

# How To Create Synthetic S-MXML Logs Using CPN Tools

Ana Karla Alves de Medeiros

Department of Technology Management, Eindhoven University of Technology  
P.O. Box 513, NL-5600 MB, Eindhoven, The Netherlands.  
`{a.k.medeiros}@tm.tue.nl`

**Abstract.** Process mining aims at automatically generating process models from event logs. The main idea is to use the discovered models as an *objective start point* to deploy systems that support the execution of business processes (for instance, workflow management systems) or as a *feedback mechanism* to check if the prescribed models fit the executed ones. In the context of the SUPER project, the event logs are going to be generated by the SUPER<sup>1</sup> engine. These event logs are going to be stored in the History Repository and (some of) its elements are going to refer to concepts of the SUPER ontologies. Eventually these logs are going to be converted to S-MXML logs. S-MXML (Semantic MXML) is an extension of the current process mining format MXML with semantic annotations. The aim is to develop process mining plug-ins that make use of the semantic information provided by the SUPER event logs while still being able to use current process mining algorithms in these same logs. However, the SUPER engine is not implemented yet. Therefore, this paper shows how to extend CP-nets to generate S-MXML event logs that can be mined by process mining tools supporting this format. This way we benefit from the simulation capabilities of CPN Tools and, therefore, we can proceed with testing and implementing our process mining algorithms in parallel with the development of the SUPER engine. The extension consisted of implementing (i) some ML functions that can be used to annotate the CP-net, and (ii) a ProM<sub>import</sub>-framework [2] plug-in that bundles up the files (generated by the CP-net simulation) into a single S-MXML file that is ready to be mined.

## 1 Introduction

Process mining targets the *automatic* discovery of information from an event log. This discovered information can be used to deploy new systems that support the execution of business processes or as a feedback tool that helps in analyzing and improving enacted business processes. The starting point of any process mining technique is an event log. This log can have lots of data, such as the tasks that are executed, their time of execution, the person/system that performed them, the data fields related to these tasks, and so forth. For instance, consider the event

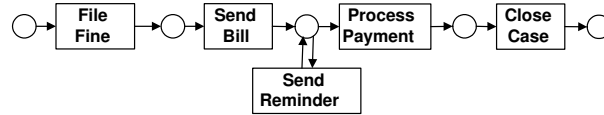
---

<sup>1</sup> More information about the European project SUPER (Semantics Utilised for Process Management within and between Enterprises) can be found at <http://www.ip-super.org/>.

log in Table 1. This log has four executions (cases) of a process that handles fines for drivers in a certain province.

Case ID	Task Name	Event Type	Originator	Timestamp	Extra Data
1	File Fine	Completed	Anne	20-07-2004 14:00:00	...
2	File Fine	Completed	Anne	20-07-2004 15:00:00	...
1	Send Bill	Completed	system	20-07-2004 15:05:00	...
2	Send Bill	Completed	system	20-07-2004 15:07:00	...
3	File Fine	Completed	Anne	21-07-2004 10:00:00	...
3	Send Bill	Completed	system	21-07-2004 14:00:00	...
4	File Fine	Completed	Anne	22-07-2004 11:00:00	...
4	Send Bill	Completed	system	22-07-2004 11:10:00	...
1	Process Payment	Completed	system	24-07-2004 15:05:00	...
1	Close Case	Completed	system	24-07-2004 15:06:00	...
2	Send Reminder	Completed	Mary	20-08-2004 10:00:00	...
3	Send Reminder	Completed	John	21-08-2004 10:00:00	...
2	Process Payment	Completed	system	22-08-2004 09:05:00	...
2	Close case	Completed	system	22-08-2004 09:06:00	...
4	Send Reminder	Completed	John	22-08-2004 15:10:00	...
4	Send Reminder	Completed	Mary	22-08-2004 17:10:00	...
4	Process Payment	Completed	system	29-08-2004 14:01:00	...
4	Close Case	Completed	system	29-08-2004 17:30:00	...
3	Send Reminder	Completed	John	21-09-2004 10:00:00	...
3	Send Reminder	Completed	John	21-10-2004 10:00:00	...
3	Process Payment	Completed	system	25-10-2004 14:00:00	...
3	Close Case	Completed	system	25-10-2004 14:01:00	...

**Table 1.** Example of an event log.



**Fig. 1.** Petri net illustrating the control-flow perspective that can be mined from the event log in Table 1.

The amount of data in the event log determines which *perspectives* of process mining can be discovered. If the log provides the tasks that are executed in the process and it is possible to infer their order of execution, the *control-flow perspective* can be mined. The log in Table 1 has this data (cf. fields “Case ID”, “Task Name” and “Timestamp”). So, for this log, mining algorithms could discover the process in Figure 1. Basically, the process describes that after a fine is entered in

the system, the bill is sent to the driver. If the driver does not pay the bill within one month, a reminder is sent. When the bill is paid, the case is archived. If the log provides information about the persons/systems that executed the tasks, the *organizational perspective* can be discovered. The organizational perspective discovers information like the social network in a process, the transferring of work etc. For instance, the log in Table 1 shows that “Anne” transfers work for both “Mary” (case 2) and “John” (cases 3 and 4), and “John” sometimes transfers work for “Mary” (case 4). Besides, by inspecting the log, the mining algorithm could discover that “Mary” never has to send a reminder more than once, while “John” does not seem to perform as good. The managers could talk to “Mary” and check if she has another approach to send reminders that “John” could benefit from. This can help in making good practices a common knowledge in the organization. When the log contains more details about the tasks, like the values of data fields that the execution of the task modifies, the *case perspective* can be discovered. So, for instance, a forecast for executing cases can be made based on previous already completed cases, exceptional situations can be discovered etc. In our particular example, logging information about the profiles of drivers (like age, gender, car etc) could help in assessing the probability that they would pay their fines on time. Moreover, logging information about the places where the fines were applied could help in improving the traffic measures in these places.

Having these three perspectives in mind, and the different mining tools<sup>2</sup> to tackle one or more of these perspectives, an effort was made in [4] to define a single XML format, called the Mining XML (MXML) format, that could be used as input to the different tools. In the context of SUPER, we extended this MXML format by adding semantic annotations to its elements. This semantic extension is called S-MXML (Semantic Mining XML). By converting simulated or real-life logs to the S-MXML format, one can use the mining techniques in multiple contexts. The S-MXML format is backwards compatible with the MXML format. This is important because this allows for directly using the current process mining techniques over the S-MXML event logs extracted from the SUPER repositories.

CPN Tools<sup>3</sup> supports the modelling, execution and analysis of *Coloured Petri nets* (CP-nets) [10]. Additionally, there is a fair amount of CPN models that can be used as input to test mining algorithms. Thus, we decided to extend CPN Tools to support the creation of S-MXML logs. The extension is based on the work in [6], where CPN Tools were extended to create MXML logs. The main idea is to create random S-MXML logs by simulating CP-nets in CPN Tools. The

---

<sup>2</sup> Examples of mining tools are InWolvE [9, 11], Process Miner [12], EMiT [8], Little Thumb [13], MiSoN [3] and ProM framework [7]. The InWolvE, Process Miner, EMiT and Little Thumb mine the control-flow perspective. The MiSoN mines the organizational perspective. The ProM framework has plug-ins that to mine all three perspectives. Actually, the mining tools EMiT, MiSoN and Little Thumb were respectively implemented as the ProM mining plug-ins “Alpha algorithm”, “Social network miner” and “Heuristics miner”.

<sup>3</sup> In this paper we assume the reader to be familiar with CPN Tools. However, if necessary, more information about this tool is provided at <http://wiki.daimi.au.dk/cpntools/cpntools.wiki>.

first part of the extension consisted of implementing the ML functions to support the logging from a CP-net. The second part consisted of extending the current implementation of the “CPN Tools” plug-in in the the ProM<sub>import</sub> framework [2] to bundle the logged files into a single S-MXML file.

In short, two steps are necessary to create S-MXML logs using CPN Tools:

1. Modify a CP-net to invoke the set of ML functions that will create logs for every case executed by the CP-net. This step involves modifying the *declarations* of the CP-net and the *input/output/action* transition inscriptions.
2. Use the CPN Tools plug-in, in the ProM<sub>import</sub> framework, to group the logs for the individual cases into a single S-MXML log.

The rest of this paper is organized as follows. Section 2 briefly explains the S-MXML format. Understanding how this format supports the different perspectives of mining helps in understanding how CPN Tools was extended. Section 3 describes how to modify a CP-net to create partial S-MXML logs during its simulation (Step 1 above). Section 4 shows how to use the ProM<sub>import</sub> framework to bundle these partial S-MXML logs into a single log that can be mined (Step 2 above). Section 5 presents some conclusions and future work.

## 2 The S-MXML format

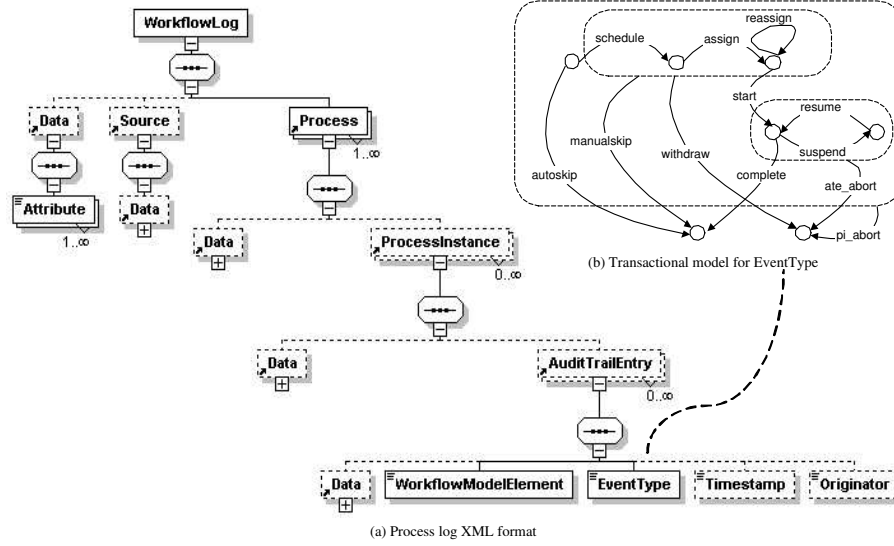
Since the Semantic Mining XML (S-MXML) format is a semantic annotated version of the MXML format, let us first introduce this format.

The Mining XML<sup>4</sup> format (MXML) started as an initiative to make different mining tools have a common input format [4]. This way, event logs could be shared among different mining tools. Actually, defining a common input format like MXML is the first step towards the creation of a repository on which process mining researchers can test their algorithms. In this section we explain the MXML format because this helps to understand the ML functions defined for the extension of CP-nets. The schema for this MXML format (depicted in Figure 2) is available at <http://www.processmining.org/WorkflowLog.xsd>.

As can be seen in Figure 2, an event log (field *WorkflowLog*) has the execution of one or more processes (field *Process*), and optional information about the source program that generated the log (field *Source*) and additional data elements (field *Data*). Every process (field *Process*) has zero or more cases or process instances (field *ProcessInstance*)<sup>5</sup>. Similarly, every process instance has zero or more tasks (field *AuditTrailEntry*). Every task or audit trail entry (ATE) should at least have a name (field *WorkflowModelElement*) and an event type (field *EventType*). The event type determines the state of the tasks. There are 13 supported event types: schedule, assign, reassign, start, resume, suspend, autoskip, manualskip, withdraw, complete, ate\_abort, pi\_abort and unknown. The other task fields are optional. The *Timestamp* field supports the logging of time for the task. The

<sup>4</sup> More information about the Extensible Markup Language (XML) can be found in [1].

<sup>5</sup> In the rest of this document, the words “execution”, “case” and “process instance” are interchangeable.



**Fig. 2.** The visual description of the schema for the Mining XML (MXML) format.

*Originator* field records the person/system that performed the task. The *Data* field allows for more logging of additional information. More details about the MXML format can be found in [7].

Mapping the MXML format to the three mining perspectives, we see that the *control-flow* perspective mainly focuses on the WorkflowModelElement, the EventType and the Timestamp<sup>6</sup> fields. The *organizational* perspective chiefly depends on the Originator field. The *case* perspective especially relies on the extra Data fields.

Note that in CPN Tools the process corresponds to the CP-net, the tasks (or ATEs) are the transitions in the CP-net, and each simulation of the CP-net corresponds to the creation of a process instance.

The S-MXML format is just like the MXML format, except for the fact that all elements have an extra attribute - called *modelReference* - that links to a *list* of concepts in ontologies. The schema for the S-MXML format is available at <http://www.processmining.org/SMXML.xsd>.

### 3 Extending a CP-net to Produce S-MXML Event Logs

This section shows how to annotate a CP-net with the ML functions that we created to log S-MXML files. The ML functions can be downloaded from the section “Tools”<sup>7</sup> in [2]. To illustrate the extension process, we use the CP-net

<sup>6</sup> When the Timestamp field is not logged, the sequence in which the tasks appear in the log is used to infer their order of execution.

<sup>7</sup> Search for the link to the file “CPNToolsConverter.zip”.

in Figure 3. The extension of this CP-net involves editing its declaration and transition inscriptions.

**CPN Declarations** The declarations of a CP-net need to be modified to import the ML functions to log transitions. These functions are in the file *loggingFunctionsMultipleFilesWithOntologySupport.sml*. The ML functions in this file use two constants: *FILE* and *FILE\_EXTENSION*. The constant *FILE* sets the location and the name prefix of the S-MXML files that the CP-net will create for every case it executes. The constant *FILE\_EXTENSION* sets the extension that these created files have. For instance, to (1) create the XML log files for every case at the subdirectory *logs* from the directory where the CP-net is located and name every log with the prefix *fines*; and (2) assign the extension *.cpnxml* to every created log, the following should be declared:

1. *val FILE = "logs/fines"*
2. *val FILE\_EXTENSION = ".cpnxml"*
3. *use "loggingFunctionsMultipleFilesWithOntologySupport.sml";*

Note that the use of the file *loggingFunctionsMultipleFilesWithOntologySupport.sml* **must be declared after** declaring the constants *FILE* and *FILE\_EXTENSION*. Table 3 shows what these declarations look like in the CP-net of Figure 3. Additionally, be aware that ML is case sensitive and the subdirectories provided in the constant *FILE* should already exist.

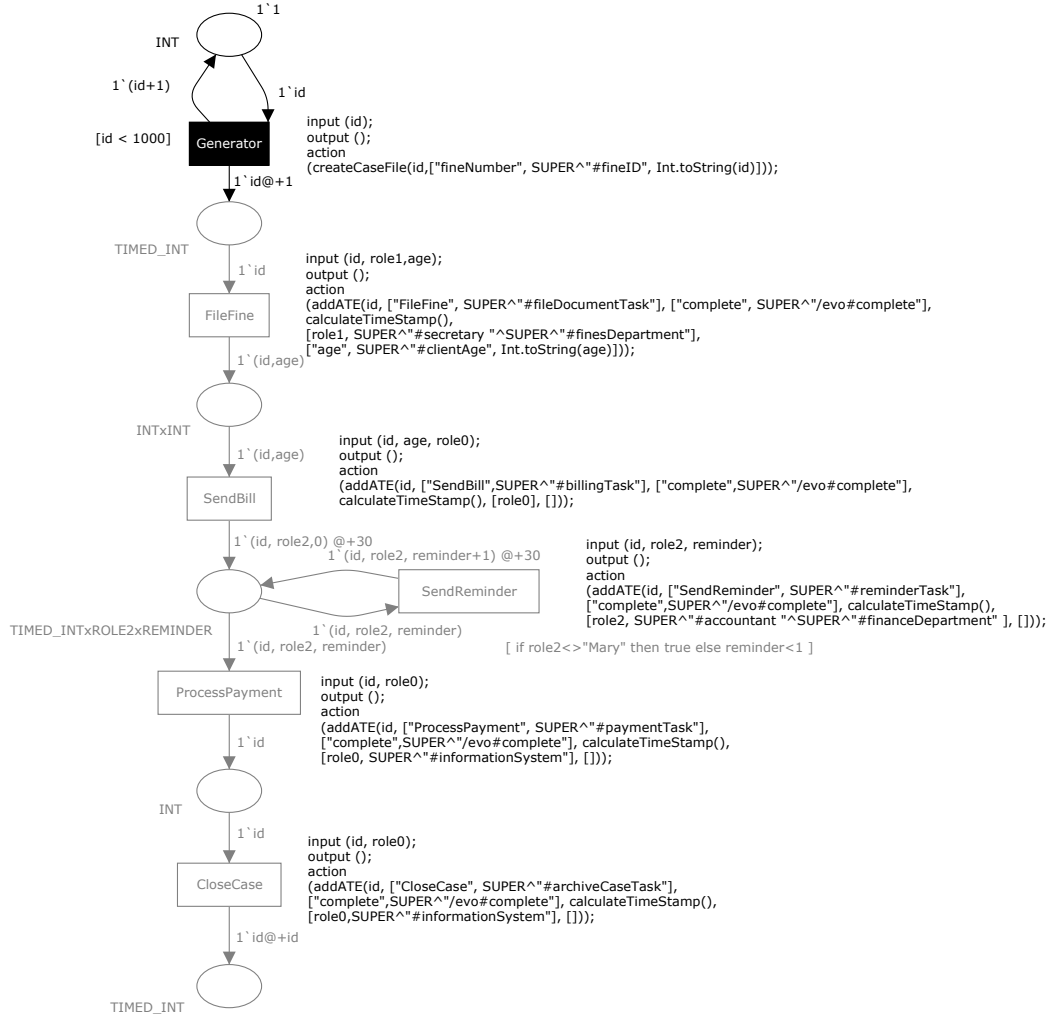
**CPN Transitions** Once the declarations of the CP-net have been updated, the *input/output/action* inscriptions of transitions can be modified to invoke the logging functions. The CP-net will create a partial S-MXML log for *every* case that it executes. In the example in Figure 3, the transition *Generator* creates the unique case identifiers. After a partial S-MXML log has been created, the transitions (or tasks) executed for a case are written to its partial S-MXML log. Thus, two ML functions are provided: *createCaseFile* and *addATE*.

The function *createCaseFile*(int caseld, ListOfStrings data) opens the log file for a case. This function should be invoked only *once per case*, and before the function *addATE* is invoked for this same case. The transition *Generator* in Figure 3 illustrates how to use the function *createCaseFile*. Note that this function receives as input (i) an *integer* (the case identifier!) and (ii) a list of data attributes. Every data attribute has three consecutive elements in the list: (1) the attribute name, (2) a list of ontological concepts<sup>8</sup> and (3) the attribute value.

The function *addATE*(int caseld, ListOfStrings transitionName, ListOfStrings eventType, StringTimestamp timestamp, ListOfStrings originator, ListOfStrings data) logs the execution of a transition to the log of a case. For instance, the transition *FileFine* in Figure 3 has an invocation of this function. The parameters of the function *addATE* are:

---

<sup>8</sup> This is an XML list, so the elements are separated by white spaces. For instance, a list with two concepts would be like "uri1#concept1 uri2#concept2".



**Fig. 3.** Example of an extended CP-net for the process described in Figure 1. The highlighted transition “Generator” and the input/output/action inscriptions were added during extension of the CP-net.

```

(* Standard declarations *)
colset E = with e;
colset INT = int;
colset BOOL = bool;
colset STRING = string;
(* Net declarations *)
val SUPER="http://www.ip-super.org/ontology"
colset TIMEDINT = int timed;
colset ROLE0 = subset STRING with ["system"];
colset ROLE1 = subset STRING with ["Anne"];
colset ROLE2 = subset STRING with ["Mary", "John"];
colset INTxINT= product INT * INT;
colset TIMED_INTxROLE2xREMINDER = product INT * ROLE2 * INT timed;
colset AGE = int with 18..75;
var id: INT;
var age:AGE;
var reminder:INT;
var role0: ROLE0;
var role1: ROLE1;
var role2: ROLE2;
fun OK(id) =
  if id < 1000 then true
  else false;
(* Log declarations *)
val FILE = "logs/fines"
val FILE_EXTENSION = ".cpnxml"
use "loggingFunctionsMultipleFilesWithOntologySupport.sml";

```

**Table 2.** Declarations for the CP-net in Figure 3. The declarations in **bold** were used to extend the model to log S-MXML files.



1. **caseId**: *integer* that uniquely identifies a case. In Figure 3, the case id is given by the variable *id*.
2. **transitionName**: *list of strings* that has the name of the task to log and the list ontological concepts to which this task is linked. Note that all strings in ML should be in quotes (""). In Figure 3, the task name for transition *FileFine* is "*FileFine*" and this task is semantically annotated with the SUPER concept "*fileDocumentTask*".
3. **eventType**: *list of strings*. If the event type is supported, the list should contain a *one* or *two* element and have the format [**name**, <**listOfOntologicalConcepts**>], where *name* in {"assign", "withdraw", "reassign", "start", "suspend", "resume", "complete", "autoskip", "manualskip", "pi\_abort", "ate\_abort"} . If the event type is *unknown*, this list should have *three* elements and the format [**"unknown"**, "**name**", <**listOfOntologicalConcepts**>], where *name* is the unknown event type name. In Figure 3, the event type for transition *FileFine* is [*"complete"*] and this event type is linked to the concept "complete" in the SUPER ontology "evo".
4. **timestamp**: *string* that represents the date and time in which the task was executed. The **timestamp** has the XML pre-defined format *dateTime* [5]. For instance, a valid timestamp string is "2005-06-30T14:55:00.000+01:00". The function `calculateTimeStamp()` is provided to automatically calculate the timestamp field based on the current time (in minutes) of a CP-net. The starting date of the function `calculateTimeStamp()` is the starting date of Unix (1-1-1970). The function `calculateTimeStamp()` is included in the file *loggingFunctionsMultipleFilesWithOntologySupport.sml*. In Figure 3, the function `calculateTimeStamp()` was used to provide the timestamp.
5. **originator**: *list of strings* that has the name of the originator (person or system) that executed the transition, plus the ontological concepts to which it belongs. In Figure 3, the user with *role1* executed the task *FileFine* and this user is linked to *two* concepts of the SUPER ontology: "secretary" and "finesDepartment".
6. **data**: *list of strings* containing the additional data fields that may be associated to a task. This list must have the format [attributeName1, attributeValue1, listOfOntologicalConcepts1, attributeName2, attributeValue2, listOfOntologicalConcepts1, ..., attributeNameX, attributeValueX, listOfOntologicalConceptsX]. In Figure 3, the data attribute *age* is logged for the task *FileFine*. Note that this attribute is linked to the concept "clientAge" and has the value set by the variable "age".

The parameters **timestamp**, **originator** and **data** can be empty (see optional fields in Figure 2). The first two are empty if the string "" is given as input. The **data** parameter is empty when [] is given as input.

The simulation of the extended CP-net will create the partial S-MXML log files whose aggregation is described in the next section.

## 4 Final Log Aggregation

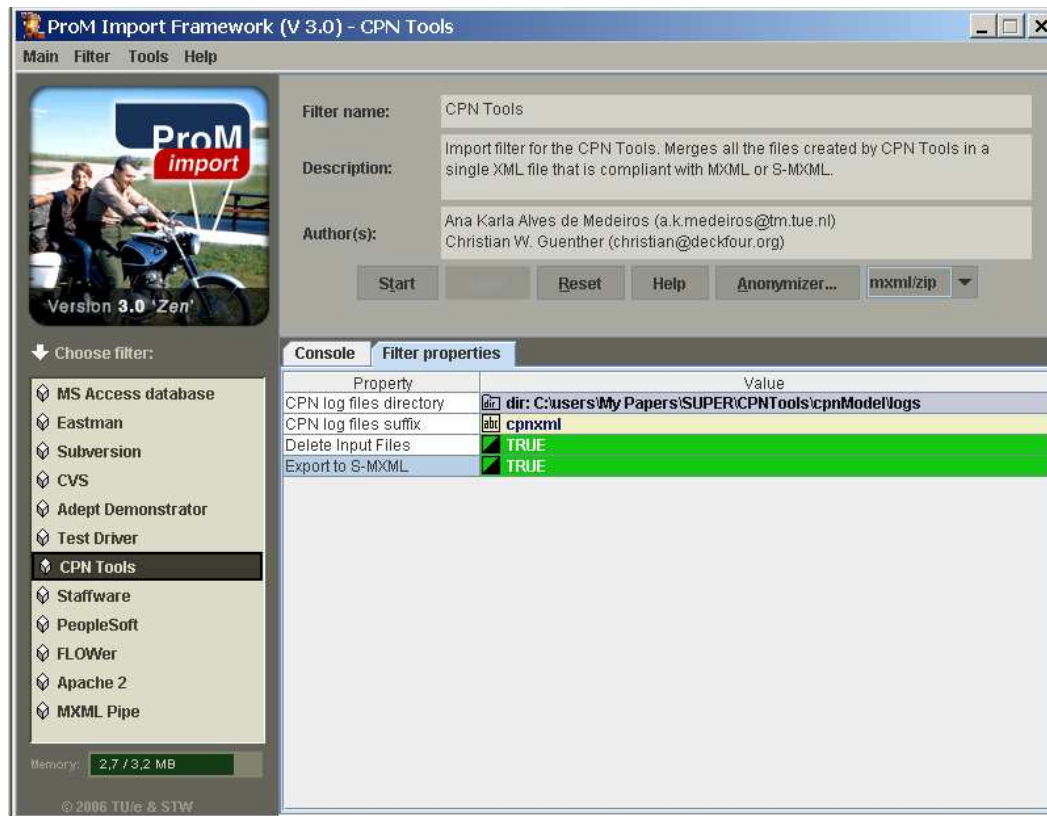
Once the CP-net has been simulated with logging extensions enabled, the generated log output has to be correctly aggregated and converted, such that it can be read and interpreted by mining tools like the ProM framework. As explained in Section 3, simulating a CP-net with logging extensions included and enabled will yield one log file per invocation of `createCaseFile`, including all logged events. The task of log aggregation is now to combine these files as process instances within one single S-MXML log file, representing all logs for the respective process model (i.e., the simulated CP-net).

This aggregation pass has been implemented as the “CPN Tools” import plug-in for the ProM<sub>import</sub> framework [2]. This framework has been developed to serve as a common environment for converting and importing logs from all kinds of information systems, and subsequently creating S-MXML compliant log files from them. The actual procedures for importing logs from a specific source system can be implemented as plug-ins, which can be dynamically loaded and removed from the running framework. The framework has a common graphical user interface for configuring and controlling import plug-ins, keeps all configuration data persistent, and provides a set of useful classes which can be used by all import plug-ins in order to ease development.

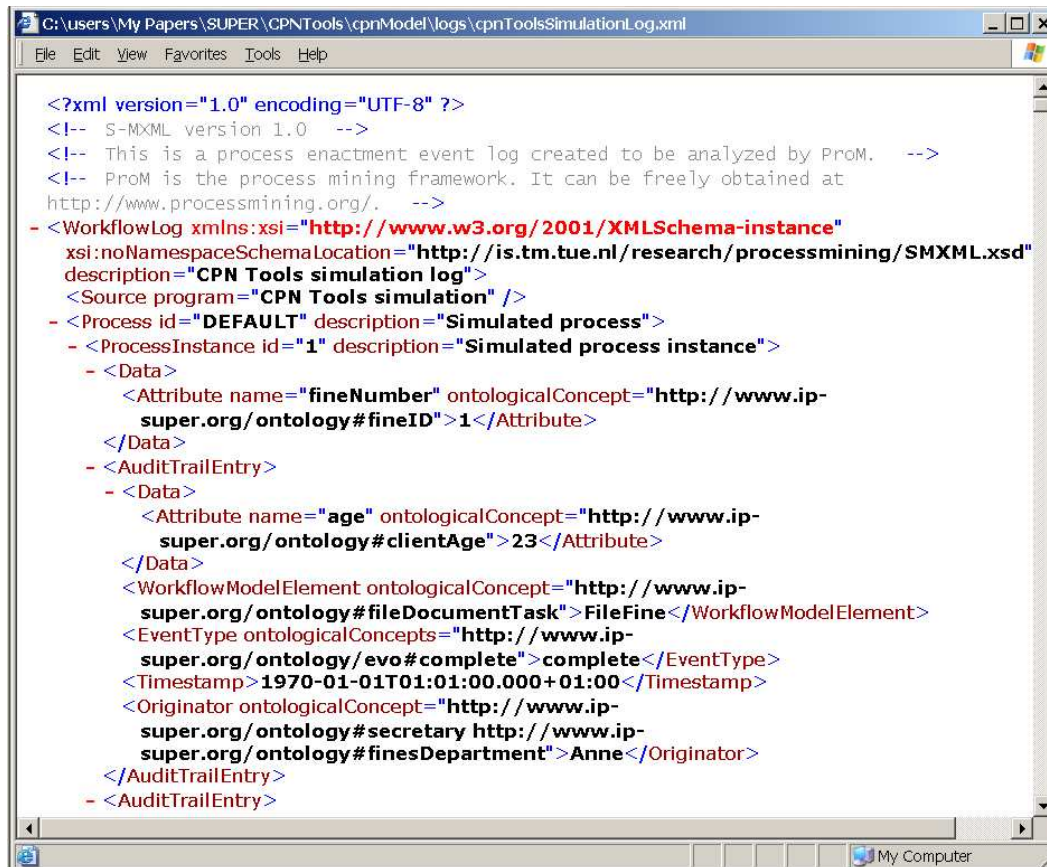
As the implemented ML functions `createCaseFile` and `addATE` (cf. Section 3) write S-MXML compliant log fractions during simulation runs, all that is left to do is to generate a common enclosing log file. For each file created by the simulation runs, a process instance is created within that common log file. In this process instance, all log events from one input file are added in their given order. In the end, the output log file thus includes all data from the simulation run logs in an aggregated manner, ready to be analyzed by mining tools like the ProM framework.

Using the “CPN Tools” import plug-in is fairly straightforward: The configuration pane (see Figure 4) has four filter properties that allow the user to (i + ii) select the input directory for the partial S-MXML files and the suffix of these files, (iii) set whether the partial S-MXML logs created during the simulation should be detected after aggregated into a single S-MXML file, and (iv) determine the target output format (MXML or S-MXML). After running the simulation passes in CPN Tools, the files created for each trace of one process model can be selected here. As this plug-in is geared towards aggregating logs that have resulted from executing one process model, it will accordingly write one single output log file. In this, the import framework provides the choice between writing the log into a compressed ZIP file, or as plain XML file. In either case, the resulting file can subsequently be loaded from within, for instance, the ProM framework, where a multitude of Process Mining plug-ins are available for analysis.

As an illustration, Figure 5 shows an excerpt of the S-MXML log that was aggregated for the simulation of the CP-net in Figure 3. Additionally, Figure 6 shows a screenshot with the results of applying four different ProM mining plug-ins for this S-MXML log. These plug-ins do not yet consider the semantic annotations while performing the mining over the S-MXML log. However, the results in Fig-

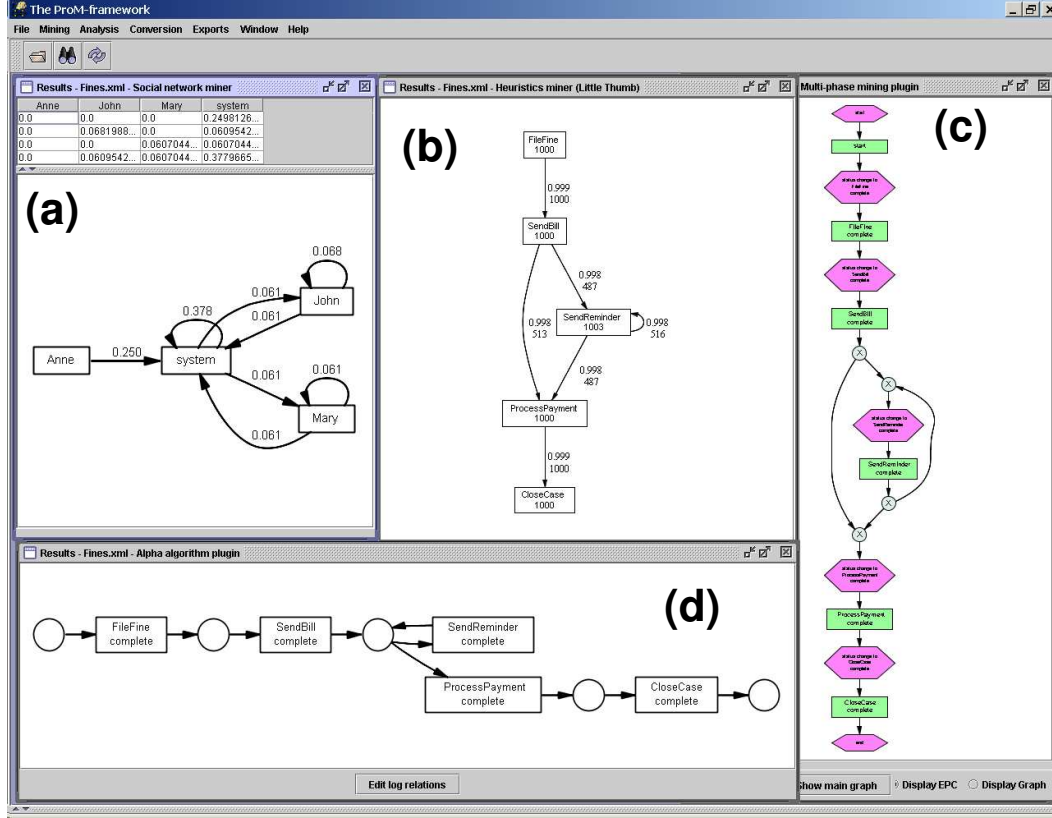


**Fig. 4.** Screenshot of the CPN Tools plug-in. This plug-in is implemented in the ProM<sub>import</sub> framework.



```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- S-MXML version 1.0 -->
<!-- This is a process enactment event log created to be analyzed by ProM. -->
<!-- ProM is the process mining framework. It can be freely obtained at
http://www.processmining.org/. -->
- <WorkflowLog xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://is.tm.tue.nl/research/processmining/SXML.xsd"
description="CPN Tools simulation log">
  <Source program="CPN Tools simulation" />
  - <Process id="DEFAULT" description="Simulated process">
    - <ProcessInstance id="1" description="Simulated process instance">
      - <Data>
        <Attribute name="fineNumber" ontologicalConcept="http://www.ip-
super.org/ontology#fineID">1</Attribute>
      </Data>
      - <AuditTrailEntry>
        - <Data>
          <Attribute name="age" ontologicalConcept="http://www.ip-
super.org/ontology#clientAge">23</Attribute>
        </Data>
        <WorkflowModelElement ontologicalConcept="http://www.ip-
super.org/ontology#fileDocumentTask">FileFine</WorkflowModelElement>
        <EventType ontologicalConcepts="http://www.ip-
super.org/ontology/evo#complete">complete</EventType>
        <Timestamp>1970-01-01T01:01:00.000+01:00</Timestamp>
        <Originator ontologicalConcept="http://www.ip-
super.org/ontology#secretary http://www.ip-
super.org/ontology#finesDepartment">Anne</Originator>
      </AuditTrailEntry>
    - <AuditTrailEntry>
```

Fig. 5. Excerpt of the S-MXML log that was aggregated for a simulation of the CP-net in Figure 3.



**Fig. 6.** Screenshot of the mined Petri net for the in Figure 5. The shown mining plug-ins are: (a) *Social network miner*, (b) *Heuristics miner*, (c) *Multi-phase mining* and (d) *Alpha algorithm*. The *Social network miner* plug-in can mine the *organizational perspective* (cf. Section 1) of an event log. Here we show the *handover of work* setting, considering only direct succession. Note that the users “John” and “Mary” never transfer work to each other. This is compatible with the CP-net in Figure 3. The other plug-ins in this figure can mine the *control-flow perspective* of the event log. As expected, all mined control-flow structures are like the CP-net in Figure 3.

ure 3 illustrate that the current process mining techniques can already be used to mine information from semantically annotated logs generated in the context of the SUPER project.

## 5 Conclusions and Future Work

This paper demonstrates how to benefit from the CPN Tools simulation capabilities to build event logs that can be used in the process mining research. The extension to CPN Tools consisted of implementing (i) a set of ML functions to create log files and (ii) a ProM<sub>import</sub> plug-in. The ML functions can be called from the input/output/action transition inscriptions of a CP-net. When the CP-net is simulated, partial logs are created. These partial logs are bundled into a single S-MXML log by using the ProM<sub>import</sub> CPN Tools plug-in. The resulting log can be mined by mining tools like the ProM framework. All the tools/files necessary for extending a CP-net to create S-MXML logs can be found at <http://www.processmining.org>.

As future work, we are going to use the framework presented in this paper to test the semantic process mining algorithms that we are going to develop in the SUPER project. Note that the use of synthetic logs will allow us to develop and test our plug-in *while* the SUPER engine is being developed. Once the engine is ready, we are going to use the logs it generates.

## Acknowledgements

The work presented in this paper was funded by the European Commission under the project SUPER (FP6- 026850).

## References

1. Extensible Markup Language (XML). <http://www.w3.org/XML/>.
2. Process mining website. <http://www.processmining.org>.
3. W.M.P. van der Aalst and M. Song. Mining Social Networks: Uncovering interaction patterns in business processes. In J. Desel, B. Pernici, and M. Weske, editors, *International Conference on Business Process Management (BPM 2004)*, volume 3080 of *Lecture Notes in Computer Science*, pages 244–260. Springer-Verlag, Berlin, 2004.
4. W.M.P. van der Aalst, B.F. van Dongen, J. Herbst, L. Maruster, G. Schimm, and A.J.M.M. Weijters. Workflow Mining: A Survey of Issues and Approaches. *Data and Knowledge Engineering*, 47(2):237–267, 2003.
5. P. V. Biron and A. Malhotra. XML Schema Part 2: Datatypes (Second Edition). <http://www.w3.org/TR/xmlschema-2/>, 2004.
6. A.K. Alves de Medeiros and C.W. Guenther. Process Mining: Using CPN Tools to Create Test Logs for Mining Algorithms. In K. Jensen, editor, *Proceedings of the Sixth Workshop on the Practical Use of Coloured Petri Nets and CPN Tools (CPN 2005)*, volume 576 of *DAIMI*, pages 177–190, Aarhus, Denmark, October 2005. University of Aarhus.

7. B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, and W.M.P. van der Aalst. The ProM framework: A new era in process mining tool support. In G. Ciardo and P. Darondeau, editors, *International Conference on Applications and Theory of Petri Nets (ATPN 2005)*, volume 3536 of *Lecture Notes in Computer Science*, pages 444–454. Springer-Verlag, Berlin, 2005.
8. B.F. van Dongen and W.M.P. van der Aalst. EMiT: A process mining tool. In J. Cortadelle and W. Reisig, editors, *International Conference on Applications and Theory of Petri Nets (ATPN 2004)*, volume 3099 of *Lecture Notes in Computer Science*, pages 454–463. Springer-Verlag, Berlin, 2004.
9. J. Herbst and D. Karagiannis. Workflow mining with inwolve. *Computers in Industry*, 53(3):245–264, 2004.
10. K. Jensen. *Coloured Petri Nets. Basic Concepts, Analysis Methods and Practical Use. Volume 1*. EATCS monographs on Theoretical Computer Science. Springer-Verlag, Berlin, 1997.
11. J. Herbst M. Hammori and N. Kleiner. Interactive workflow mining. In B. Pernici J. Desel and M. Weske, editors, *International Conference on Business Process Management (BPM 2004)*, volume 3080 of *LNCS*, pages 211–226. Springer Verlag, January 2000.
12. G. Schimm. Mining exact models of concurrent workflows. *Computers in Industry*, 53(3):265–281, 2004.
13. A.J.M.M. Weijters and W.M.P. van der Aalst. Rediscovering Workflow Models from Event-Based Data using Little Thumb. *Integrated Computer-Aided Engineering*, 10(2):151–162, 2003.