



**International
Standard**

ISO 24613-6

**Language resource management —
Lexical markup framework (LMF) —**

**Part 6:
Syntax and semantics**

*Gestion des ressources linguistiques — Cadre de balisage
lexical (LMF) —*

Partie 6: Syntaxe et sémantique

**First edition
2024-04**

STANDARDSISO.COM : Click to view the full PDF of ISO 24613-6:2024



COPYRIGHT PROTECTED DOCUMENT

© ISO 2024

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword.....	iv
1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions.....	1
4 Description of the syntax and semantics module (SynSem) model.....	2
4.1 General principles of the SynSem model.....	2
4.2 Class description.....	4
4.2.1 <i>SyntacticBehaviour</i> class.....	4
4.2.2 <i>SubcategorizationFrame</i> class.....	4
4.2.3 <i>SyntacticArgument</i> class.....	4
4.2.4 <i>SubcategorizationFrameSet</i> class.....	5
4.2.5 <i>SynArgMap</i> class.....	5
4.2.6 <i>SenseRelation</i> class.....	5
4.2.7 <i>SemanticPredicate</i> class.....	5
4.2.8 <i>PredicativeRepresentation</i> class.....	5
4.2.9 <i>SemanticArgument</i> class.....	6
4.2.10 <i>ArgumentRelation</i> class.....	6
4.2.11 <i>SynSemArgMap</i> class.....	6
4.2.12 <i>SynSemCorrespondence</i> class.....	6
4.2.13 <i>PredicateRelation</i> class.....	6
Annex A (informative) Syntax-semantics examples.....	7
Bibliography.....	11

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 4, *Language resource management*.

A list of all parts in the ISO 24613 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Language resource management — Lexical markup framework (LMF) —

Part 6: Syntax and semantics

1 Scope

This document specifies the syntax and semantics (SynSem) module of the lexical markup framework (LMF), a metamodel for representing data in monolingual and multilingual lexical databases used with computer applications. The SynSem module allows for the description of specific syntactic and semantic properties of lexemes, as well as the complex interactions between them. More specifically, the syntax part of the module describes the properties of a lexeme when combined with other lexemes in a sentence. When recorded in a lexicon, these properties make up the syntactic description of a lexical entry instance. The semantics part of the module, on the other hand, describes the sense of a lexeme and its relationship with other senses belonging to the same language. The SynSem interface describes the predicates and the mapping between syntactic and semantic arguments.

This serialization covers the classes of ISO 24613-1 (Core model), ISO 24613-2 (Machine-readable dictionary (MRD) model) and ISO 24613-4 (TEI serialization).

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 24613-1, *Language resource management — Lexical markup framework (LMF) — Part 1: Core model*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 24613-1 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1

syntactic behaviour

one of the possible alternations that a lexeme can show, at the syntactic level

EXAMPLE A verb can have different types of syntactic behaviours for *subcategorization frame* (3.2) alternations, such as the active voice, the passive voice, reflexive, etc.

Note 1 to entry: A syntactic behaviour is described in terms of subcategorization frames.^{[9][12]}

3.2

subcategorization frame**valence****valency**

set of restrictions on a lexeme indicating the properties of the *syntactic arguments* (3.4) that can or must occur with this given lexeme

3.3

semantic predicate

formal semantic unit that represents a semantic relation between one or more *semantic arguments* (3.5) in a predicate-argument structure

[SOURCE: ISO 24617-4:2014, 2.1.2, modified — term “predicate” replaced by “semantic predicate”. “argument” replaced by “semantic argument” within definition. Note 1 to entry deleted.]

3.4

syntactic argument

one of the essential and functional constituents in a clause that identifies the participants in the process referred to by a lexeme

EXAMPLE Alfred (syntactic argument) reads a book (syntactic argument) today (adjunct).

3.5

semantic argument

formal semantic unit that is an essential constituent of a predicate-argument structure and can have variable instantiations depending on the utterance

[SOURCE: ISO 24617-4:2014, 2.1.1, modified — term “argument” replaced by “semantic argument”. Notes to entry deleted.]

4 Description of the syntax and semantics module (SynSem) model

4.1 General principles of the SynSem model

The modelling of the syntactic and semantic properties of lexicons can differ considerably from lexicon to lexicon, due to different linguistic theories.

The present modelling of syntax and semantics builds on the previous one, given in ISO 24613:2008^[2], which contained two sub-modules for syntax and semantics. In ISO 24613:2008, experts identified a nucleus of elements that are sufficiently generic to constitute a common ground. The ISO 24613:2008 syntactic and semantics models were the basis for other modelling of the syntax/semantics interface, such as the W3C Ontolex “Syntax and Semantics Module”.^[13] For these reasons, the overall structure of the previous model, which was used in various projects such as the PAROLE^[11] and SIMPLE^[8] projects, was not changed in this document, but some simplifications/modifications were introduced. In particular, the previous semantic module contained elements that were entirely dedicated to the modelling of WordNet-like lexicons, introducing a *Synset* class. This is judged to be unnecessary in this document since the *Sense* and *SenseRelation* classes can be used instead. Another difference with respect to ISO 24613:2008 is the lack of a *feat* class, which was used to make up for specific elements which a lexicographer can potentially want to introduce but which were not generic enough to be included in the model. In the present UML-based standardization, only the core features of the syntax/semantics interface are described. However, the lexicographer can extend the model to extend other features.

[Figure 1](#) gives a class diagram for SynSem with the module specific classes in pink. [Figure 2](#) shows the subclasses of CrossREF (described in ISO 24613-1) which are referred to in the class diagram in [Figure 1](#).

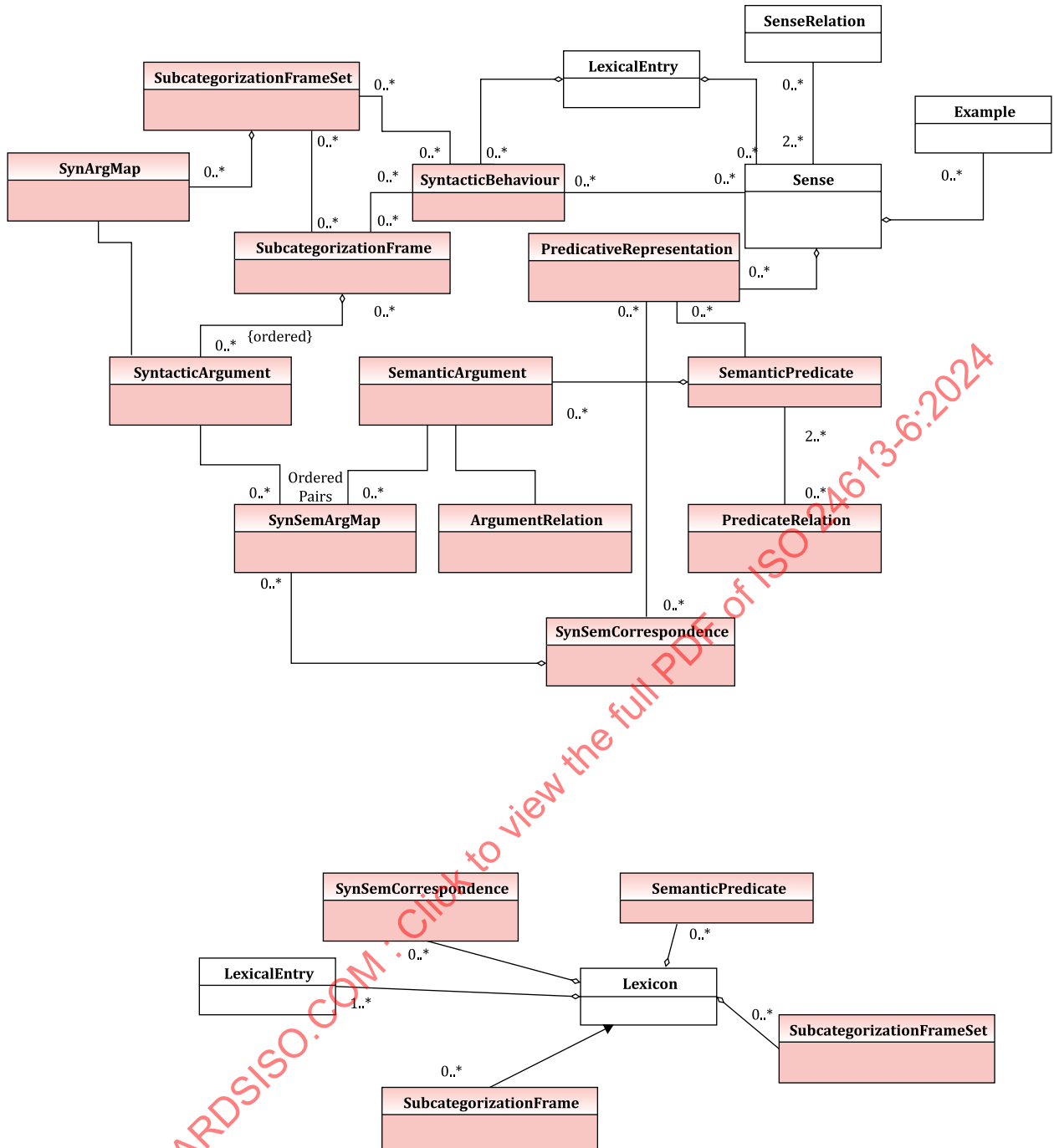


Figure 1 — SynSem module

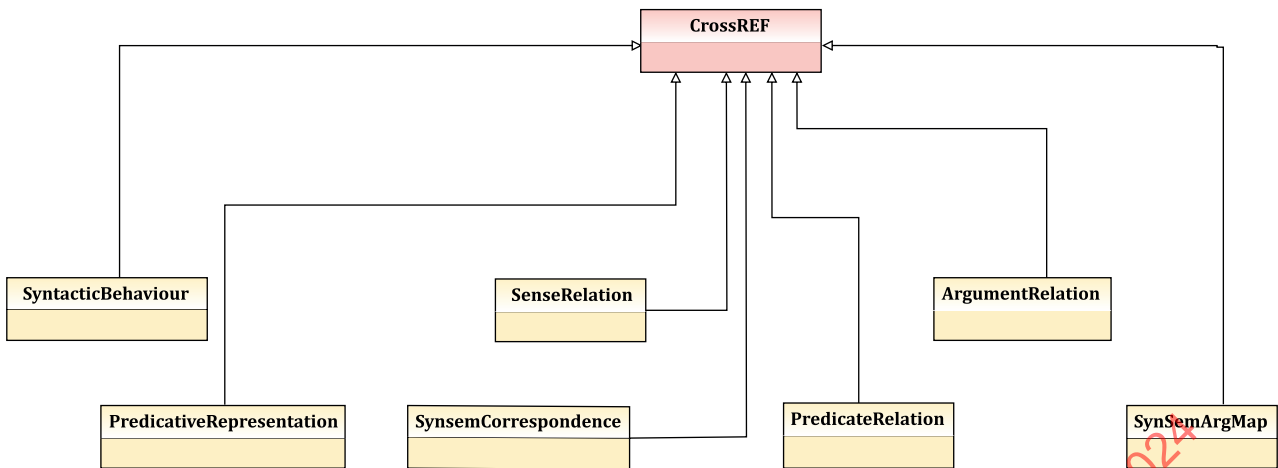


Figure 2 — Subclasses of CrossREF

Regarding the modelling of syntax in [Figure 1](#), a *LexicalEntry* may have one or more instances of *SyntacticBehaviour*, associated with separate *SubcategorizationFrame* instances, each described with *SyntacticArgument*. Regarding the modelling of semantics, it applies to senses. The *Sense* class is specified in the core package. The *Sense* class is aggregated in the *LexicalEntry* class. A *PredicativeRepresentation* serves to connect a *Sense* with one or more instances of *SemanticPredicate*, which are described in terms of *SemanticArgument* instances. The linking between syntax and semantics is done by the *SynSemArgMap* component, which links a *SemanticArgument* with a *SyntacticArgument*.

[Clause A.1](#) illustrates the full model, using the French verb “aider” (“help”) with its syntactic and semantic properties.

4.2 Class description

4.2.1 SyntacticBehaviour class

SyntacticBehaviour is a class representing one of the possible syntactic behaviours of a lexeme. It is a subclass of *CrossREF*, linking a *LexicalEntry* instance with a *SubcategorizationFrame* instance. The presence in a given lexicon of one *SyntacticBehaviour* instance for a lexical entry means that this entry can have this behaviour in the language under description. Syntactic description is optional, so it is possible to describe morphology and semantics without any syntactic description. *LexicalEntry*, *SyntacticBehaviour* and *Sense* instances form a triangle representing morphology, syntax and semantics. A detailed description of the syntactic behaviour of a lexical entry is defined by the *SubcategorizationFrame* instance.

4.2.2 SubcategorizationFrame class

SubcategorizationFrame is a class representing one syntactic construction. A *SubcategorizationFrame* instance is shared by all *LexicalEntry* instances that have the same syntactic behaviour in the same language.

EXAMPLE In a *LexicalEntry* for the Italian verb “amare”, a *SyntacticBehaviour* instance can be created and associated with a *SubcategorizationFrame* instance called *regularSVOAvere*. This latter instance describes the regular subject, verb, object structure with a verb using the auxiliary “avere”.

4.2.3 SyntacticArgument class

SyntacticArgument is a class representing an argument of a given *SubcategorizationFrame*. A *SyntacticArgument* can be linked recursively to a *SubcategorizationFrame* instance to describe complex

arguments. *SyntacticArgument* allows the connection with a *SemanticArgument* by means of a *SynSemArgMap* instance.

EXAMPLE For instance, the aforementioned *regularSVOAvere SubcategorizationFrame* contains two syntactic arguments, one for the subject and one for the object. Syntactic arguments within a *SubcategorizationFrame* are ordered.

4.2.4 *SubcategorizationFrameSet* class

SubcategorizationFrameSet is a class representing a set of syntactic constructions and possibly the relationship between these constructions. A *SubcategorizationFrameSet* groups various syntactic constructions that appear frequently for certain sets of lexemes. The objective is to factorize syntactic descriptions and to maintain a minimum of *SyntacticBehaviour* instances in the lexicon.

EXAMPLE For instance, the English verb “boil” presents an anti-causative syntactic alternation: “John boils the water”, “The water boils”. A possible modelling of this with this *SubcategorizationFrameSet* is achieved as follows: a single *SyntacticBehaviour* for the *LexicalEntry* of *boil* is instantiated; this *SyntacticBehaviour* links to a *SubcategorizationFrameSet* called *anticausativeVerbType1*, and this in turn contains two instances of *SubcategorizationFrame*, *regularSVO* and *regularSV*. See [Clause A.3](#) for a further example.

4.2.5 *SynArgMap* class

The *SynArgMap* is a class representing the relationship that maps various *SyntacticArgument* instances of the same *SubcategorizationFrameSet* instance.

EXAMPLE In the previous example, *SynArgMap* can be used to map the object of the first *SubcategorizationFrame* to the subject of the second one.

4.2.6 *SenseRelation* class

SenseRelation is a class representing the oriented relationship between *Sense* instances.

EXAMPLE In the PAROLE-SIMPLE-CLIPS lexicon^[6], the two senses of the word “violetta” (“violet”, in English) – USem65914violetta and USemD63723violetta referring respectively to the flower and the whole plant – are linked by the *PolysemyPlant-Flower* relation.

4.2.7 *SemanticPredicate* class

SemanticPredicate is a class representing an abstract meaning together with its association with the *SemanticArgument* class. A *SemanticPredicate* instance may be used to represent the common meaning between different senses that are not necessarily fully synonymous. These senses may be linked to *LexicalEntry* instances whose parts of speech are different. A *SemanticPredicate* instance pertains to a given *Lexicon* instance.

EXAMPLE In a *LexicalEntry* instance for “buy” in the sense of “get something by paying money for it”, a *SemanticPredicate* instance can be defined with two semantic arguments: one for the person who buys and one for what is bought. Another *LexicalEntry* instance can be recorded for “buyer” and linked to the same predicate.

4.2.8 *PredicativeRepresentation* class

PredicativeRepresentation is a class representing the link between the *Sense* and the *SemanticPredicate* classes.

EXAMPLE In the example given in 4.2.7, the link between the sense of the verb (i.e. “buy”) and the predicate can be marked as *Verb*. The link between the sense of the noun (i.e. “buyer”) and the predicate can be marked for instance as *VerbNominalization*. [Clause A.2](#) provides an illustration of the way in which a verb and its nominalization can link to the same *SemanticPredicate*.

4.2.9 *SemanticArgument* class

SemanticArgument is a class representing an argument of a given *SemanticPredicate*. A *SemanticArgument* instance may be described in terms of data categories such as “agent”, “patient”, and typed by means of a relation to one or many external resources (such as an ontology). Semantic arguments are ordered.

EXAMPLE In the example given in [4.2.7](#), the predicate can have two *SemanticArgument* instances: one for the person who buys and one for what is bought. The first one can be described as “agent”.

4.2.10 *ArgumentRelation* class

ArgumentRelation is a class representing an oriented relationship between *SemanticArgument* instances of the same *Predicate* instance. This is typically used to express a semantic relation which must necessarily exist between two arguments in a given predicate.

EXAMPLE In the French “je me lave les mains” (reflexive construct), the underlying semantic predicate contains two arguments: Argument 1 has semantic type “human” and Argument 2 has semantic type “body part”, with the requirement that it is a body part of Argument 1.

4.2.11 *SynSemArgMap* class

SynSemArgMap is a class representing the links between a semantic argument and a syntactic argument.

EXAMPLE See *SynSemCorrespondence* class example in [4.2.12](#).

4.2.12 *SynSemCorrespondence* class

SynSemCorrespondence is a class representing a set of *SynSemArgMap* instances for a given *SubcategorizationFrame* instance.

EXAMPLE A lexical resource can contain the “giving” semantic predicate with three arguments: a Donor, a Recipient and a Theme. The same resource can contain a subcategorization frame for the double accusative construction (Subject, Object1, Object2), such as it is used in the English verb “give”. A *SynSemCorrespondence* would then contain three *SynSemArgMap* instances: one to map the Donor with the Subject, one to map the Recipient with Object1, and one to map Theme with Object2.

4.2.13 *PredicateRelation* class

PredicateRelation is a class representing the oriented relationship between instances of *SemanticPredicate*.

EXAMPLE When modelling resources such as FrameNet^[7], it is possible to indicate for instance that the Activity_“finish” predicate is a subframe of Activity.

Annex A
(informative)

Syntax-semantics examples

A.1 Illustration of the full model

[Figure A.1](#) illustrates the full model, using the French verb “aider” (“help”) with its syntactic and semantic properties.

STANDARDSISO.COM : Click to view the full PDF of ISO 24613-6:2024

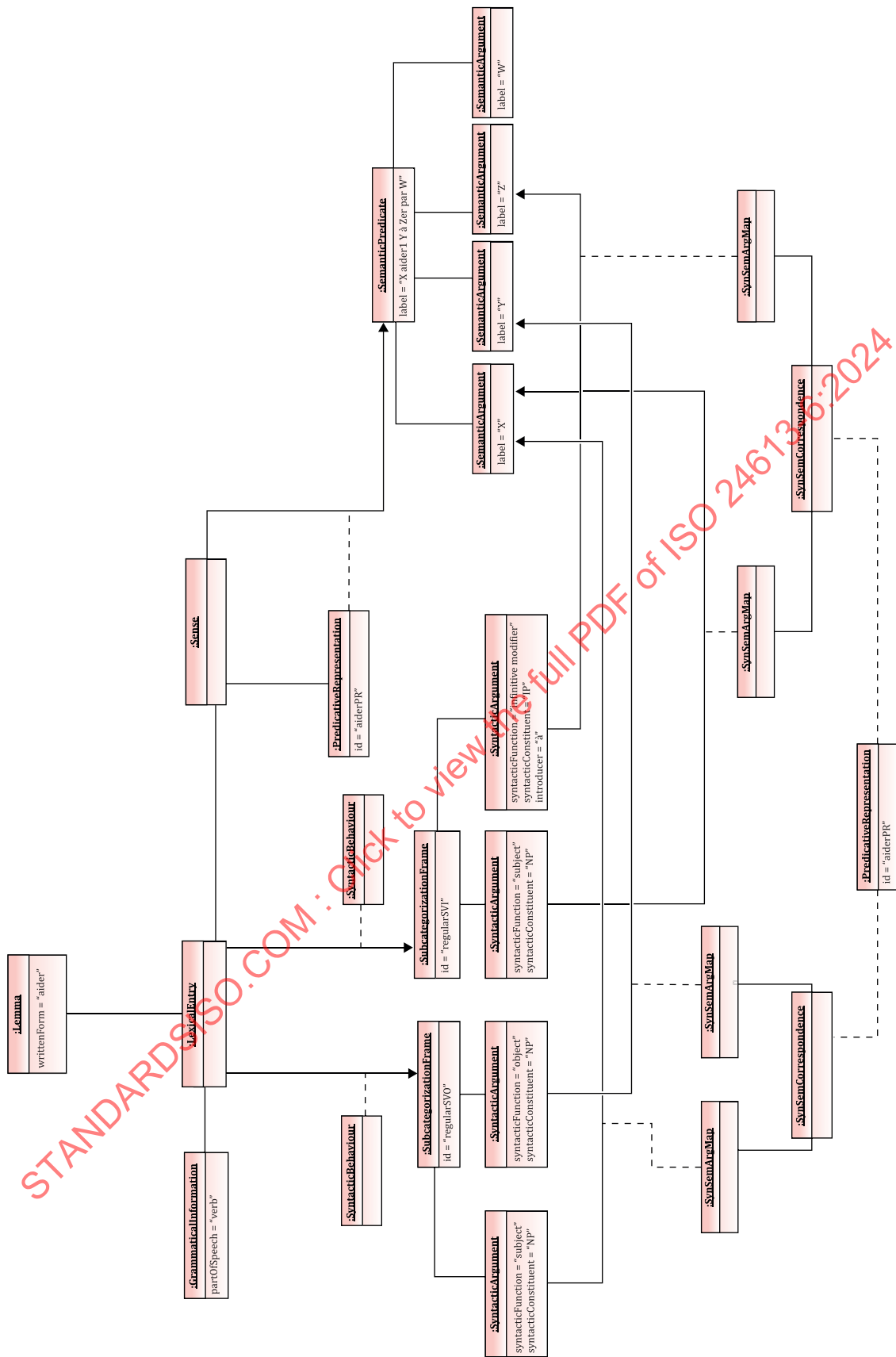


Figure A.1 — Example: “aider”