Formal Approaches to Business Processes through Petri Nets

Nick Russell
Arthur H. M. ter Hofstede
Acknowledgement

These slides summarize the content of Chapter 2 of the book:


These slides have been prepared by or inspired by slides of the following people:

- Wil van der Aalst, TUE & QUT
- Michael Adams, QUT
- Lachlan Aldred, QUT
- Bartek Kiepuszewski, Moreton, BMS, Cutter Consortium
- Marcello La Rosa, QUT
- Petia Wohed, SU/KTH
- Moe Wynn, QUT
• One of the frequent criticisms of modeling notations is that they are *imprecise* and, as a consequence, *subject to varying interpretations by different parties*.

• Describing a candidate modeling notation in terms of a **formal technique** provides an effective means of minimizing the potential for ambiguity.

• To do so, it is necessary to describe both the *syntax* and *semantics* of the modeling formalism using a well-founded technique.

• Suitable techniques for doing so generally stem from mathematical foundations and include general-purpose modeling approaches such as **Petri nets**.
Petri Nets

• Petri nets have proven to be a particularly effective mechanism for modeling the dynamic aspects of processes.
• Petri Nets they have three specific advantages:
  – Formal semantics despite the graphical nature.
  – State-based instead of event-based.
  – Abundance of analysis techniques.
• Many variant exist: we will analyze workflow nets and reset nets.
• Petri Nets have been chosen as the formal underpinning for the YAWL language (it will be discussed in an upcoming lecture).
Overview

• Formal foundations for modeling languages in BPM
  – Petri nets
  – Some fundamental results
  – Workflow nets
  – Mapping workflow concepts to Petri nets
  – Reset nets
Petri Nets

- They were originally conceived as a technique for the description and analysis of concurrent behaviour in distributed systems.
- Based on a few simple concepts, yet expressive.
- They have a simple graphical format and a precise operational semantics that makes them an attractive option for modeling the static and dynamic aspects of processes.
- Many analysis techniques exist.
- Many extensions and variants have been defined over the years.
Applications

• Applications in many different areas, such as databases, software engineering, formal semantics, etc.
• There are two main uses of Petri nets for workflows:
  – Specifications of workflows.
  – Formal foundation for workflows (semantics, analysis of properties).
A Petri Net takes the form of a **directed bipartite graph** where the nodes are either **places** or **transitions**.

**Places** represent **intermediate states** that may exist during the operation of a process. Places are represented by circles.

Places can be input/output of **transitions**. Transitions correspond to the **activities** or **events** of which the process is made up. Transitions are represented by rectangles or thick bars.

**Arcs** connect places and transitions in a way that places can only be connected to transitions and vice-versa.
Petri Nets: Definition

• Formally a Petri net \( N \) is a triple \((P, T, F)\) where
  – \( P \) is a finite set of places
  – \( T \) is a finite set of transitions where \( P \cap T = \emptyset \)
  – \( F \subseteq (P \times T \cup T \times P) \) is the set of arcs known as the flow relation

• A directed arc from a place \( p \) to a transition \( t \) indicates that \( p \) is an input place of \( t \). Formally:
  – \( p \cdot \{t \in P \mid (p, t) \in F\} \)

• A directed arc from a transition \( t \) to a place \( p \) indicates that \( p \) is an output place of \( t \). Formally:
  – \( t \cdot = \{p \in P \mid (t, p) \in F\} \)
  – With an analogous meaning, we can define:
    – \( p \cdot = \{t \in T \mid (p, t) \in F\} \) and \( \cdot p = \{t \in T \mid (t, p) \in F\} \)
Petri Net: Example

\[ P = \{p_1, p_2, p_3, p_4\} \]
\[ T = \{t_1, t_2, t_3\} \]
\[ F = \{(p_1, t_1), (p_2, t_1), (t_1, p_3), (p_2, t_2), (t_2, p_4), (p_4, t_3), (t_3, p_2)\} \]
\[ t_1 \bullet = \{p_3\}; t_1 = \{p_1, p_2\}; p_2 = \{t_3\}; p_1 = \emptyset; p_2 \bullet = \{t_1, t_2\} \]
Petri Nets: Example

\[ P = \ldots \]
\[ T = \ldots \]
\[ F = \ldots \]

\[ t_1 \bullet = \ldots ; \bullet t_1 = \ldots ; \bullet p_2 = \ldots ; p_2 \bullet = \ldots. \]
Markings

- The operational semantics of a Petri Net is described in terms of particular marks called **tokens** (graphically represented as black dots ●).
- Places in Petri Nets can contain **any number of tokens**. The distribution of tokens across all of the places in a net is called a **marking**. For a Petri net an **initial marking** $M_0$ needs to be specified.
- Marking assigns tokens to places; formally, a marking $M$ of a Petri net $N = (P,T,F)$ is a function $M: P \rightarrow \text{NAT}$.

- The marking below is formally captured by the following marking $M = \{(p_1,1),(p_2,2),(p_3,0)\}$.
State of a Petri Net

A state can be compactly described as shown in the following example:

- \(1p_1 + 2p_2 + 0p_3\) is the state with one token in place p1, two tokens in p2 and no tokens in p3.
- We can also represent this state in the following (equivalent) way: \(p_1 + 2p_2\).

We can also describe an ordering function \(\geq\) over the set of possible states such that, given a Petri net \(N = (P,T,F)\) and markings \(M\) and \(M'\), \(M \geq M'\) iff for all \(p\) in \(P\): \(M(p) \geq M'(p)\). \(M > M'\) iff \(M \geq M'\) and \(M \neq M'\).
Enabled Transitions

• The operational semantics of Petri nets are characterized by the notion of a transition executing or "firing". A transition in a Petri net can "fire" whenever there are one or more tokens in each of its input places.

• The execution of a transition occurs in accordance with the following firing rules:

  1. A transition $t$ is said to be **enabled** if and only if each input place $p$ of $t$ contains at least one token. **Only enabled transitions may fire.**

     – Formally, a transition $t$ is enabled in a marking $M$ iff for each $p$, with $p \in \bullet t$, $M(p) > 0$. *(see definition 2.7 of [DE95]*)

  2. If transition $t$ fires, then $t$ **consumes one token** from each input place $p$ of $t$ and **produces one token** for each output place $p$ of $t$.

  ![Diagram 1](T1 is enabled and may fire)

  ![Diagram 2](T1 is not enabled!)

  ![Diagram 3](T1 fires! When a transition fires, the marking and the state of the Petri Net change.)
Firing a Transition: Example

BEFORE

AFTER
Firing Transitions: Further Examples

- It is assumed that the firing of a transition is an atomic action that occurs instantaneously and cannot be interrupted.
- If there are multiple enabled transitions, any one of them may fire; however, for execution purposes, it is assumed that they cannot fire simultaneously.
- An enabled transition is not forced to fire immediately but can do so at a time of its choosing.
- These features make Petri nets particularly suitable for modeling concurrent process executions.
Firing Transitions

- Given a Petri Net \((P,T,F)\) and an initial state \(M\), we have the following notations that characterize the firing of a given transition \(t\):
  - \(M \rightarrow^t M'\) indicates that if transition \(t\) is enabled in state \(M\), then firing \(t\) in \(M\) results in state \(M'\). Formally, notation \(M \rightarrow^t M'\), is defined by:
    - \(M'(p) = M(p)\) if \(p \notin \bullet t \cup t\bullet\) or \(p \in \bullet t \cap t\bullet\)
    - \(M'(p) = M(p) - 1\) if \(p \in \bullet t\) and \(p \notin t\bullet\)
    - \(M'(p) = M(p) + 1\) if \(p \in t\bullet\) and \(p \notin \bullet t\)
  - \(M \rightarrow M'\) indicates that there is a transition \(t\) such that \(M \rightarrow^t M'\).
  - \(M \rightarrow^\sigma M'\) denotes the firing sequence \(\sigma = t_0 t_1 \cdots t_{n-1}\) that leads from state \(M\) to state \(M'\), such that
    \[
    M = M_0 \rightarrow^{t_0} M_1 \rightarrow^{t_1} M_2 \cdots M_{n-1} \rightarrow^{t_{n-1}} M_n = M'
    \]
    Note that the transitions do not have to be different!
  - A state \(M'\) is called **reachable** from state \(M\) (we write \(M \rightarrow^* M'\)) iff there is a firing sequence \(\sigma\) that leads from state \(M\) to state \(M'\).
    - Informally, a marking is reachable from another marking if there is a sequence of transitions that can fire from the first marking to arrive at the second marking.
Petri nets: Order Fulfillment Example

First, a **take order** task is executed.

When **pack order** and **check account** tasks have been both completed, the **credit check** task is executed.

If the customer has not sufficient credit the **decline order** runs and, finally, the **return stock** task ensures that the items from the order are returned to the warehouse.

Then, **pack order** and **check account** tasks are executed in parallel.

If the customer has sufficient credit remaining, the order is **despatched**.
Petri nets: Example of a vending machine (source [DE95] p. 4)

candy storage

ready for insertion

insert coin

refill

dispense candy

reject coin

holding coin

accept coin

request for refill

ready to dispense

© 2009, www.yawlfoundation.org
Elevator (1)

The number of tokens represents the number of vertical movements an elevator can make upwards or downwards in a certain state of the system.

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Petri net example: Elevator 2

Elevator (2)

The token represents the elevator.

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Petri net example: Elevator 3

Elevator (3)
The token is a person using the elevator.

Go back to the elevator

© 2009, www.yawlfoundation.org

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
We want to model with a Petri Net the behaviour of two traffic lights at an intersection, in a way that they cannot be green or yellow at the same time. Conversely, they are allowed to signal red at the same time.
Solution Traffic Lights

Two traffic lights

There are two traffic lights, which are not allowed to signal green at the same time.

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Homeworks

• Two traffic lights at an intersection. If one is red, the other should be green etc. *(many discussions on modelling traffic lights through Petri nets can be found on the internet).*

• A producer and a consumer producing and consuming (resp.) indefinitely. The consumer cannot consume more than the producer has produced thus far. How does your model change if the buffer between them is of limited size? *(this is a well-known concurrency problem)*

• Two parallel processes with two critical sections. If one of the two processes is in its critical section, the other process should not be able to enter its critical section and vice versa. *(this is also a well-known concurrency problem)*
Coverable Markings

- Coverability is a weaker notion than reachability. A marking $M$ is **coverable** iff a reachable marking $M'$ exists such that $M' \geq M$ (see e.g. Definition 5 in [HAAR09]).
- *Example*: Given the Petri net and marking in figure, $p_1+p_2+p_3$ is a reachable marking, while $p_1+p_3$ is a coverable marking (but not reachable).

- To decide whether a given marking $M$ is reachable is a $\text{DSPACE}(\text{exp})$-hard problem.
Properties

- A Petri net $N$ with initial marking $M_0$ is **live** iff for every reachable marking $M$ and every transition $t$ there exists a marking $M'$ reachable from $M$ which enables $t$. (see definition 2.16 of [DE95]).

- More informally: A Petri net with initial marking $M_0$ is live if, no matter what marking has been reached from $M_0$, it is possible to ultimately fire any transition by progressing through some further firing sequence.

- The notion of **liveness** is important since it demonstrates that at least one transition can fire in every reachable state. A live Petri net guarantees **deadlock-free operation**.

- A Petri net $N$ with initial marking $M_0$ is **deadlock free** iff every reachable marking enables some transition (see definition 2.16 of [DE95]).
Properties

• A Petri net $N$ with initial marking $M_0$ is **k-bounded** iff for every reachable marking $M$, $M(p) \leq k$ ($k$ is the minimal number for which this holds). (see definition 2.20 of [DE95]).
  
  – A 1-bounded net is called **safe**.
  
  – The property of **boundness** ensures that the number of tokens cannot grow arbitrarily.

• A Petri net $N$ is **strongly connected** iff for every pair of nodes (places or transitions) $x$ and $y$ there is a path from $x$ to $y$ and vice-versa.
Is this Petri Net live and bounded?

A bounded but non-live Petri net

M0 = (1,0,0,1)
M1 = (0,1,0,1)
M2 = (0,0,1,0)
M3 = (0,0,0,1)
Is this Petri Net live and bounded?

An unbounded but live Petri net

M0 = (1, 0, 0, 0, 0)
M1 = (0, 1, 1, 0, 0)
M2 = (0, 0, 0, 1, 1)
M3 = (1, 1, 0, 0, 0)
M4 = (0, 2, 1, 0, 0)
Exercise

• Is the vending machine live?
• Is it deadlock free?
• Is it bounded?
• Is it strongly connected?
• Can a marking be reached with tokens both in “ready for insertion” and “ready to dispense”?
• Give an example of a marking that is coverable but not reachable.
Free Choice Petri nets

- Many verification problems in Petri nets have a high complexity.
- **Free Choice Petri nets** are a subclass of Petri nets with a “nice” tradeoff between expressiveness and analyzability (see e.g. [DE95]).
- All elementary workflow concepts are essentially free choice.
- In a Free Choice Petri net “the result of the choice between two transitions can never be influenced by the rest of the system” [DE95]
Example of a Conflict
• In a Free Choice Petri net, every pair of transitions either share all their input places, or they share none.

• Formally, a Petri net $N = (P,T,F)$ is free choice iff for all transitions $t,t'$:
  - $\bullet t \cap \bullet t' \neq \emptyset \Rightarrow \bullet t = \bullet t'$
  - For any free-choice net, a $t'$ in conflict with an enabled transition $t$, is also enabled.
Wil van der Aalst has proposed the use of Petri nets for workflow modelling. In [Aalst96] three benefits are argued:

- Petri nets are formally defined;
- Petri nets support the notion of being “in between” performing tasks through the notion of place;
- Petri nets have associated analysis techniques.

He proposes a particular subclass of Petri nets, called Workflow nets (WF-nets) for this purpose.

In a workflow net, transitions represent the tasks that comprise a business process and places represent the conditions preceding and following the tasks.
Workflow nets:
Definition

• A workflow net has a single **start place** and a single **end place**.
  – This means that workflow nets closely correspond to real-life processes
    that tend to have a specific starting point and a specific end point.

• Every transition in the workflow net is on a **path from the start to the end place**.
  – This ensures that each transition in a workflow net contributes to the
    progression of an executing instance towards its end state.

• **Definition [AH02, p271-272]** A Petri net $PN = (P, T, F)$ is a
  WF-net (Workflow net) if and only if:
  • There is one source place $i \in P$ such that $\cdot i = \emptyset$
  • There is one sink place $o \in P$ such that $o \cdot = \emptyset$
  • Every node $x \in P \cup T$ is on a path from $i$ to $o$. 
Exercise: Candidate WF-nets?
Workflow nets:
Definition

• It is important to note that the previous definition traces the minimal requirements for a workflow net.

• However, it does not guarantee (by itself) that a candidate workflow net will not potentially be subject to deadlock or livelock.

• To ensure that any given process instance behaves in a predictable way, in [AH02] a number of so-called soundness criteria are formulated.
Definition [soundness]: A procedure modeled in the form of a WF-net $PN = (P, T, F)$ is sound if and only if:

- **[Option to Complete]** Given an initial marking $i$, from every marking $M$ reachable from $i$, there exists a firing sequence leading from state $M$ to state $o$. Formally:

  $$\forall M \left( (i \rightarrow M) \Rightarrow (M \rightarrow o) \right)$$

  - Basically, this means that the any executing instance of the workflow net must eventually terminate, i.e., net is free of deadlock and infinite loops.

- **[Proper Completion]** State $o$ is the only state reachable from state $i$ with at least one token in place $o$. Formally:

  $$\forall M \left( (i \rightarrow M \land M \geq o) \Rightarrow (M = o) \right)$$

  - When the workflow terminates no other tasks are still running and termination is signalled only once. At the moment of termination, there must be one token in the end place $o$ and all other places in the WF-net must be empty.

- **[No Dead Tasks]** For every transition $t$, a marking $M$ reachable from $i$ ($i \rightarrow^* M$) can be found that enables $t$.

  - The workflow does not contain any superfluous parts that can never be activated. In a nutshell, dead transitions are not allowed.
**Workflow Net Constructs**

**Automatic tasks** execute as soon as they are enabled.

**User tasks** are passed to human resources for execution once enabled.

In WF-net there are some notational enhancements (often termed “syntactic sugar”) for **split** and **join** constructs that simplify the specification of a workflow net.

**External tasks** only proceed once they are enabled and a required message or signal is received from the operating environment.

**Time tasks** only proceed once they are enabled and a specified (time-based) deadline occurs.

The basics of Petri nets can be used to understand the semantics of some elementary modeling concepts in WF-nets.
Parallelism: AND-split

- According to the WfMC [WfMC], an **AND-split** is “a point within the workflow where a single thread of control splits into two or more threads which are executed in parallel within the workflow, allowing multiple activities to be executed simultaneously.”

- The execution of A enables both task B and task C. As a result, task B and task C are executed in parallel (in an arbitrary order).

- In WF-nets, a special construct for AND-split is introduced.
According to the WfMC [WfMC], an **AND-join** is “a point in the workflow where two or more parallel executing activities converge into a single common thread of control.”

Task D is enabled after execution both B and C, i.e., D is used to synchronize two subflows.

In WF-nets, a special construct for AND-JOIN is introduced.
According to the WfMC [WfMC], a **XOR-split** is “a point within the workflow where a single thread of control makes a decision upon which branch to take when encountered with multiple alternative workflow branches.”

Note that the exclusive nature of the choice, i.e. only one of the outgoing branches can be chosen (i.e., either task B or C can be executed).

In WF-nets, a special construct for XOR-split is introduced.
According to the WfMC [WfMC], a **XOR-join** is “a point within the workflow where two or more alternative activity(s) workflow branches re-converge to a single common activity as the next step within the workflow.”

- As no parallel activity execution has occurred at the join point, no synchronization is required.
- Therefore, D is enabled when B or C complete.
- In WF-nets, a special construct for XOR-JOIN is introduced.
The **take order** task is externally triggered when an order request is received.

The **decline order** task runs automatically with the customer receiving a notification either by email or fax.

Most tasks are undertaken by human resources (i.e., staff).
In [Aalst97] it was shown that soundness for a WF-net could be determined in terms of liveness and boundedness. In [AH02] p.276 this is explained as determining that a workflow net PN is sound is equivalent to determining to whether the net PN’ which is constructed through the addition of an extra transition t, where •t = {o} and t• = {i}, is live and bounded.

As pointed out in [AH02] p.277, the computational complexity of determining whether a WF-net is sound may be quite high. Restrictions (e.g. requiring the net to be free choice) can be imposed to make this more tractable, see the discussion in [AH02] p277-286.

At Eindhoven University of Technology the Workflow Analyzer (WOFLAN) was developed which is freely available for download.
Workflow Animation – Erroneous WF

Incorrect workflow 1

This is an incorrect workflow model

As an exercise do the following things:

1. Find the error
2. Give a scenario showing the error
3. Give a scenario without any errors (i.e. how to reach the state with just one token in end)

Press the instructions button to close this panel

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Incorrect Workflow 2

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Workflow Animation – Correct WF

Workflow (2)

Animation by Wil van der Aalst, Vincent Almering and Herman Wijbenga
Reset nets

• While Petri nets and Workflow nets provide an effective means of modeling a business process, **they are not able to capture the notion of cancelation.**
  – This is a significant omission, given the relative frequency of cancelation in real-life processes.

• **Reset nets** extend Petri nets with a special type of arc, the **reset arc**, to capture the notion of cancelation.

• Like normal input arcs, reset arcs **connect places to transitions**.

• However, they operate in a different way: when a transition to which a reset arc is connected fires, **any place connected to the transition by reset arcs is emptied of any tokens that it may contain.**

• Note that the tokens in places attached to a transition by reset arcs **play no part in the enablement of the transition.**
  – This makes reachability **undecidable**.
Reset nets

place  transition  arc  reset arc
Reset nets: formal definition

- Formal definitions are based on [DFS98, FRSB02, FS01].
- Syntactically a **Reset net** is a tuple \((P, T, F, R)\) where
  - \((P, T, F)\) is a Petri net with a finite set of places \(P\), a finite set of transitions \(T\), and a flow relation \(F \subseteq (P \times T \cup T \times P)\);
  - \(R : T \to 2^P\) is a function associating reset places with transitions.
- A reachable marking \(M'\) is defined by:
  - first removing the tokens needed to enable \(t\) from its input places \(\bullet t\);
  - then removing all tokens from reset places associated with \(t\) (i.e., \(R(t)\));
  - and finally by adding tokens to its output places \(t\bullet\).
- Let \(N = (P, T, F, R)\) be a Reset net and \(M\) a marking.
  - A transition \(t \in T\) is enabled iff \(\bullet t \leq M\).
  - An enabled transition \(t\) can fire thus changing the state to \(M'\), denoted \(M \rightarrow^{t} M'\), with \(M' = (M - \bullet t)[P \backslash R(t)] + t\bullet\).
- The definition of occurrence sequence extends naturally from Petri nets.
Reset nets
Reset nets: the order fulfillment process
Sources and References


- www.workflowcourse.com (among others for the animations)