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

iTeh Standards
(https://standards.iteh.ai)
Document Preview

PROOF

iTeh Standards
(https://standards.iteh.ai)
Document Preview

ISO/PRF 24613-6
https://standards.iteh.ai/catalog/standards/iso/5c264942-cede-4a74-9a35-41ea81a144a3/iso-prf-24613-6

Contents

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 describesspecifies the syntax and semantics (SynSem) module of the Lexicallexical 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*valency* (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).[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

[SOURCE: ISO 24613:2008, 3.41]

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).

[SOURCE: ISO 24613:2008, 3.43]

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