WordNet Web Navigation Interface: A fast interface to navigate EuroWordNet hierarchies

Eneko Agirre, Olatz Ansa, Xabier Arregi, Kike Fernandez
IXA NLP Group
University of the Basque Country
649 pk. 20.080
Donostia. Spain
eneko@si.ehu.es

Abstract

This paper introduces WWNI, a new web interface for multilingual WordNets. The main features of this interface are the following: all items shown are clickable, multilingual information can be shown if desired, and the user can navigate across the hypernymy and meronymy hierarchies in a straightforward way. The multilingual WordNet database is implemented in mSql, and the cgi's in Perl that produces dynamic html pages based on the W3C DOM model and javascript. Because of the DOM model, it requires versions equal or higher than Netscape 6 or Internet Explorer 5. As far as we know, it is the first web interface that allows for fast hierarchy navigation. It is accessible at the following URL: http://ixa.si.ehu.es/tresnak/wwni/index.html.

Introduction

Our research team has been working with WordNet for many years, and has been lately involved in the construction of the Basque WordNet. The Basque WordNet is being constructed using an interface developed in UPC (Benitez et al. 1998), but the standard WordNet tcl/tk interface is also used to consult the English WordNet. Linguists have often complained that these interfaces don't allow for easy navigation across the relations and hierarchies that involve words, word senses and synsets. For instance, if they were working on a word and wanted to check the meaning of its synonyms, they had to type each of the synonyms in turn. Even worse, the only way to view the parts of the hypernymy hierarchies was either asking for the whole tree (sometimes very slow, and producing enormous trees) or asking for only the direct hyponyms of the root. In the former case it can be very slow and produce meaningless huge trees. In the later, if the user wants to go one level deeper he/she needs to type all the hyponyms in turn.

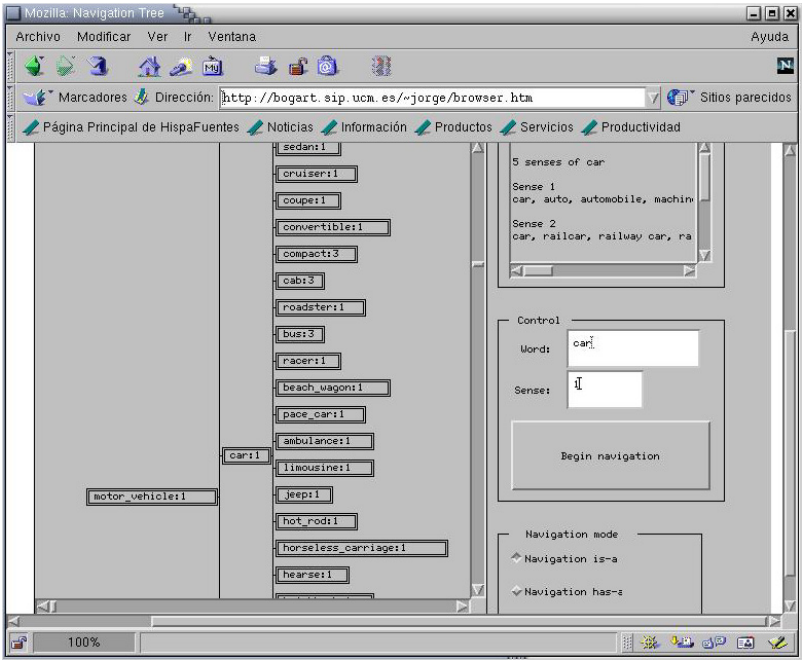
In view of these limitations we started to look for an interface that had the following features:

- Web-accessible, i.e. the actual multilingual WordNet database was in a remote machine
- Easy to navigate:
  - Hierarchies shown as trees that were expanded on demand
  - All items were clickable
- Fast and nice look and feel.
- Easy and intuitive to use. For instance, the tree could look like directory trees.
- No linked to any specific web browser

1 Review of other interfaces

We made an analysis of interfaces (others than the standard tcl/tk and the UPC one) listed in WordNet's official web page[1]. We found out that all but two of them do not allow the user to click on the data (e.g. words, synsets) or navigate in the hierarchies:

- "The visual navigation tree" does allow for both things, but is rather slow, and the way to show the trees was very awkward.
- "The WordNet treewalk" allows for easier navigation, but it is far from intuitive (there relations are represented by symbols) and it is a standalone program (not web based).



Due to the inexistence of a interface as desired we decided to design and build our own.

Figure 1: The visual navigation interface.

2 Design

Figure 2 shows the main window of the WWNI interface. It is divided in three frames: the *control frame*, the *word sense frame* and the *hierarchy tree frame*.

The upper part is the *control frame*. The *target word* is input here, alongside the *desired WordNet* (English, Basque or other). The *multilingual check box* can be used when we wish to show the information for one language or all the languages in the database.

After pressing the *search button* the information about the senses of the target word are shown in the *word sense frame* (left), including the gloss, and the words in the synset. If the multilingual box is checked the information for all equivalents in the languages in the database are also shown in the word sense frame. Each language uses a different color. For each of the word senses we can choose *info* to show the complete list of relations for the synset in a separate window (see figure 3), or to start a *hyponymy/meronymy tree* in the third frame.

[1] The list of interfaces checked is the following: Kyoto One Touch, Oxford SQL, Visual Navigation Tree, TreeWalk, Greaves SQL, WN Python, Enst interactive, Vancouver One Touch, the UPC euroWordNet interface Check the official WordNet site for more detailes.

At the beginning the *hierarchy tree* is rooted in the word sense chosen, and only direct hyponyms, if any, are shown. Each concept in the hierarchy allows for expansion or collapsing clicking the + / - *box*, i.e. we can expand / cut the subtree below any synset. Only three variants are shown for each synset, but clicking … we can see all. The *bulb* represents one synset and clicking on it a *menu* is displayed with the following choices: *synset info* for opening a window with all the info of the synset (cf. figure 3), *become root* for deleting all the tree except this synset and its direct hyponyms, *open / close* for opening or closing the hyponymy tree (same functionality as clicking the + / icons), and finally *hyper / holo* for showing the hypernym(s) / holonym

The hierarchies in WordNet are rather lattices than strict trees. As there are not straightforward algorithms to control lattices in the way we can control trees, we decided to use trees. This means that expanding a tree upwards and downwards we could display the same subtree in different places. In order to control the display properly hyponyms of the root synset can only be expanded downwards, and hyperrnyms of the root can only be expanded upwards. The root is the only synset that allows for downward and upward expansion. The *become root* option allows for quickly setting a different synset as root.

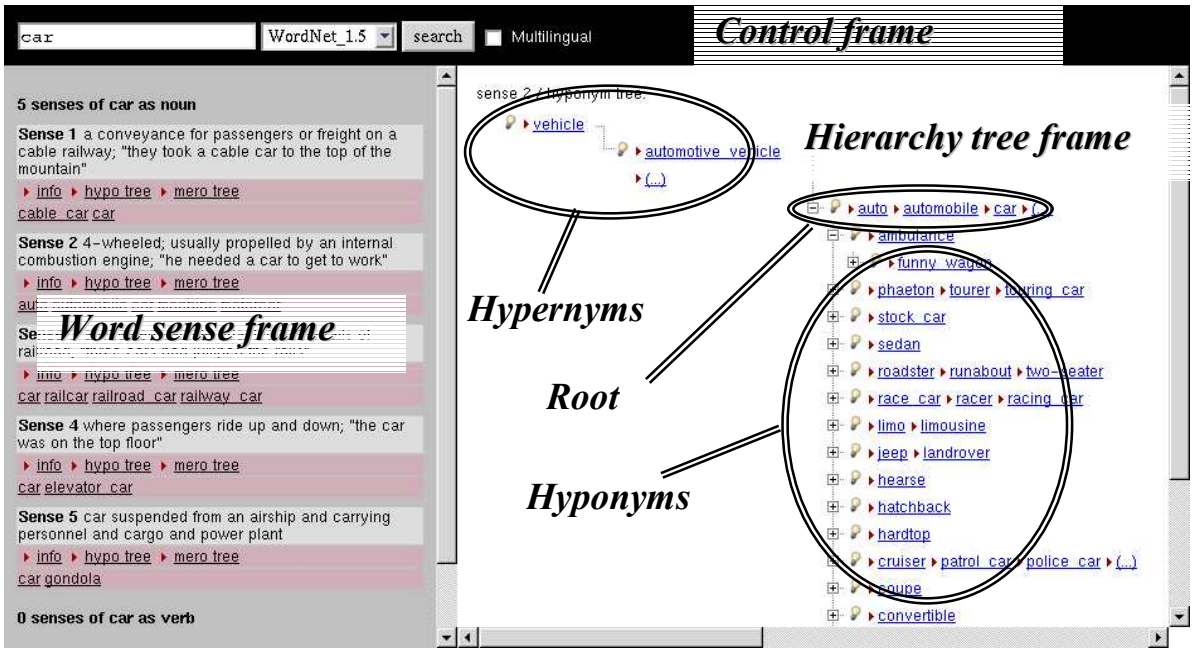


Figure 2: Main window of WWNI.

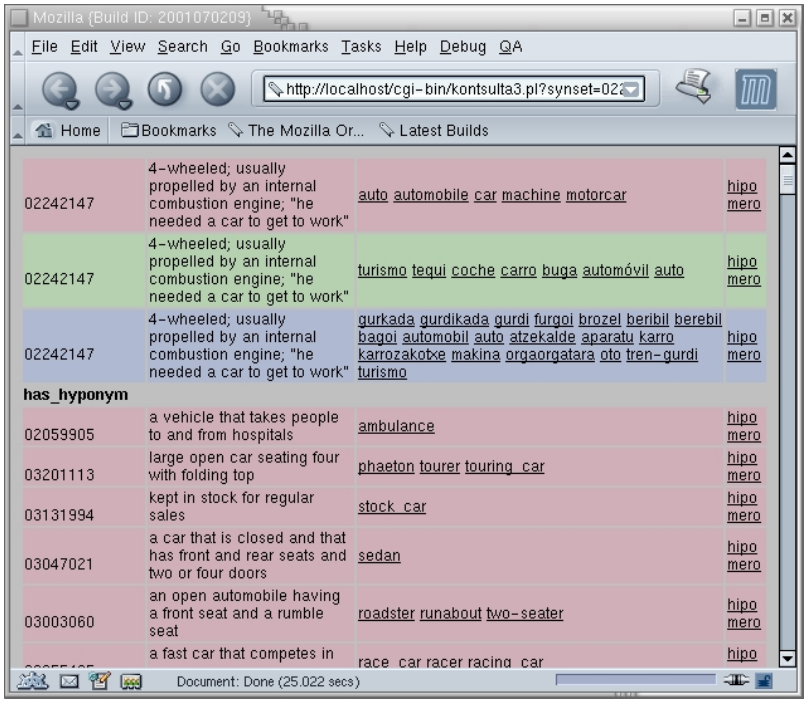


Figure 3: Information window.

All relations and available information for a given synset can be checked using the *info* option either in the word sense frame or in the menu of the hierarchy tree frame. An example is shown in figure 3. As in the word sense frame, whenever the multilingual box is checked, the information for all languages is shown, using different colors to differentiate the languages.

All words (variants) and synsets in any frame or window are clickable. For words, the word sense frame is updated with the corresponding word sense information. For synsets, the option in the menu (hierarchy tree frame) or the link (word sense frame) determines the action to be taken.

3 Architecture and implementation details

In the server side, the multilingual WordNet database is implemented in mSql, and the cgi's in Perl. The Perl scripts access the database using the DBI and MSql modules. The Internet server used is Apache. Each time the web browser requests for a page, the respective Perl cgi is run. For all frames except the hierarchy tree frame, the Perl cgi returns static html code.

For the hierarchy tree frame, there is a set of javascript functions that are loaded with the main page. These functions allow for the manipulation and display of the tree, using the W3C DOM model. Every time the user changes the tree, a request to the corresponding cgi in the server is made. The cgi, in this case, returns an empty frame with javascrit function calls that include the information needed to update the tree. When the function calls are executed in the client side, the tree is changed accordingly. This process simulated that whenever the user changes the tree, the javascript functions with the requested information are executed (red lines in figure 4).

Because of the need of a browser with DOM, it requires versions equal or higher than Netscape 6 or Internet Explorer 5.

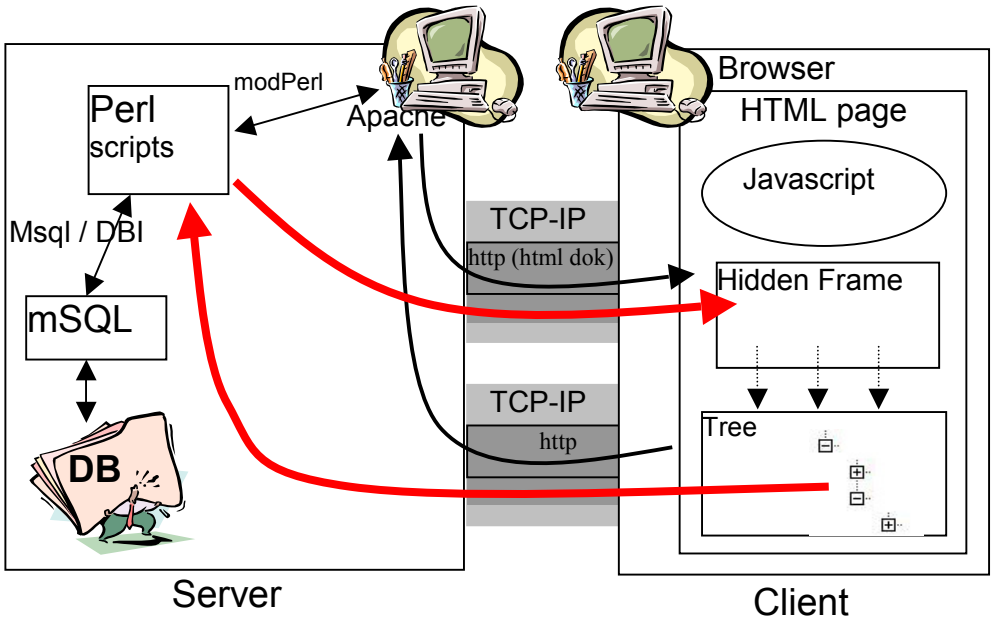


Figure 4: architecture of the interface (client) and the server

4 Conclusions

The WWNI web interface for multilingual WordNets is an amiable interface that has as a special feature the easiness to navigate across the hypernymy and holonymy hierarchies. It also allows the user to easily access any word or synset, as all items in the interface are clickable. It is accessible at the following URL: http://ixa.si.ehu.es/tresnak/wwni/index.html.

The interface is built over a relational database representation of the EuroWordNet design, but could easily be ported to other WordNet implementations that provide the convenient API's for Perl.

References

Benítez, L. ;Escudero, G. ; Farreras, J. & Rigau, G. 1998. WWI: A Multilingual WordNet Interface using the Web. Technical Report LSI-98-6-T. LSI Department, Universitat Politécnica de Catalunya.