Petri Nets at TU/e

Introduction

The Architecture of Information Systems (AIS) research group at Eindhoven University of Technology (TU/e) investigates *methods, techniques and tools for the design and analysis of Process-Aware Information Systems* (PAIS), i.e., systems that support business processes (workflows) inside and between organizations. We are not only interested in these information systems and their architecture, but also model and analyze the business processes and organizations they support. We use *Petri nets* for the modeling, analysis, discovery, and enactment of processes.

Our mission is to be one of the worldwide leading research groups in *process modeling and analysis*, *process mining*, and *PAIS technology*. We aim at results that are highly original and applicable in real-life situations. Our motto is "Process Technology that Works".

The Architecture of Information Systems (AIS) at TU/e group is chaired by prof.dr.ir. Wil van der Aalst. The permanent staff of AIS involved in research currently consists of dr.ir. Boudewijn van Dongen, prof.dr. Kees van Hee, prof.dr.ir. Wim Nuijten, dr.MSc. Natalia Sidorova, and dr.ir. Eric Verbeek. Until recently dr. Marc Voorhoeve was also in our group but, sadly, he passed away on 7-10-2011. There are about eight postdocs and seven PhD's. Postdocs very active in the Petri net area are dr. Christian Stahl, dr. Dirk Fahland, dr. Michael Westergaard, and dr. Jan Martijn van der Werf.

The group got the highest possible ranking in the last evaluation of Dutch Computer Science research (5-5-5-5-5). The work is highly cited and the software tools (co-)developed by the group (e.g., ProM, Declare, CPN Tools, and YAWL) are widely known and used.

History and Scope of Research

Until 2004 the Petri net group at Eindhoven University of Technology (TU/e) was named *Specification and Modeling of Information Systems* (SMIS). In 2004 the group was renamed to *Architecture of Information Systems* (AIS). Within the SMIS/AIS group there is a long-standing tradition in modeling and analyzing *Process-Aware Information System* (PAISs) using *Petri nets*. A PAIS is a software system that manages and executes operational processes involving people, applications, and/or information sources on the basis of process models. Example PAISs are workflow management systems, case-handling systems, middleware platforms, enterprise information systems, etc.

In the nineties the group worked on *ExSpect* (Executable Specification Tool, `www.exspect.com`), a specification language and corresponding toolset based on Petri nets extended with data, time, and hierarchy. ExSpect has been used to specify a wide variety of systems (from entire supply chains to embedded software in e.g. copiers) and its simulation engine turned out to be very useful in all kinds of practical situations. In the late nineties the development moved to Bakkenist consultancy (now part of Deloitte), because most of the research challenges related to the execution and simulation of high-level Petri nets had been addressed. Despite several successful applications, Deloitte discontinued the development of ExSpect. However, the simulation engine of ExSpect is still being used in the most widely used business process modeling tool in the Netherlands: Protos of Pallas Athena.

CPN Tools is by far the most widely used Petri-net tool having more than 10,000 licenses in 150 countries. The CPN group at Aarhus University in Denmark has developed and supported CPN Tools (group of Kurt Jensen). However, in 2010, CPN Tools was transferred to our group. Michael Westergaard, who moved to TU/e around the same period, plays a key role in the development and support of CPN Tools. In the autumn of 2010, we released version 3.0. The current version is 3.2.

Although tools such as ExSpect and CPN Tools are highly generic and can be applied to a variety of systems and processes (e.g., embedded systems, supply chains, etc.), the main focus since the late nineties has been on PAISs (in particular workflow management systems). The AIS group was among the first groups to formalize workflow concepts in a systematic manner. Van der Aalst introduced the so-called *WorkFlow nets* (WF-nets) and a correctness criterion called *soundness*. WF-nets are a subclass of Petri nets tailored towards workflow modeling and analysis. The modeling of WF-nets and the analysis of soundness are supported by tools such as Woflan and Yasper. WF-nets and soundness have been widely adopted within the academic community and these techniques are increasingly used in all kinds of commercial software products (Protos, IBM WebSphere, etc.). Members of the AIS group worked on alternative soundness notions, verification techniques, and also applied these techniques to large sets of real-life process models. For example, we showed that more than 20 percent of the 600 process models in SAP's well-known Reference Model contain errors.

In September 2006, Van der Aalst moved from the Information Systems group in the Department of Industrial Engineering & Innovation Sciences (IEIS) to the AIS group. Because of the move of Van der Aalst, Verbeek, and Van Dongen from IEIS to AIS, the focus of AIS was extended to also include process mining. To explain the relation between process mining and the earlier work of AIS on Petri nets and workflow verification, let us consider Figure 1. This figure shows the role of (process) models in the PAIS context. Process models can be used to describe and analyze processes and to specify, configure, or implement information systems.

The left-hand-side of Figure 1 shows some examples of *design-time analysis*: *validation* (i.e., testing whether the process behaves as expected), *verification*, (i.e., establishing the correctness of a process definition), and traditional (i.e., non-log based) *performance analysis* (e.g., using simulation to evaluate the ability to meet requirements with respect to throughput times, service levels, and resource utilization).

Traditionally, most of AIS's research focused on design-time analysis. How-

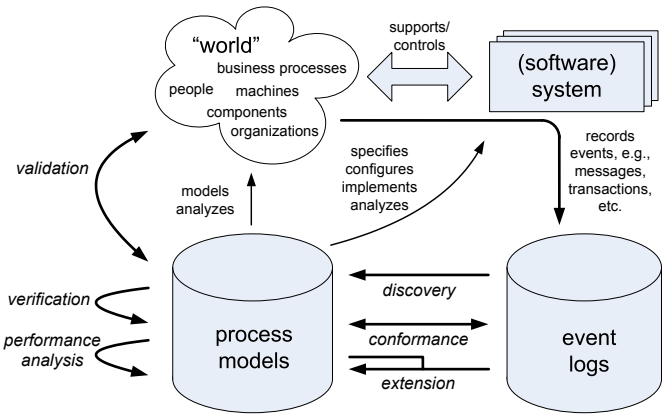

Figure 1: Positioning AIS's research.

ever, more and more information about (business) processes is recorded by information systems in the form of so-called "event logs". IT systems are becoming more and more intertwined with these processes, resulting in an "explosion" of available data that can be used for analysis purposes. The goal of *process mining* is to extract process-related information from event logs, e.g., to automatically *discover* a process model by observing events recorded by some information system. However, process mining is not limited to discovery and also includes *conformance checking* (investigating whether reality conforms to a given model and vice versa) and *extension* (augmenting an existing model with additional insights extracted from some event log).

Since 2006, AIS is also focusing on process mining. In fact, Van der Aalst and his colleagues established process mining as a research field and developed the influential ProM framework. Currently, process mining is seen as one of the main innovations in business process management, and the ideas are rapidly being incorporated in commercial products (BPM|one, Futura Reflect, ARIS PPM, etc.). The combination of knowledge about PAIS technology, process modeling notations, Petri net theory, process verification, and data mining turns out to be an excellent basis for process mining research. Classical techniques in the field of data mining and the so-called Business Intelligence (BI) tools used in industry do not explicitly focus on process models. As a result, the scope is limited to data dependencies and performance measurements. AIS's unique set of capabilities can be used to bridge the gap between process modeling and analysis on the one hand and data mining and BI on the other. This provides an ideal starting point for scientific and technological breakthroughs in process mining.

Research Lines

Current AIS research concentrates on formalisms for modeling and methods to discover and analyze models. On the one hand formal methods are being used, e.g., the group has a long tradition in Petri-net modeling and analysis. On the other hand, we are interested in modeling languages widely used in industry

(EPCs, UML, BPMN, BPEL, etc.). In contrast to many other research groups in this area, we do not accept a model as an objective starting point, i.e., we also try to discover process models through process mining and check the conformance of models based on reality.

The goal for the next five years is to further develop the main three research lines of AIS:

- **Research line 1: Process Modeling/Analysis.** While various types of process notations are used in industry, formal models such as Petri nets are more suitable for analysis purposes. Driven by questions from the other two research lines (process mining and PAIS technology), particular models (e.g., WF-nets, WF-nets with data and resources, history-dependent nets, open nets, nested nets, etc.) are used to answer questions related to correctness and performance. The main techniques that are used are model checking, structural techniques (invariants, etc.), simulation, and Markov (decision) processes. Moreover, quite some efforts are made to translate industry standards and proprietary languages (EPCs, UML, BPMN, BPEL, etc.) to formal models (typically Petri nets). One of the main goals in Research line 1 is to further improve *verification techniques to check various properties such as soundness, data/resource soundness, accordance, controllability, and selected temporal properties.* Here there is a need for more *empirical* research, i.e., analyzing large repositories of models like the SAP Reference Model. Moreover, pattern-based approaches can be used for correctness-by-design. Another goal is to develop *innovative simulation approaches* that better reflect reality and that can be used in an operational setting while using process mining results. A prerequisite for the above analysis approaches is the consistent integration of the different perspectives.

- **Research line 2: Process Mining.** Process mining techniques are used to extract process-related information from event logs, e.g., to automatically discover models, check conformance, and augment existing models with additional insights extracted from some event log. The main difference with Research line 1 is that *event logs play a central role* (rather than predefined process models). One goal is to significantly *improve the state-of-the-art in process discovery.* A particular challenge is to deal with less structured processes and incomplete event logs. For this we want to improve our approaches based on region theory, fuzzy models, and genetic mining. Another goal is to *advance the state-of-the-art in conformance checking*, e.g., by refining our replay strategies and to allow for on-the-fly checking. Related is the challenge to *predict problems*, i.e., provide warnings based on historic information (e.g., a case will be late or an error is likely to occur). To achieve this, we plan to use pattern analysis, correlation analysis, and trace clustering.

- **Research line 3: PAIS Technology.** PAISs are used to *manage and execute operational processes involving people, applications, and/or information sources.* Examples are WFM (Workflow Management), BPM (Business Process Management), and ERP (Enterprise Resource Planning) systems. Increasingly, these systems are driven by models (connection to Research line 1) and produce high-quality event logs (connection to Re-

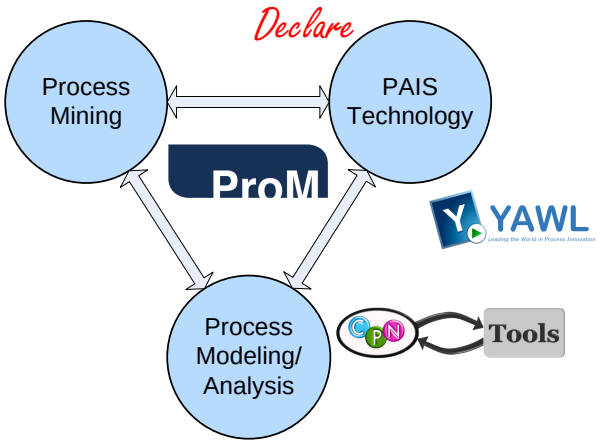

Figure 2: The three main research areas and the most relevant tools.

search line 2). We are interested in the artifacts used and produced by these systems (i.e., models and logs) as these are essential for testing the techniques developed in the two other research lines. For example, it is interesting to convert and verify process models expressed in some particular industry language. This enables empirical research and triggers new questions. The same holds of course for event logs. We are also studying *PAIS architectures*. Note that service-orientation plays an important role here and this new architectural style poses new research questions. Although most PAISs are used in a business setting (governments, banks, insurance companies, supply chains, etc.), we are also interested in scientific computing and grid architectures. Note that the "process of process mining" can be seen as a scientific workflow. Moreover, for large scale process mining experiments, we are using our own grid with a dedicated grid architecture. In Research line 3 we heavily rely on the workflow patterns. This helps us to understand and characterize PAISs.

The three research lines are interconnected in various ways. PAISs are process-aware, use models, and provide for event logs. These models and logs can be analyzed using the techniques developed in Research line 1 and Research line 2. Moreover, these analysis results can be used by the PAIS, e.g., for recommendations, predictions, and diagnosis.

Tooling

Tools play a crucial role in the development of the three research lines. As Figure 2 shows, *ProM* is the central tool for realizing and evaluating our ideas. New techniques related to process mining and process modeling/analysis will be realized in ProM. ProM subsumes the functionality of *Woflan* (workflow verification) and *Yasper* (workflow modeling and analysis). Other relevant tools are *CPN Tools*, *Declare* and *YAWL*. These are three open-source workflow man-

agement systems that are (partly) developed within AIS. Declare is a system aiming at more flexibility using a declarative style of modeling grounded in temporal logic. YAWL is a highly expressive workflow management system based on the workflow patterns and a result of our joint research with QUT. Both Declare and YAWL are tightly connected to ProM, e.g., ProM can analyze the logs and models of YAWL and Declare. Figure 2 also mentions some software systems not developed within AIS. In our teaching and research we are heavily using *CPN Tools*. CPN Tools is a standard Petri net tool for modeling and analyzing high-level nets. We are mainly using CPN Tools for simulation and conceptualizing ideas. *BPM|one* of Pallas Athena and *Websphere* of IBM are two commercial PAISs we are frequently using. Both BPM|one and Websphere have adopted results from our research and are interesting commercial platforms for testing research ideas.

In out view, tools such as ProM, CPN Tools, and Declare are essential for conducting relevant research. We welcome other groups to collaborate with us on the further development of these tools!

For more information, we refer to:

- AIS: `www.win.tue.nl/ais/`
- Process mining and ProM: `www.processmining.org`
- CPN Tools: `www.cpntools.org`
- Declare: `www.win.tue.nl/declare/`