Workflow Simulation for Operational Decision Support

A. Rozinat, M.T. Wynn, W.M.P. van der Aalst, A.H.M. ter Hofstede, and C.J. Fidge

Anne Rozinat

BPM’08, 03.09.2008, Milan
Motivation

- Simulation enables “what-if” analysis (flow time, service level, utilization, ..., waiting time)
- Exploration of process redesigns by simulating their effects
Motivation

- Simulation **enables “what-if” analysis** (flow time, service level, utilization, ..., waiting time)
- Exploration of process redesigns by simulating their effects

“Despite the interest in business process simulation - the actual use by end-users is limited. Why is this?”
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
4. Discussion
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
4. Discussion
1. Three Common Pitfalls

1) *Modeling from scratch*
2) *Incorrect modeling of resources*
3) *Focus on design rather than operational decision making*

➡ Here we address 1) and 3) by

- integrating existing artifacts that can be extracted from a workflow system
- incorporating the current state of a workflow system

---

Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
4. Discussion
2. Our Approach

- Simulation engine
  - Simulation logs
    - records
    - analyze
  - Simulation model
    - specifies configures
    - models
- Workflow & organizational model
  - specifies configures
  - models
- Workflow system
  - records
- Real-world process
  - supports / controls
- Simulated process
  - simulates
  - specifies configures
  - models

- Design information
- Historic information
- Current state information
2. Our Approach

<table>
<thead>
<tr>
<th>Design information</th>
<th>Historic information</th>
<th>State information</th>
</tr>
</thead>
<tbody>
<tr>
<td>(obtained from the workflow and organization model used to configure the workflow system)</td>
<td>(extracted from event logs containing information on the actual execution of cases)</td>
<td>(based on information about cases currently being enacted using the workflow system)</td>
</tr>
<tr>
<td>• control and data flow (activities and causalities)</td>
<td>• data value range distributions</td>
<td>• progress state of cases (state markers)</td>
</tr>
<tr>
<td>• organizational model (roles, resources, etc.)</td>
<td>• execution time distributions</td>
<td>• data values for running cases</td>
</tr>
<tr>
<td>• initial data values</td>
<td>• case arrival rate</td>
<td>• busy resources</td>
</tr>
<tr>
<td>• roles per task</td>
<td>• availability patterns of resources</td>
<td>• run times for cases</td>
</tr>
</tbody>
</table>
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
   1. Architecture
   2. Extracting simulation-relevant information
   3. Generating the simulation model
   4. Loading the current state
   5. Analyzing simulation logs
4. Discussion
3.1 Architecture

- YAWL file
- OrgM. file
- MXML file
- WFS file

- Import new YAWL
- Import OrgModel
- Analyse Log
- Import WorkFlow State

- Merge
- Convert

- Export CPN
- Export WorkFlow State

- CPN file
- SML file

- CPN Tools
- MXML files
- Gnuplot scripts
- Other logs

Type of simulation-relevant information:
- design
- historic
- current state
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
   1. Architecture
   2. Extracting simulation-relevant information
   3. Generating the simulation model
   4. Loading the current state
   5. Analyzing simulation logs
4. Discussion
3.2 Extracting Simulation-relevant Information

- Export information:
  (a) design
  (b) historic
  (c) current state

- Using interfaces to YAWL engine: R, B, and E
3.2 Extracting Simulation-relevant Information

- Export information:
  (a) design
  (b) historic
  (c) current state

- Using interfaces to YAWL engine: R, B, and E
3.2 Extracting Simulation-relevant Information

- Export information:
  (a) design
  (b) historic
  (c) current state

- Using interfaces to YAWL engine: R, B, and E
3.2 Extracting Simulation-relevant Information

- Export information:
  (a) design
  (b) historic
  (c) current state

- Using interfaces to YAWL engine: R, B, and E
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
   1. Architecture
   2. Extracting simulation-relevant information
   3. Generating the simulation model
   4. Loading the current state
   5. Analyzing simulation logs
4. Discussion
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log

2. Merge

3. Convert

4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export

Diagram:
- YAWL file
- OrgM. file
- MXML file
- Import new YAWL
- Import OrgModel
- Analyse Log
- Merge
- Convert
- Export CPN
- CPN file
- SML file
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log
2. Merge
3. Convert
4. Export
3.3 Generating the Simulation Model

1. Import
   - wf model
   - org model
   - event log

2. Merge

3. Convert

4. Export
### 3.3 Generating the Simulation Model

1. **Import**
   - wf model
   - org model
   - event log

2. **Merge**

3. **Convert**

4. **Export**
3.3 Generating the Simulation Model

- **ProM**
  - newYAWL Import
  - OrgModel Import
  - Log Analysis

- **YAWL**
  - YAWL file
  - Resource DB
  - YAWL Engine
  - YAWL Users
  - YAWL Logs

- **ProM**
  - import plug-in
  - Organizational Model Extractor
  - MXML file

- **OrgM.**
  - + roles and their corresponding resources in the whole organization

- **Merge**

- **Convert**

- **CPN Tools**
  - CPN file

- **Current State Extractor**
  - SML file

- **EBR**
  - + roles and resources per role
  - + data attributes (value range)

- **CPN Export**
  - + case arrival rate
  - + time

- **A**
  - + role
  - + data
  - + link condition

- **B**
  - + role
  - + data

- **C**
  - + role
  - + data
3.3 Generating the Simulation Model

- **ProM**
  - newYAWL Import
  - OrgModel Import
  - Log Analysis

- **YAWL Editor**
  - Process Designer
  - Resource DB

- **YAWL Engine**
  - YAWL Users

- **YAWL Logs**
  - ProM import plug-in

- **Organizational Model Extractor**
  - OrgM. file

- **MXML file**

- **Current State Extractor**
  - SML file

**Roles and their corresponding resources**:
- + roles and resources in the whole organization
- + roles and resources per role

**Data attributes**:
- + initial value
- + value range

**Condition**:
- + link condition

**Time**:
- + case arrival rate
- + time

**Data**:
- + data
- + role
- + link condition

**Conversion**:
- Convert

**CPN Tools**
- CPN file

**Export**
- CPN Export
3.3 Generating the Simulation Model

- **ProM**
  - `newYAWL Import`
  - `OrgModel Import`
  - `Log Analysis`

- **YAWL Editor**
  - `Process Designer`
  - `Resource DB`
  - `YAWL Engine`

- **YAWL file**
  - `Resource DB`

- **YAWL Logs**
  - `ProM import plug-in`

- **Organizational Model Extractor**
  - `OrgM. file`
  - `MXML file`

- **Current State Extractor**
  - `SML file`

- **Converter**
  - `CPN file`
  - `CPN Tools`

- **Log Analysis**
  - `case arrival rate` + data attributes
    - initial value
    - value range

- **Process Analyst**
  - `case arrival rate` + data attributes
    - initial value
    - value range
  - `roles and resources per role`

- **CPN Export**
  - `...`
  - `A` + data attributes
    - initial value
    - value range
  - `B` + role
  - `C` + role
3.3 Generating the Simulation Model

ProM

newYAWL Import

OrgModel Import

Log Analysis

Merge

Convert

CPN Export

+ data
+ role

+ link condition

+ data
+ role

+ data attributes (initial value)
+ roles in process

+ roles and their corresponding resources in the whole organization

+ roles and resources per role

+ case arrival rate
+ data attributes (initial value and value range)

+ case arrival rate
+ data attributes (initial value and value range)
+ roles and resources per role

+ case arrival rate
+ data attributes (value range)
3.3 Generating the Simulation Model

ProM

newYAWL Import

OrgModel Import

Log Analysis

Merge

Convert

CPN Export

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value)
+ link condition

+ link condition

+ role

+ role

+ role

+ data

+ role

+ role

+ data

+ data

+ role

+ role

+ data

+ role

+ data

+ role

+ role

+ time

+ time

+ time

+ case arrival rate
+ data attributes (initial value and value range)
+ roles and resources per role

+ case arrival rate
+ data attributes (initial value range)
+ roles and resources per role

+ role

+ role

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ role

+ data

+ role

+ data

+ role

+ role

+ data

+ role

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data

+ role

+ data
3.3 Generating the Simulation Model

ProM

newYAWL Import

OrgModel Import

Log Analysis

Merge

Convert

CPN Export

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value and value range)
+ roles and resources per role

+ time
+ data attributes (initial value)
+ roles in process

+ time
+ data attributes (initial value and value range)
+ roles and resources per role

+ case arrival rate
+ data attributes (initial value and value range)
+ roles and resources per role

+ link condition

+ case arrival rate
+ data attributes (initial value and value range)
+ roles and resources per role

+ link condition

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value and value range)
+ roles and resources per role

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value and value range)
+ roles and resources per role

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value and value range)
+ roles and resources per role

+ data attributes (initial value)
+ roles in process

+ data attributes (initial value and value range)
+ roles and resources per role
3.3 Generating the Simulation Model

- **ProM**
- **newYAWL Import**
- **OrgModel Import**
- **Log Analysis**

---

**ProM**

**newYAWL Import**

- **+ data attributes**
  - initial value
  - roles in process

**OrgModel Import**

- **+ roles and their corresponding resources in the whole organization**

**Log Analysis**

- **+ case arrival rate**
- **+ data attributes**
  - value range

---

**Merge**

**Convert**

**CPN Export**

- **+ case arrival rate**
- **+ data attributes**
  - initial value and value range
- **+ roles and resources per role**

---

**+ data**

**+ role**

**+ link condition**

**+ data attributes**

**+ role**

**+ data**

**+ role**

**+ time**

**+ link condition**

**+ data attributes**

**+ role**

**+ data**

**+ role**

**+ time**

**+ data attributes**

**+ role**

**+ data**

**+ role**

**+ time**

**+ data attributes**

**+ role**

**+ data**

**+ role**

**+ time**
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
   1. Architecture
   2. Extracting simulation-relevant information
   3. Generating the simulation model
   4. Loading the current state
   5. Analyzing simulation logs
4. Discussion
3.4 Loading the Current State

- Current state can be updated without changing the simulation model
3.4 Loading the Current State

- Current state can be updated without changing the simulation model
3.4 Loading the Current State

- Current state can be updated without changing the simulation model
3.4 Loading the Current State
3.4 Loading the Current State
3.4 Loading the Current State
3.4 Loading the Current State

getInitialCaseData()

Case data

CASE_IDxDATA

(c, data) → (c, modifiedData)

getInitialTokensExePlace("TASK_check_for_completeness_4_E") @ round(normal(1800.0, 519.42)) div 2

(c, t, clerk) @ round(normal(1800.0, 519.42))

1. (39, {loanAmt=500, completeApp=false, decideApp=false}) @ 0

1. (40, {loanAmt=0, completeApp=false, decideApp=false}) @ 0

1. (41, {loanAmt=1500, completeApp=false, decideApp=false}) @ 0

E 1

TASK_check_for_completeness_4_start

_CASE_IDxSTART_TIMExCLERK

clerk

(c, t, clerk)

TASK_check_for_completeness_4_complete

CASE_IDxSTART_TIME

clerk

input (data);
output (modifiedData);
action (DATA.set_completeApp data (randomcompleteApp()));

FREE.all()

"AnneR"++
"GoldB"++
"JohnsH"++
"JonesF"++
"LewisC"++
"LewisF"++
"MoeW"

Resources

ANYBODY

(anybody)
Outline

1. Three Common Pitfalls
2. Our Approach

3. Realization through YAWL and ProM
   1. Architecture
   2. Extracting simulation-relevant information
   3. Generating the simulation model
   4. Loading the current state
   5. Analyzing simulation logs

4. Discussion
3.5 Analyzing Simulation Logs

- MXML files
- CpN file
- SML file
- Gnuplot scripts
- Other logs

- CPN Tools
- Design historic current state
- Type of simulation-relevant information
- MXML files
- Merge
- Convert
- Import Workflow State
- Export Workflow State
3.5 Analyzing Simulation Logs

*Example:* 4 different simulation scenarios:

1. An **empty** initial state (‘empty’)
2. After loading the **current state** file with the 150 applications currently in the system (‘as is’)
3. After loading the current state file but adding **four extra resources** (‘to be A’)
4. After loading the current state file and adding **eight extra resources** (‘to be B’)


3.5 Analyzing Simulation Logs

Number of applications that are in the system for four different scenarios:

1) 'as is'
2) 'to be A'
3) 'to be B'
4) 'empty'

Time horizon: two weeks (in seconds)
95% Confidence Intervals Average Throughput Time in Min for the Four Simulation Scenarios (50 Replications each)

- 'as is': 5.88 days
- 'to be A': 4.91 days
- 'empty': 3.86 days
- 'to be B': 4.72 days
Outline

1. Three Common Pitfalls
2. Our Approach
3. Realization through YAWL and ProM
4. Discussion
5. Discussion

• Faithful simulation models
  ★ making use of existing artifacts (e.g., process history)
  - But: modeling human behavior remains challenging

• Short-term simulation
  ★ operational decision making based on current state
  - But: tool integration can be improved (feedback)

• Viewing real and simulated process in unified manner
  ★ simulation analysis results can be more easily related to initial properties of the process