, we start by collecting information about the workflow processes as they occur. We assume that it is possible to record events such that (i) each event refers to a process i.e., a well-defined step in the workflow, (ii) each event refers to a case i.e., a workflow instance, and (iii) events are ordered i.e., in the monitor events are recorded in succession even though they may be executed in parallel.

Any information system using transactional systems or workflow management systems will offer this information in some form. Here we assume that it is possible to collect workflow logs with event data. These workflow logs are used to construct a process specification which adequately represents the behavior registered. We use the term process mining for the method of finding an ordered process description from a set of real executions.

Case No.	Process-ID
Case 1	Process A
Case 2	Process A
Case 3	Process A
Case 6	Process A
Case 5	Process A
Case 3	Process B
Case 1	Process B
Case 1	Process C
Case 6	Process E
Case 2	Process B
Case 4	Process A
Case 2	Process D
Case 2	Process C
Case 5	Process F
Case 4	Process F
Case 1	Process D
Case 3	Process C
Case 3	Process D
Case 4	Process G
Case 5	Process G
Case 5	Process H
Case 5	Process I
Case 4	Process H
Case 1	Process I
Case 6	Process I
Case 4	Process I
Case 2	Process I
Case 3	Process I

To illustrate the principle of process mining, we consider the workflow log shown in Table Above. This log contains information about six cases (i.e., workflow instances). The log shows that for Three cases (1, 2, 3, and 4) the Processes A, B, C, D and I have been executed. For the fourth and fifth case also five processes are executed: A, F, G, H and I. For the Sixth case only 3 processes are executed: A, E and I. Each case starts with the execution of A and ends with the execution of I. If Process B is executed, then also Process C, D are executed. However, for some cases Process C is executed before Process D, also the Process E is executed alone. Also if Process F or G is executed then Process H is also executed no matter the order in which F and G are

executed. Based on the information shown in Table and by making some assumptions about the completeness of the log (i.e., assuming that the cases are representative and an adequate subset of possible behaviors is observed), we can deduce for example the process model shown in Figure 1. The model is represented in terms of a Petri net. The Petri net starts with Process A and finishes with Process I. These Processes are represented by transitions. After executing A there is a choice between either executing B with C and D in parallel or just executing task E or executing F and G in parallel, along with H. To execute C and D in parallel and F and G in parallel, two non-observable Processes (AND-split and AND-join) have been added. These Processes have been added for routing purposes only and are not present in the workflow log. Here we assume that two Processes are in parallel if they appear in any order. However, by separating between start and end events for Processes it is possible to explicitly detect parallelism. Start events and end events can also be used to indicate that Processes take time. However, to simplify the structure we assume processes to be atomic without losing generality. But, it is important to understand that such an approach only works if events are recorded at the time of their occurrence.

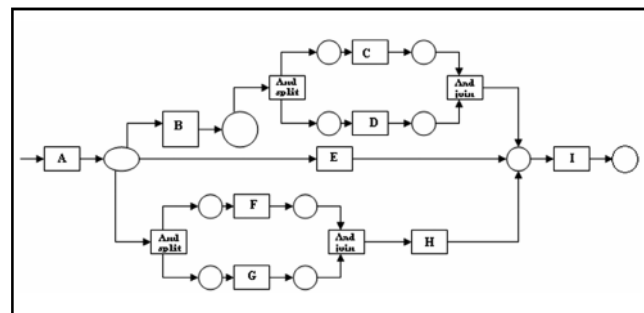


Fig.2: A Process Model Analogous to the Workflow Log

The basic idea behind process mining, also referred to as workflow mining, is to construct a flow diagram from the information given in Table.

5. CONCLUSION

Here we have shown that Software Process concepts can be applied to Workflow development in order to construct better Workflow products. Specially, an integrated view of the activities related with Workflow management has been introduced. The definition of the Workflow lifecycle as the composition of five main processes (namely modeling, defining and implementation, ratification, verification and analysis) subsuming the well known tasks of Workflow modeling and ratification plus prototyping and investigation is the first step in the development of new frameworks for Workflow development. Future work includes the definition of a homogeneous model to be used along the whole Workflow life cycle to improve the traceability of Workflow specifications and minimize the

impedance mismatches. Moreover, this model should be expressive enough to integrate evolution aspects in a seamless way. Through examples we tried to show that the model provides exciting analysis results for workflow processes. In the future, we hope to present the mining algorithm so that it is able to rediscover an even larger class of workflow problems. At this point in time, there is possibility of two improvements. First, it should be possible to deal with short loops of a particular form. Second, the rediscovery problem could be relaxed to take behaviorally equivalent workflows into account. The over all goal is to be able to analyze any workflow log without any knowledge of the underlying process. Future work also includes the mechanism for dealing with loops and for the detection of dependencies.

REFERENCE

- [1] "Workflow Mining: Which Processes Can be Rediscovered?", W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster.
- [2] "Workflow Mining: Discovering Process Models from Event Logs", W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster.
- [3] "Interactive Workflow Mining—Requirements, Concepts and Implementation Markus Hammori", Joachim Herbst, Niko Kleiner.
- [4] "A Workflow Specification Environment", Pierre Azema, Francois Vernadat, Pierre Gradiat.
- [5] "Workflow Management Coalition The Workflow Reference Model Document Number TC00-1003", David Hollingsworth.