Introduction

This article describes the structure of the EngValLex data format.

EngValLex is an english valency lexicon based on the formats and standards of the Prague Dependency Treebank 2.0. It was automatically created from annotations of PropBank lexicon, and consequently manually corrected.

The format was developed following the Czech PDT-ValLex. Support for references to the PropBank was added.

Elements of the EngValEval XML

valency_lexicon

Root element of the lexicon.

Contains head and body.

head

Defines meta information about the lexicon.
Contains global_history and list_of_users.

**global_history**

Currently not in use.

**list_of_users**

List of users - currently not in use.

**user**

User - currently not in use.

**body**

The body of the lexicon composed of word entries.

**word**

Entry of the lexicon defining valency for a single word (lemma).

Parameters:
- **lemma**
  
  Lemma of the lexical entry
- **id**
  
  Automatically generated identifier
- **POS**
  
  Part of speech of the word - for the time being only "V" (verb)

Contains a valency_frames tag and an examples tag.

**valency_frames**

Section defining valency frames for a single word.

Contains zero or more frames.

**frame**

Definition of one valency frame.

Contains an external link, a frame_elements section and an examples sections.

Parameters:
- **status**
Status of the node.

• roleset_name

   Name of the PropBank role set which this frame comes from.

• id

   Automatically generated identifier.

**externallink**

Defines a reference to a external data source. For the time being, it is used only for references to PropBank lexicon.

Parameters:

• type

   Type of the external link: PB roleset for a reference to PropBank role set, PB role for a reference to PropBank role, PB example for a reference to PropBank example.

• fileref

   File name of the refereed item.

• rolesetref

   Identifier of the PropBank role set.

• n

   Identifier of the PropBank argument name or number (for PB role references only).

• f

   Identifier of the PropBank function name (for PB role references only).

• name

   Identifier of the PropBank example name (for PB example references only).

**frame_elements**

Introduces a list of elements of the frame.

Contains zero or more elements.

Parameters:

• status

   Status of the node for annotator's use.

• id
Currently not in use.

status id

element

Defines one element of a frame. It corresponds to a functor of the frame.

Parameters:

• type
  Type of the element - eg. oblig.

• status
  Status of the node (active).

• descr
  Description of the element originally taken from PropBank description of the role.

• n
  The original argument name in the ProbBank.

• f
  The original function name in the ProbBank.

• id
  Automatically generated identifier.

elements

Introduces the section of examples of the frame.

Contains zero or more examples.

element

Defines one example of a frame. It correspond to functor of the frame.

Contains an externallink, a text element, and zero or more arg elements.

text

Defines the text of the example.

arg

Defines an argument used in the example.
Parameters:

• n
  The original argument name in the ProbBank.

• f
  The original Identifier function name in the ProbBank.

Currently not in use (empty content).