Quality of modeling languages

John Krogstie
Professor, IDI, NTNU
Overview of presentation

- Brief review of the quality framework
- Quality of modeling languages as means for creating good models
- Based on
  - chapter 5
SEQUAL

Goals of modelling

Modeling Domain

Model externalization

Technical actor interpretation

Language extension

Social actor explicit knowledge

Social actor interpretation

Perceived semantic

Deontic (learning)

Physical

Pragmatic (human understanding)

Syntaxic

Pragmatic (tool understanding)

Empirical

Semantic

Deontic (action)

Deontic
Physical vs social aspects

Model-Based Development and Evolution of Information Systems

**Physical**

**Modeling Domain** $D$

**Language Extension** $L$

**Model externalization** $M$

**Goals of Modelling** $G$

**Social**

**Stakeholder explicit knowledge** $K$

**Modeller explicit knowledge** $K$

**Social actor Interpretation** $I$

**Technical actor Interpretation** $T$

**Language Extension**

- **Semantic**
  - Domain appropriateness
  - Participant appropriateness
  - Modeller appropriateness
  - Comprehensibility appropriateness
  - Physical appropriateness
  - Social pragmatic appropriateness

- **Syntactic**
  - Empirical pragmatic
  - Technical pragmatic

**Modeling Domain**

- Domain appropriateness

**Model externalization**

- Tool appropriateness

**Goals of Modelling**

- Knowledge appropriateness

**Social actor Interpretation**

- Tool comprehension support

**Technical actor Interpretation**

- Tool comprehension support
Overall structure of framework

- Quality type (physical, empirical…)
  - One or more quality goals per quality type
    - Means to achieve this goal
      - Beneficial existing quality
      - Model properties
      - Language properties
      - Modeling activities
      - Tool-support
Why focus on language quality?

- A modeling language is a mean to be able to express knowledge in a model
- A modeling language is meant to help focusing on the important aspects of a situation
- The (modeling) language you use forms your perception of reality
- It is possible to make good models in a poor language
- It is possible to make poor models in a good language.
- It will always be possible to identify problems in any language/tool
Some background

- Bunge Wand Weber as a reference model
- Sindre
- Bertin vs. Visual variables
- Moody – Physics of Notation
- Gestalt psychology
Bunge-Wand Weber as reference framework
Model-Based Development and Evolution of Information Systems

Bunge-Wand-Weber

Legend

<table>
<thead>
<tr>
<th>RF</th>
<th>Set of semantics prescribed in the Reference Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML</td>
<td>Set of semantics expressible in the Modeling Language</td>
</tr>
<tr>
<td>●</td>
<td>Chunk of semantics prescribed in the Reference Framework</td>
</tr>
<tr>
<td>○</td>
<td>Chunk of semantics expressible in the Modeling Language</td>
</tr>
</tbody>
</table>
Sindre 1990 – language quality criteria

- Criteria for the conceptual basis of the language (what we often call the meta-model)
- Criteria for the external representation of the language (what we often call the notation of the language)

For each of these kinds, Sindre identified four groups of criteria:

- Perceptibility: how easy is it for human beings to understand the language?
- Expressive power: what is it possible to express in the language (expressiveness)?
- Expressive economy: how efficiently can things be expressed in the language?
- Method/tool potential: how easily does the language lend itself to proper method and tool support?
Bertin 1983

<table>
<thead>
<tr>
<th>PLANAR VARIABLES</th>
<th>RETINAL VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal Position</td>
<td>Shape</td>
</tr>
<tr>
<td>Vertical Position</td>
<td>Brightness</td>
</tr>
<tr>
<td></td>
<td>Size</td>
</tr>
<tr>
<td></td>
<td>Orientation</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
</tr>
<tr>
<td></td>
<td>Texture</td>
</tr>
<tr>
<td></td>
<td>Colour</td>
</tr>
<tr>
<td></td>
<td>Texture</td>
</tr>
</tbody>
</table>

- **Size**
- **Value**
- **Texture**
- **Color**
- **Orientation**
- **Shape**

**Association**
- The marks can be perceived as similar

**Selection**
- The marks can be perceived as different

**Order**
- The marks can be perceived as ordered

**Quantity**
- The marks can be perceived as proportional
Principles from gestalt psychology

- A closed contour in a node-link diagram generally represent a concept of some kind.
- The shape of a closed contour is frequently used to represent a concept type.
- The color of an enclosed region represent a concept type.
- The size of an enclosed region can be used to represent the magnitude of a concept.
- Lines that partition a region within a closed contour can delineate subparts of a concept.
- Closed-contour regions may be aggregated by overlapping them. The result is readily seen as a composite concept.
- A number of closed-contour regions within a larger closed contour can represent conceptual containment.
- A linking line between concepts represents some kind of relationship between them.
- A line linking closed contours can have different colors, or other graphical qualities such as waviness, and this effectively represents an attribute or type of relationship.
- The thickness of a connecting line can be used to represent the magnitude of a relationship (a scalar attribute).
- A contour can be shaped with tabs and sockets that can indicate which components have particular relationships.
Moody – ’Physics’ of notation

- **Semiotic clarity (SC)**
- **Perceptual discriminability (PD)**
- **Semantic transparency (ST)**
- **Complexity management (CM)**
- **Cognitive integration (CI):**
- **Visual expressiveness (VE):**
- **Dual coding (DC):**
- **Graphic economy (GE):**
- **Cognitive fit (CF):**
Moody – ’Physics’ of notation - 2009

- Semiotic clarity (SC)
- Perceptual discriminability (PD)
- Semantic transparency (ST)
- Complexity management (CM)
- Cognitive integration (CI):
  - Visual expressiveness (VE):
  - Dual coding (DC):
  - Graphic economy (GE):
  - Cognitive fit (CF):
Model-Based Development and Evolution of Information Systems

SEQUAL – language quality

K: Social actor explicit knowledge
G: Goals of modelling
D: Modeling domain
M: Model externalization
T: Technical actor interpretation
I: Social actor interpretation
L: Language extension

Participant appropriateness
Modeller appropriateness
Comprehensibility appropriateness
Domain appropriateness
Tool appropriateness
Organizational appropriateness
Aspects of language quality

- Is the language appropriate to the domain?
- Is the language appropriate for making comprehensible model?
- Is the language appropriate for the participants’ knowledge of modelling languages?
- Is the language appropriate to express the knowledge of the modeller?
- Is the language appropriate for providing tool-support?
- Is the language appropriate for the chosen/standardised tools and modelling languages within the organization?

Differentiate criteria for the conceptual basis of the language (language-model/meta-model), and the external representation of the language (notation)
Domain appropriateness

- $D \setminus L = \emptyset$

- **Basis**
  - Must be able to express anything in the domain
    - Can be based on ontological analysis
    - Can be based on state of the art and state of practice in the area
  - Should not be able to express concepts not in the domain

- **Notation:** Must be able to express everything in the conceptual basis in a distinguishable way.

- **Domain appropriateness for a generic domain = expressiveness**

- Support achievement of semantic quality.
Expressiveness of languages based on ontological analysis

- Bunge-Wand-Weber (BWW) as an example
- Representation model as a basis for evaluating IS analysis and design languages relative to their ability to create models that are a proper representation of the world.
- Based on Mario Bunge’s ontology
  - Well developed
  - Formalised
- BWW does not discuss if one is modelling an objective reality or a socially constructed one
Representation model as a basis for language quality

- Ontological completeness (vs. domain appropriateness)

- Ontological clarity
  - Construct overload (Comprehensibility appropriateness)
  - Construct redundancy (Comprehensibility appropriateness)
  - Construct excess (vs. domain appropriateness)
Ontological completeness

Ontological concepts

Modeling-concepts
Construct excess

Ontological concepts

Modeling-concepts
Expressiveness based on what is found useful within the domain of modeling

- Behavioral
- Functional
- Structural
- Rule-oriented
- Object-oriented
- Language acts
- Actor/roles-oriented
- Topological

Not only look on individual concepts, but look on how concepts can be put together to represent larger structures/patterns, e.g. use of workflow patterns for evaluation of process modeling languages.
Comprehensibility appropriateness

- $L \setminus I = \emptyset$

- Support achievement of empirical and through this pragmatic quality
Aspect of comprehensibility appropriateness....

- **Basis**
  - Limited number of building blocks
  - The language is well-defined
  - Easy to differentiate between different parts of the language
  - Possible to relate concepts freely ➔ new semantics through composition
  - Hierarchically structured
  - Uniform use of building block
  - Flexibility in detail
  - Expressive economy:
    - Frequent and important statements can be expressed concisely
Construct overload

Ontological concepts

Modeling-concepts
Construct redundancy

Ontological concepts

Modeling-concepts
Improvement of comprehension continues...

- **External representation**
  - Semiotic clarity (SC), including uniform use of symbols and syntactic disjointness
  - Perceptual discriminability (PD): Easy to differentiate different symbols. Appropriate use of color and other retinal variables
  - Semantic transparency (ST):
  - Complexity management (CM):
  - Graphic economy (GE)
  - Emphasis of what is important
    - Symbol size, solidity, color, differences, movement, number of edges.
  - Cognitive integration (CI):
Different ways of illustrating associations between classes (but better than predecessor)
Use of emphasis

- Size (the big is more easily noticed than the small), given that size ratios are predefined.

- Solidity (e.g. bold letters vs. ordinary letters, full lines vs. dotted lines, thick lines vs. thin lines, filled boxes vs. non-filled boxes)

- Difference from ordinary pattern (e.g. slanted letters, a rare symbol will attract attention among a large number of ordinary ones)

- Foreground/background differences (if the background is white, things will be easier noticed the darker they are)

- Color (red attracts the eye more than other colors).

- Change (blinking or moving symbols attract attention)

- Pictures vs. text (pictures usually having a much higher perceptibility, information conveyed in pictures will be emphasized at the cost of information conveyed textually)

- Connectivity (objects able to connect to many others (having a high degree) will attract attention compared to objects making few connections)
Example of misguided emphasis
Examples of intuitive symbols

Intersection

Sequence/Causality

Subclass/Subset

Hierarchy
Semantically ’perverse’ notation in UML component diagram
Participant appropriateness

- The stakeholders must be able to use (understand) the language efficiently

- Basis: Should be according to how people perceive reality. No standard answer on this.
  - Can judge based on general experiences from learning of modeling languages
  - Clarify the participants experience with modeling languages (and adapt if possible the languages and views to models for this)

- Possible to represent inconsistencies

- External representation
  - Intuitive

- Support the achievement of pragmatic quality (relative to comprehension)
Modeler appropriateness

- \( K \setminus L = \emptyset \)
- For a prescriptive model is it possible to capture a situation in a model, or will it always be a post-hoc rationalization?
- Use of metaphors etc. ala Nonaka to support making tacit knowledge explicit
- Differences between novice and expert modelers
- Need to also be able to model vague, incomplete knowledge
- Support the achievement of semantic quality
Modeling of metaphors and analogies

“Leading a research groups is like herding cats”
Tool appropriateness

- The language is appropriate for automatic reasoning and execution.

- Formal syntax

- Formal semantics (operational and/or mathematical)
  - Discover inconsistencies and ambiguity
  - Formal proofs possible
  - Translation into executable models
  - Ensure that different participants interpret the model in the same way.

- Analyzability/executability

- Specific requirements for languages that are to be used for modelling ‘by hand’ (enterprise modelling)

- Support the achievement of syntactical, semantic and pragmatic quality.
Organizational appropriateness

- **Language standardization (independent of tools)**
  - Increase the potential for reuse
  - Support organizational learning
  - Enable a common methodology

- **Tool standardization**
  - Decrease cost of modeling tools
  - Easier to transfer models between different modeling task (within a project) and modeling activities (between projects)
Trade-off between different criteria

- The different criteria influence each other, and partly works against each other.
  - Intuitivity vs. the number of symbols: Chinese vs. English
  - Intuitivity vs. Simplicity of symbols: Classes vs. instances in ERAE
  - Expressiveness vs. Symbol simplicity: EEML vs. DFD
  - Expressive economy vs. Number of symbols: DFD – bidirectional arrows

- A perfect language is not possible to make, weaknesses in a language can be made less acute through method and tool support.
Language representation as a model

- A language model typically contains:
  - Meta-model (for language semantic)
    - Conceptual meta-model (for human understanding)
    - Logical meta-model (e.g. class diagrams for logical consistency)
    - Physical meta-model (for tool developers)
  - Notation (for syntax and visual appearance)

- Might also be in the form of tutorials etc.

- How a language is supported in a tool can also be evaluated using the quality levels.
Quality of a language model

- **Physical**
  - The language model is available

- **Empirical**
  - As any other model

- **Syntactic**
  - All examples follow the syntax of the language. The meta-model follows the syntax of the meta-modeling language

- **Semantic**
  - The whole language is described correctly. The meta-model and the notation guide is consistent.

- **Pragmatic**
  - Indexes, cross-references, tutorials etc.

- **Social**
  - vs. Language standardization, and agreement related to meta-model extensions

- **Organizational**
  - Minimize training time
Aspects of language quality – analytical evaluation

- Is the language appropriate to the domain?
- Is the language appropriate for the participants’ knowledge of modeling languages?
- Is the language appropriate to express the knowledge of the modeller?
- Is the language appropriate for making comprehensible model?
- Is the language appropriate for tools?
- Is the language appropriate for the chosen/standardized tools and modeling languages within the organization?
Quality of modeling languages

John Krogstie
Professor, IDI, NTNU