

Web Application for Semantic Network Editing

Adam Rambousek and Tomáš Hrušo

Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
xrambous@fi.muni.cz, 396240@mail.muni.cz

Abstract. Semantic network editor, DEBVisDic [1], has been used to create more than 20 national wordnets. The editor was developed in Mozilla Development Platform, as the extension for Mozilla-based web browsers.

However, the development of the web-related technologies took a step from the browser-based extensions to rich web applications, usable in any browser. We decided to rewrite the editor from scratch and create multi-platform web-based application for general semantic networks editing. In the first phase, the editor will be used to build Open Dutch Wordnet, under the Cornetto project. In case of successful deployment and evaluation, the editor will be enhanced to build any wordnet-like semantic network.

Key words: semantic network, ontology, editor, web application, DEB-VisDic

1 Introduction

The original wordnet, Princeton WordNet, is one of the most popular lexical resources in the NLP field [2]. It was followed by multilingual EuroWordNet 1, 2 projects (1998-99) [3] and Balkanet project (2001-4) [4] in which the wordnets for 13 languages have been developed (English, Dutch, Italian, Spanish, French, German, Czech, Estonian, Bulgarian, Greek, Romanian, Serbian and Turkish). In the course of this work the software tools for browsing and editing wordnets have been designed and implemented, without whose the job could hardly have been performed. Within the EuroWordNet project the Polaris (and Periscope) tools have been implemented and used [5].

For Balkanet project the browser and editor VisDic has been prepared at the NLP Laboratory at the Faculty of Informatics Masaryk University [6] since the development of the Polaris tool has been closed by 1999.

In comparison with the previous tools VisDic exploits XML data format thus making the wordnet-like databases more standard and exchangeable. Not only that, thanks to the XML data format used and to its dictionary specific configurability VisDic can serve for developing various types of dictionaries, i.e. monolingual, translational, thesauri and multilingually linked wordnet-like databases. The experience with the VisDic tool during Balkanet project has been positive [7] and it was used as the main tool with which all Balkanet wordnets were developed.

2 DEB platform and DEBVisDic editor

VisDic, however, has its disadvantages, particularly it is designed for off-line use by single user, and team coordination is really difficult.

Based on the experience with VisDic, we designed and implemented more universal dictionary writing system that could be exploited in various lexicographic applications to build large lexical databases. The system has been called Dictionary Editor and Browser (further DEB) [8] and has been used in many lexicographic projects, i.e. for development of the Czech Lexical Database [9], or currently running Pattern Dictionary of English Verbs [10], and Family names in UK [11].

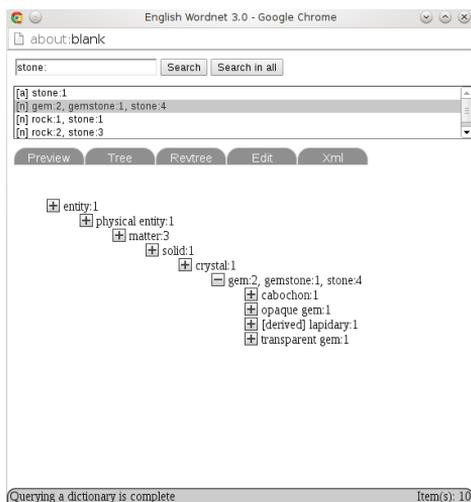


Fig. 1: Example of hypero-hyponymic tree.

The DEB platform is based on client-server architecture, which brings along a lot of benefits. All the data are stored on the server and considerable part of functionality is also implemented on the server, while the client application can be very lightweight.

This approach provides very good tools for team cooperation; data modifications are immediately seen by all the users. Server also provides authentication and authorization tools.

The design of the DEB allows us to modify it also for building wordnet-like databases. For this purpose, VisDic tool was re-implemented on top of the DEB platform, as the DEBVisDic editor[1].

DEBVisDic editor was designed as a client application for the DEB server, and created using the Mozilla Development Platform[12], which was at the time the best option to design and build cross-platform GUI applications, utilizing open standard.

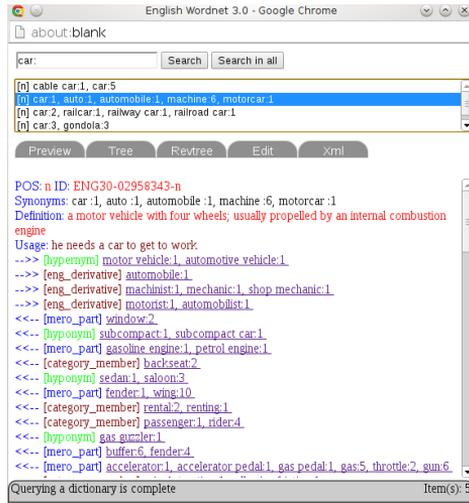


Fig. 2: Example of synset preview.

However, the applications based on Mozilla Development Platform are limited only to Mozilla-based browsers (mainly Firefox), while users prefer many different web-browsers. Since the development of DEBVisDic, Firefox browser introduced several major changes to application interface, limiting DEBVisDic to be used only in specified versions of Firefox browser. As a result, the editor would need major changes to work with recent Firefox versions.

Fortunately, the standards for web-based application supports much more features and are implemented by all the major web browsers. Considering all the options, we decided to re-implement DEBVisDic editor as a general web application, not limited to single web browser and without need to install specific extensions.

3 DEBVisDic 2

Thanks to the client-server architecture of DEB platform, no changes were needed on the server side. Only the client side application needed to be reimplemented, reusing the existing DEB interface. Main feature requests when designing the new version, were reimplementing all DEBVisDic features, and to provide application working in all major web browsers.

Similar to previous version, *DEBVisDic 2* aims primarily on wordnet-type semantic network browsing and editing, but supports different types of dictionaries. The application consists of main window with settings and separate windows for each dictionary that user want to edit. Single dictionary window includes the list of entries (synsets) and a set of tabs with several views on selected entry: basic preview, XML representation, hypero-hyponymic tree, and editing form. Context (right-click) menu provides functions for displaying and

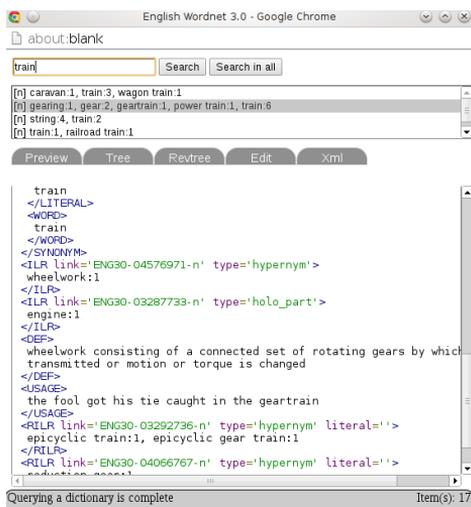


Fig. 3: Example of synset XML representation.

creating inter-dictionary links (i.e. display all synsets using selected ontology term).

DEBVisDic 2 utilizes Model-view-controller architecture and design follows this principle. Current open standards are used in the application: HTML and CSS for data presentation (view), and Javascript for application logic scripting (model, controller). The application is modular, with separate core shared by all the dictionaries, and a plugin with specific functionality for each type of dictionary.

Because the implementation of web-related standards (mainly Javascript) may vary in different browsers, several frameworks and libraries provide unified environment on top of the browser interface. After reviewing several frameworks, we decided to use jQuery library[13], that is versatile Javascript library for simpler document and data manipulation, but doesn't add unnecessary features, thus staying lightweight and not slowing down the application.

One of the most challenging features, was the implementation of the context menu functions, because of the huge differences in different browsers. In the end, we were able to implement the context menu to behave the same as in *DEBVisDic*, with the help of jQuery contextMenu plugin¹. Pretty printing of entry in XML format is provided by Prettify plugin².

Apart from complete reimplementing of *DEBVisDic* tool, new version comes with several new features. For example, saving user settings (opened dictionaries and window positions, with the possibility to store more information) on the server, thus allowing user to switch browsers and computers, and continue in work.

¹ <http://medialize.github.io/jquery-contextMenu/>

² <http://google-code-prettify.googlecode.com/svn/trunk/README.html>

Another major new feature are more generalized links and relations between dictionaries. It is possible to use any part of XML entry to build inter-dictionary search queries. For example, selecting all lexical units in a synset, automatically view details of ontology term for selected synset, all synonym or near synonym synsets between two wordnet languages.

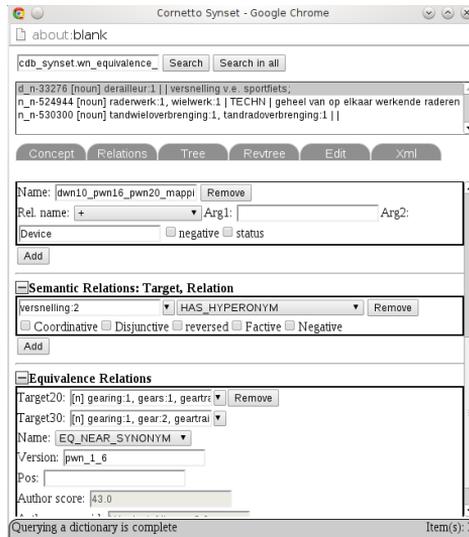


Fig. 4: Example of editing form.

4 Testing

DEBVisDic 2 editor was developed for the creation of Open Dutch Wordnet, based on the Cornetto project [14]. The data from the Cornetto project are not just a simple wordnet-like semantic network. It contains separate lexical database with detailed information about lexical units, and semantic network with synsets linked to each other, and also with links to corresponding lexical units and to several versions of English Wordnet.

Because of the database design, specific dictionary modules were needed for Cornetto Synsets, Cornetto Lexical Units, Open Dutch Wordnet, and English Wordnet, all inter-connected together.

Lexicographers' feedback after a few weeks of intensive editing is highly positive, and we are gathering comments to incorporate in future updates.

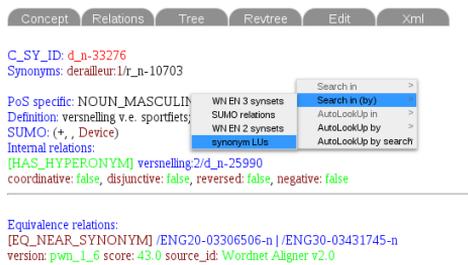


Fig. 5: Example of context menu with inter-dictionary links.

5 Future work

After the end of initial editing phase, *DEBVisDic 2* will be updated based on user feedback. We plan to add support for the editing of general wordnet-type semantic networks, thus encouraging the creation of more national wordnets.

Acknowledgements This work has been partly supported by the Ministry of Education of CR within the LINDAT-Clarin project LM2010013.

References

1. Horák, A., Pala, K., Rambousek, A., Povolný, M.: First version of new client-server wordnet browsing and editing tool. In: Proceedings of the Third International WordNet Conference - GWC 2006, Jeju, South Korea, Masaryk University, Brno (2006) 325–328
2. Fellbaum, C., ed.: WordNet: An Electronic Lexical Database. MIT Press (1998)
3. Vossen, P., ed.: EuroWordNet: a multilingual database with lexical semantic networks for European Languages. Kluwer (1998)
4. Christodoulakis, D.: Balkanet Final Report, University of Patras, DBLAB (2004) No. IST-2000-29388.
5. M., L.: Polaris user's guide. Technical report, Belgium (1998)
6. Horák, A., Smrž, P.: VisDic – wordnet browsing and editing tool. In: Proceedings of the Second International WordNet Conference – GWC 2004, Brno, Czech Republic (2003) 136–141 <http://nlp.fi.muni.cz/projekty/visdic/>.
7. Horák, A., Smrž, P.: New features of wordnet editor VisDic. In: Romanian Journal of Information Science and Technology. Volume 7. (2004) 1–13
8. Horák, A., Rambousek, A.: DEB Platform Deployment – Current Applications. In: RASLAN 2007: Recent Advances in Slavonic Natural Language Processing, Brno, Czech Republic, Masaryk University (2007) 3–11
9. Rangelova, A., Králík, J.: Wider Framework of the Research Plan Creation of a Lexical Database of the Czech Language of the Beginning of the 21st Century. In: Proceedings of the Computer Treatment of Slavic and East European Languages 2007, Bratislava, Slovakia (2007) 209–217
10. Hanks, P.: Corpus pattern analysis. In: Proceedings of the Eleventh EURALEX International Congress, Lorient, France, Universite de Bretagne-Sud (2004)

11. Hanks, P., Cullen, P., Draper, S., Coates, R.: Family names of the United Kingdom. (2014)
12. Oeschger, I., et al.: Creating Applications with Mozilla. O'Reilly and Associates, Inc., Sebastopol, California (2002)
13. jQuery Foundation: jQuery. <http://jquery.org> (2013)
14. Horák, A., Vossen, P., Rambousek, A.: The Development of a Complex-Structured Lexicon based on WordNet. In: Proceedings of the Fourth Global WordNet Conference, Szegéd, Hungary, University of Szegéd (2008) 200–208