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)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

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.

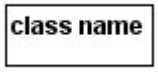


Figure 1 — Class

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).

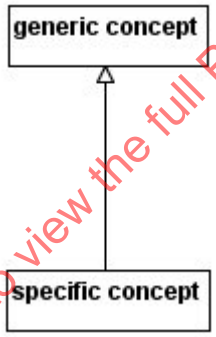


Figure 2 — Generic relation

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

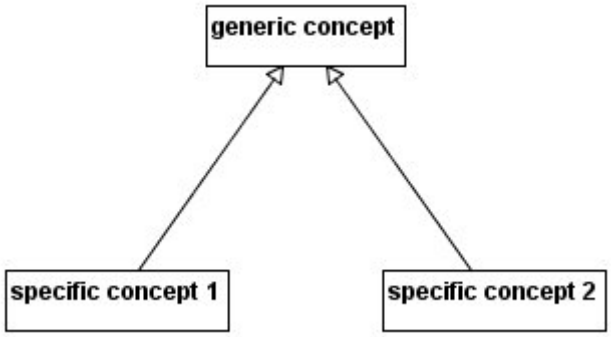


Figure 3 — Multiple generic relation

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

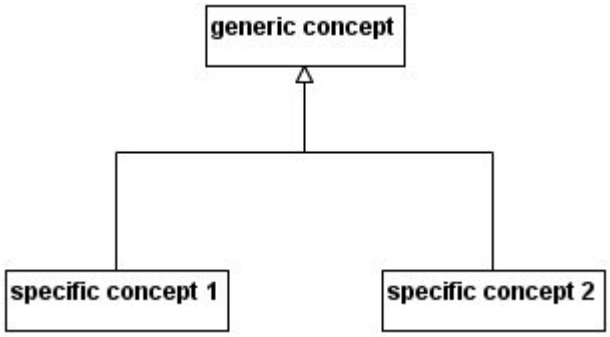


Figure 4 — Multiple generic relation depicted with forked arrow

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.

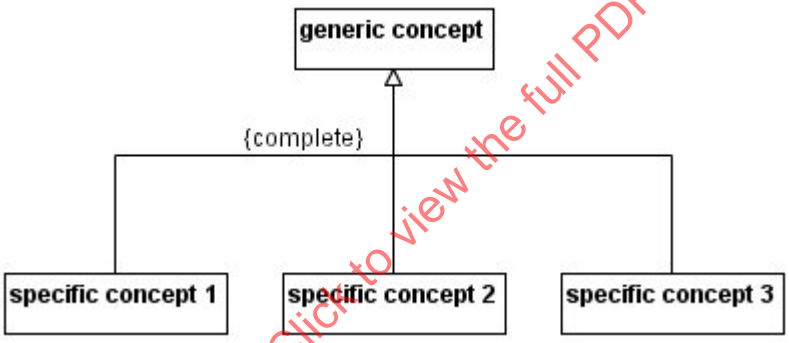


Figure 5 — Generic relation with all existing specific concepts presented

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

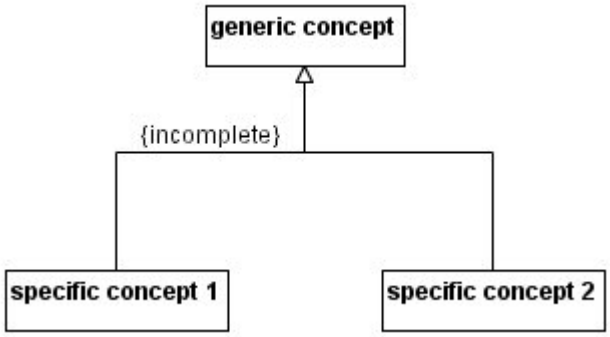


Figure 6 — Generic relation where not all specific concepts are presented

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

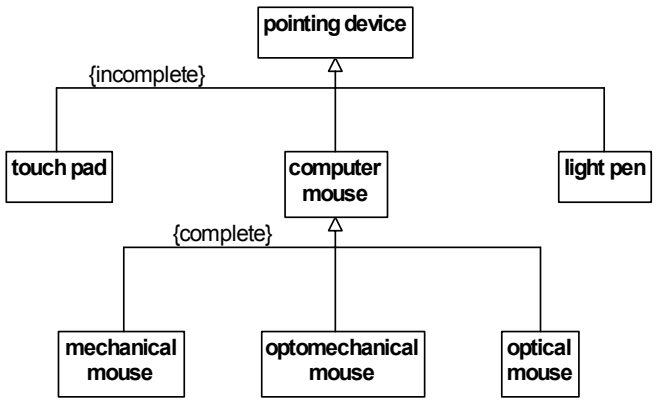


Figure 7 — Generic relations between pointing devices

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.

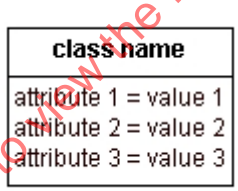


Figure 8 — Attribute compartment

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

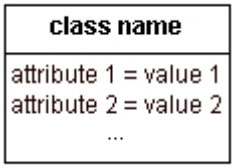


Figure 9 — Hidden attribute

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).

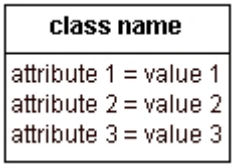


Figure 10 — Attribute with value determining the subdivision

5.7 Criterion of subdivision

In a multidimensional concept system, a generic concept may be specified by different criteria of subdivision, e.g. the concept “computer mouse” may be subdivided by “means of movement detection” or “kind of computer connection”. Criteria of subdivision may be modelled in three different ways.

a) A class symbol may indicate the type of characteristics, which is a criterion of subdivision. In the UML such a class, specifying the criterion for specialisation, is called a powertype class. It is modelled as a class symbol connected to the generalisation arrow with a dotted line, which should cross the arrow line. When more than one criterion of subdivision is shown in the same generic relation view, forked arrows for each subdivision are recommended (see Figure 11).

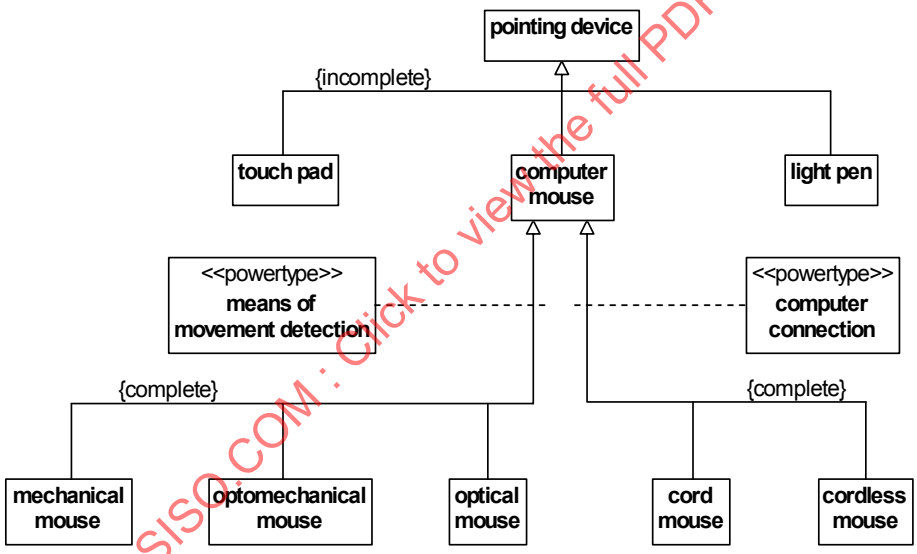


Figure 11 — Powertype

The powertype class does not represent any *delimiting characteristic* [ISO 1087-1:2000] of the generic concept. Its specialisations, however, represent delimiting characteristics of the specific concepts (see Figure 12).

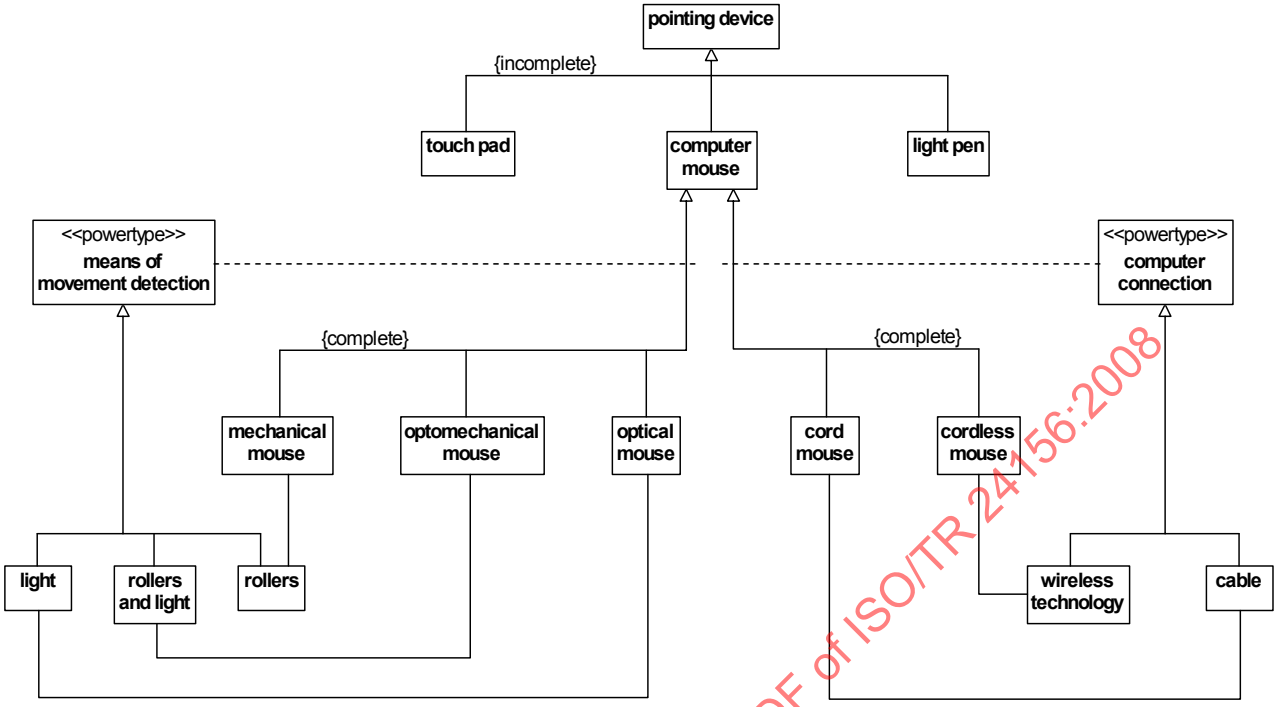


Figure 12 — Powertype with specialisations

b) The criteria of subdivision may be included as attributes in the generic concept class. These attributes do not represent delimiting characteristics of the generic concept. In the specific concept classes, the attributes get values, and these attribute-value pairs are parts of the delimiting characteristics of the specific concepts (see Figure 13).

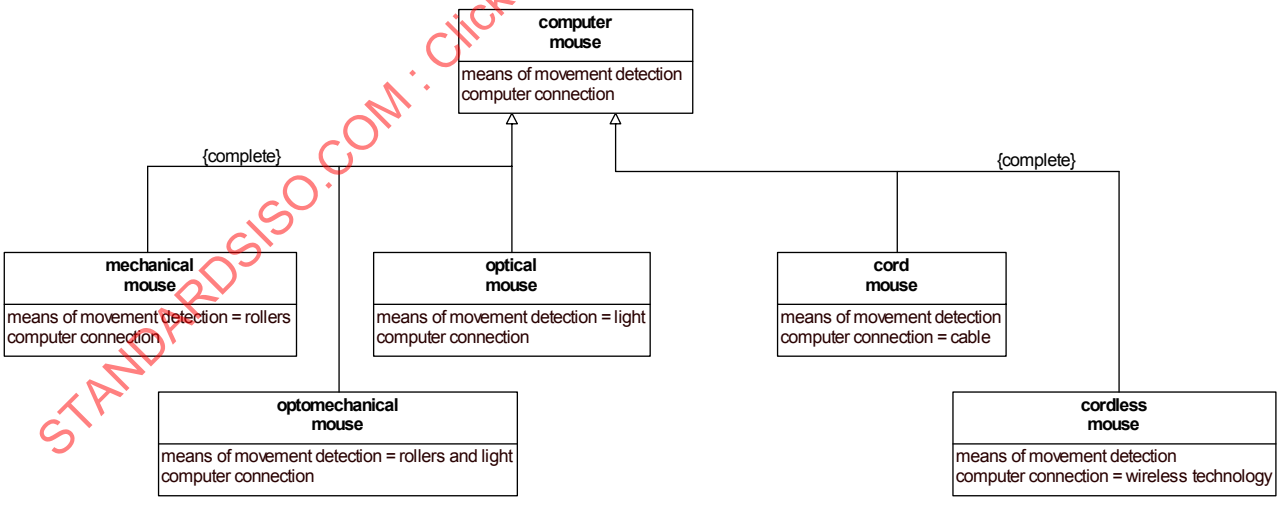


Figure 13 — Attribute representing a criterion of subdivision

With this method, multiple inheritances can be clearly shown (see Figure 14).

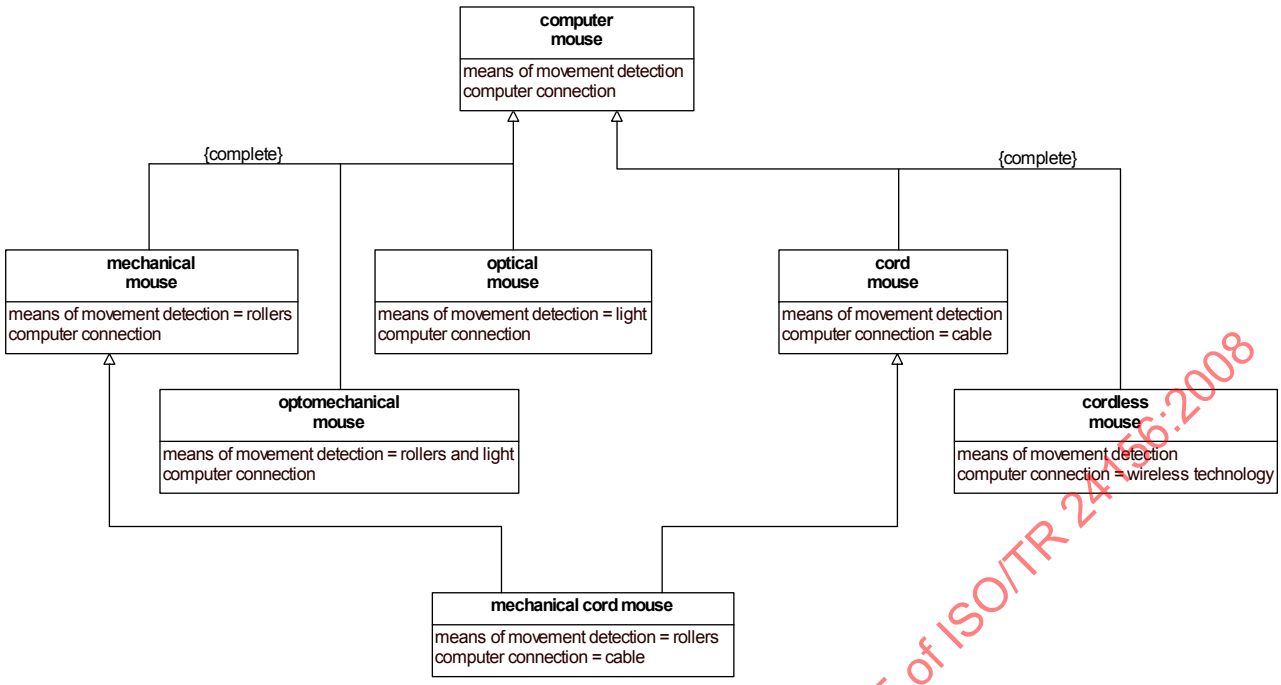


Figure 14 — Multiple generic relation

c) The criteria of subdivision may be represented by the UML discriminator symbol which is text written along the [forked] lines, see Reference [7] (see Figure 15).

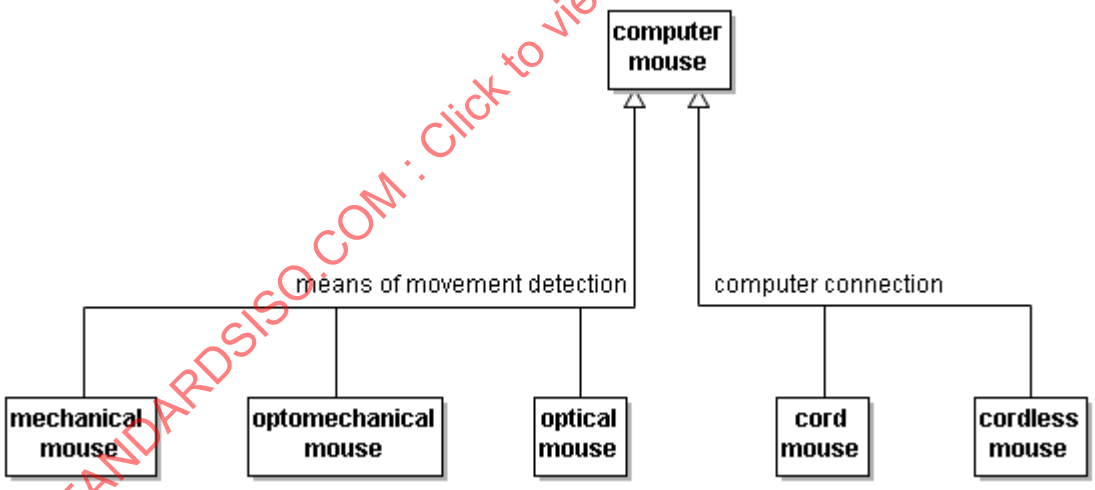


Figure 15 — Discriminator in generalisation

5.8 Partitive relation

A *partitive relation* [ISO 1087-1:2000] should be depicted using a UML aggregation symbol. The symbol is an arrow with a solid line pointing towards the *partitive concept* [ISO 1087-1:2000] and an open diamond end pointing towards the *comprehensive concept* [ISO 1087-1:2000] (see Figure 16).

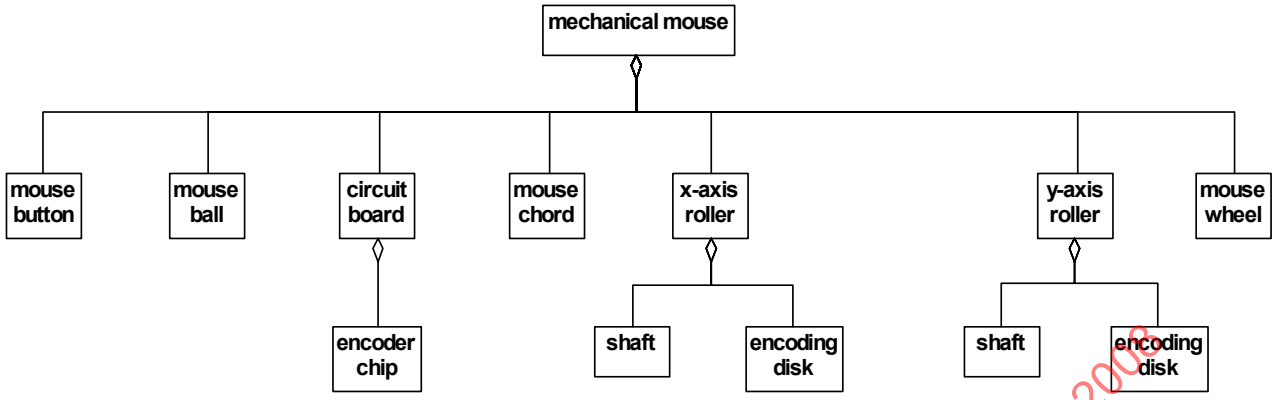


Figure 16 — Partitive relation

If a comprehensive concept is composed of more than one group of partitive concepts, this grouping should be clearly marked by means of generalisation (see Figure 17).

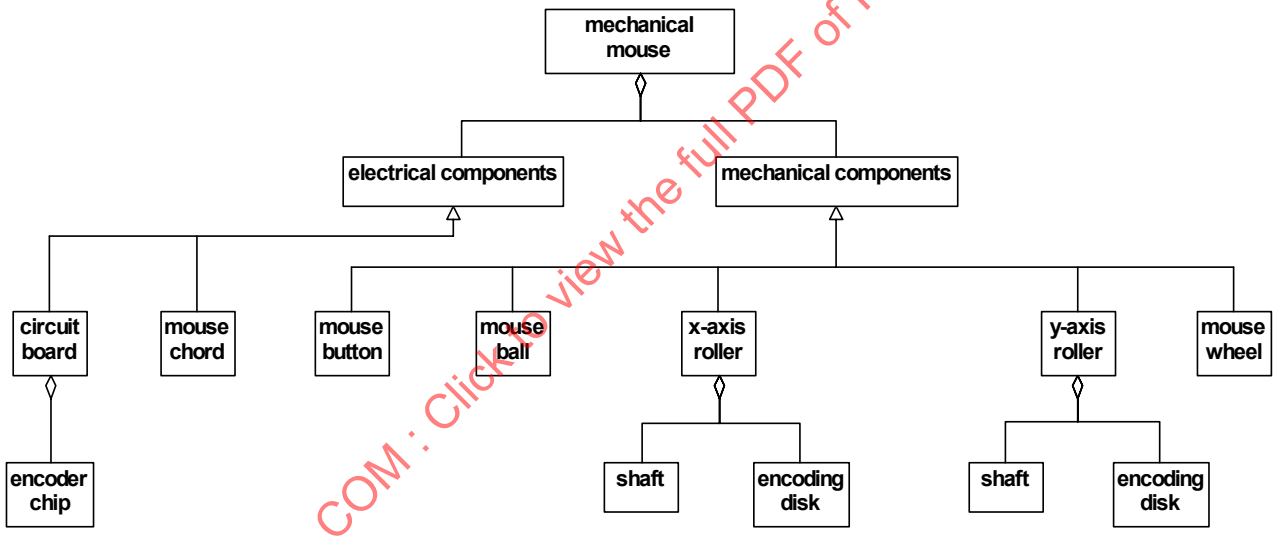


Figure 17 — Groups of components in a partitive relation

An ellipsis may indicate the possibility that there are partitive concepts in the relation which are not depicted (see Figure 18).

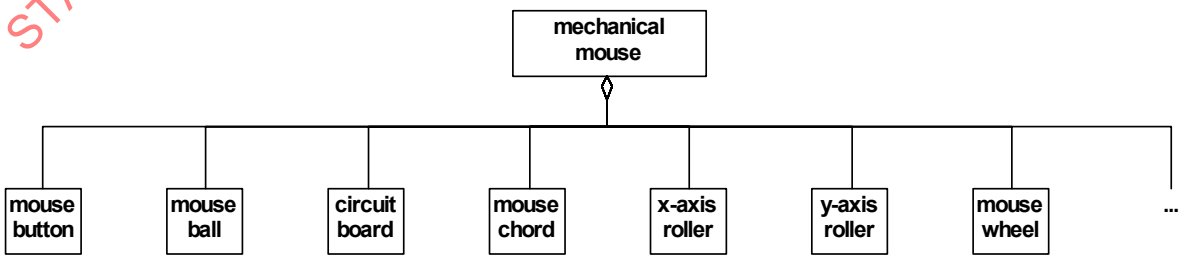


Figure 18 — Incomplete partitive relation

5.9 Associative relation

An *associative relation* [ISO 1087-1:2000] is represented by a UML association symbol. The association symbol is a line connecting the two concepts. An association may have a name, which is a description of the association. The name of an association is written adjacent to the association symbol. It should be written in italics (see Figure 19).

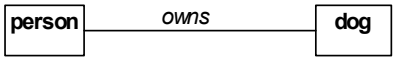


Figure 19 — Named associative relation

UML associations can be represented in an object relational schema either as unidirectional relationships or as bidirectional relationships. A unidirectional association means that the association can be crossed only in one direction, whereas a bidirectional association can be crossed in both directions. Associative relations between concepts are always bidirectional. An association symbol navigable only in one direction may therefore not be used in terminological concept modelling. The association has consequent names in both directions (reverse of the semantic link: “a person owns a dog” – “a dog is owned by a person”). If it is actually the same associative relation that is described in both directions, only one direction has to be named. When a name is added to the association, it may be adorned with a triangular arrow pointing in the name direction (see Figure 20).

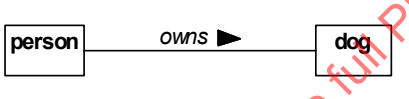


Figure 20 — Associative relation with direction of name

There may be two different associative relations to be modelled, and in that case each of them will have a name, possibly in different directions (“a dog bites a person” – “a person owns a dog”). Two different association lines should be drawn to clearly state that there are two different associations, both of them bidirectional (“a dog is owned by a person”, “a person is bitten by a dog”) (see Figure 21).

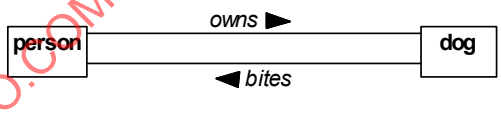


Figure 21 — Multiple associative relations between a pair of concepts

6 Common features of the UML used to extend terminological concept modelling

6.1 General

The UML provides a set of special features which can enrich a **terminological concept model** (3.1). The terminological **concept model view** (3.2) may, by means of those features, become more detailed than a common concept diagram, and with those features specified, the migration to an information model will be safe and avoid loss of information and need for supplementary input. This clause describes two features which specify certain properties of a concept system: **multiplicity** (3.6) and **constraint** (3.3).

6.2 Multiplicity

When modelling associative and partitive relations, it may prove practical to state **multiplicity** (3.6). The ends of an associative or partitive relation symbol may be adorned with multiplicity stating how many objects in the extension of the concept at that end can be related to one single object in the extension of the concept at the other end. The multiplicity could be any range of numbers from 0 to *, where * designates any number larger than 1. The notation is “n1..n2” where n1 is the lower limit of the range and n2 the upper limit of the range. If the occurrences can be only one number, e.g. 0 or 1, only that number is written. If there is always more than one occurrence, only * is written. Discrete series of numbers are allowed and should be separated by commas.

EXAMPLE 1 “0”, “1”, “5”, “0..1”, “0..*”, “1..*”, “3..*”, “10..15”, “*”, “0..3, 8, 11..*”.

EXAMPLE 2 A bicycle has 2 wheels; a wheel does not have to be mounted on a bicycle and cannot be mounted on more than 1 bicycle at a time (see Figure 22).

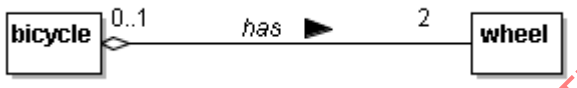


Figure 22 — Multiplicity in partitive relations

EXAMPLE 3 A book has to be written by at least 1 author and may be written by several authors. An author does not have to write books (may be a playwright or journalist) but may of course write more than 1 book (see Figure 23).

Figure 23 — Multiplicity in associative relations

6.3 Constraint

A **constraint** (3.3) represents a delimiting rule to an element or a set of elements in a concept model. One concept may be associated to two other concepts, but only one of the relations may be present at a time. Such a constraint is mathematically an “xor” constraint, but it may also be written “either”. A generic relation may be complete, meaning that there is no possibility that more specific concepts occur than those depicted. It may be incomplete, stating that there are or may be more specific concepts which are not depicted. The second example equals the notation in a concept diagram where one specific concept is shown as “..”.

Equally, a partitive relation may be complete or incomplete, the first alternative meaning that all partitive concepts making up the comprehensive concept are shown.

A **constraint** (3.3) is modelled in either of two ways. One is within curly brackets {constraint}, and the other is as a comment connected to the model element or model elements which it constrains. Both notations shown in Figures 24 and 25 mean that only one of the two relations may be present.

Multiplicity (3.6) is actually a kind of constraint, as it constrains the number of occurrences that is allowed of the model element addressed by the multiplicity.

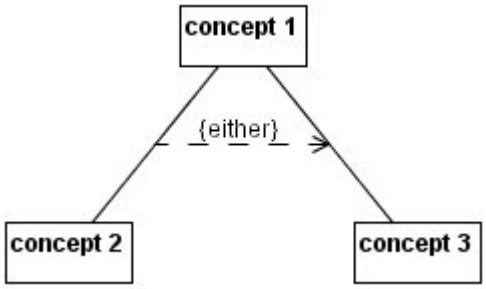


Figure 24 — Constraint between associative relations, notation option 1

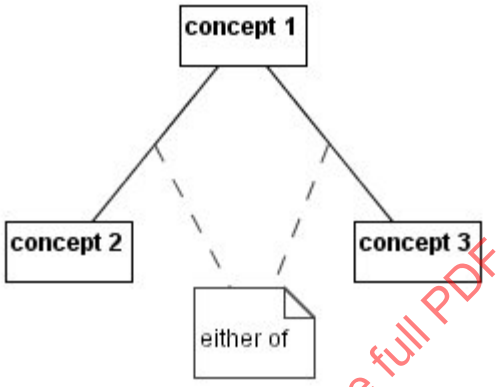


Figure 25 — Constraint between associative relations, notation option 2

Annex A
(informative)

Conversion table between concepts in ISO 1087-1:2000 and their representations in the Unified Modeling Language

A.1 General

This Technical Report suggests the use of UML modelling elements for representation of concept systems. Table A.1 below lists a set of concepts from ISO 1087-1:2000 and their corresponding concepts in the UML. Concepts in ISO 1087-1:2000 which are not listed have no corresponding representation of their own. For example, different kinds of associative relations are not represented by different kinds of UML elements. Instead, the UML association can be named in such a way that the kind of associative relation is clarified.

A.2 Table A.1 explanations

For each row, the term in ISO 1087-1:2000 is given in the first column. In the second column, the name of the model element in the UML is given, together with the explaining entry in the glossary clause of ISO/IEC 19501:2005. If one model element is mentioned more than once, the glossary entry is entered only in the first row where it occurs. If a UML model element is generally recognised, described in ISO/IEC 19501:2005 but not listed in the glossary, only the model element name is entered in the second column. The third column contains comments, explanations and graphic examples. Specifically, explanations of model elements which are not explained in the second column are given in the third column.

Table A.1 — Conversion between concepts in ISO 1087-1:2000 and their representations in the Unified Modeling Language

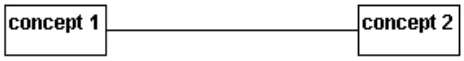
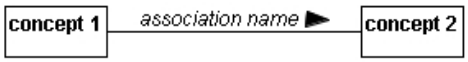
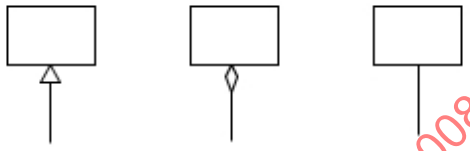
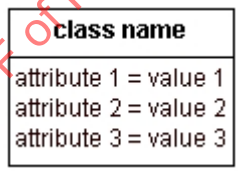
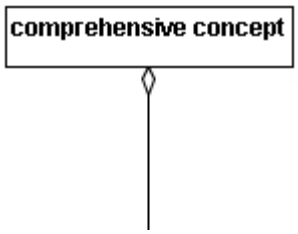
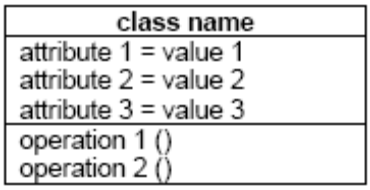
ISO 1087-1:2000	UML (ISO/IEC 19501:2005)	UML symbol/notation, comments
associative relation 	**association**	The association is modelled as a straight line between the classes representing the related concepts.  In the UML, the association may be given a name describing the kind of associative relation. This name may be adorned with a triangle acting as an arrow pointing in the direction in which the name is valid.  In concept systems, all associative relations are bidirectional. In the UML it is possible to model unidirectional associations. Such a notation is not applicable in terminological concept modelling.

Table A.1 (*continued*)

ISO 1087-1:2000	UML (ISO/IEC 19501:2005)	UML symbol/notation, comments
characteristic	**A:** **class**	A: A characteristic is a concept, and may therefore be represented in the UML as a class and its relationship to the concept which it characterises. 
	B: **attribute-value**	B: The characteristic may be represented in the UML as an attribute-value pair of the class representing the concept having the characteristic. If attributes are listed, they should be located in a separate compartment. 
comprehensive concept	**aggregate**	The aggregate is modelled as any class and identified by its relationship being the class at the open diamond end of the aggregation symbol. 
concept	**class**	Classes are represented by rectangles which show the name of the class and, optionally, the names of the operations, attributes and values. Compartments are used to divide the class name, attributes and operations. In terminological concept modelling, operations are never shown, and the class may have only one or two compartments. 
concept diagram	**terminological concept model view**	The concept “terminological concept model view” is not defined in the UML. Whilst a concept diagram is defined as any graphic representation of a concept system, a terminological concept model may be presented graphically as an image of the entire model or a part of the model.