BPMN Modeling Guidelines
Table of Contents

- Introduction
  - Learn PAd
  - Understandable models
  - Literature

- What we did
  - Guidelines
  - Measures
  - Thresholds

- BP modeling guidelines
  - Examples
  - Guidelines Application EPBR scenario

- Validation
  - Questionnaire
The context...

**Learn PAd**
Model-Based Social Learning for Public Administrations

**EU Programme**
FP7-ICT-2013.8.2 Technology-enhanced learning
Learn PAd main goal

To provide a social and collaborative learning platform for civil servants.

Key aspects:

- e-Learning platform
- Wiki
- Business Process Models (BPMN)
- Web 2.0
Our job (in collaboration with the CNR of Pisa)

To guarantee that models used in the Learn PAd platform are UNDERSTANDABLE.
What happens?

It usually happens that someone studies the BPMN notation and “becomes a modeler”. So he/she starts using the BPMN notation to model everything (and this is good) but in the wrong way.

He will probably design models that are too large, with too many BPMN elements and maybe they are used in the wrong way. Or there may be too much details, too much annotations which doesn’t allow for a nice view of the process.

**Note**: if you use the BPMN syntax as it is described in the BPMN specification your model cannot be considered wrong. But if you want the model to be understandable by others you can say that there are “errors” in the model or at least that they could be improved.
So...
How can we guarantee that a model is **UNDERSTANDABLE**?

&

How can we help the modeler to design **UNDERSTANDABLE** models?
Literature - (1)


Literature - (2)

**Guidelines** - general rules which the modeler should follow to ensure the model he designs results to be understandable.

**Metrics or Measures** - what we use to refer to the amount or the size of something (e.g. the amount of BPMN elements used in a BP Model).

**Thresholds** - values that measures should not exceed to guarantee the BP Model is understandable.
What we did...

We collected and refined

- Guidelines (50)
- Measures
- Thresholds
# Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measures of Rolón [72]</strong></td>
<td></td>
</tr>
<tr>
<td>TNSF</td>
<td>Total Number of sequence flows</td>
</tr>
<tr>
<td>TNE</td>
<td>Total Number of events</td>
</tr>
<tr>
<td>TNG</td>
<td>Total Number of gateways</td>
</tr>
<tr>
<td>NSFE</td>
<td>Number of sequence flows from events</td>
</tr>
<tr>
<td>NMF</td>
<td>Number of message flows</td>
</tr>
<tr>
<td>NSFG</td>
<td>Number of sequence flows from gateways</td>
</tr>
<tr>
<td>CLP</td>
<td>Connectivity level between participants</td>
</tr>
<tr>
<td>NDOOut</td>
<td>Number of data objects which are outputs of activities</td>
</tr>
<tr>
<td>NDOIn</td>
<td>Number of data objects which are inputs of activities</td>
</tr>
<tr>
<td>CLA</td>
<td>Connectivity level between activities</td>
</tr>
<tr>
<td><strong>Measures of Cardoso [11]</strong></td>
<td></td>
</tr>
<tr>
<td>CFC</td>
<td>Control flow complexity. Sum over all gateways weighted by their potential combinations of states after the split</td>
</tr>
<tr>
<td><strong>Measures of Mendling [52]</strong></td>
<td></td>
</tr>
<tr>
<td>Number of nodes</td>
<td>Number of activities and routing elements in a process model</td>
</tr>
<tr>
<td>Gateway mismatch</td>
<td>Sum of gateway pairs that do not match each other, e.g. when an AND-split is followed by an OR-join</td>
</tr>
<tr>
<td>Depth</td>
<td>Maximum nesting of structured blocks in a process model</td>
</tr>
<tr>
<td>Connectivity coefficient</td>
<td>Ratio of total number of arcs in a process model to its total number of nodes</td>
</tr>
<tr>
<td>Density</td>
<td>Ratio of total number of arcs in a process model to the theoretically maximum number of arcs</td>
</tr>
<tr>
<td>Sequentiality</td>
<td>Degree to which the model is constructed from pure sequences of tasks</td>
</tr>
</tbody>
</table>

U*: Understandability, M*: Modifiability
## Thresholds

<table>
<thead>
<tr>
<th>Model Metric</th>
<th>Very Inefficient</th>
<th>Rather Inefficient</th>
<th>Rather Efficient</th>
<th>Very Efficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nºnodes</td>
<td>65</td>
<td>50</td>
<td>37</td>
<td>31</td>
</tr>
<tr>
<td>GatewayMismatch</td>
<td>29</td>
<td>16</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Depth</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Coefficient of connectivity</td>
<td>1,7</td>
<td>1,1</td>
<td>0,6</td>
<td>0,4</td>
</tr>
<tr>
<td>Sequentiality</td>
<td>0,1</td>
<td>0,35</td>
<td>0,6</td>
<td>0,7</td>
</tr>
<tr>
<td>TNSF</td>
<td>72</td>
<td>49</td>
<td>34</td>
<td>20</td>
</tr>
<tr>
<td>TNE</td>
<td>20</td>
<td>12</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>TNG</td>
<td>17</td>
<td>10</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>NSFE</td>
<td>28</td>
<td>13</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>NMF</td>
<td>27</td>
<td>15</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>NSFG</td>
<td>40</td>
<td>22</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>CLP</td>
<td>7,5</td>
<td>4,23</td>
<td>2,2</td>
<td>0,2</td>
</tr>
<tr>
<td>NDOIN</td>
<td>31</td>
<td>14</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>NDOOUT</td>
<td>23</td>
<td>11</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>CFCxor</td>
<td>30</td>
<td>17</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>CFCor</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CFCand</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Guidelines

- **General**: they impact on different aspects of the overall BPMN modeling practice
- **Notation**: they focus on the usage of the BPMN Syntax
- **Labeling**: the correct use of names/labels, assigned to BPMN elements
- **Patterns**: specific arrangements of BPMN elements
- **Appearance**: refers to a clear disposition of the BPMN elements in the entire model
BP Modeling Guidelines - General (1)

- Minimize model size

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize model size</td>
<td>2</td>
</tr>
</tbody>
</table>

Description:
The modeler should try to keep models as small as possible. Large process models are difficult to read and comprehend. Additionally, they tend to contain more errors. Defining the correct scope of tasks and level of detail of processes is key to reduce the overage of information.

Source:
[7, 8, 27, 28, 29, 30, 31, 32, 26, 33]

Associated Metrics and Thresholds:

\[
MinimizeModelSize(x) = \begin{cases} 
0 & \text{if } SN \leq 31 \\
1 & \text{otherwise}
\end{cases}
\]

where:
\( x \in \text{Nodes of BPMN Model} \land 
SN \text{ is the number of nodes: number of activities and routing elements in a process model} \).
BP Modeling Guidelines - General (2)

- Apply hierarchical structure with SubProcesses

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimize model size</td>
<td>2</td>
</tr>
</tbody>
</table>

**Description**

The modeler should try to keep models as small as possible. Large process models are difficult to read and comprehend. Additionally, they tend to contain more errors. Defining the correct scope of tasks and level of detail of processes is key to reduce the overage of information.

**Source**

[7, 8, 27, 28, 29, 30, 31, 32, 26, 33]

**Associated Metrics and Thresholds**

\[
\text{MinimizeModelSize}(x) = \begin{cases} 
0 & \text{if } SN \leq 31 \\
1 & \text{otherwise}
\end{cases}
\]

*where:*

\[x \in \text{Nodes of BPMN Model} \land \]

\[SN \text{ is the number of nodes: number of activities and routing elements in a process model.}\]
BP Modeling Guidelines - Notation (1)

- Explicit usage of gateways

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explicit usage of gateways</td>
<td>16</td>
</tr>
</tbody>
</table>

**Description**

The modeler should split or join sequence flows always using gateways. The modeler should not split or join flows using activities or events. This includes that an activity can have only one incoming sequence flow and only one outgoing sequence flow.

**Source**

[36, 11, 7, 8, 26, 38]

**Convention on the modeling**

**Bad Modeling**

**Good Modeling**
BP Modeling Guidelines - Notation (2)

- Consistent usage of pools

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistent usage of pools</td>
<td>10</td>
</tr>
</tbody>
</table>

Description:
The modeler should define as many pools as processes and/or participants. Use a black-box pool to represent external participant/processes. The modeled pools need to be in relation with each other and have to be linked to the main process through message exchange.

Source:
[36, 11]

Convention on the modeling:

Bad Modeling

Good Modeling
• Labeling Activities

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling Activities</td>
<td>30</td>
</tr>
</tbody>
</table>

Convention concerning the name

Label activities with one verb, and one object. The verb used should use the present tense and be familiar to the organization. The object has to be qualified and also of meaning to the business. Multiple activities should not be labeled with the same name, except for same Call Activities used many time in the process.

Source

[9, 11, 7, 8, 39, 32, 26, 40, 41, 38, 33]
BP Modeling Guidelines - Labeling (2)

- Labeling XOR Gateway

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling XOR Gateway</td>
<td>34</td>
</tr>
</tbody>
</table>

**Convention concerning the name**

Label XOR split gateways with an interrogative phrase (do not label XOR join-gateways). Sequence flows coming out of diverging gateways of type exclusive, inclusive and complex should be labeled using their associated conditions stated as outcomes.

**Source**

[11, 26, 40, 41]
BP Modeling Guidelines - Patterns

- Use subprocesses to scope attached events

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use subprocesses to scope attached events</td>
<td>42</td>
</tr>
</tbody>
</table>

**Description**

A subprocess with attached event enables to clearly define the scope of an event. If the response to the handling of an exception (in the use of boundary events) is the same for every activity within a contiguous segment of the process, the modeler should not attach the same boundary event to each of those activities and he should not represent the same exception flows multiple times. The correct way to model it is to enclose that segment in a subprocess and attach a single boundary event to the subprocess boundary.

**Source**

[38]
BP Modeling Guidelines - Patterns (example)

- **Before**

  ![Before Diagram](image)

- **After**

  ![After Diagram](image)
BP Modeling Guidelines - Appearance (1)

- Absence of overlapping elements

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of overlapping elements</td>
<td>44</td>
</tr>
</tbody>
</table>

**Description**

The BPMN elements should not overlap one another. Which means, avoid overlapping, or crossing, tasks, control flows, message flows etc.

**Source**

[36, 11, 26, 33]

**Convention on the modeling**

**Bad Modeling**

**Good Modeling**
BP Modeling Guidelines - Appearance (2)

- Linear Message flows

<table>
<thead>
<tr>
<th>Guideline Name</th>
<th>Guideline ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear message flows</td>
<td>46</td>
</tr>
</tbody>
</table>

**Description**
Linear message flows without useless foldings help to maintain the model clear.

**Source**
[36, 11, 33]

**Convention on the modeling**

**Bad Modeling**

**Good Modeling**
Activity referring to the same topic can be aggregated in a subprocess. In this way we reduce the model size **(guideline 2)**.

*EPBR: European Project Budget Report.*
Data object should have a proper label (the data object states goes between square brackets []) (guideline 37)
• Activities require proper labels with essential information, details can go into the activity description (guidelines 26, 27, 30 and 50).
• The model should be as structured as possible, gateways should be balanced, Xor gateways should have a marker (guidelines 4, 17, 19 and 34).
• If multiple end states are present they should be labeled, if they represent the same state, they should be merged. (guidelines 14 and 32).
Guidelines Application - EPBR scenario (e.3)

- After
Validation

We defined a questionnaire to investigate the importance of the modelling guidelines for the design of understandable BP models. We had 76 participants including: students, BPMN experts, companies and civil servants. An example of question and answers is reported below.

16. Observing the following Exclusive Gateway, do you think that the marker improves the understandability of the model?

- No, A - I prefer the Exclusive Gateway without marker
- Yes, B - I prefer the Exclusive Gateway with marker
- I don't know

- 76.4%
- 22.2%
Validation - results

We analysed the overall answers to the questionnaire and we came to the conclusion that the answers by different profiles are quite similar.

The answers to the questionnaire comply with the vision that led us to the definition of the BPMN modelling guidelines.

After this questionnaire we can confirm that the usage of the defined modeling guidelines leads to the design of understandable BPMN models.
Technical Report


URL

BEBoP - understandability verifier for Business Process models

Our collaborators from the CNR of Pisa, developed a webservice which enables us to upload a BPMN model, designed with Eclipse or Signavio BPMN modeler, and to test which guidelines the model respects and which not.
BEBoP - Exercises and Testing

To take practice with the Guidelines introduced, you can access the BEBoP webservice.

BEBoP website:
http://understandabilitybpmn.isti.cnr.it:8080/JSPUIUnderstandability/contentform.jsf
The End...

Thank You!

My contact: fabrizio.fornari@unicam.it