state-of-art in simulation

1. Problems
   - Mismatched models
   - Overlooked artifacts
   - Wrong objectives
   - Inadequate visualization
2. Partial solution: Chunks
3. Partial solution: Short-term simulation
4. Conclusion
Acknowledgements


Problems

a) Mismatched models
b) Overlooked artifacts
c) Wrong objectives
d) Inadequate visualization
Some of the problems related to BPS

- Mismatched models
- Overlooked artifacts
- Wrong objectives
- Inadequate visualization
Problem 1: Mismatched models
Example: ASML test process

37.5% OK
62.5% NOK
Example: My first real-life simulation

- ExSpect simulation model of supply chain of DAF to Spain.
- Interesting problems such as 200,000+ different spare parts.
  - Taking the characteristic ones or just the fast-movers does not work!
  - Consider for example truck loads, warehouse, etc.
- Key Performance Indicators (KPIs) in initial simulations dramatically different from reality.
Example: Numerous master projects

- After measuring times and routing probabilities, a faithful simulation model is constructed.
- However, KPIs in initial simulations dramatically different from reality.
  - Simulated flow times of hours correspond to real-life flow times of weeks.
  - Simulated utilizations of 30 percent in processes where employees complain about workload (burnout, boreout, or simout?).
- Simulation model is "massaged" until reality and simulation match.
- Observation: processes adapt based on context (when busy; skip checks, work longer, etc.).
How to Lie With Statistics?
How to Lie With Simulation?

- **M/M/1 queue**: arrival rate $\lambda$, service rate $\mu$, utilization $\rho = \lambda/\mu$.
- Flow time $= 1/(\mu-\lambda)$, # in system $= \rho/(1-\rho)$
Problem 2: Overlooked artifacts

- Simulation models are typically built from scratch ignoring a wealth of information:
  - process models (e.g., workflow models implicit or explicit)
  - historic data (event logs, information about arrival rates, service rates, etc.)
  - current state
Example: How to turn a workflow model into a simulation model?

- arrival process
- service times
- probabilities for choices
- subrun settings
- priorities
- number of resources

Use historic data!
Problem 3: Wrong objectives

- Traditional steady state analysis focusing on long-term average behavior

- Classical focus of simulation (tools)
Steady state analysis

- warm-up period
- independent runs
- warm-up period
- subruns

(re-)design-time analysis, i.e., not for operational decision making
Transient analysis

Steady-state may not exist and may not be relevant!
Problem 4: Inadequate visualization

- Simulation tools show "tables" and "graphs", and, if they are advanced also "token game animations".
- Management dashboards show "tables" and "graphs", and, if they are advanced also "speedometers", etc.
- Problems:
  - Observing the "simulated world" and the "real world" should be unified to allow for a better interpretation of the results.
  - Management dashboards are not looking "inside the process".
  - Process visualization is rather primitive and tries to show design artifacts rather than the process itself.
Trip facts are released for accounting.

Planned trip is approved.

Travel Expenses

Advance payment

Need to correct planned trip is transmitted.

Unrequested trip has taken place.

Trip facts and receipts have been released for checking.

Approved trip has taken place.

Trip costs statement is transmitted.

Entry of trip facts

Entry of a travel request

Accounting date is reached.

Payment amount transmitted to bank/payee.

Cancellation

Trip costs must be included in cost accounting.

Amounts liable to employment tax transmitted to payroll.

Amounts relevant to accounting transmitted to payroll accounting.

Need for trip has arisen.

Payments must be released.

Trip is requested.

Approval of trip facts is transmitted.

Approval of travel request

Trip expenses reimbursement is rejected.

Planned trip must be canceled.

Trip advance is transmitted/paid.

Trip expenses reimbursement must be canceled.

Trip is canceled.

Trip costs cancelation statement is transmitted.
More significant nodes are emphasized.
More to learn from maps...

**Aggregation**
Clustering of coherent, less significant structures

**Abstraction**
Removing isolated, less significant structures
Fuzzy miner (Christian Güther)
Showing reality  (Christian Güther)
Overview problems

- Mismatched models
- Overlooked artifacts
- Wrong objectives
- Inadequate visualization

real world  simulated world
Partial solution: Chunks

Mismatched models because of inadequate resource modeling
Chunks?

- Remember: Simulated flow times of hours while real-life flow times are weeks!
- Hypothesis: Primary cause is inadequate resource modeling.
- People:
  - are working part-time, have breaks, holidays, sick leaves, etc.
  - are involved in multiple processes and need to assign priorities dynamically,
  - do not work at a constant speed,
  - etc.
$5 \times 0.2 \neq 1$
Classical simulation assumptions

- A resource is:
  - eager to start working,
  - dedicated to a single process,
  - works at a constant speed,
  - does not work in batches,
  - does not have coffee breaks,
  - etc.
- Do you know this person?
- Chunks: towards a more accurate modeling of resource availability
Avoid modeling the world in a detailed manner

Goal: Characterize resource availability with just a few parameters
Chunks: Basic Idea

resource:
- inactive
- ready
- busy
- $\rho = \frac{\lambda}{\mu} \leq a$, i.e., utilization is less than availability
- $c \leq h$, i.e., chunk size cannot be larger than the horizon
- $(a \times h) \mod c = 0$ in experiments to avoid unusable availability
CPN model
Fig. 9. Graph showing availability against flow time ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $c = 200$, and $h = 1000$). The flow time reduces as the availability increases. (The straight line shows the trend using linear regression.)
Fig. 10. Graph showing chunk size against flow time ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $a = 0.2$, and $h = 1000$). The flow time increases as the chunk size increases.
Fig. 11. Graph showing the horizon against the flow times ($\lambda = \frac{1}{100}$, $\mu = \frac{1}{15}$, $\rho = 0.15$, $c = 200$, and $a = 0.8$). The flow time decreases as the horizon increases.
**Effect of utilization ($\rho$)**

Fig. 12. Graph showing utilization against flow time ($\mu = \frac{1}{15}$, $c = 200$, $a = 0.8$, and $h = 1000$). The flow time increases as utilization increases.
Experiment: Note multiple resources and potential accumulation of effects
CPN model
## Some findings

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Flow Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Base Case Scenario ((c = 5, h = 2000, \lambda = \frac{1}{50} \text{ and } a = 0.4, \text{ see Appendix B for all other parameters}))</td>
<td>757.6 ±65.0</td>
</tr>
<tr>
<td>b) i) Divide the horizon by 20 ((h = 100))</td>
<td>1218.9 ±72.3</td>
</tr>
<tr>
<td>ii) Divide the horizon by 40 ((h = 50))</td>
<td>1247.8 ±51.8</td>
</tr>
<tr>
<td>c) i) Multiply the chunk size by 5 ((c = 25))</td>
<td>1158.7 ±47.2</td>
</tr>
<tr>
<td>ii) Multiply the chunk size by 20 ((c = 100))</td>
<td>1698 ±139</td>
</tr>
<tr>
<td>iii) Multiply the chunk size by 80 ((c = 400))</td>
<td>1950 ±83.7</td>
</tr>
<tr>
<td>iv) Multiply the chunk size by 160 ((c = 800))</td>
<td>2025 ±99</td>
</tr>
<tr>
<td>d) i) Decrease availability and arrival rate by 2 ((a = 0.2, \lambda = \frac{1}{100}))</td>
<td>1634 ±105</td>
</tr>
<tr>
<td>ii) Decrease availability and arrival rate by 4 ((a = 0.1, \lambda = \frac{1}{200}))</td>
<td>3420.32 ±252</td>
</tr>
</tbody>
</table>
"Chunks Conclusion"

- It is important not to assume that people are always available and eager to work when cases arrive.
- The assumptions heavily impacts flow time, e.g., the bigger the chunk size, the longer the flow times of cases.
- The "chunk model" is rather simple, however, the typical assumptions made in today's simulation tools (i.e. \( a = 1, c = 0, \) and \( h=\text{inf} \)), may result in flow times of minutes or hours while with more realistic settings for \( a, c, \) and \( h \) the flow time may go up to weeks or months and actually coincide with the actual flow times observed.
Partial solution: Short-Term Simulation

Focus: transient analysis using design, historic, and current state information.

- Wrong objectives
- Overlooked artifacts

Real world

Simulated world
Overview: Short-Term Simulation

Simulated process

Simulation engine
- simulates
- records
- specifies configures

Simulation model
- models
- analyze

Workflow & organizational model
- models
- specifies configures
- records

Real-world process

Workflow system
- supports / controls
- records

Design information

Historic information

Current state information
Implementation using YAWL, ProM, and CPN Tools
ProM: Merging and converting models covering different aspects

YAWL 2.0 Import

OrgModel Import

Log Analysis

merge models

CPN Export

+ data + role
+ data + role
+ data + role
+ data + role
+ data + role
+ data + role
+ data + role
+ data + role
+ data + role
+ case arrival rate
+ 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
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’)

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

- 'as is'
- 'to be A'
- 'to be B'
- 'empty'

Time horizon: two weeks (in seconds)
Confidence intervals

95% Confidence Intervals Average Throughput Time in Min for the Four Simulation Scenarios (50 Replications each)

Confidence Interval

Simulation Scenarios

'as is' 5.88 days
'to be A' 4.91 days
'empty' 3.86 days
'to be B' 4.72 days
Conclusion Short-Term Simulation

- **Transient analysis** is essential for operational decision making!
- The **initial state** matters!
- **Artifacts** (design, historic, and current state information) from a workflow management systems like YAWL can be used!
- **Interesting side effect of the YAWL, ProM, CPN Tools integration:** the real and simulated process can be viewed in a unified manner using process mining!
Problems and some solutions ...

- Wrong objectives
- Mismatched models
- Overlooked artifacts
- Inadequate visualization
Relevant WWW sites

- http://www.processmining.org
- http://www.workflowpatterns.com
- http://www.workflowcourse.com
- http://www.vdaalst.com