TECHNICAL REPORT

ISO/TR 24156

First edition
2008-11-15

Corrected version
2009-02-15

Guidelines for using UML notation in terminology work

Lignes directrices pour l'application de la notation UML dans le travail terminologique

Reference number
ISO/TR 24156:2008(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO/TR 24156 was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 1, *Principles and methods*.

This corrected version of ISO/TR 24156:2008 incorporates the following corrections.

— In 5.4, in the sentence after Figure 3, "(see Figure 1 and also 5.7)" has been changed to "(see Figure 4 and also 5.7)".

— In 6.2, the UML notation for multiplicity has been corrected from (...) to (..) in the first paragraph and in Example 1.

— In 6.3, the UML notation for multiplicity has been corrected from (...) to (..) at the end of the first paragraph.

Introduction

Terminology work combines elements from many theoretical approaches which concern processing, ordering, and presentation of knowledge. The basic method of terminology work is concept analysis, which aims to achieve a comprehensive description and presentation of concepts in a subject field. Traditionally the results of concept analysis in terminology are presented in the form of one or more concept diagrams and a set of terms with textual definitions.

In object-oriented programming, graphic techniques are used to describe entity types which are characterised by certain properties and behaviour. The Unified Modeling Language (UML) is a widely spread language which can be used for all kinds of object modelling (information modelling, data modelling, etc.).

This Technical Report describes the application of UML graphical notation by creating a UML profile for the presentation of terminological concept analysis. This UML profile uses TC 37 terminology semantics to extend and partly replace UML semantics. This is not meant to become a replacement for traditional concept diagrams, but should be considered as an alternative and supplementary notation. This Technical Report is meant to promote the use of terminological concept analysis when developing information models and data models.

The core text describes the recommendations for use of the UML. Annex A contains a conversion table between concepts of ISO 1087-1:2000 and suggested representation in the UML.

ISO/IEC 19501:2005 is referenced in this Technical Report. In ISO/IEC 19501:2005 there is no clause “Terms and definitions”. Instead, every UML concept is described in the normative text and in a glossary. When a reference to ISO/IEC 19501:2005 is given in the term list, the definition given in this Technical Report is adapted from the descriptive text in ISO/IEC 19501:2005. Therefore, the definition is noted “Adapted from ISO/IEC 19501:2005”.

Guidelines for using UML notation in terminology work

1 Scope

This Technical Report gives guidelines for using a subset of the Unified Modeling Language, UML, to represent the results of terminological concept analysis. It describes how object modelling techniques can be used for this purpose. The representation of concepts and concept relations used in terminology work by corresponding entities in the UML is described.

This Technical Report does not describe the UML and its general use in depth. These matters are covered in e.g. ISO/IEC 19501.

This Technical Report does not describe the principles and methods of terminology work. This is covered in ISO 704[4].

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 1087-1:2000, *Terminology Work — Vocabulary — Part 1: Theory and Application*

ISO/IEC 19501:2005, *Information technology — Open Distributed Processing — Unified Modeling Language (UML) Version 1.4.2*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 1087-1 and the following apply.

3.1
terminological concept model
representation of a *concept system* [ISO 1087-1:2000] by means of a **formal language** (3.5)

3.2
concept model view
image of a defined part of a **terminological concept model** (3.1)

3.3
constraint
semantic restriction of model elements

NOTE 1 Adapted from ISO/IEC 19501:2005.

NOTE 2 A constraint is used to restrict the possible options for a class or a relationship. In terminological concept modelling, constraint can be used to show how relationships interact and how they are delimited.

EXAMPLE 1 There are two associative relations from a concept, but if one of them is present the other one is impossible [constraint {either}].

EXAMPLE 2 In a generic relation, no more specific concepts than those stated are possible [constraint {complete}].

3.4
core concept
concept [ISO 1087-1:2000] that has focus of interest in a group of related concepts

3.5
formal language
language with a defined set of allowed symbols and a precise definition of which strings composed from these symbols are considered syntactically correct and interpretable in formal logic

NOTE A formal language is not meant to be spoken. Its purpose is to assure exact communication of information, e.g. between computer systems, and between man and computer.

EXAMPLE OWL.

3.6
multiplicity
specification of the range of allowable numbers that a set may assume

NOTE 1 Adapted from ISO/IEC 19501:2005.

NOTE 2 In terminological concept modelling, multiplicity specifies how many objects a certain concept may correspond to. If the range of numbers specified by the multiplicity includes zero (0), the object is optional.

EXAMPLE 1 A characteristic of a month is that it is a period of 28-31 days (28..31).

EXAMPLE 2 A mouse (pointing device) may or may not have a ball, depending on whether it is a mechanical or optical mouse. It has zero balls or one ball (0..1). In that case the multiplicity itself is a criterion of subdivision, as a mechanical mouse has exactly one ball (1).

4 Abbreviated terms

OWL Web Ontology Language

UML Unified Modeling Language

5 Mapping of terminological items to UML

5.1 General

This clause describes how terminological concepts defined in ISO 1087-1:2000 can be represented in terminological concept modelling by means of a limited set of symbols in the UML. Features that are not described in this clause are outside the scope of this Technical Report.

In each paragraph, the principles of how the UML symbols can be used in terminological concept modelling are described.

Table A.1 (Annex A) describes ISO 1087-1 terms and the corresponding UML modelling elements.

5.2 Concept

A concept should be represented by a class symbol. The *designation* [ISO 1087-1:2000] of the concept is represented by the name of the class. This applies for *individual concepts* [ISO 1087-1:2000] as well as for *general concepts* [ISO 1087-1:2000].

The class symbol used to represent a concept normally has only one compartment, which contains the name of the class (see Figure 1). The class name should be written conformant with the rules in ISO 10241. The name should be boldface and not in italics.

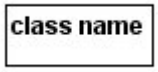

Figure 1 — Class

5.3 Concept system

A **terminological concept model** (3.1) is meant to depict and represent a concept system. A graphic tool may store the terminological concept model in a **formal language** (3.5), making it possible to communicate, in a machine-readable format, the model with conceptual data modelling, information modelling and system development.

5.4 Generic relation

A *generic relation* [ISO 1087-1:2000] is represented by the UML generalisation symbol.

The generalisation symbol is a line with an arrow ending with a triangle pointing towards the *generic concept* [ISO 1087-1:2000] (see Figure 2).

Figure 2 — Generic relation

There may be one arrow for each *specific concept* [ISO 1087-1:2000] (see Figure 3).

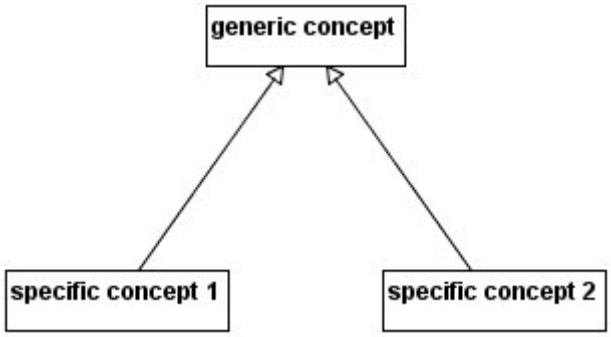

Figure 3 — Multiple generic relation

If there is more than one specific concept, a forked arrow may alternatively be used (see Figure 4 and also 5.7).

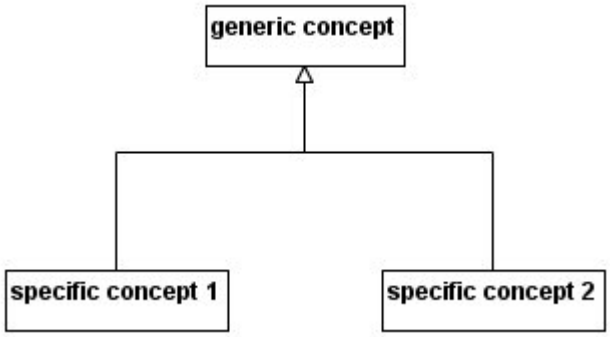

Figure 4 — Multiple generic relation depicted with forked arrow

A symbol may be provided which shows that there may be more specific concepts in a special subdivision than those depicted. This can be done by means of a **constraint** (3.3), named complete or incomplete (see also 6.3).

If all existing specific concepts are presented, the generic relation may be marked {complete} as shown in Figure 5.

Figure 5 — Generic relation with all existing specific concepts presented

If there are specific concepts that are not presented in the graph, the generalisation can be marked {incomplete} as in Figure 6.

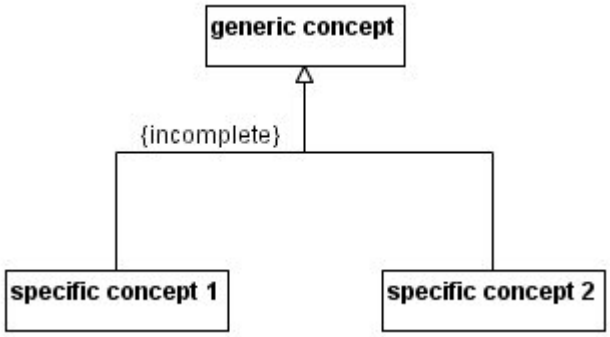

Figure 6 — Generic relation where not all specific concepts are presented

EXAMPLE Pointing devices and subtypes of them (see Figure 7).

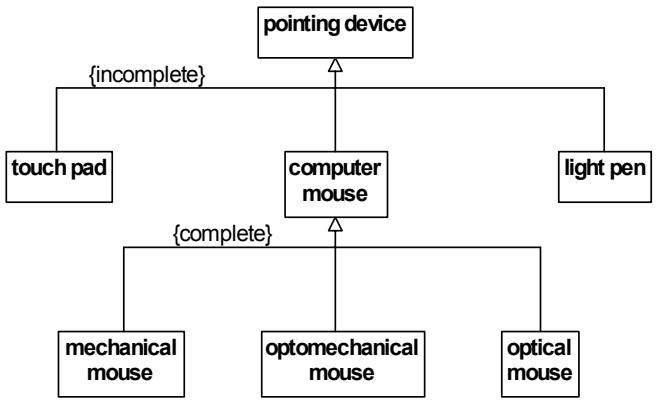

Figure 7 — Generic relations between pointing devices

5.5 Characteristic

A *characteristic* [ISO 1087-1:2000] can be modelled in two ways. One way is as a separate concept with its relation to the **core concept** (3.4). Semantically, a characteristic usually consists of the relation and the related concept. The other way is as an attribute-value pair of the class representing the concept having the characteristic (see Figure 8). If attributes are listed, they should be located in a separate compartment.

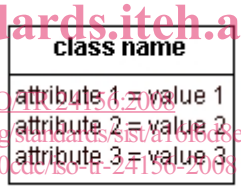

Figure 8 — Attribute compartment

An ellipsis (...) may indicate that there are attributes which are not shown in the graph (see Figure 9).

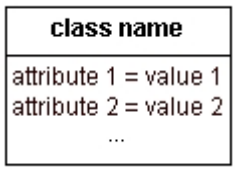

Figure 9 — Hidden attribute

5.6 Type of characteristics

In generic relations, the *type of characteristics* [ISO 1087-1:2000] serving as a criterion of subdivision may be modelled either as a class symbol following certain rules (*powertype class* [ISO 19501:2005]), or as an attribute in the class representing the generic concept with expressed values in the specific concepts (see Figure 10).