

Hydra

User Manual

Table of Contents

1. Introduction.....	2
1.1. Overview.....	2
1.2. Wordnet.....	3
1.3. Wordnet representation in Hydra.....	3
2. Getting started.....	5
2.1. Starting Hydra.....	5
2.2. The Search window.....	6
2.2.1. Search.....	6
2.2.2. Dictionary management.....	7
2.3. The Dictionaries.....	8
2.4. Synchronisation.....	9
2.4.1. Synchronisation between the wordnets and the Search window.....	10
2.4.2. Synchronisation between the wordnets.....	11
2.5. Working with the Wordnet data.....	13
2.5.1. Dictionary views.....	13
2.5.2. Editing.....	17
2.5.2.1. Editing an existing object.....	17
2.5.2.1.1. General description.....	17
2.5.2.1.2. Operations in the Edit mode.....	19
2.5.2.2. Creating a new synset.....	31
2.5.2.3. Cloning a synset.....	34
2.5.2.3.....	37
2.5.3. Searching in Hydra.....	37
2.5.3.1. Simple search.....	37
2.5.3.2. Regular expression search.....	38
2.5.3.3. Formula search.....	38
2.5.3.4. Formula query tips.....	41
References.....	43

1. Introduction

1.1. Overview

Hydra is an OS-independent system designed for wordnet development, validation and exploration. It represents Wordnet as a relational structure and embeds a modal language for searching in the wordnet data.

The wordnet data are represented as a relational database. Information retrieval and management is handled by means of a relational database management system and SQL.

The system enables users to edit and browse any number of monolingual wordnets at a time. It provides a user-friendly GUI with different options for data display. The individual wordnets are synchronised, so that equivalent synsets in the different wordnets may be viewed and explored in parallel.

An important feature of the system is the multiple-user concurrent access.

The changes made to the database are updated immediately, so that all the users are able to access the updated data at once.

The system performs automatic data consistency and completeness verifications. The completeness checks performed for the obligatory elements of a synset are described in the relevant sections below. User specified validation queries are also enabled. Examples of such queries are given in section 2.5.3.4., subsection **Validation queries**.

Hydra is coupled with the corpus annotation tool Chooser and has been successfully employed in the annotation of the Bulgarian Sense-Annotated Corpus, as well as extensively used in the development of the Bulgarian wordnet.

This manual provides a description of the user interface, the different types of operations maintained by the system with relevant instructions, as well as a brief description of the search language and useful query examples.

1.2. Wordnet

A wordnet is a lexical-semantic database modeled after the Princeton WordNet (Fellbaum 1998, Miller et al. 1993). It represents the words in a language as groups of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations (<http://wordnet.princeton.edu/>).

Wordnets have a synset-centric organisation. A **synset** is defined as “a set of words that are interchangeable in some context without changing the truth value of the preposition in which they are embedded” (<http://wordnet.princeton.edu/>). The simple words and multiword expressions that represent synonyms in a synset are called **literals**. The meaning of the synset is represented by an explanatory **definition**.

The synsets may also contain:

(i) **usage examples** – sentences or phrases illustrating the use of the synset members;

(ii) **synset notes (snotes)** – grammar, pragmatic, or technical notes pertaining to a synset, for instance its register (colloquial, formal, etc.);

(iii) **literal notes (lnotes)** – grammar, pragmatic, or technical notes pertaining to a literal in a synset, for example – the aspect of a verb, etc.

The database that represents the linguistic data in the individual wordnets will be referred to as Wordnet.

1.3. Wordnet representation in Hydra

The Wordnet is represented as a relational structure (binary-tuple of a set of objects and a set of binary relations). There are three sorts of objects in the database:

- (i) objects of type **Synset** – represent the synonym sets in a Wordnet structure;
- (ii) objects of type **Literal** – represent the members of a synonym set;
- (iii) objects of type **Note** – represent text data in a Wordnet structure such as usage examples and explanatory notes;

These objects are referred to as linguistic units (LUs). The objects in the Wordnet structure are related with one another by means of a number of binary relations:

(i) linguistic relations - the conceptual-semantic and lexical relations defined in the Princeton WordNet, as well as all other types of relations between words and concepts that might be defined in a wordnet. A list of the linguistic relations in the Wordnet database is given in the relevant section;

(ii) structure-organising relations – the relations between the sorts of objects:

(a) relations of type **literal** – connect Literals with the Synsets to which they pertain;

(b) relations of type **lnote** – connect Note objects with the Literals to which they pertain;

(c) relations of type **snote** – connect Note objects with the relevant objects of type Synset;

(d) relations of type **usage** – connect Note objects representing usage examples with the relevant Synsets;

(e) relations of type **ili** – connect the synsets in the different wordnets that denote equivalent senses.

Every LU is associated with a single synset. A Synset is associated with itself.

2. Getting started

Hydra is available from <http://dcl.bas.bg/Tools/Hydra/hydra.zip>. For the initial setup of Hydra and the MySQL database consult the installation manual:

<http://dcl.bas.bg/Tools/Hydra/Hydra-InstallationManual.pdf>

2.1. Starting Hydra

To launch Hydra using a command line, run the following command (provided you are in Hydra's directory)

```
python hydra.py
```

The following examples show how to run Hydra in a Linux environment, assuming that Hydra is located in /home/user/hydra on machine 'machine':

(1) from the local directory where the Hydra executable file **hydra.py** is stored:

```
user@machine:~/hydra$ python hydra.py
```

(2) using the full path to the executable file:

```
user@machine:~$ python /home/user/hydra/hydra.py
```

(3) using a relative path:

```
user@machine:~$ python hydra/hydra.py
```

Windows users can launch Hydra by double-clicking on the executable hydra.py icon.

If Hydra starts properly, connection to the database is established and the **Search window** of the application appears on the screen (Fig. 1).

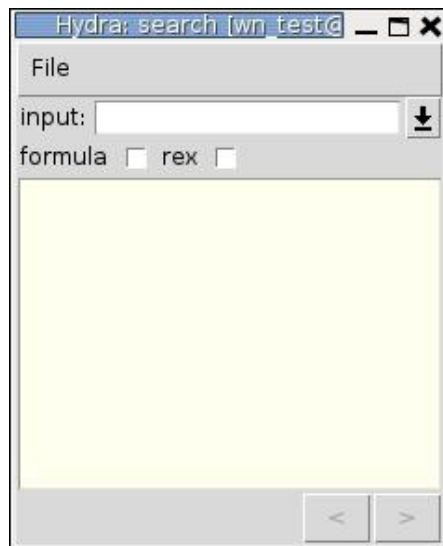


Fig. 1. The Search window on launching Hydra

In case of database connection failure, the system goes into interactive mode and asks the user to provide the host and database name or username and password. Default values (taken from the configuration) are suggested. The user can confirm the default values by hitting **Enter**.

2.2. The Search window

The main **Search window** serves two tasks – browsing the database and dictionary management.

2.2.1. Search

The **Search window** provides the main search tool of the system – the entry point to the Wordnet data. For instructions on how to submit a query, and the types of queries supported by the system, confer the General description of the **Editing** section. A detailed

description of the query language is given in the section **Searching in Hydra (2.5.3.)**.

2.2.2. Dictionary management.

In order to access the wordnet data available for a given language, you need to open a dictionary for this particular language. The dictionary is a collection of synset view controls and is associated to a single language, so it visualizes only synsets in that language.

(1) To do that, click on the File menu of the Search window. A list of the wordnets available in the database will be displayed (Fig. 2).



Fig. 2. The Search window with the list of the wordnet languages available in the database

(2) Select a language name from the menu by clicking on it.

(3) A window (containing the dictionaries) appears on the screen. The name of the wordnet to be visualised in it is displayed in the left upper area (Fig. 3, circled in red).

(4) Select as many dictionaries for the wordnet languages in the database as you need in the same way, one at a time. Each of them is displayed in a separate pane of the Dictionaries window.



- You may open any number of dictionaries for a particular language.
- Always close the programme by closing the Search window – from the Quit option in the Search window menu or the standard close button.

2.3. The Dictionaries

The dictionary panes are arranged from left to right in the order of selection. The screenshot in Fig. 3 shows a Bulgarian and an English wordnet, where the former was loaded first.

The panes are separated by sliders that enable users to resize the panes' width. In case multiple dictionaries are opened, it is possible that not all of them are visible. To fix this, use the sliders to expand the panes. Move the mouse pointer anywhere over the divider (visible as a purple line in Fig. 3), so that an arrow pointer appears, and drag it to the left or to the right.

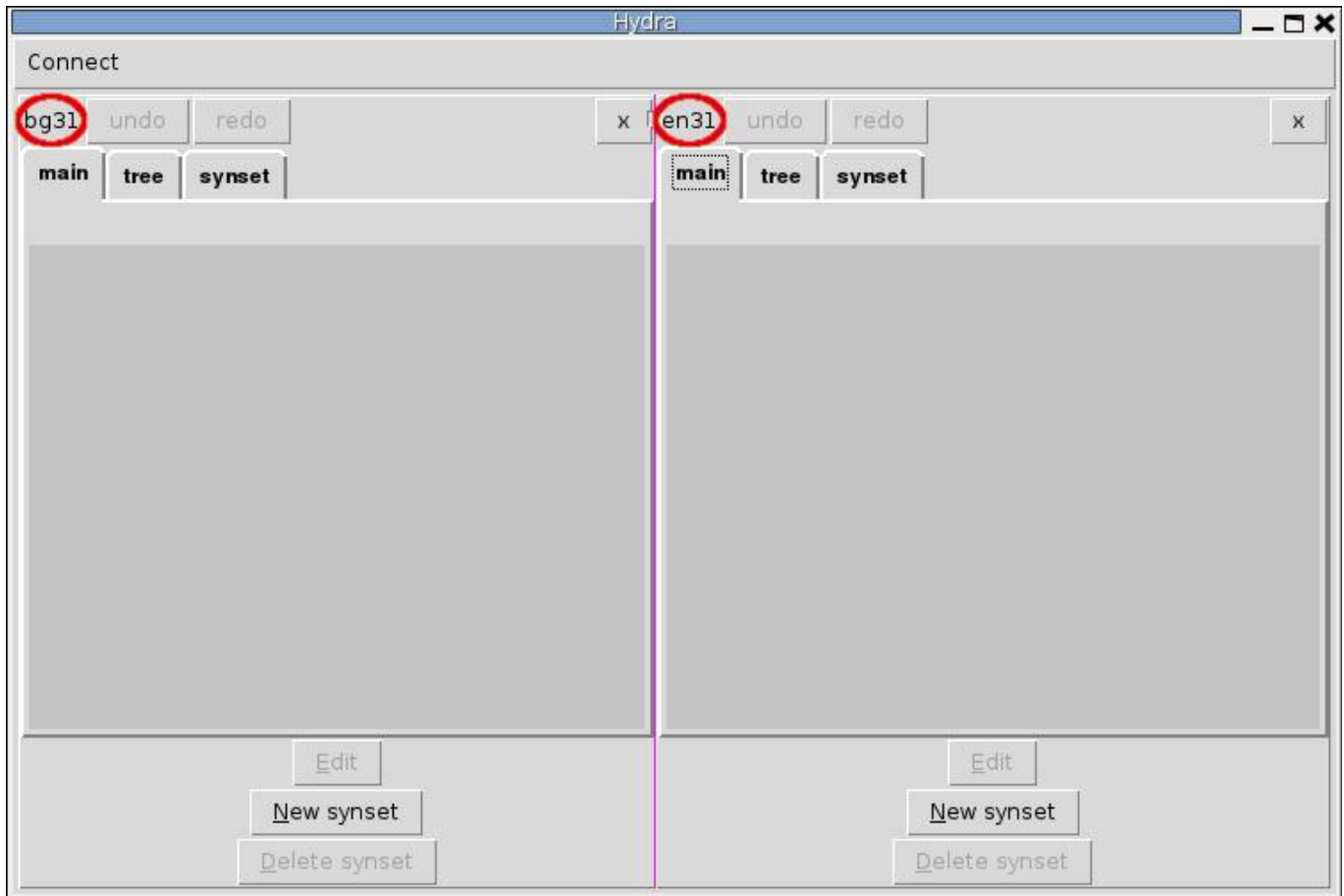


Fig. 3. The Dictionaries window in which a Bulgarian and an English wordnet (bg31 and en31, respectively) are opened in each pane

2.4. Synchronisation

This section deals with the synchronisation of the synsets displayed in the opened dictionaries. The system provides synchronisation between each pair of the opened wordnet dictionaries and between each of them and the search tool.

2.4.1. Synchronisation between the wordnets and the Search window

The synchronisation between the **Search window** and the wordnets is enabled by default, while the wordnet-to-wordnet synchronisation is explicitly specified by the user.

When a Wordnet object is invoked, a clone of the corresponding synset in the respective wordnet is created and displayed.

For instance, when a query is submitted in the **Search window** (Fig. 4), the synsets that match the query are displayed in the area below the **input** field. The results are paginated.

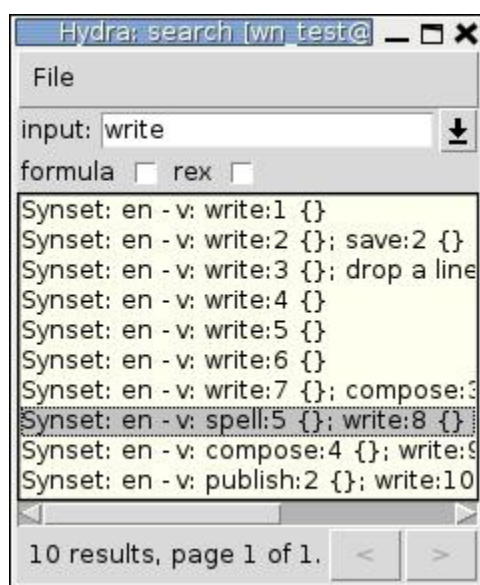


Fig. 4. List of the synsets containing the literal 'write'. The gray one is selected by the user and displayed in the Main view of the WN window

To display a synset from the list in the WN window (Fig. 5), select the synset by clicking on it. The synset is then highlighted (coloured gray in Fig. 4.)

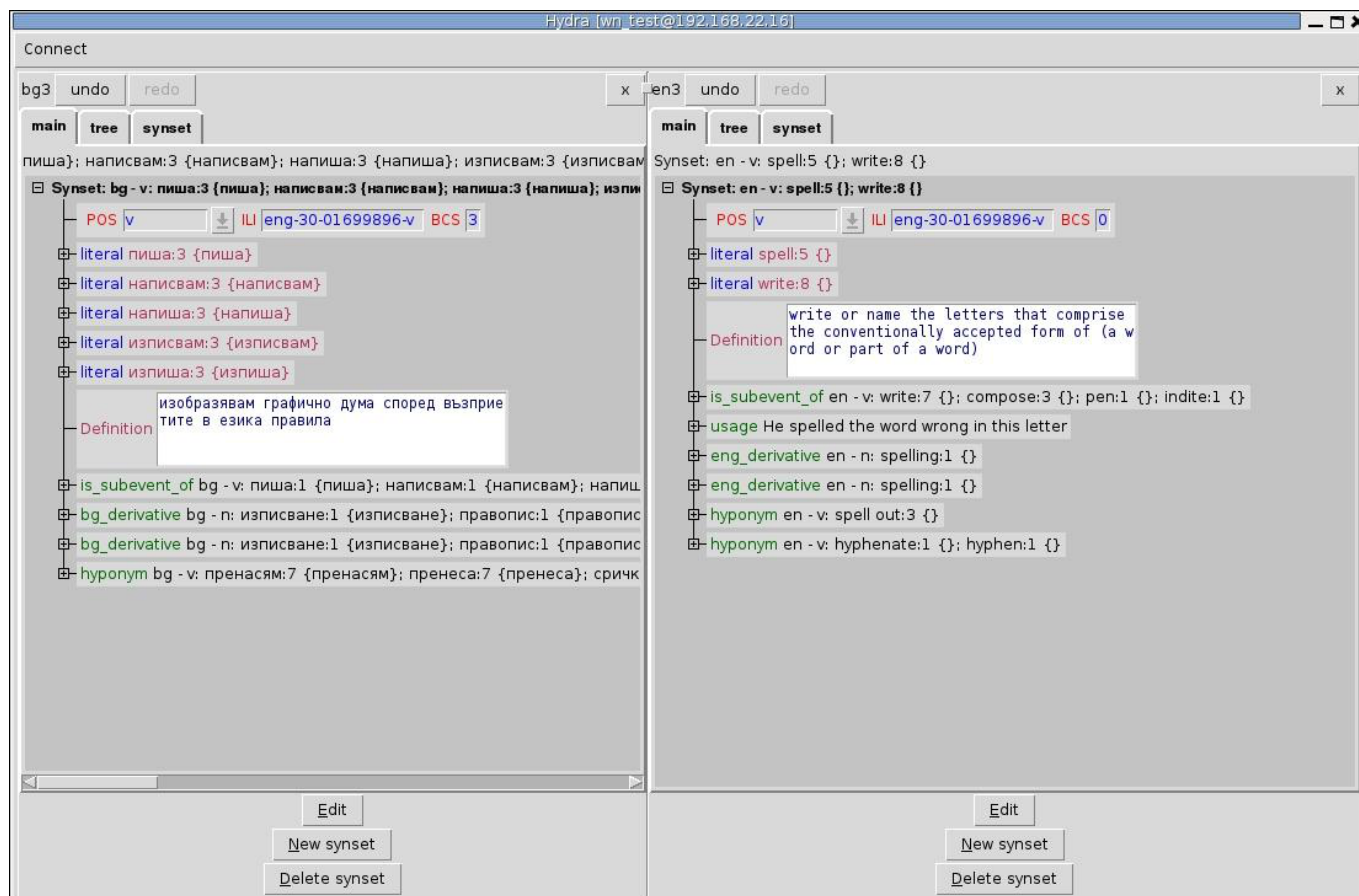


Fig. 5. The WN window with the synset {spell:5; write:8} and its equivalent in the Bulgarian wordnet.

2.4.2. Synchronisation between the wordnets

The equivalent synsets in the different wordnets are synchronised by means of unique synset identifiers. The equivalence is encoded in the symmetric ‘**ili**’ relation. It allows the users to view and browse the data in the different wordnets simultaneously.

To synchronise a pair of wordnets, click on the **Connect** menu of the WN window.

- (1) A list of the synchronisation options with check boxes will be displayed.

(2) Check the boxes for the pairs you want to synchronise, bg31 → en31 (Fig. 6).

In case only the box bg31 → en31 is checked, when the user browses the Bulgarian wordnet, the English wordnet will be synchronised with it, but not vice versa.

(3) To have a pair of wordnets symmetrically synchronised, check the boxes corresponding to both directions, e.g. bg31 → en31, en31 → bg31. (Fig. 6).

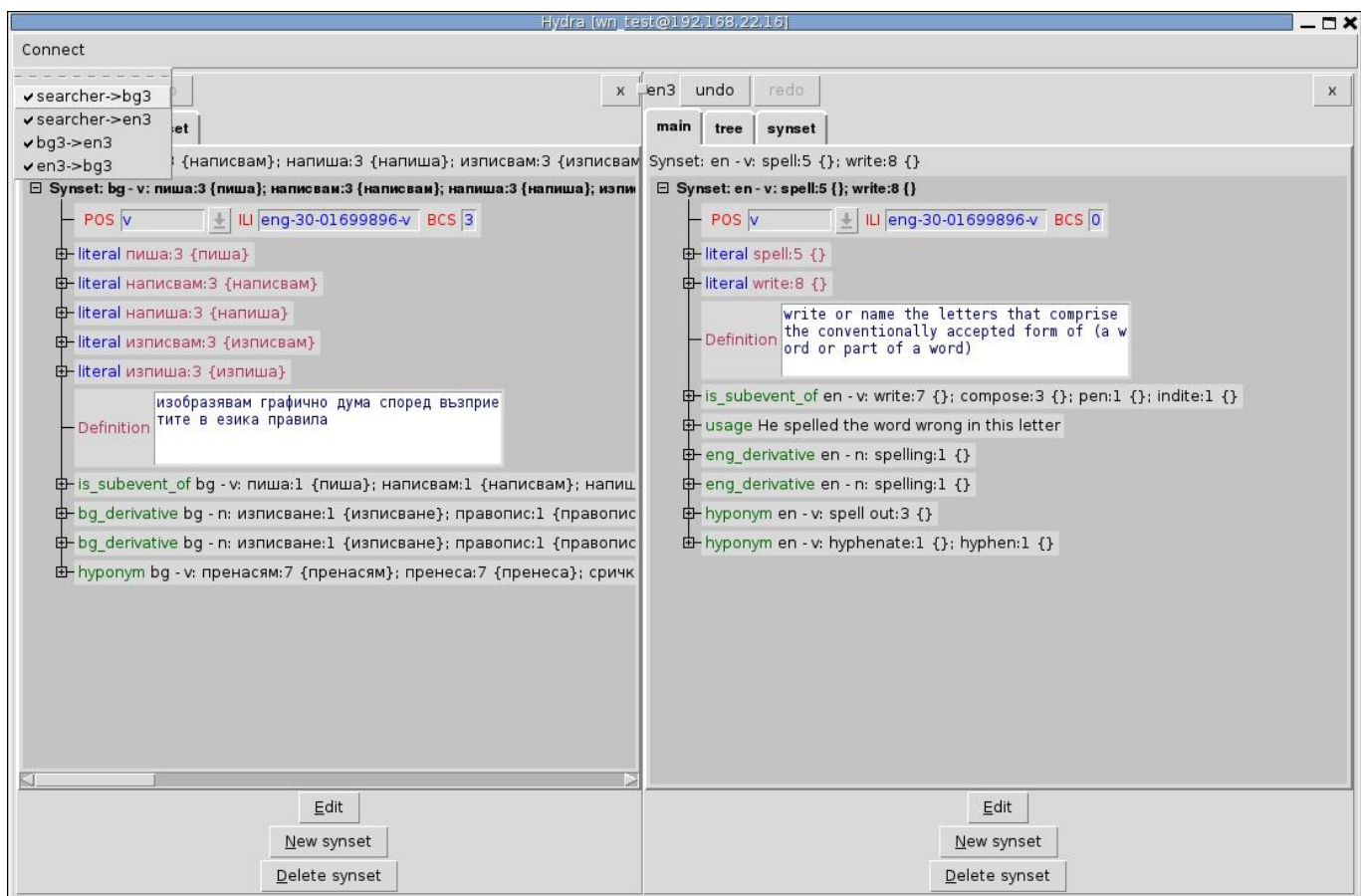


Fig. 6. Synchronisation of the English and the Bulgarian wordnets

(4) In case the default synchronisation between the search tool and the dictionary is

disabled, check the boxes searcher → name of wordnet, e.g. searcher → bg31, searcher → en31 (Fig. 6) to restore it. Otherwise the selected objects in the Search window will not be displayed in the WN window.

2.5. Working with the Wordnet data

There are three types of views for the display of LUs in any dictionary.

2.5.1. Dictionary views

A. The Main View

The **MainView** (Fig. 5, Fig. 6) provides a number of functions:

- (i) edit LUs;
- (ii) add and remove relations;
- (iii) create and delete LUs;
- (iv) clone synsets from other available wordnets.

An important feature of the **Main View** is the recursive representation of the Wordnet relational structure. It is visualised as a tree structure in which the wordnet objects are represented as expandable nodes. The data and relations associated with a node are displayed by clicking on the plus sign on its left. The edges represent the relations between LUs.

This view has configurable 'look and feel' through an XML configuration file (unit_view.xml, where data visualisation properties such as order, colour, size, control types (combobox or list view) - may be configured).

B. The Tree View

The **Tree View** displays the relations as tree structures. It visualises only acyclic relations. If **R** is such relation, a successor of a node **I** in the tree is each neighbour **LU** **x**, so that **IRx**. The **Tree view** pane is divided into two columns (Fig. 7). The tree on the right shows the position of a selected node in the graph structure of the particular relation and the path to the topmost synset starting from the first antecedent. The user may view the path to the bottom of the tree (the node's successors) by expanding the node.

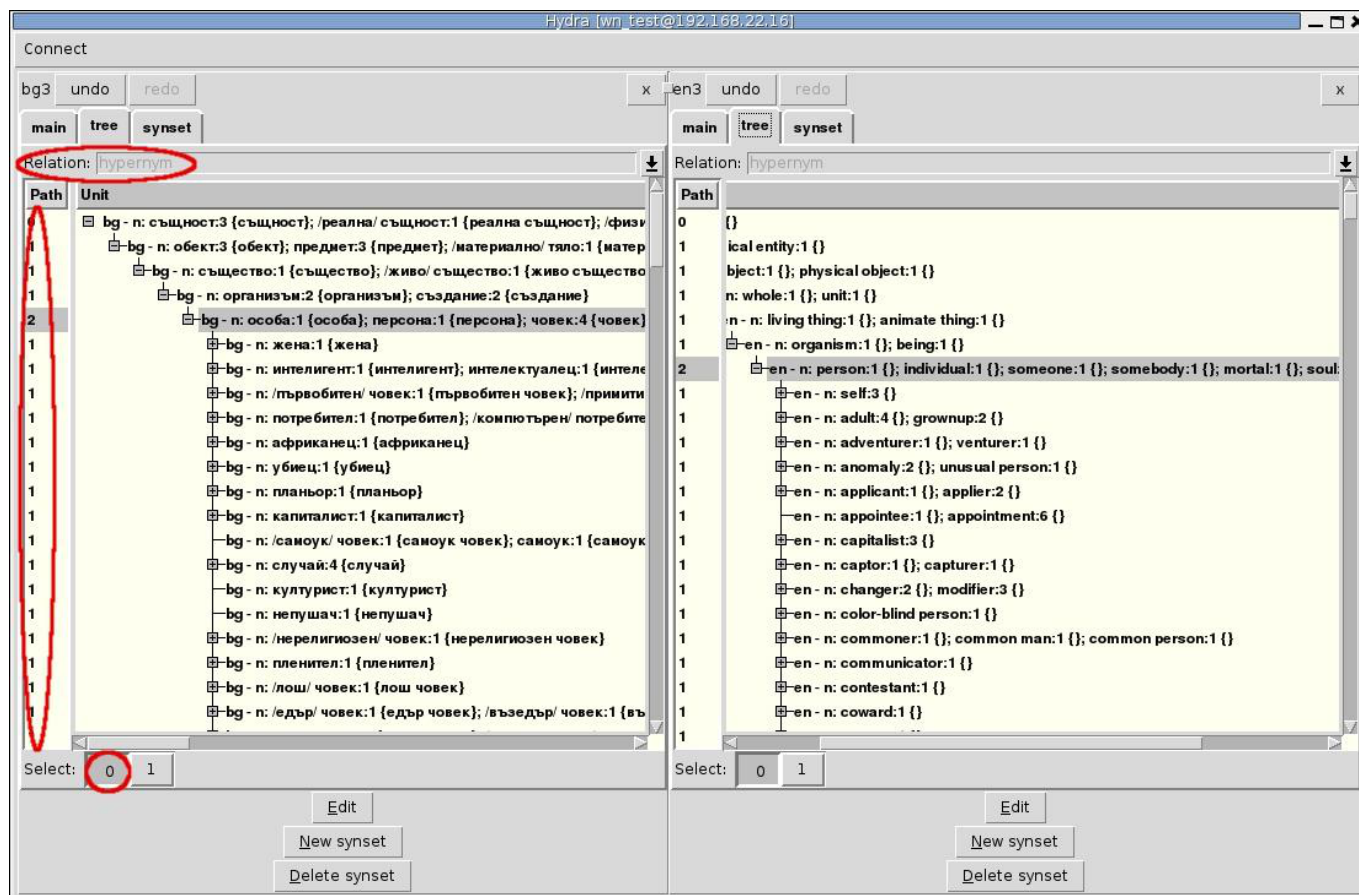


Fig. 7. Tree view with the hypernym tree for the selected node {person:1; individual:1; someone:1; somebody:1; mortal:1; soul:1} (highlighted in gray).

The left column displays the number of antecedents for the corresponding LUs, as well as for its antecedents and immediate successors. If the antecedents of a LU are more than one (the selected node in Fig. 7), the user may specify which antecedent's path to the root to view by pressing the respective **Select** button below the Tree pane. The topmost synsets have 0 antecedents.

Fig. 7 shows the central synset {person:1; individual:1; someone:1; somebody:1; mortal:1; soul:1} with the path to the topmost node following the first hypernym {organism:1; being:1} (the button circled in red). The list of synsets below the central node represents its immediate successors (left column, circled in red).

The default relation is hypernymy. To choose another relation, use the combobox (Fig. 8).

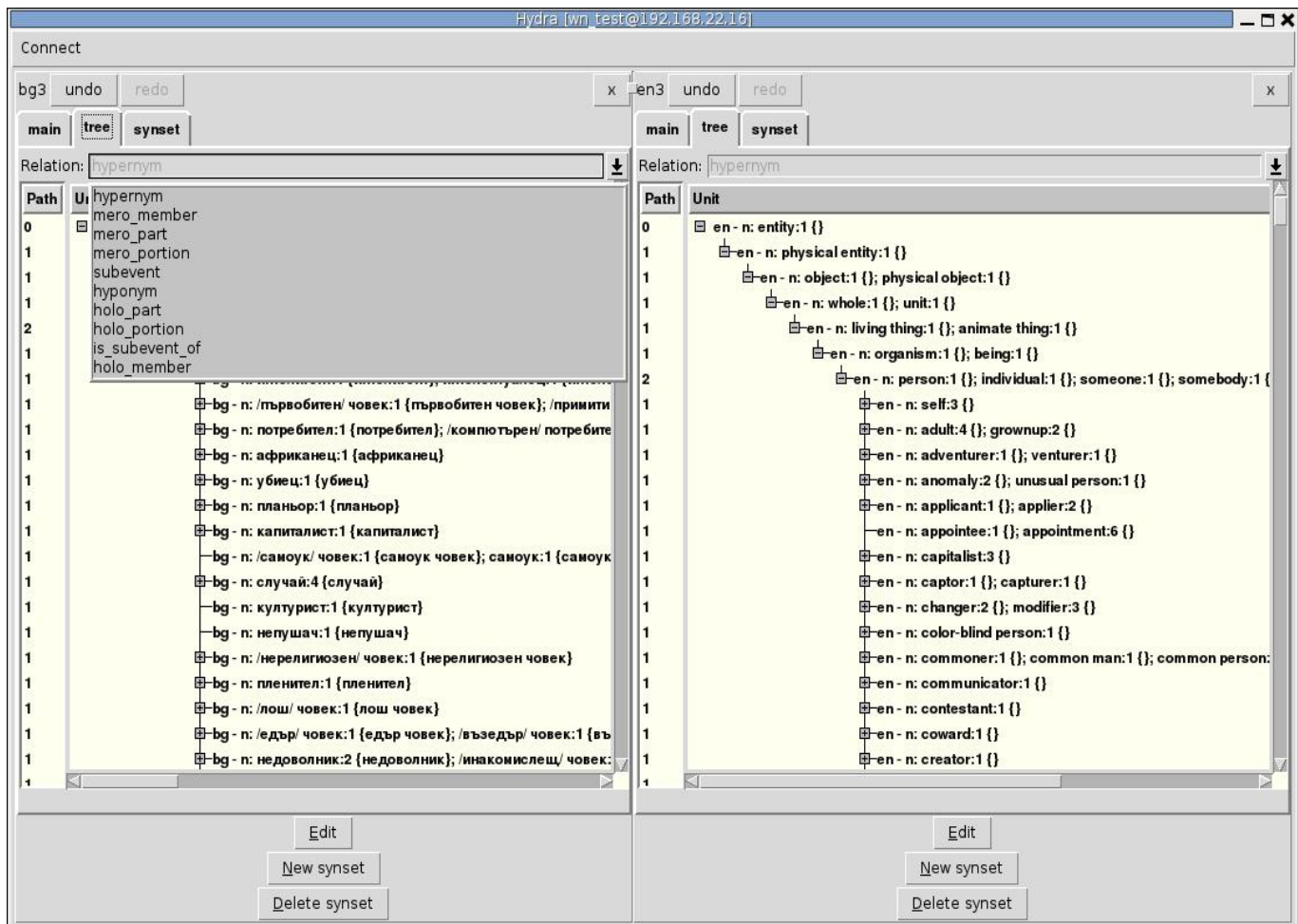


Fig. 8. Tree view with the relations combobox.

C. The Synset View

The **Synset View** displays the characteristic attributes of a synset such as **Pos** (part of speech), **ID** (unique interlingual identifier), literals with their attributes – **word** and **lemma**. The immediate neighbour nodes of the synset are shown, as well.

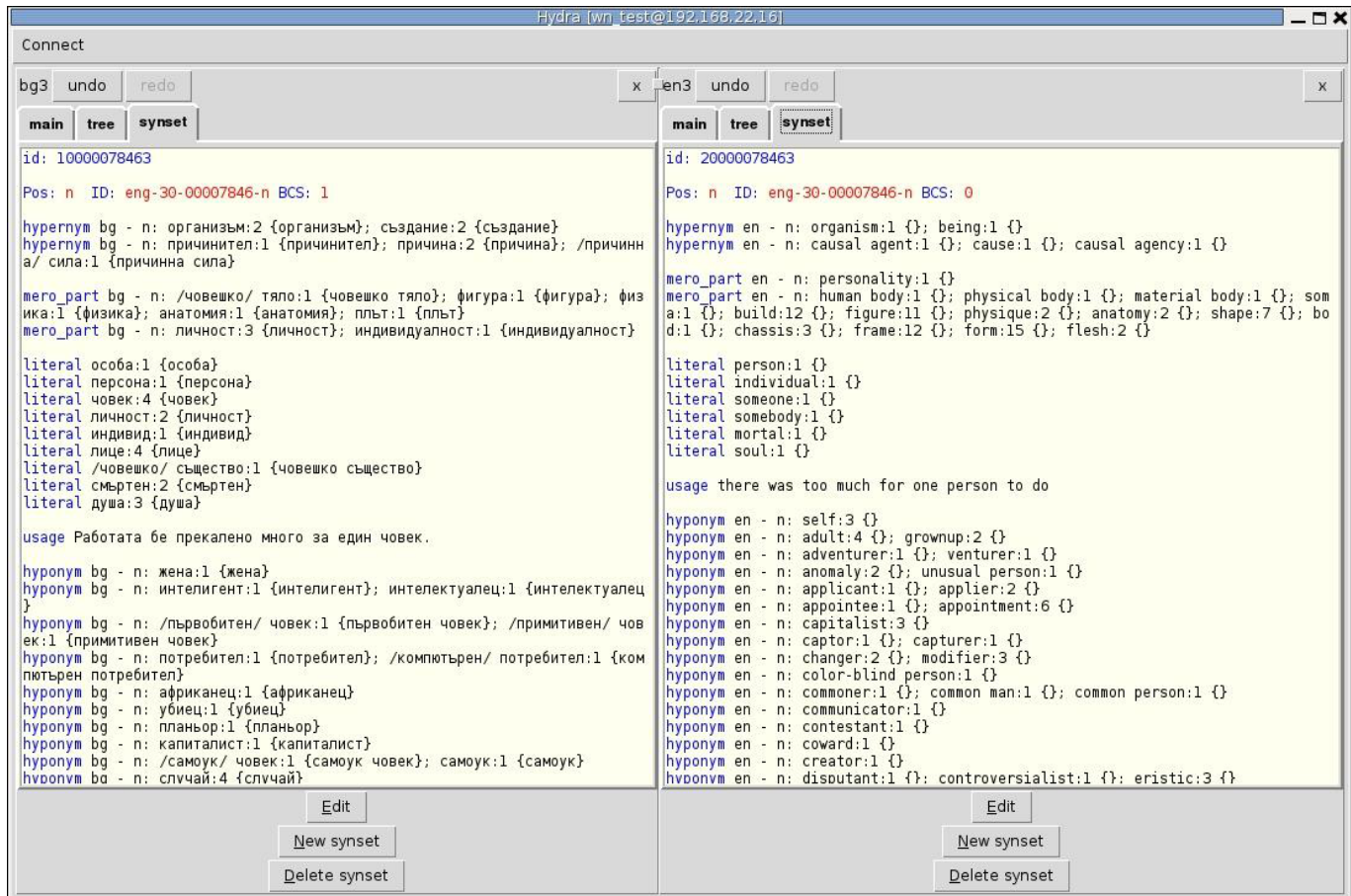


Fig. 9. Synset view for the synset {person:1; individual:1; someone:1; somebody:1; mortal:1; soul:1}

2.5.2. Editing

2.5.2.1. Editing an existing object

2.5.2.1.1. General description

In order to make an object editable, you need to activate it. There are two ways to do that:

- (1) Select an object from the **Search window**:

1.1 Type a query in the input field of the Search window and hit Enter. This is the usual way of looking up synsets that contain a particular literal. A list of the objects that match

the query is displayed in the area below the input field.

1. 2. Select the relevant synset from the list by clicking on it. Editing is performed in the **Main View**.

1.3. If either of the other views is active, switch to **Main View**.



- The search query may consist of word(s), regular expressions or formulae (see the section **Searching in Hydra**). The default search option is a word or a combination of words (simple search). To submit a regular expression query (the same as the simple search but using regular expressions) or a formula, check the rex or formula box.
- The **Search window** must be synchronised with the respective wordnet's pane (Synchronisation with the search tool is the default option, but you might need to enable it from the **Connect** menu, see the section on **Synchronisation** above).

(2) Invoke objects from the **Main View**, **Tree View** or **Synset View**

1.1. If the **Main View** is active

Expand or collapse the object by pressing the plus or minus sign on its left.

In case another object is active and the changes made to it are not saved, you need to save the data first by pressing **Save** in order to be able to activate another object.

1.2. If you are in the **Tree View**

Select the object in the tree. The active object is highlighted.

1.3. If you are in the **Synset View**

The Synset view shows the synset associated with the current (selected) object. You can

use the view to make sure which the current synset is. To select a different object you must switch to **Main View** or **Tree View**.

If either the **Tree View** or the **Synset View** is active, switch to **Main View** to start editing.

To enable the edit mode, click on the Edit button in the bottom area of the **Main View**.

2.5.2.1.2. Operations in the Edit mode

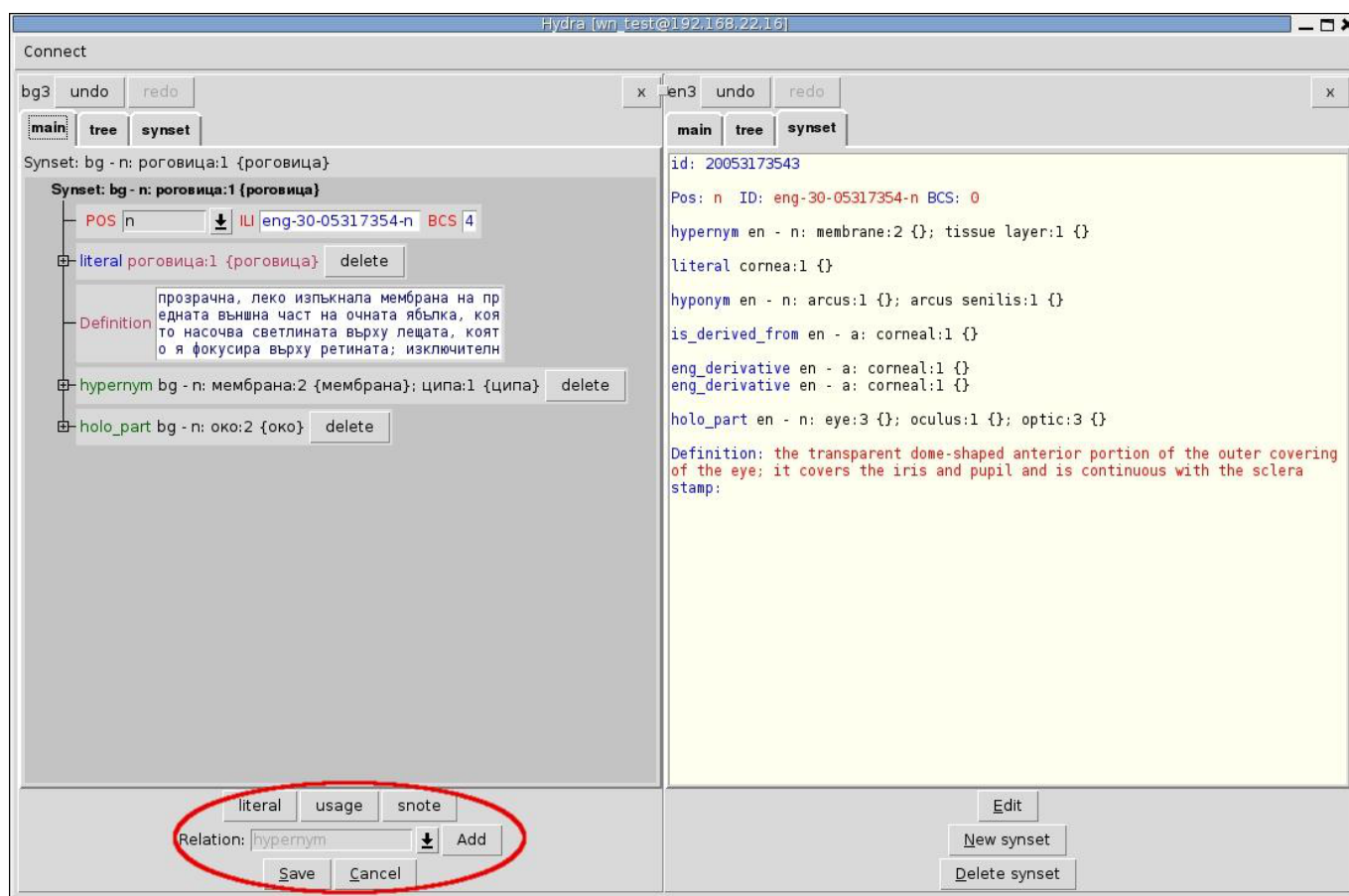


Fig. 10. The edit mode of the Main view

The **Edit** mode is where the actual creation and correction of the Wordnet data take place.

The operations are enabled by clicking on the respective button in the bottom area of the **Main view** (Fig. 10).

A. Adding objects (literal – Literal; usage, snote, lnote – Note)

1. Adding literals

In Hydra's approach, a literal is a word that stands in the **Literal** relation with the corresponding synonym set.

To add a new literal to an existing synset:

- (1) Make sure the relevant synset is active.
- (2) Press the **Edit** button in the bottom area of the **Main View**.
- (3) Click on the literal button to add a new literal. A pair of empty fields named **Word** and **Lemma** is created. Simultaneously three buttons - **Lnote**, **Save** and **Cancel** appear in the bottom area of the pane (Fig. 11)
- (4) Type the simple word or MWE you wish to add in the **Word** field.
- (5) Type the lemma in the **Lemma** field.

This field is optional and is only required if the particular wordnet has adopted manual encoding or validation of lemmas.

- (6) Press the **Save** button at the bottom to save the literal.

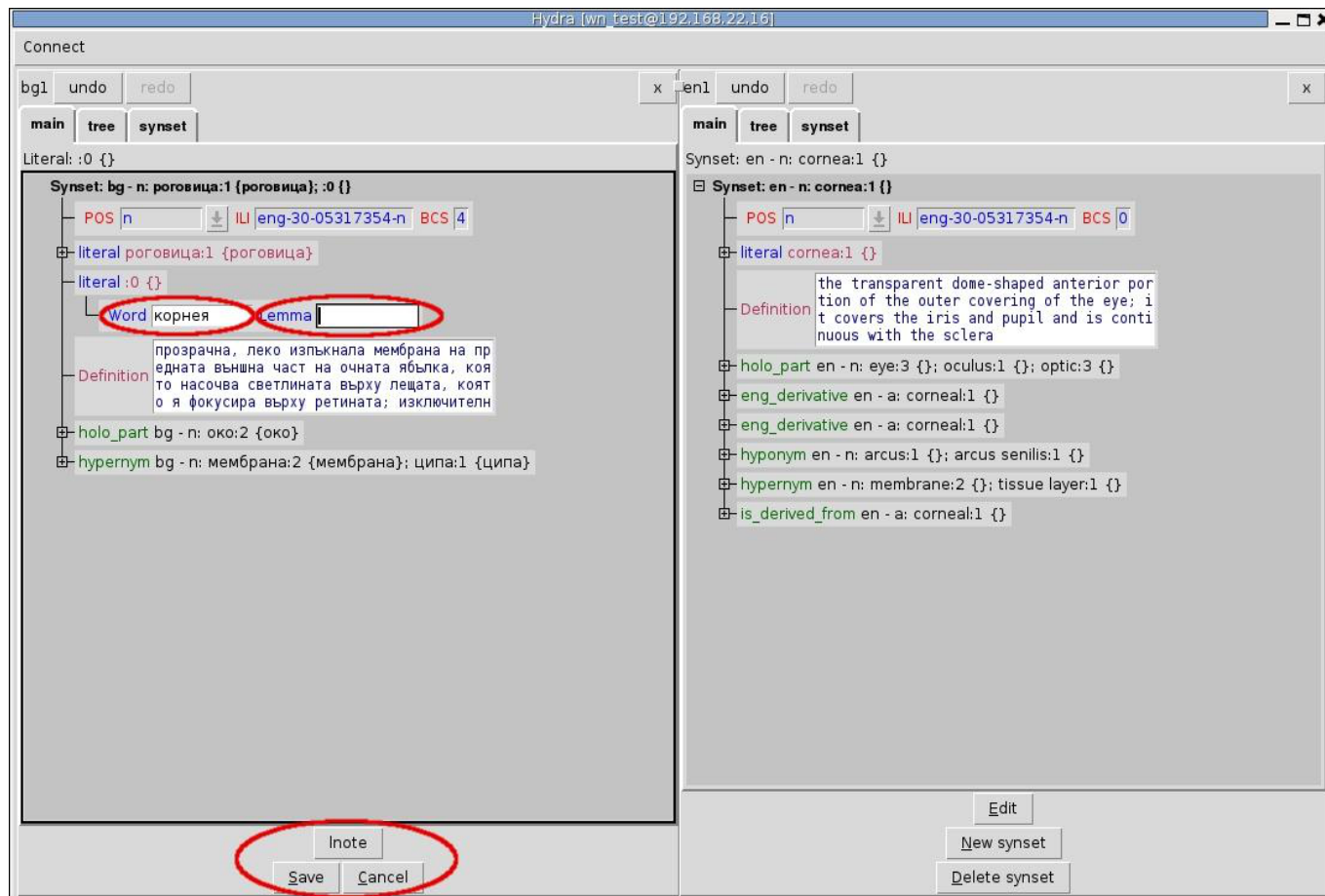


Fig. 11. Creation of a literal with the relevant fields and buttons highlighted in red.

Hydra incorporates a number of completeness checks. In case you try to save an empty literal a warning dialog pops up (Fig. 12). To discard it, press **ok**. You will not be able to proceed to create or edit other objects unless you fill in the **Word** field of the literal, or cancel it entirely.

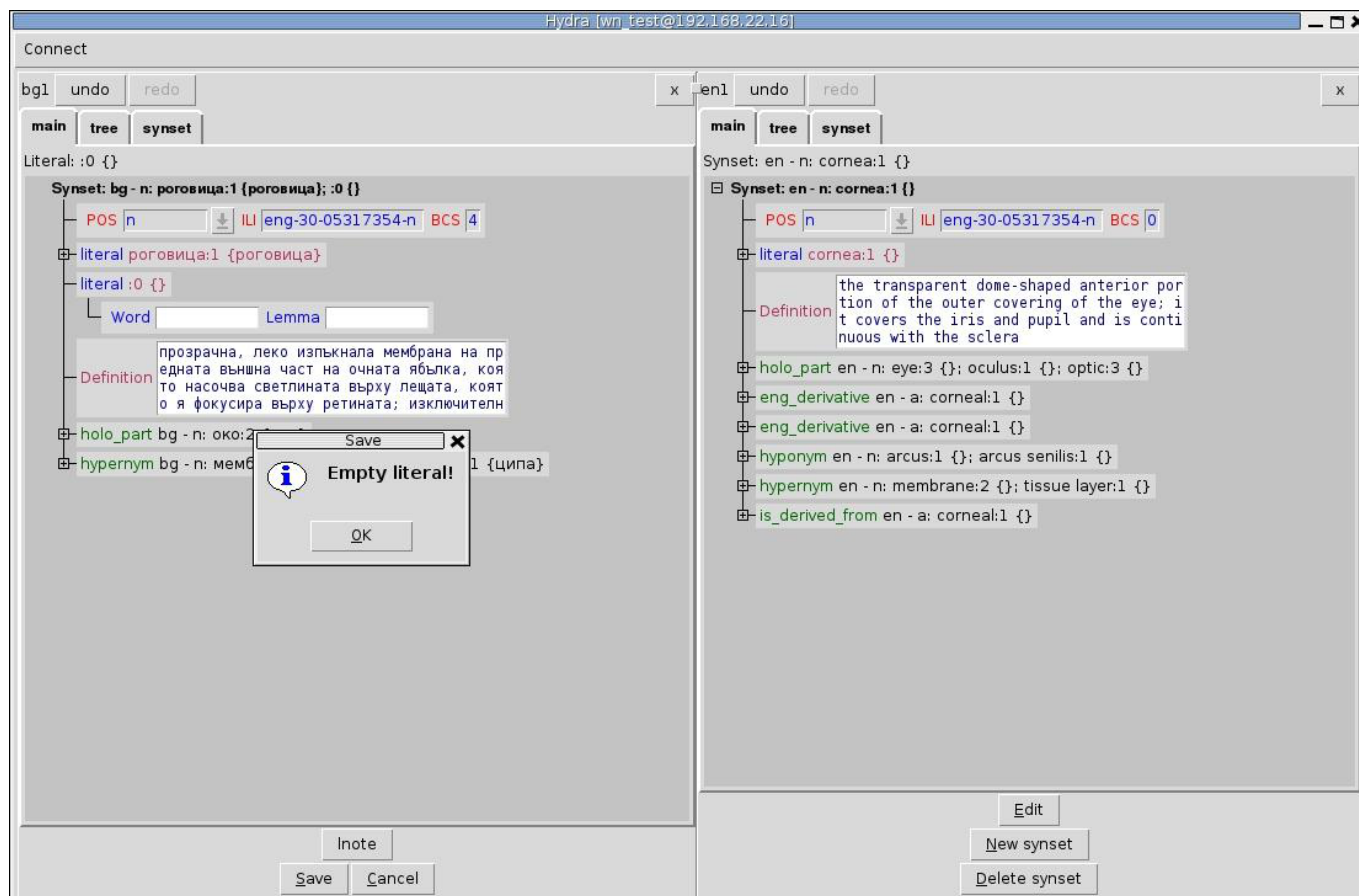


Fig. 12. The warning message notifying that you are trying to save an empty literal.

2. Adding usage examples, snotes and lnotes

Usage examples, snotes and lnotes are created in a similar way as literals.

To create usage examples and snotes:

- (1) Make sure the relevant synset is active.
- (2) Press the **Edit** button.
- (3) Open an empty **Usage** or **Snote** field by clicking on the respective button at the bottom of the pane.

(4) Type or copy and paste a usage example/note in the field (Fig. 13), then press **Save**.

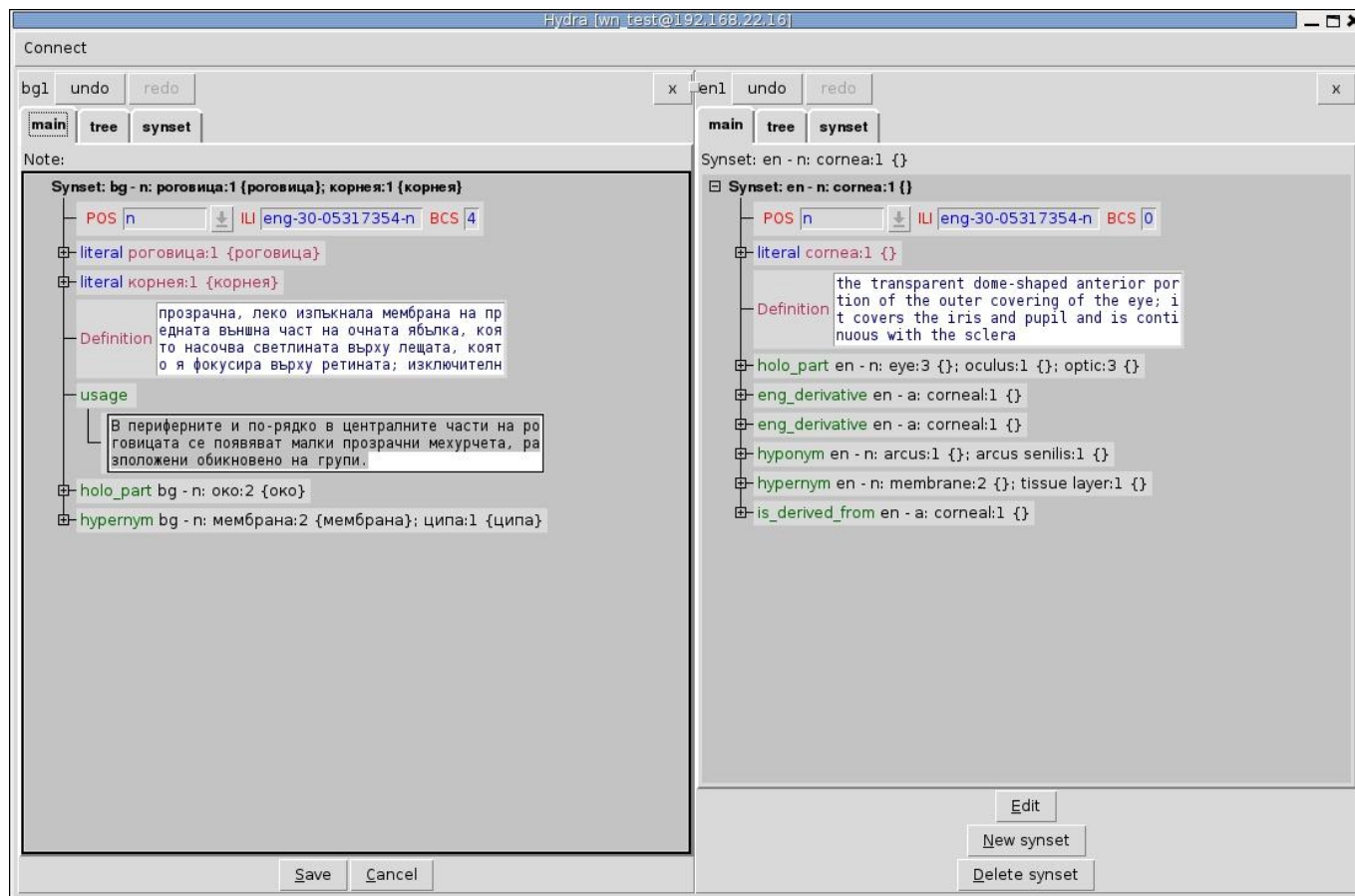


Fig. 13. Creation of a Usage

To add an Inote

- (1) Make sure the relevant literal is active.
- (2) Press the **Inote** button in the bottom area (Fig. 11) to create an **Inote** field.

(3) Type the info in the field (Fig. 14), then press **Save** to save the lonte.

(4) **Save** the literal.

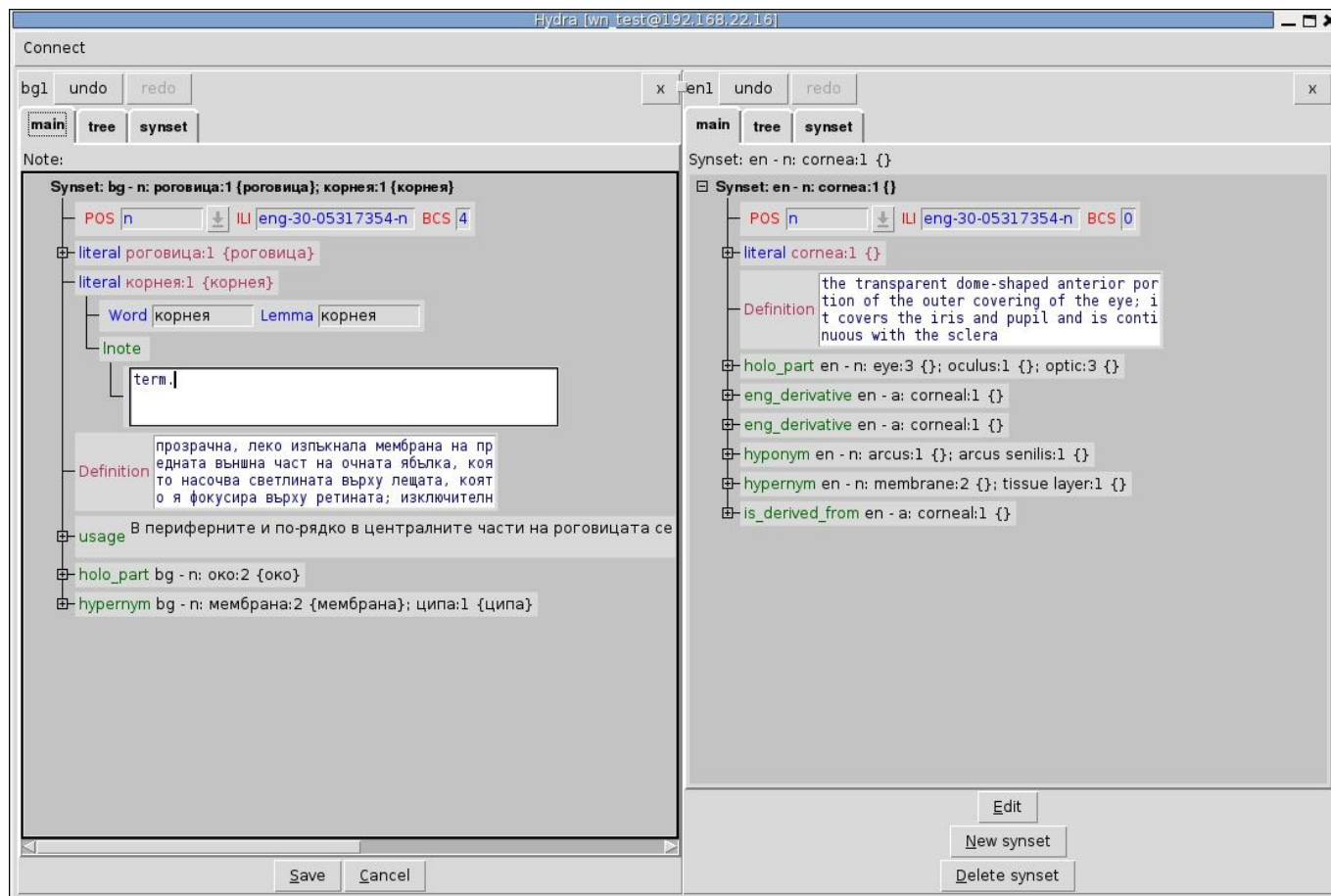


Fig. 14. Creation of an Lnote for the literal *корнея* (cornea).

In Fig. 14 the code **term.** indicates that the literal is used as a term in a specific domain unlike the neutral word *роговица*. When a new object of the type **literal**, **usage** or **note** is created, it is automatically linked to the synset or literal to which it belongs by means of

the respective relation.

3. Editing existing objects

To edit existing literals, notes, usage examples

- (1) Enable the **Edit** mode of the respective object
- . (2) Type/correct/add information
- . (3) **Save** the object.

B. Deleting objects

To delete a literal, a usage example or an snote, use the **Delete** button on its right. For Hydra to perform the command, you need to be in the synset **Edit** mode.

- (1) To delete an lnote, press the **Delete** button on its right; the literal must be active.
- (2) To delete a synset, press the **Delete synset** button in the bottom area of the **Main/Tree/Synset view** pane. The synset **Edit** mode must be disabled.

C. Adding relations

Hydra allows users to add new relations to existing objects or to newly created ones.

To add a relation:

- (1) Enable the **Edit** mode by hitting the Edit button.
- (2) Click on the **Add** button in the bottom area of the WN window.
- (3) A combobox with a list of the relations appears (Fig. 15).

The default relation for a synset is hypernymy. Only the relations available for the edited object type (synset, literal or note) are displayed.

- (4) To select another type of relation, click the pointer on the name of the relation.

The name of the selected relation appears in the **Relations** field (Fig. 15).

(5) Press the **Add** button.

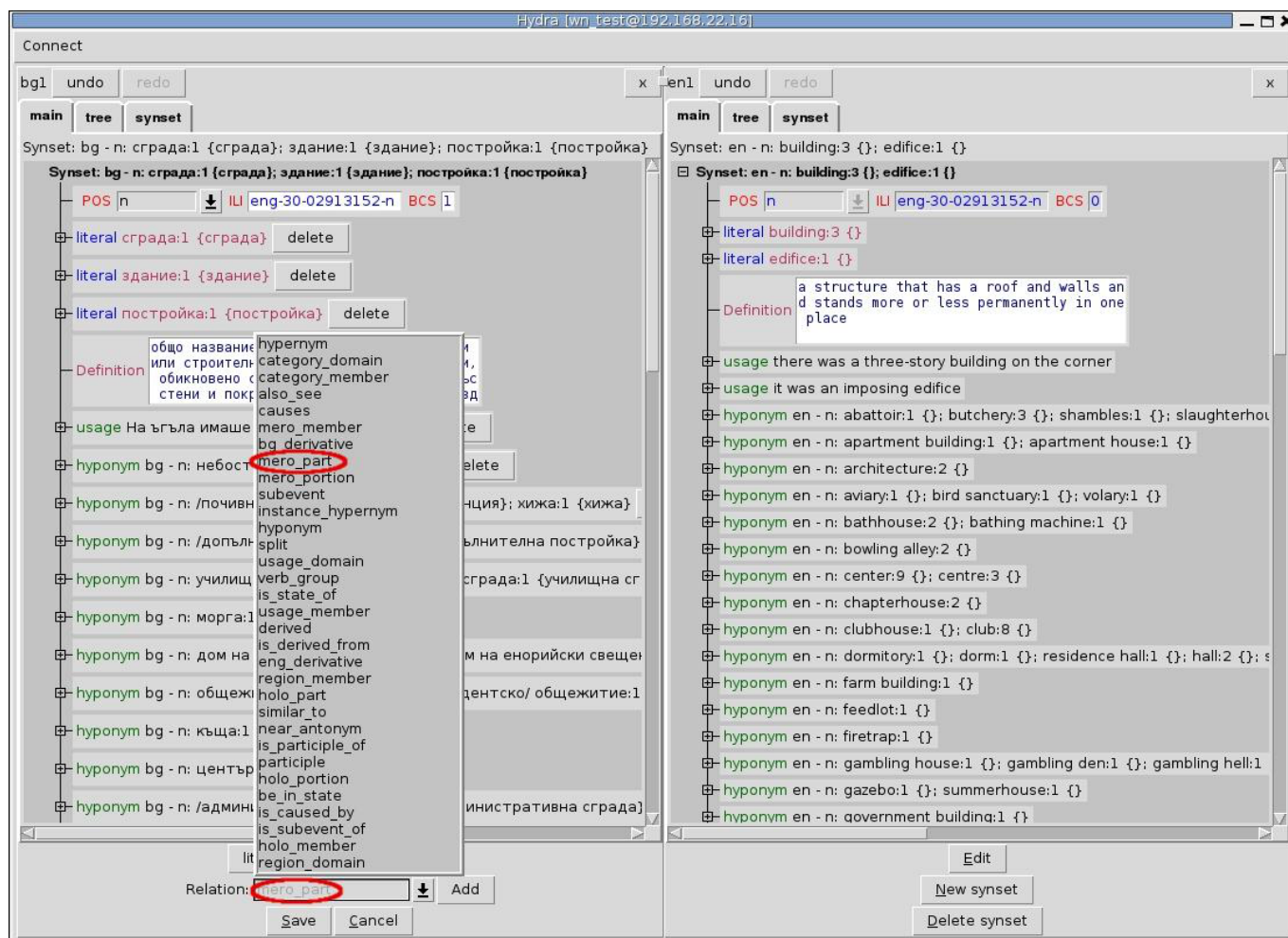


Fig. 15. The relations combo box.

Fig. 15 shows the synset {сграда:1; здание:1; постройка:1} (building:3; edifice:1) which is being connected with a meronym through the **mero_part** relation.

(6) A search tool similar to the main **Search window** is displayed in the bottom area of the WN pane (Fig. 16). Use the **input** field to type a query for the synset, literal or note to which the relation should point and press the **Enter** key.

(7) A list of the objects that match the query is displayed in the area below the **input** field.

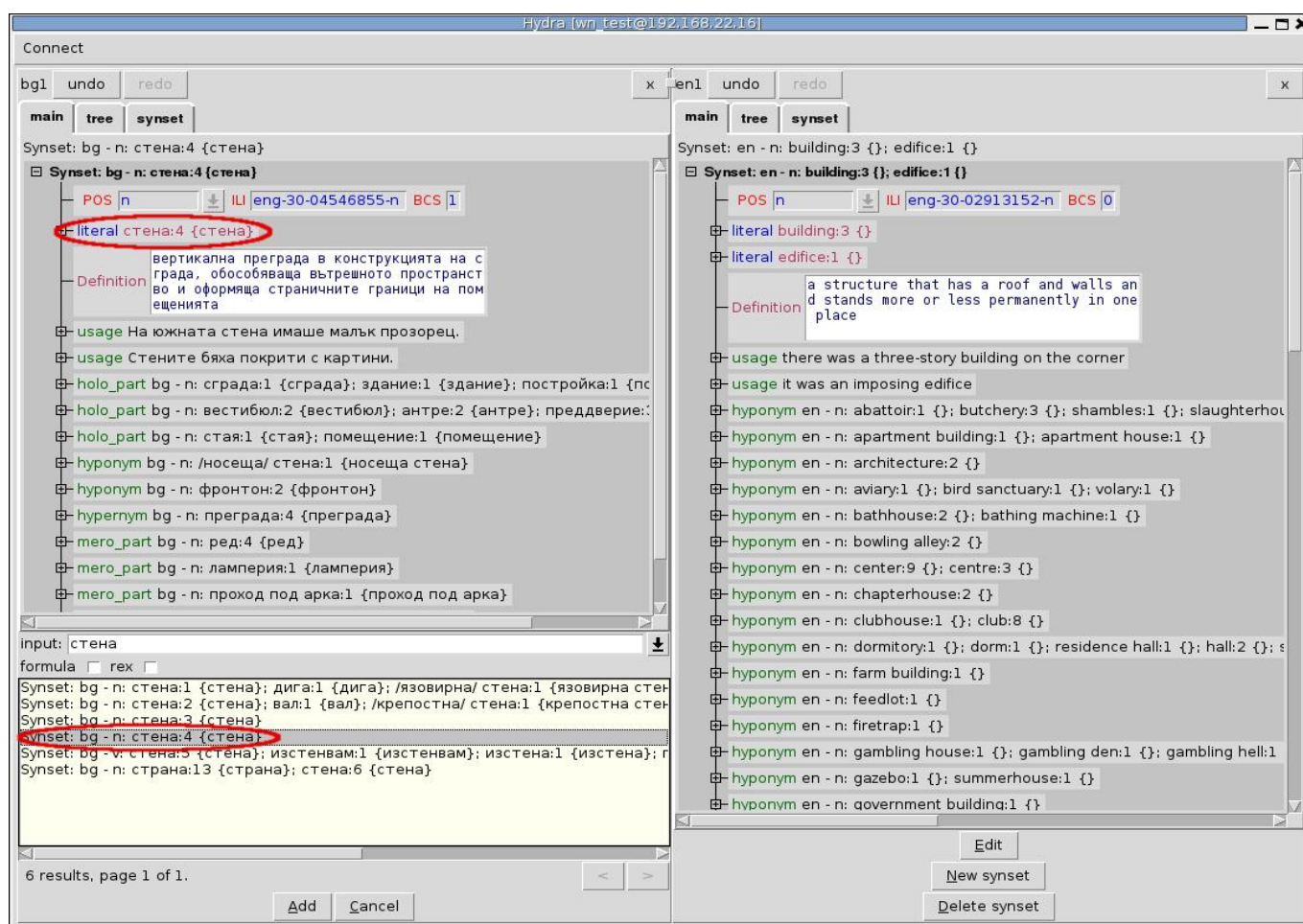


Fig. 16. Creating a relation between two synsets

(8) To view the synset associated with an object from the list, select it from the list

by clicking on it (Fig. 16). The corresponding synset - {стена:4} ({wall:3}) is displayed in the list is displayed in the upper part of the pane.

(9) To add a relation to an object, select the object from the list by clicking on it, then press **Add**. The WN window reverts to its regular **Edit** mode. The added object is displayed in the list of relations of the target object (Fig. 17).

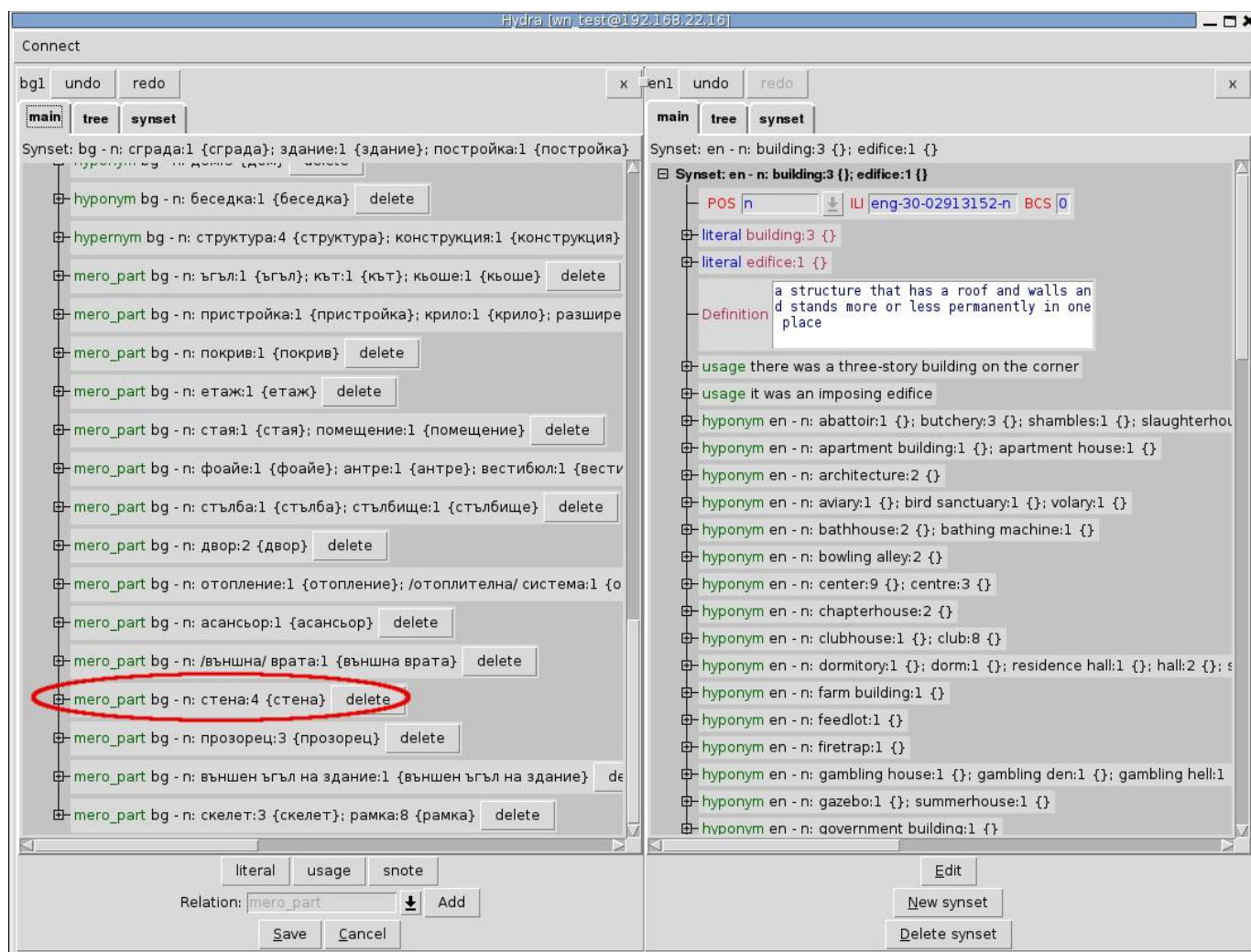


Fig. 17. The newly added relation (circled in red).

Fig. 17 shows that the synset {стена:4} ({wall}) selected as a meronym of {сграда:1;

здание:1; постройка:1} ({building:3; edifice:1}) is added to the relations of the synset {сграда:1; здание:1; постройка:1}.

Adding symmetric and asymmetric synset relations

The lists of the relations currently used in the Wordnet database are shown in Table 1 and Table 2. For the definition of the relations see the documentation of the Princeton WordNet, the EuroWordNet and the BalkaNet project, although there are some differences.

Many relations are asymmetric. The complete list of the asymmetric relations together with the corresponding inverse relations is shown in Table 2.

Relation (R)	Inverse Relation (~R)
hypernym	hyponym
instance_hyponym	instance_hyponym
holo_part	mero_part
holo_portion	mero_portion
holo_member	mero_member
causes	is_caused_by
be_in_state	is_state_of
derived	is_derived_from
participle	is_participle_of
category_domain	category_member
region_domain	region_member
usage_domain	usage_member

Table 1. Asymmetric relations



- A pair of synsets may be connected in either direction. It is important to

use the appropriate relation R or $\sim R$. In terms of the representation in Table 1, to connect an element in the first column, e.g. a hypernym or a holonym, to the second element of the respective relation, e.g. a hyponym or a meronym, use the inverse relation ($\sim R$), and vice versa - to link an element in the second column, e.g. a hyponym or a meronym, to the second element of the relation (a hypernym or a holonym), use the respective relation (R).

Table 2 shows the symmetric relations in the Wordnet database.

verb_group	verb_group
similar_to	similar_to
near_antonym	near_antonym
also_see	also_see
eng_derivative	eng_derivative
bg_derivative	bg_derivative

Table 2. Symmetric relations



- Unlike asymmetric relations there are no specific requirements with respect to the direction in which a symmetric relation is assigned.
- For both types of relations assign a relation in one of the directions only.
- Hydra provides for new relations to be defined.

D. Deleting relations

To delete a relation, press the **Delete** button on its right. To be able to do that, you need to enable the synset **Edit** mode.

E. Changing relations

There is no specific operation for changing the type of a relation. In order to do that, delete the relation, and add an appropriate one in the way described in **Adding relations**.

F. Editing a definition

Definitions are edited in the **Definition** field. The synset **Edit** mode should be enabled.

G. Editing the Part of speech value

The part of speech is corrected in the **POS** combo box in the left upper corner of the **Main View**.

H. Saving a synset

To save a synset after creating or editing it, press the **Save** button located in the bottom area of the wordnet pane.



- You are not allowed to proceed to edit or create another synset before saving the current one, or cancelling the operations performed.

2.5.2.2. Creating a new synset

New synsets are ones that do not exist in any of the languages. The creation of a new synset consists of several steps. In order for it to be completed successfully, a minimum of attributes must be supplied. A synset is minimally complete if it has a **POS** value, at least one literal and a definition. If any of the obligatory fields is left empty, the synset is ill-formed and therefore cannot be saved. A warning message pops up when pressing **Save** and further operations are disallowed as shown in Fig. 12 and Fig. 19.

To create a new synset press the **New synset** button located in the bottom area of the **Main View** pane. The new synset is automatically assigned a unique identifier (**ILI**) (Fig. 18).

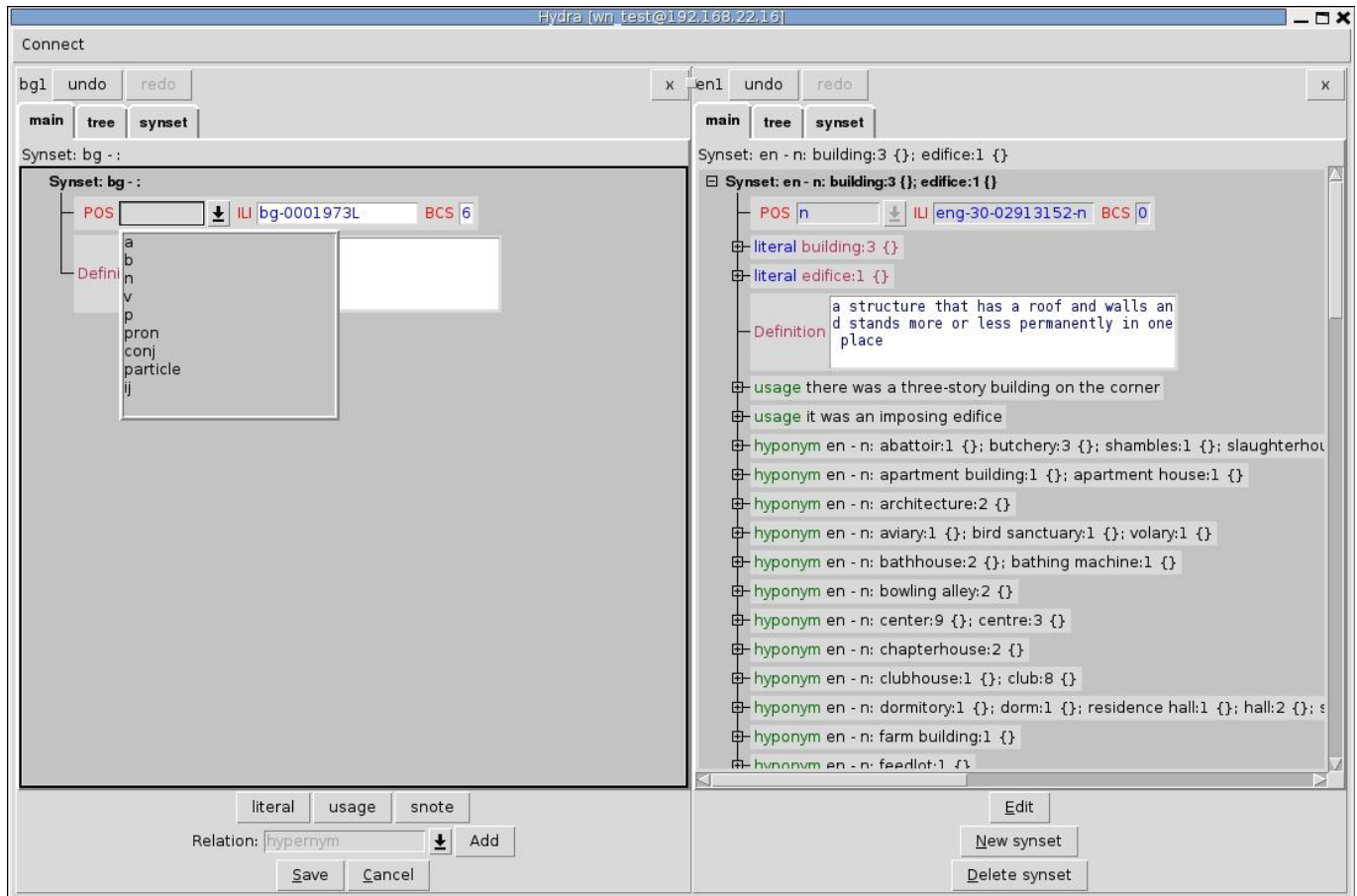


Fig. 18. A newly created synset

A. Part of speech

The part of speech of newly created synsets is assigned by manual selection from the **POS** combo box in the left upper corner of the **Main View** (Fig. 18). The part of speech codes of the content-word classes correspond to those in the Princeton WordNet:

a - adjective

b - adverb

n - noun

v - verb

Since the Bulgarian wordnet has been expanded with synsets from the other parts of speech, five additional codes have been configured:

p - preposition

pron - pronoun

conj - conjunction

particle - particle

ij - interjection

B. The definition

An explanatory definition is required in the **Definition** field.

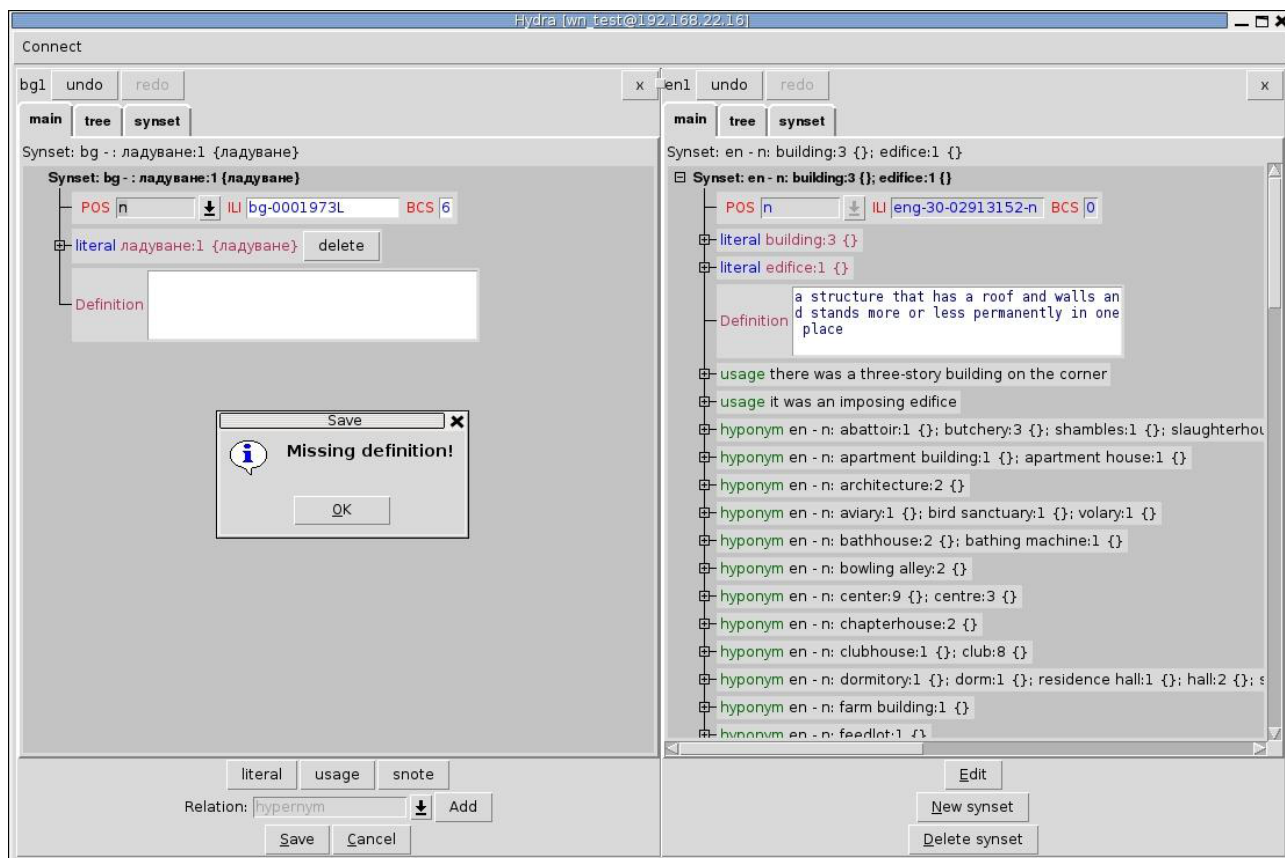


Fig. 19. Missing definition

C. Literals, usage examples, notes

To add, edit, delete and save literals, usage examples, snotes and lnotes of a new synset, proceed in the way described in the **Editing section** above.

D. Relations

To add, change or delete a relation of a new synset, follow the steps described in the relevant part of the **Editing section**.

E. Saving a new synset

To save the created synset, click on the **Save** button at the bottom of the **Main View**.

2.5.2.3. Cloning a synset

Cloning is an operation whereby a copy of a synset available in a wordnet (source wordnet) is created in another (target) wordnet. The clone inherits some attributes of the original synset: ILI, BCS, part of speech and the relations associated with it. Cloning is used to create synsets by preserving the original wordnet structure from the source wordnet into the target wordnet.

To clone a synset from a source wordnet into a target wordnet:

- (1) Right click inside the bottom area of the target wordnet, next to the **Edit**, **New synset** and **Delete synset** buttons, but not directly over them. Note, that pressing any of the buttons will enable the respective operation.

- (2) A list box with the possible source wordnets will appear (Fig. 20).

- (3) Select a wordnet to clone a synset from by placing the pointer over the wordnet name and then release the right button.

- (4) A clone synset appears in the target wordnet (Fig. 21).
- (5) Proceed to edit the clone synset in the same way as described for existing objects.

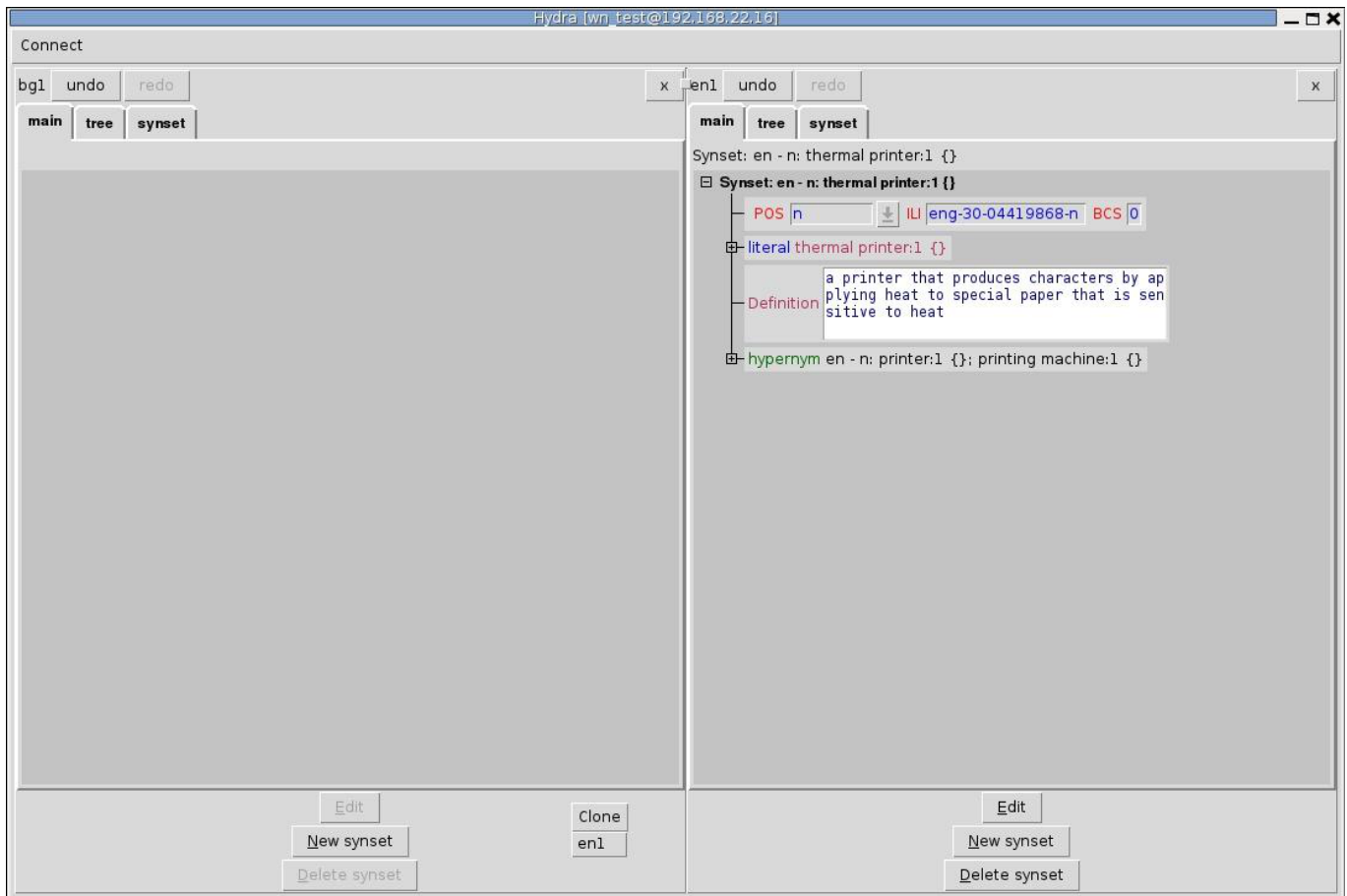


Fig. 20. Cloning a synset

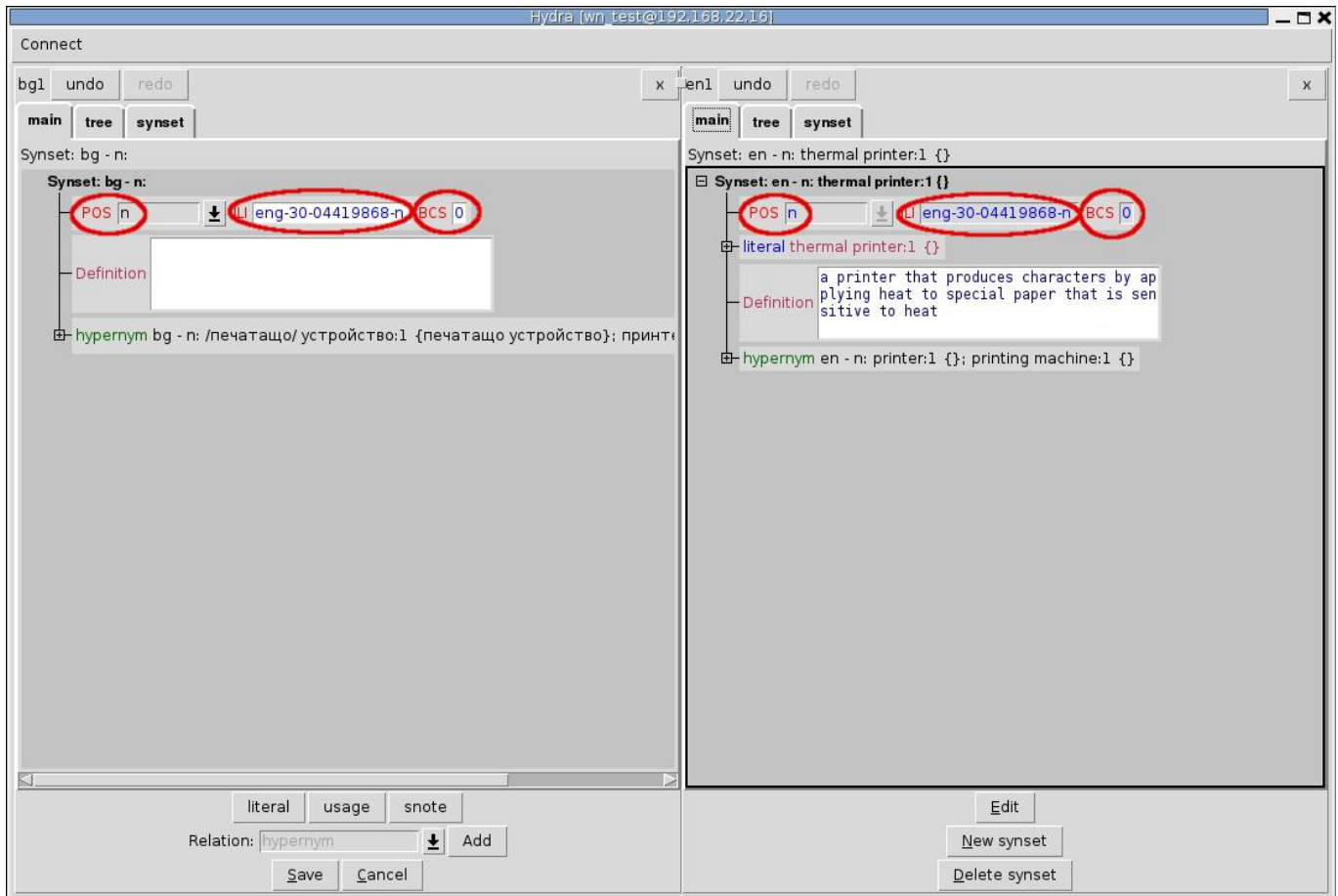


Fig. 21. A synset clone for the synset {thermal printer:1} (red circle). The clone inherits the POS value, the ILI and the relations (circled in red) of the original synset.



- Make sure the synset to be cloned is active in the source wordnet's pane. For a description of how to make an object active confer the Editing section above.
- Make sure the pointer is placed in the bottom area of the target wordnet pane.

2.5.2.3

Undo and Redo

Hydra has a high quality system for managing user operations in a way that can be easily undone, cancelled and redone. While the undo and redo operations are performed in atomic steps, the cancel operation applies undo to a whole group of operations – all the atomic operations performed after enabling the editing of an object.

To undo or redo an operation press the **Undo** or **Redo** button in the area above the views.

To cancel a group of operations, press the **Cancel** button in the area above the view.

2.5.3. Searching in Hydra

The search tool comes with three search modes:

- (i) using word(s): *cat, name, human*;
- (ii) using MySQL regular expressions: *cat, ^cat, cat\$*;
- (iii) using formulae in the Modal Language for WordNet.

The search in the first two modes retrieves the synsets that contain literals matching the user's input query. The formulae return the objects in Wordnet for which the formula is true. The default search option is the simple search mode.

2.5.3.1. Simple search

The simple search allows the users to search for a whole word or a combination of words that represent literals. For instance:

the query **material** - retrieves all the synsets that contain a literal **material**

the query **plant material** - retrieves the synsets containing a literal **plant material**

2.5.3.2. Regular expression search

This search mode uses MySQL regular expressions. To submit a regular expression query, check the **rex** box below the **input** field of the Search window.

Regular expressions:

cat - returns all the words containing the string **cat**

^cat - all the words starting with **cat**

cat\$ - all the words ending in **cat**

^c[au]t\$ - retrieves **cat** or **cut**

^cat.*cat\$ - all the words beginning with **cat** and ending in **cat** with any number of symbols in between.

For a detailed description of the MySQL regular expression syntax confer section 12.5.2 of the MySQL 5.1. Reference Manual.

2.5.3.3. Formula search

The formula search is a complex search mode based on the Modal Language for WordNet (Rizov 2008a, Rizov 2008b).

Modal Language for WordNet

A. Atomic formulae

The atomic formulae are Nominals or Boolean constants

Nominals – 1, 1235123412, 100323

Every object in Wordnet has a unique identifier, which is a natural number. This identifier is also an atomic formula in the language.

Boolean constants

\$s – true in every synset; the query retrieves all the synsets in the database

\$l – true in every literal; retrieves all the literals in the database

\$n – true in every note; retrieves all the notes in the database

type('value') -- retrieves the objects that have a data field 'type' with the value 'value'

type(#regex) – objects that have a data field 'type' with a value matching the regular expression 'regex'

The three sorts of objects are associated with the following types:

Synset: pos, ili, definition, lang, bcs, frequency

Literal: word, lemma, sense

Note: note

Simple and regular expression queries are defined and implemented as formulae.

For instance, the simple search **cat** retrieves all the synsets that have a literal **cat**, such as:

{cat, true cat} (feline mammal usually having thick soft fur and no ability to roar: domestic cats; wildcats)

{Caterpillar, cat} (a large tracked vehicle that is propelled by two endless metal belts; frequently used for moving earth in construction and farm work)

{big cat, cat} (any of several large cats typically able to roar and living in the wild)

Hydra converts the query into the following formula:

<literal>word('cat')

where:

word('cat') retrieves all the literals that contain a **word** field with the value **cat**.

To retrieve the synsets associated with these literals, the modal operator **<literal>** is used (see below), so that the formula **<literal>word('cat')** is formed.

In a similar way, the regular expression **cat** is converted to the formula:

<literal>word(#'cat').

B. Formulae

The formulae are defined by induction.

1. The Atomic formulae are formulae

2. Modal operators

Let **q** be a query and **R** – a relation:

$\langle R \rangle q, \langle R, n \rangle q, \langle R, n, m \rangle q$

$[R]q, [R, n]q, [R, n, m]q$ are formulae

The inverse relation of **R** is denoted by **~R**

The modal operators are interpreted in the standard Kripke's semantics.

Boolean combinations

If **p** and **q** are queries

$!p, p \ \& \ q, p \ | \ q, p \ = \> \ q, p \ < = \> \ q$ are formulae.

To submit a formula, check the **formula** box below the **input** field of the Search window.

2.5.3.4. Formula query tips

A. Searching by ILI, wordnet and Pos

ili('eng-30-01815628-v') - returns the synset with the ILI **eng-30-01815628-v** in every wordnet in the wordnet database in which it is found

ili('#15628') – matches all the synsets whose ILI contains the sequence of numbers **15628**

lang('bg') - retrieves all the synsets in the wordnet **bg**

lang('bg')&ili('#15628-v\$') - retrieves all the synsets in the wordnet **bg** whose ILI ends in **15628-v**

lang('bg')&pos('v') - retrieves all the verb synsets in the wordnet **bg**

B. Searching in literals and definitions

word('play') - matches the literals containing the word **play**

<literal>word('play') - matches all the synsets that contain the literal **play**

<literal>word('play')&pos('n') - matches all the noun synsets that contain the literal **play**

definition('#play') - matches all the definitions containing the word **play**

C. Searching in notes

<snote>\$n - retrieves all the synsets that have at least one Snote

<snote>note('pl. t.') - synsets containing an Snote **pl. t.** (pluralia tantum)

<literal><lnote>note('pl. t.') - synsets that contain literals having an lnote **pl. t.**

D. Searching in synset-to-synset relations

<hypernym>ili('eng-30-02396716-v') - matches all the synsets that share a hypernym

with the ILI 'eng-30-02396716-v'

<participle>\$s - matches all synsets that stand in the **participle** relation with (an)other synset(s)

pos('b')&<near_antonym>\$s - matches all the synsets with the POS=b (adverbs) that stand in the **near_antonym** relation with (an)other synset(s)

<hypernym, 4>\$s - matches all the synsets that have more than four hypernyms

pos('v')&<hypernym, 1>\$s - matches all the verb synsets with more than one hypernyms

pos('v')&<causes>\$s - matches all verbs that have a **causes** relation

pos('v')&<causes, 1>\$s matches all verbs that contain more than one **causes** relation

<ili><literal>word('човек') matches the translations (synsets) of the word 'човек'.

<ili>(lang('bg')&bcs('1')) matches the synsets that are equivalent of the synsets in language bg and having bcs with value 1.

E. Combining different types in complex queries

The formula: **pos('a')&<literal>word('good')&<near_antonym><literal>word('bad')** matches all the adjective synsets containing the literal **good** that stand an **near_antonym** relation with synset(s) containing the literal **bad** (i.e. finds the pairs of adjective synsets in which **good** and **bad** are antonyms)

F. Validation queries

Empty tags may occur when the operating system shuts down unexpectedly. The following formulae perform checks for empty tags:

<literal>word('') - retrieves all the synsets that contain an empty literal

definition('') - all the synsets that contain an empty definition

<usage>note('') - all the synsets that contain an empty usage

References

Fellbaum 1998: Fellbaum, C., ed. WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press, 1998, pp. 423.

Miller et al. 1993: Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross and K. Miller. Introduction to WordNet: An On-line Lexical Database. www: <http://wordnet.princeton.edu/wordnet/publications/>

MySQL 5.1. Reference Manual. www: <http://dev.mysql.com/doc/refman/5.1/en/index.html>

Rizov 2008b: Rizov, B. Hydra: A Modal Logic Tool for Wordnet Development, Validation and Exploration, Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, 2008, European Language Resources Association (ELRA) electronic publication. ISBN 2-9517408-4-0

Stamou et al. 2002: Stamou, K. Oflazer, K. Pala, D. Christodoulakis, D. Cristea, D. Tufis, Sv. Koeva, G. Totkov, D. Dutoit, M. Grigoriadou. BalkaNet: A Multilingual Semantic Network for Balkan Languages, In: Proceedings of the first Global WordNet Conference (GWC), Mysore, India, 2002, pp. 12-14. www: http://www.ceid.upatras.gr/Balkanet/pubs/GWA_paper_03.pdf

Vossen 2002: Vossen, ed. EuroWordNet General Document. EuroWordNet Project LE2-4003 & LE4-8328 report, University of Amsterdam, 2002. www: <http://www.vossen.info/docs/2002/EWNGeneral.pdf>