Process Modelling from Insurance Event Log

P.V. Kumaraguru
Research scholar,
Dr.M.G.R Educational and Research Institute University
Chennai- 600 095 India

Dr. S.P. Rajagopalan
Professor Emeritus,
Dr. M.G.R Educational and Research Institution University
Chennai-600095. India

ABSTRACT
Converting the mysterious mind process in to a most tangible, structured and understandable, process model is not only a great challenge of the day but also the need of the hour for many industries. Since 2005 the evolution of modern business industry has taken one step forward from business intelligence to business optimization. In the resent past all the business industries are in search of the means and ways to handle the data explosion of the digital universe. Machine learning and data mining are the only solutions to enable the business industries, not only to tackle the data explosion but also to convert the vital data to optimize the potential business resource. This paper has made an attempt to convert the event logs of the insurance process in to process model using petri net.

Keywords
Business intelligence, Data explosion, Digital universe, Event logs, process model, four eye, a-priori, token game, stochastic.

1. INTRODUCTION
The down pore of business data which are generated by the events in the business process has grown and become unimaginably huge and triggers many challenges for data mining, particularly for process mining. The main goal of process mining is to use event logs to extract process related information. For example, by modeling a business process and analyzing it, management may get ideas on how to reduce costs while improving service levels.

Data mining concepts have huge scope in insurance industry because it handles rich data sources. Till recent years many insurance industries suffered to handle obese databases. The evolution of Data mining discipline has turned the burden of handling database into a vital asset for the insurance industry. Data mining yields predictive models with which insurance industry can handle policy issuance, claims management, fraud detection, database segmentation, identify target customers, process optimization, new product development, and marketing strategies.

2. PROCESS MINING
Process mining is a bridge between data mining and business process management[1]. Process mining is applicable to a wide range of systems. These systems may be pure Information systems or systems where the hardware plays a more prominent role. The only requirement is that the system produces Event Logs, thus recording the actual behavior. Process mining is practically relevant and the logical next step in Business Process Management. Process mining provides many interesting challenges for scientists, customers, users, managers, consultants, and tool developers.

The goal of process mining is to extract information from event logs i.e; process mining describes a family of aposteriori analysis techniques exploiting the information recorded in the event logs. Typically these approached assume that it is possible to record sequential events such that each event refers an activity (well defined step in the process) and is related to a particular process instance. Furthermore, some mining techniques use additional information such as the performer or the originator of the event, the timestamp of the event or data elements recorded with the event.

Process mining (Fig-1) also aims to discover, monitor and improve real process by extracting knowledge from event logs. The three basic types of process mining are (1) discovery (2) conformance (3) extension

Discovery: Traditionally, process mining has been focusing on discovery i.e., deriving information about the organization context, and execution properties from enactment logs. An example of a technique addressing the control flow perspective is the alpha algorithm, which constructs Petri net model describing the behavior observed in the event logs. Process mining is not limited to process models and resent process mining techniques are more and more focusing on other perspectives e.g., the organizational perspective, performance perspective or data perspective. For example, there are approaches to extract social networks from event logs and analyze them using social network analyzer. This allows organizations to monitor how people, groups or software/system components are working together. Also there are approaches to visualizes performance related information e.g., there are approaches which graphically shows the bottlenecks and all kinds of performance indicators e.g., total flow time or the time spent between two activities.

Conformance: There is an a-priori model. This model used to check if reality conforms to the model. For example there may be a process model indicating that purchase orders for very huge amount require two checks . Another example is the checking of the so-called "four-eyes" principle. Conformation checking may be used to detect deviation, to locate and explain these deviations, and to measure the severity of these deviations.

Extension: There is an a-priori model. This model is extended with a new aspect or perspective, i.e., the goal is not to check conformance but to enrich the model with the data in the event log. An example is the extension of a process model with performance data, i.e., some a-priori process model is used on which bottlenecks are projected.

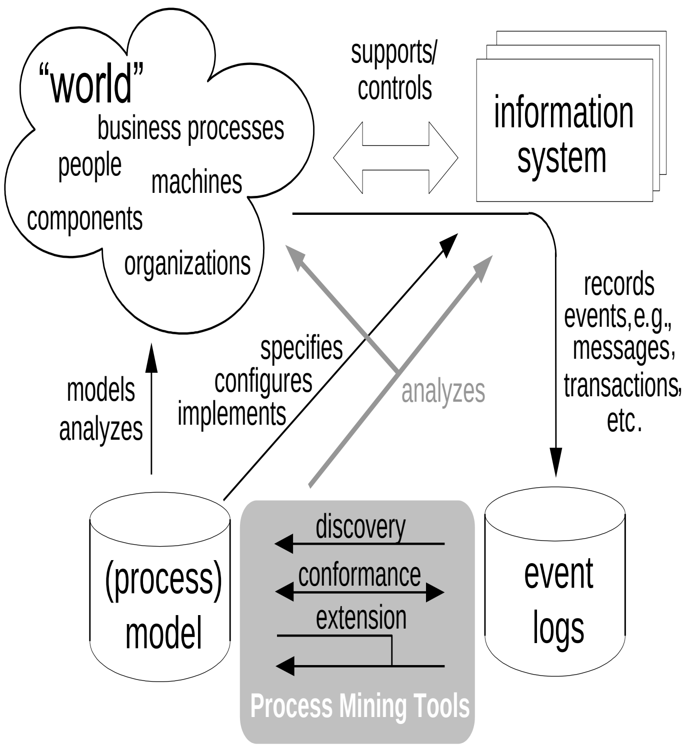

Fig- 1 . The components of process mining tools

3. PETRI NETS

Petri Nets are graphical and mathematical modeling notations. A Petri Net consists of places, transitions and arcs that connect them. Places are drawn as circles; Transitions are rectangles and arcs as arrows. In put arcs connect transitions with places. Places are passive components and are modeling the system state. They can contain TOKENS, depicted as black dots. The current state of the Petri net is given by the number of tokens on each place. Transitions are active components modeling activities which can occur and cause a change of the state by a new assignment of token to places. Transitions are only allowed to occur if they are enable, which means that there is at least on token on each input place. By occurring, the transition removes a token from each input place and adds a token on each output place. The repeated occurrence of transitions and the resulting sequence of marking is called the token game.

Petri nets are traditionally use for describing and analyzing systems that are characterized as concurrent, asynchronous, distributed, parallel nondeterministic and/or stochastic. Due to their graphical nature, Petri Nets can be used as a visualization technique like flow charts oar block diagrams but with much more scope on concurrency aspects. As a strict mathematical notation, it is possible to apply formal concepts like linear algebraic equations or probability theory for investigation the behavior model led system. In the last decade, Petri Nets have become a powerful concept in the area of business process modeling and work flow management.

Petri Net model consist of two parts

(1) The net structure that represents the static part of the system and

(2) A marking that represents the overall state on the structure. The token distribution among the places of a Petri net is called its marking. When one or more tokens reside in a place, the place is said to be marked, otherwise it is unmarked.

Petri nets have been successfully used to model and analyze processes from many domains such as software and business processes especially work flow processes.

The classical Petri net is a directed bipartite graph with two node types called *places* and *transitions*. The nodes are connected via directed *arcs*. Connections between two nodes of the same type are not allowed. Places are represented by circles and transitions by rectangles.

4. WORK FLOW MNAGEMENT

The term workflow management refers to the domain which focuses on the logistics of business processes. There are also people that use the term *office logistics*. The ultimate goal of workflow management is to make sure that the proper activities are executed by the right person at the right time.
Workflows are *case-based*, i.e., every piece of work is executed for a specific *case*. Examples of cases are a mortgage, an insurance claim, a tax declaration, an order, or a request for information. Cases are often generated by an external customer. However, it is also possible that a case is generated by another department within the same organization (internal customer). The goal of workflow management is to handle cases as efficiently and effectively as possible. A workflow process is designed to handle similar cases. Cases are handled by executing *tasks* in a specific order. The *workflow process definition* specifies which tasks need to be executed and in what order. Since tasks are executed in a specific order, it is useful to identify *conditions* which correspond to causal dependencies between tasks. A condition holds or does not hold (true or false). Each task has pre- and post conditions: the preconditions should hold before the task is executed, and the post conditions should hold after execution of the task.

Many cases can be handled by following the same work flow process definition. As a result, the same task has to be executed for many cases.

A task which needs to be executed for a specific case is called a *work item*. Most work items are executed by a *resource*. A resource is either a machine (e.g. a printer or a fax) or a person (participant, worker, and employee). In most offices the resources are mainly human. However, because workflow management is not restricted to offices, we prefer the term resource. Resources are allowed to deal with specific work items. To facilitate the allocation of work items to resources, resources are grouped into classes. A *resource class* is a group of resources with similar characteristics.

A work item which is being executed by a specific resource is called an *activity*. Work items link cases and tasks. Activities link cases, tasks, and resources.

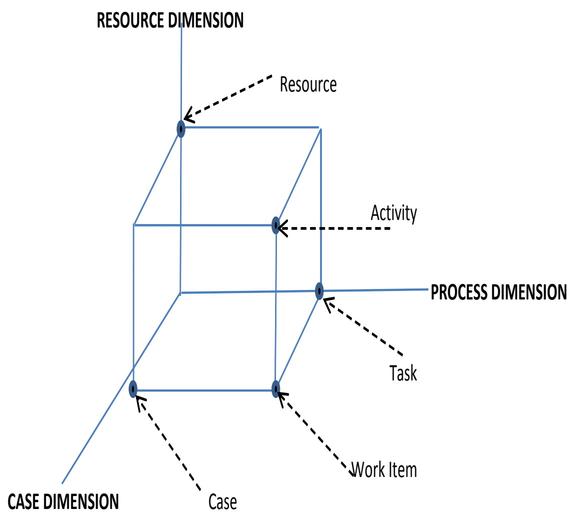

Fig- 2. A three dimensional view of a workflow

Figure 2 shows that a workflow has three dimensions: (1) the case dimension, (2) the process dimension and (3) the resource dimension.

The case dimension signifies the fact that all cases are handled individually.

In the process dimension, the workflow process, i.e., the tasks and the routing along these tasks, is specified.

In the resource dimension, the resources are grouped into roles and organizational units.

Figure 2 shows that work flow management is the glue between the cases, the tasks, and the organization.

5. MAPPING MOTOR CLAIM PROCESSING ONTO PETRI NETS

To illustrate the mapping of workflow management concepts onto Petri nets we consider the processing of motor claims in an insurance company [7].

First the customer registers the claim with the insurance company (task *register the claim*). The loss is recorded in the claim register (task *examine*) and a claim form is issued to the customer. The insurance company then verifies the policy records to see if the policy is in force and also confirms 64VB clause of the Insurance Act, 1938 of India (task *Check 64 VB confirmation*). The insurance company then decides (task *decide*) and either accepts the claim (task *Honour the claim*) or rejects the claim (task *Repudiate the claim*).

The insured is required to submit a detailed estimate of repairs from any repairer of his choice. Generally, these repairs are acceptable but at times the insurance company asks the customer to obtain repair estimate from another repairer, if they have reason to believe that the competence, moral hazard or business integrity of the first repairer is not satisfactory (task *Panel assessment*).Verification of all the claim documents by the approved loss assessor or the surveyor is the task *Check documents.*

Note: Section 64 VB of the Insurance Act, 1938 of India stipulates that no risk can be assumed without prior payment of full premium except when

(i) The entire amount of the premium is guaranteed to be paid by a bank before the end of the first calendar month after the month in which the risk is assumed, or

(ii) An advance deposit is made with the insurer to the credit of the Insured sufficient to cover the payment of the entire amount of premium

Table 1: A fragment of insurance event log :each line corresponds to an event

		PROPERTIES			
		TIMESTAMP			
CASE ID	**EVENT ID**	**DATE**	**TIME**	**ACTIVITY**	**RESOURCE**
1	4587320	11/8/2011	15:05	Register claim request	Raj
	4587321	11/8/2011	15:30	Examine	Sam
	4587322	11/8/2011	15:45	Check 64VB confirmation	Mano
	4587323	11/8/2011	16:30	Decide	Ravi
	4587324	11/8/2011	16:45	Repudiate the claim	Mani
2	4587331	26/04/2011	11:30	Register claim request	Sam
	4587332	26/04/2011	11:45	Check 64VB confirmation	Mano
	4587333	26/04/2011	12:00	Examine	Raj
	4587334	26/04/2011	12:30	Decide	Ravi
	4587335	26/04/2011	12:45	Honor the claim	Mani
3	4587341	29/06/2011	18:20	Register claim request	Raj
	4587342	30/06/2011	9:45	Examine	Ravi
	4587343	30/06/2011	10:15	Check 64VB confirmation	Mani

	4587344	30/06/2011	10:40	Decide	Mano
	4587345	30/06/2011	15:20	Panel assessment	Sam
	4587346	30/06/2011	16:00	Examine	Mano
	4587347	1/7/2011	10:00	Check documents	Sam
	4587348	1/7/2011	11:00	Decide	Mano
	4587349	1/7/2011	13:15	Honor the claim	Ravi
4	4587351	3/3/2011	10:00	Register claim request	Ravi
	4587352	3/3/2011	11:00	Examine	Mani
	4587353	3/3/2011	11:15	Check 64VB confirmation	Sam
	4587354	3/3/2011	12:10	Decide	Mano
	4587355	3/3/2011	15:25	Panel assessment	Sam
	4587356	3/3/2011	17:00	Check documents	Mano
	4587357	3/3/2011	17:45	Examine	Sam
	4587358	4/3/2011	9:30	Decide	Mano
	4587359	4/3/2011	10:15	Honor the claim	Mano
5	4587391	17/09/2011	11:30	Register claim request	Sam
	4587392	17/09/2011	11:45	Examine	Mano
	4587393	17/09/2011	12:30	Check 64VB confirmation	Mano
	4587394	17/09/2011	12:45	Decide	Ravi
	4587395	17/09/2011	15:30	Panel assessment	Mani
	4587396	17/09/2011	14:10	Check documents	Raj
	4587397	17/09/2011	15:00	Examine	Ravi
	4587398	17/09/2011	15:30	Decide	Ravi
	4587399	17/09/2011	15:45	Repudiate the claim	Mani

The above table represents the information in an event log. The bare minimum requirements for process mining are that an event can be related to both a case and an activity.

With the information from the above table we obtain the more compact information representation of the event log as shown below.

Table 2 :Trace of the event log

CASE ID	TRACE
1	(a,b,c,d,h)
2	(a,c,b,d,g)
3	(a,b,c,d,e,b,f,d,g)
4	(a,b,c,d,e,f,b,d,g)
5	(a,b,c,d,e,f,b,d,h)

In the above table, the case is represented by a sequence of activities referred to as trace and the activity names are represented by single letter labels. The various labels mentioned in the table denotes as

- a – Register claim request
- b – Examine
- c - Check 64VB confirmation
- d – Decide
- e - Panel assessment
- f – Check documents
- g – Honor the claim
- h – Repudiate the claim

The information given in the above table can be transformed into a process model.

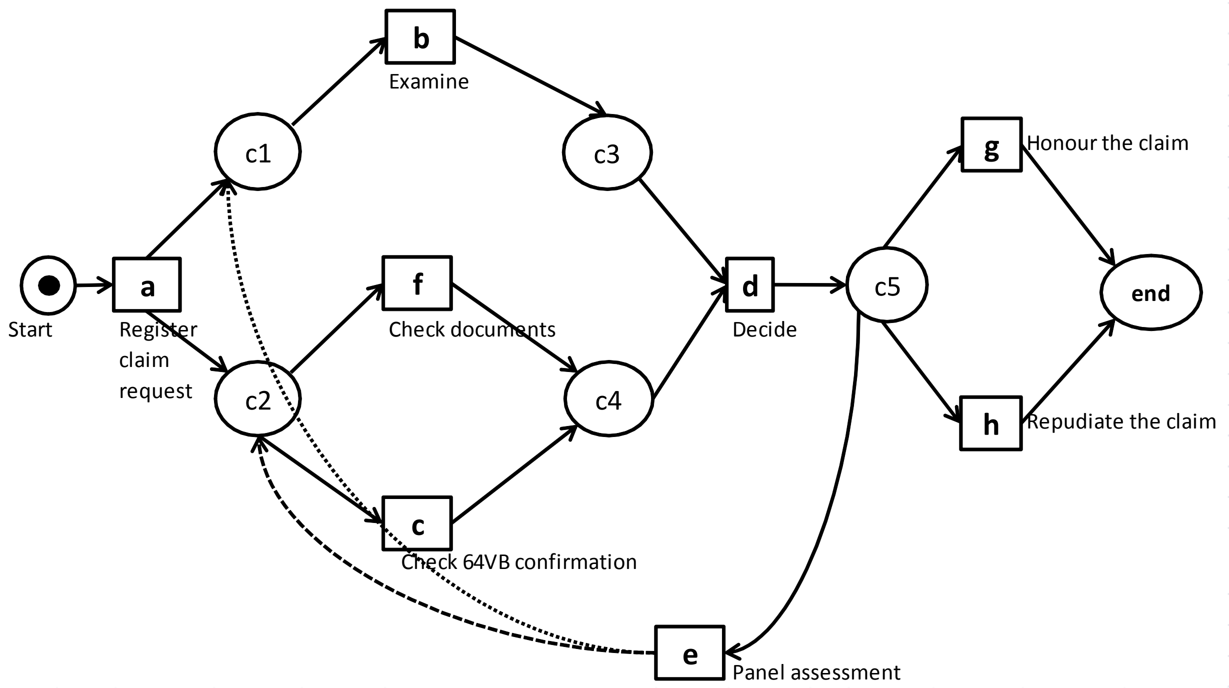

Fig – 3 Process model

Let us relay the trace of the first case – (a, b, c, d, h) to show that the trace fits to the model.

With the help of the token (Fig-3) in **Start, a** is indeed enabled. After firing **a** places c1 and c2 are marked with tokens. **b** is then enabled at this marking and its execution results in the tokens c2 and c3. The next event **c** is enabled and this results in marking enabling **d** and further firing **d** results in the marking with one token in c5. This marking enables the final event **h** in the trace.

After executing **h** the case ends in the final marking with just a token in the place **end.** All the other four traces can also be checked in the model and all these traces result in marking with just a token in the place **end.**

6. CONCLUSION

In this paper we discussed the application of Petri nets to the insurance claim process workflow. Workflow management turns out to be an application domain which could benefit from the features of Petri nets. Workflow management is a very rewarding application domain for Petri nets. Flaws in handling the claims process of the customer by the insurance company will lead to very low service levels and subsequent retention of the customer would become very expensive. It has further possibility to identify bottlenecks, anticipate problems, record policy violations, recommend countermeasures, remodel and streamline the processes.

7. REFERENCES

[1] Wil M.P. Van der Aalst, 2010,Process Mining, Springer

[2] W.M.P. Van der Aalst 1998 The applications of Petri Nets to wotk flow management , The journal of Circuits, Systems and computers

[3] W.M.P. Van der Aalst and K.M.Van Hee. Workflow Management: Models, Methods , and system. MIT press, Cambridge, MA, 2002.

[4] S.Kumaran and K.Raja, Modeling and simulation of Projects with Petri Nets. American journal of applied Sciences, 2008.

[5] A.J.M.M Weijters and W.M.P Van der Aalst, Process Mining : 2001 Discovering workflow Models from Event based Data, 13th Belgium- Netherlands Conference on Artificial intelligence(BNAIC 2001).

[6] W.M.P van der Aalst A.J.M.M. Weijters, and L. Maruster. 2004, Work flow Mining: Discovering process models from Event Logs. IEEE Transactions on Knowledge and Data Engineering

[7] H.Ananathakrishnan,V.A.Pai 2004, Motor Insurance, Insurance Institute of India.