Towards Understanding the Understandability of UML Models

Chintan Amrit
IEBIS Department
University of Twente
Enschede, Netherlands
c.amrit@utwente.nl

Niek Tax
EWI Faculty
University of Twente
Enschede, Netherlands
n.tax@student.utwente.nl

ABSTRACT
Metrics for conceptual diagram understandability are generally based on a particular modelling notation and do not consider multiple perspectives in arriving at measurable metrics. This paper aims to be a step towards integration of multiple perspectives of conceptual model understandability. We demonstrate how one can arrive at a suite of metrics for conceptual diagram understandability. We use a questionnaire based on a few selected metrics to gain insights on the relative impact of different perspectives (namely, the notation and aesthetic perspectives) on conceptual diagram understandability. The main result of this study is that given a particular modelling notation (UML) and novice modellers, certain aesthetic aspects seem to have a greater impact on perceived diagram understandability than the notation aspects. We conclude with a call for further research in order to draw valid conclusions about the overall impact of the aesthetic and notation perspectives on diagram understandability.

Categories and Subject Descriptors
D.3.3 [Programming Languages]: Language Constructs and Features – abstract data types, polymorphism, control structures.

General Terms
Measurement, Design, Languages.

Keywords
UML, Diagram understandability, Model syntax

1. INTRODUCTION
Communication is critical for efficient workplace coordination [1]. For efficient stakeholder communication it is of importance that the conceptual models used for communication are easily understandable [2]. Lindland et al [3] state that the quality of a conceptual model is determined largely by how easily all the concerned parties understand the model.

The understandability and the ability to process information in a conceptual diagram depend on the appearance of the diagram. Aesthetic properties like the number of Line Crossings and Line Bends have been known to influence the degree to which stakeholders are able to read and cognitively process the information in a conceptual diagram [4]. Other non-aesthetic factors like diagram structure (diagram complexity) [5] and modelling notation choice [6] have been shown to influence the degree to which stakeholders can read and cognitively process the information in a conceptual diagram.

Literature on Model Understandability can be broadly categorised into three perspectives: (i) Notation Perspective: Studies in the notation perspective focus on the fit between a chosen modelling notation and the people working with the diagram and the MOSE '14, May 31 - June 07, 2014, Hyderabad, India. Aesthetic properties like the number of Line Crossings in the diagram are examples of a diagram’s aesthetics, and (iii) Structure Perspective: Studies in the structure perspective focus on structural properties of a diagram, like the number of entities and the number of relationships in a diagram. The structure of a diagram is related to the semantic concepts that are modelled in the diagram.

We base this three perspective categorisation on Moody’s [6] categorisation into sentence level and language level perspectives. Moody (2009) defines language perspective as the fit among a chosen modelling notation, the people working with the diagram and the diagramming task (our notation perspective). Sentence level can be described as all aspects that describe the actual model visualisation, such as line bending, line crossings as well as the number of instances of particular tokens (which is both our aesthetics perspective and structure perspective).

Although several models exist that look at different perspectives of conceptual diagram understandability, they mostly focus on one of these perspectives (for example, aesthetic or non-aesthetic) to analyse diagrams. As an example, we could not find any papers that looked at sentence level as well as language level understandability by considering and comparing the different aspects and metrics and critically validating them. To be able to improve conceptual diagram understandability, multiple perspectives are relevant and cannot be neglected. This leads to the following research question:

Which of the perspectives (notation, aesthetics or structure) has a higher impact on understandability of a particular modelling notation for a particular stakeholder?

This question would help us gain some insight into what conceptual modellers can do to improve conceptual model understandability for a particular stakeholder. In this paper we attempt at a partial answer of this question by combining two perspectives to analyse model understandability; namely the notation and the aesthetic perspective. We then test which of these diagram characteristics impact the understandability of UML more, through a case study performed on students.

The next section details the literature background; this is followed by a description of our questionnaire design and the survey performed. In the following section 4 we analyse the results of the questionnaire and end in section 5 with a discussion of the results.

2. MODEL APPEARANCE
The following four aspects are understood to have an influence on the understandability of a diagram:

- Modelling notation [6]
- Model semantics [4]
• Sentence level design choices [4, 7]
• Modelling tool [8]

Table 1 describes the particular perspective for which each of the four aspects fit.

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notation perspective</td>
<td>Modelling notation</td>
</tr>
<tr>
<td></td>
<td>Modelling tool</td>
</tr>
<tr>
<td>Aesthetics perspective</td>
<td>Sentence level design choices</td>
</tr>
<tr>
<td></td>
<td>Modelling tool</td>
</tr>
<tr>
<td>Structure perspective</td>
<td>Model semantics</td>
</tr>
</tbody>
</table>

In the following sections, we explain the aspects influencing diagram appearance. We then apply the categorisation of perspectives from Table 1 to discuss the aspects in a structured manner.

2.1 Notation perspective

Moody [6] describes eight principles that determine the degree to which a chosen conceptual modelling language fits the characteristics of the task that it is used for, and also fits the problem solving skills of the stakeholders [2].

In order to have an optimal fit, different modelling notations or visual dialects may be required for different tasks and audiences. One of the reasons for considering fit is the big difference in the understanding of conceptual diagrams between novices and experts [9].

Moody defines the following eight principles:

- Perceptual Discriminability
- Graphic Economy
- Semantic Transparency
- Semiotic Clarity
- Visual Expressiveness
- Cognitive Integration
- Complexity Management
- Dual Coding

In order to measure the impact of these principles on diagram understandability, we need a way to measure them. Hence, we can arrive at metrics for measuring the extent to which each principle is met in a given diagram (sentence level). These metrics are shown in Table 2 (# indicates number). In this paper we propose and then validate some of these metrics, as well as use them to draw conclusions on the objective application of Moody’s [4] principles to diagrams. Note that we provide metrics for all the principles at the sentence level; hence, the metric for Graphic Economy principle cannot be used, as it is defined at the level of the language. The metrics can be used to gain some insight on how well some of Moody’s [4] principles can be applied in a particular diagram.

At the language level, a tool can influence the graphical appearance; especially in the way it fits the notation. E.g., a tool may not support all the graphical symbols of a particular modelling notation. This would limit the number of graphical symbols the modeller can use and may thereby influence the understandability of a diagram.

2.2 Aesthetics perspective

2.2.1 Sentence Level Design Choices

The modelling notation and semantics are not the only influencers for diagram appearance. For a given modelling notation and semantics, the graph designer can make a lot of design decisions that can affect the diagram’s understandability. These decisions include size of graphical elements, place of graphical elements, and font selection.

2.2.2 Diagramming Tool influencing Aesthetics

At the sentence level, a tool can influence the appearance of a diagram in the following ways:

- By applying an Automatic Graph Layout algorithm.
- By using the different graphical symbols offered by the tool

Automatic Graph Layout algorithms make some aesthetic choices independently of the diagram designer. At the same time, it is conceivable that an independently made aesthetic choice of the Automatic Graph Layout algorithm persuades the designer to make different sentence level design choices. The Automatic Graph Layout algorithm assists the diagram designer in several ways. These include automatic alignment of graphical elements and in avoiding Line Crossings and Line Bends. Hence, sentence level graph aesthetics are usually an interplay between sentence level design choices and the Automatic Graph Layout algorithm of the tool.

A modelling notation determines the appearance of a model largely by the graphical symbols it uses. However, there are still
some free variables available for the tool to vary. Examples of free aesthetic variables often implemented by the tool are line thickness and font choices within graphical symbols.

Currently, most diagram modelling tools do not implement state-of-the-art Automatic Graph Layout algorithms and are not capable of generating very understandable diagrams [10].

2.2.3 Purchase Metrics

Purchase [11] defines a set of metrics defining understandability for several sentence level diagram aesthetics. She defines seven principles that reach optimal understandability [11]. For each principle, a metric is defined to identify the degree to which a diagram fulfils this principle.

The following seven principles are defined by Purchase [11]:

- Minimizing edge crossings
- Minimizing edge bends
- Maximizing symmetry
- Maximizing the minimum angle between edges leaving a node
- Maximizing Edge Orthogonality
- Maximizing Node Orthogonality
- Maximizing consistent flow direction (directed graphs only)

2.3 Structure perspective

The real-world objects and events that are being modelled have influence on the appearance of the graph. This is self-evident, as a different real-world model implies different semantic constructs that result in different symbol instances. The metrics that can be used to determine the understandability of the conceptual diagram structure depend on the notation in which the diagram is created. Genero (2005) created a set of metrics which can be used to determine understandability of ER diagrams [5]. Several metrics exist for UML diagrams, but they are still in early stages of development and are mostly not validated [12].

The real-world objects and events being modelled are almost in all cases not changeable by the diagram designer. Due to this reason we do not consider the influence of diagram semantics in our questionnaire design. Instead, we focus on Notation and the Aesthetic perspectives for designing our questionnaire.

3. QUESTIONNAIRE DESIGN

In order to provide answers to our research questions, we created a questionnaire\(^1\) containing eleven conceptual models modelling the same real-world scenario.

We varied these eleven conceptual models both on aspects of notation and aesthetic perspectives. Among the many perspectives we chose four aspects of aesthetics perspective and two aspects of notation perspective. By testing the understandability of the diagrams while varying both the aesthetic perspective and notation perspectives, we could gain some insights on the relative impact of these particular perspectives on conceptual diagram understandability.

We left out the structure perspective from the questionnaire as we felt that diagram designers are often not able to modify the semantic aspects within the structure perspective.

### 3.1 Aesthetics perspective aspects

Within the aesthetics perspective we chose Line Crossings, Line Bends, Node Orthogonality and Edge Orthogonality to vary the diagrams on, as these aspects have been validated to make a difference on understandability of UML class diagrams [4].

In the following sections, the metrics for Line Crossings, Line Bends, Node Orthogonality and Edge Orthogonality as defined by Purchase [11] will be explained briefly. In these sections, the following definitions are applicable:

- \( n \) the number of nodes in the diagram
- \( m \) the number of edges in the diagram
- \( u_i \) the \( i \)th node in the diagram
- \( h \) the number of vertical grid point in the diagram
- \( w \) the number of horizontal grid point in the diagram
- \( n' \) the number of nodes in a diagram if all bends in the edges are promoted to nodes
- \( m' \) the number of edges in a diagram if all bends in the edges are promoted to nodes
- \( \theta_i \) the positive angle between the \( i \)th edge and the x-axis

#### 3.1.1 Line Crossings

Purchase [11] defines an approximation for the upper bound of the number of edge crossings. The number of edge crossings is calculated assuming every edge of the graph crosses every other edge

\[
c_{\text{all}} = \frac{\sum_{i=1}^{m'} (i - 1)}{2} = \frac{m'(m'-1)}{2}
\]

Where \( m' \) is the total number of edges (each denoted by \( i \)). This is an overestimate that includes the total number of edge crossings, known to be impossible in a straight-line diagram. In straight-line drawings of connected graphs with at most one edge between nodes, adjacent edges cannot cross. The total number of such impossible crossings is therefore:

\[
c_{\text{impossible}} = \frac{1}{2} \sum_{i=1}^{n'} \text{degree}(u_j)(\text{degree}(u_j) - 1)
\]

Where \( \text{degree}(u_j) \) denotes the degree of the node \( u_j \). The upper bound on the number of edge crosses in a diagram is therefore:

\[
c_{\text{mx}} = c_{\text{all}} - c_{\text{impossible}}
\]

The actual number of crosses in a diagram is scaled against the maximum number of crossings. To reach a scale in which 1 represents maximum ‘crosslessness’, the scaled measurement of crosses is subtracted from 1.

\[
\text{Line crossings} = 1 - \begin{cases} 
\frac{c}{c_{\text{mx}}} & \text{if } c_{\text{mx}} > 0 \\
0 & \text{otherwise}
\end{cases}
\]

#### 3.1.2 Line Bends

It is impossible to scale against an upper bound of the number of bends, which is infinite, so we divide it by the total number of edge segments:

\[
b_{\text{avg}} = \frac{m' - m}{m'}
\]

So that 1 represents maximum ‘bendlessness’, this scaled measurement of bends is subtracted from 1.

\[
\text{Line bends} = 1 - b_{\text{avg}}
\]

---

\(^1\) [http://www.thesistools.com/web/?id=236082](http://www.thesistools.com/web/?id=236082)
3.1.3 Edge Orthogonality
The edge deviation factor of the $i$th edge segment ($\delta_i$) represents how far away from an orthogonal angle the edge segment has deviated. It is computed as a proportion of the angular deviation of the $i$th edge segment from the horizontal or vertical gridlines:

$$\delta_i = \frac{\min (\theta_i, 190^\circ - \theta_i, 180^\circ - \theta_i, 360^\circ - \theta_i)}{45^\circ}$$

So that 1 represents drawings with optimal edge deviation with respect to the orthogonal grid (i.e., assumed easier to read), the average edge deviation factor over all edge segments is subtracted from 1.

$$\text{Edge orthogonality} = 1 - \frac{1}{m} \sum_{i=1}^{m} (\delta_i)$$

3.1.4 Node Orthogonality
The number of available grid-point intersections in the imaginary unit grid is:

$$A = (w + 1)(h + 1)$$

Purchase defines the Node Orthogonality metric as the extent to which the drawing makes maximal use of the grid area:

$$\text{Node orthogonality} = \frac{n'}{A}$$

3.2 Notation perspective aspects
For the notation perspective we choose Perceptual Discriminability and Semiotic Clarity (Moody [6]) to modify for the different diagrams. We choose these particular Moody principles, as both the principles directly or indirectly influence the cognitive fit of the diagram. While, at the same time, they are the most independent of other Moody principles [6] (Figure 36, page 772). Furthermore, the reasons for leaving out the other aspects of Notational perspectives from our questionnaire are as follows:

- Graphic Economy – is hard to measure
- Visual Expressiveness – impacts the language level
- Cognitive Integration – impacts the language level
- Complexity Management - by choice we wanted to use the same Objects and Relationships
- Dual Coding – always possible with UML
- Cognitive Fit – is an aspect of the dependent variable (diagram understandability)

In the following sections a brief explanation of Perceptual Discriminability and Semiotic Clarity is given. Metrics for both aspects can be found in Table 2.

3.2.1 Semiotic Clarity
Symbol redundancy was added to several of the diagrams in order to decrease the Semiotic Clarity. We expected that the introduction of symbol redundancy to the UML diagrams would have a negative influence on diagram understandability.

3.2.2 Perceptual Discriminability
The other notation perspective aspect that we varied in the diagrams was the Perceptual Discriminability between symbols within the notation. Perceptual Discriminability is the ease and accuracy with which different graphical symbols can be differentiated from each other. Accurate discrimination between symbols is a necessary prerequisite for accurate interpretation of diagrams.

Perceptual Discriminability is primarily determined by the visual distance between symbols, which is defined by the number of visual variables on which models differ and the size of these differences. In general, the greater the visual distance between symbols used to represent different constructs, the faster and more accurately they will be recognised. Bad Perceptual Discriminability has been found to affect modelling novices more than the experts [13].

The UML class diagram uses only the visual shape variable to differ between types of symbols, and even within this visual variable, rectangular shapes are mostly used [14]. The use of colour as visual variable is specifically avoided in UML:

“UML avoids the use of graphic markers, such as colour, that present challenges for certain persons (the colour blind) and for important kinds of equipment (such as printers, copiers, and fax machines).” [15]

In several of the questionnaire diagrams, colour is added to the UML notation by assigning each graphical symbol a different colour for better symbol discriminability. We expect that this increased symbol discriminability will lead to a better understanding of the diagram, especially with users that are less familiar with conceptual modelling.

3.3 Variance of aspects in the questionnaire diagrams
To measure the relative impact of Line crossings, Line Bends, Node Orthogonality, Edge Orthogonality, Semiotic Clarity and Perceptual Discriminability on the perceived understandability of a diagram, we created several diagrams that differed on these aspects. We then used the metrics as defined by Purchase [11] (explained in section 5.1) to measure the extent of each aspect.

The metric for Perceptual Discriminability is defined as the visual distance in a diagram (Table 2). In UML, class diagrams only use shape as visual variable. When colour is added as extra visual variable to discriminate between symbols, the Perceptual Discriminability metric of that UML class diagram increases from 1 to 2. Moody states that the more visual variables the graphical symbols differ on, the more visual variables on which they are, harder to measure [16].

The metric for Semiotic Clarity can be defined as the number of instances of symbol redundancy, symbol overload and symbol excess in a diagram (Table 2). Most UML class diagrams do not include instances of symbol redundancy, symbol overload and symbol excess in a diagram. Those diagrams have a Semiotic Clarity score of 0. In several diagrams an instance of symbol redundancy is added, those diagrams have a score of 1 on the Semiotic Clarity metric. Table 3 provides an overview of how well each of the 11 diagrams scores on each metric. The aesthetic perspective aspects (Line Crossings, Line Bends, Node Orthogonality and Edge Orthogonality) have a score on a scale from 0 to 1. Purchase perceives a higher score to have a positive effect on diagram understandability [11]. The dependent variable “perceived diagram understandability” is measured on a five point Likert scale.

The properties of the diagrams were chosen in such a way that most diagrams differed on one or more of the defined aesthetic perspective properties, one or more of the defined notation perspective properties, or a combination of both.

We need to compare the perceived understandability of these diagrams to the perceived understandability of a clean UML diagram without the defined flaws.
4. RESULTS & DISCUSSION

In total 47 students filled the questionnaire and we obtained 517 data points, with 11 diagrams per subject. We start by using a one-sample T test to determine if the familiarity of the model subject throughout the questionnaire has an influence on the diagram understandability.

### 4.1 One-sample T test

<table>
<thead>
<tr>
<th>Diagram</th>
<th>Line Crossings</th>
<th>Line Ends</th>
<th>Node Orthogonality</th>
<th>Edge Orthogonality</th>
<th>Perceptual Discriminability</th>
<th>Semantic Clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.57</td>
<td>1.00</td>
<td>0.29</td>
<td>0.58</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>0.80</td>
<td>0.32</td>
<td>1.00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.79</td>
<td>1.00</td>
<td>0.32</td>
<td>0.81</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1.00</td>
<td>0.33</td>
<td>0.04</td>
<td>0.49</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.80</td>
<td>0.28</td>
<td>1.00</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0.79</td>
<td>1.00</td>
<td>0.28</td>
<td>0.49</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>0.50</td>
<td>0.06</td>
<td>0.81</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0.57</td>
<td>1.00</td>
<td>0.29</td>
<td>0.58</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>0.33</td>
<td>0.04</td>
<td>0.49</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1.00</td>
<td>0.17</td>
<td>0.20</td>
<td>1.00</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1.00</td>
<td>1.00</td>
<td>0.38</td>
<td>0.41</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

As described earlier, diagram 1 and diagram 8 are identical. If both diagrams score equally in their perceived understandability, we can conclude that the increase in semantic model familiarity throughout the questionnaire does not influence diagram understandability.

In table 4, it can easily be seen that the mean perceived diagram understandability and standard deviation of diagram 1 and diagram 8 are almost identical. It can therefore be concluded that the increasing familiarity of the various aspects of the diagrams (while answering the questionnaire) does not significantly influence perceived diagram understandability.

### 4.2 Multiple regression analysis

As previous research showed significant correlation of the differed factors in the diagrams and diagram understandability [4, 6], we presume all independent variables correlate with diagram understandability.

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandard Coeff.</th>
<th>Standard Coeff.</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.081</td>
<td>.598</td>
<td></td>
<td>1.808</td>
<td>.071</td>
</tr>
<tr>
<td>Line Crossings</td>
<td>1.472</td>
<td>.448</td>
<td>.230</td>
<td>3.285</td>
<td>.001</td>
</tr>
<tr>
<td>Line Ends</td>
<td>.600</td>
<td>.350</td>
<td>.173</td>
<td>1.715</td>
<td>.087</td>
</tr>
<tr>
<td>Node Orthogonality</td>
<td>-.946</td>
<td>.906</td>
<td>-.104</td>
<td>-1.044</td>
<td>.297</td>
</tr>
<tr>
<td>Edge Orthogonality</td>
<td>1.743</td>
<td>.263</td>
<td>.359</td>
<td>6.618</td>
<td>.000</td>
</tr>
<tr>
<td>Perceptual Discriminability</td>
<td>-.110</td>
<td>.166</td>
<td>-.045</td>
<td>-6.60</td>
<td>.510</td>
</tr>
<tr>
<td>Semiotic Clarity</td>
<td>-.619</td>
<td>.114</td>
<td>-.256</td>
<td>-5.444</td>
<td>.000</td>
</tr>
</tbody>
</table>

The questionnaire can be viewed online at: [http://www.thesistools.com/web/?id=236082](http://www.thesistools.com/web/?id=236082)
Using a multiple regression analysis, we gained insight into the impact of the independent variables on diagram understandability. Studies [16, 17] have shown there is little harm in treating Likert-type scales as continuous variables when five or more categories are used. Table 5 shows the results of running a multiple regression on the different variables with “diagram understandability” as the dependent variable.

From Table 5, we see that the predictor variables Line Bends, Node Orthogonality and Perceptual Discriminability are not statistically significant (p>0.01). As a result, the coefficients (Beta in Table 5) of these metrics provided by the multiple regression analysis can be expected to be unreliable. Therefore, we do not elaborate on these metrics.

The Line Crossings metric (p=0.001) and the Edge Orthogonality metric (p=0.00) have a very significant and a reasonably strong impact (Beta=0.23 and -0.104 respectively) on diagram understandability. We find the impact of Edge Orthogonality (Beta = 0.359) on diagram understandability to be more than the impact of Line Crossings (Beta = 0.230).

The impact of Semiotic Clarity (Beta=-0.256) on diagram understandability is very significant and is equal to the impact of Line Crossings. It seems, however, that Semiotic Clarity has a negative impact on diagram understandability, which contradicts the findings of Moody [6]. This may be caused by the familiarity of the test group with a colourless UML diagram, a colourful UML class diagram may conflict their expectations of a UML class diagram that might influence their perceived understandability.

5. CONCLUSIONS

Our research question in the introduction of this paper was the following:

Which of the perspectives (notation, aesthetics or structure) has a higher impact on understandability of a particular modelling notation for a particular stakeholder?

6. REFERENCES