WordnetLoom – a Multilingual Wordnet Editing System Focused on Graph-based Presentation

Tomasz Naskręt[1], Agnieszka Dziob[1], Maciej Piasecki[1], Chakaveh Saedi[2], António Branco[2]
[1]G4.19 Research Group, Department of Computational Intelligence
Wrocław University of Science and Technology, Wrocław, Poland
[2]NLX-Natural Language and Speech Group, Department of Informatics
University of Lisbon, Faculty of Sciences, Portugal
{maciej.piasecki, agnieszka.dziob, tomasz.naskret}@pwr.edu.pl
{chakaveh.saedi, Antonio.Branco}@di.fc.ul.pt

Abstract

The paper presents a new re-built and expanded, version 2.0 of WordnetLoom – an open wordnet editor. It facilitates work on a multilingual system of wordnets, is based on efficient software architecture of thin client, and offers more flexibility in enriching wordnet representation. This new version is built on the experience collected during the use of the previous one for more than 10 years of plWordNet development. We discuss its extensions motivated by the collected experience. A special focus is given to the development of a variant for the needs of *MultiWordnet of Portuguese*, which is based on a very different wordnet development model.

1 Introduction

A wordnet is a complex graph of several types of nodes (e.g. lexical units[1], synsets) and edges (e.g. lexical relations, synset relations). Initially Princeton WordNet development was based on manual editing of text files storing wordnet representation (Fellbaum, 1998). Such an approach was error prone and the files edited manually required a lot of error verification and maintenance. At the beginning of the plWordNet project in the year 2005, we developed a wordnet editing system, called WordnetLoom in order to avoid problems with manual editing of wordnet representation. It was based on a database and Graphical User Interface (GUI), and separated users from the internal representation of the wordnet. As plWordNet was developed by a team of linguists, it was important to provide distributed access to the system. WordnetLoom has been constructed in a way providing support for the corpus-based wordnet development method used for plWordNet (Maziarz et al., 2013); i.e. enabling close association between editors' decisions and language data, the use of substitution tests and application of semi-automatic methods as tools for editors. An unique feature of WordnetLoom is the possibility to simultaneously browse and edit wordnet graphs directly on the screen. Nevertheless, WordnetLoom was based on a quite inefficient thick client model, as well as it had restricted expressiveness of the applied wordnet representation and limited possibilities to adapt UI to the format extensions. Moreover, WordnetLoom was initially designed to support a monolingual wordnet. It was successfully used for editing plWordNet onto Princeton WordNet mapping, but the simultaneous presentation and editing of the two wordnets was due to a trick: introduction of additional 'English' PoS.

Our goal is to present a new re-built and expanded, version of WordnetLoom 2.0 facilitating work on a multilingual system of wordnets, based on an efficient software architecture of a thin client, and offering more flexibility in enriching wordnet representation. This new version originates from the experience collected during the use of the previous one that has clearly motivated the extensions. We will also discuss its applications and variants, with a special focus on the *MultiWordnet of Portuguese*.

2 Related Works

The first popular wordnet editor was probably *VisDic* (Horák and Smrž, 2004). In VisDic the relation definitions were still written in text windows, but an XML based format was utilised. VisDic was a monolithic application directly working on XML files, contrary to its descendant DEBVisDic (Horák et al., 2006) – a client-server, lexical database editor, based on a general platform for dictionaries called DEB (Horak et al.,

[1]A triple: lemma, Part of Speech, sense id.

Tomasz Naskręt, Agnieszka Dziob, Maciej Piasecki, Chakaveh Saedi and António Branco (2018) WordnetLoom – a Multilingual Wordnet Editing System Focused on Graph-based Presentation, *Proceedings of the 9th Global WordNet Conference (GWC 2018)*, pp 191–200, Singapore, ISBN 978-981-11-7087-4

2008). DEBVisDic reimplemented and extended the functionality of VisDic, and offered also more flexibility in adapting XML representation structures. Data presentation was limited and there was no means for visual editing the relation structure. Several other wordnet editors also do not provide elaborated visualisation for wordnet structures, e.g. *Hydra* Rizov (2014) or *OMWEdit* (Morgado da Costa and Bond, 2015).

A web-based system *sloWTool* (Fišer and Novak, 2011, Fišer and Sagot, 2015) offers good UI and visual wordnet browsing and editing. However, presentation is always limited to a small fragment of the wordnet graph (up to two links distance) and there is no means for neither viewing larger parts, nor comparing different parts.

Visualisation of wordnet graphs in most tools follows a radial pattern: a synset in focus is presented in the middle and all links, irrespectively of their types are placed radially around the central element, e.g. *sloWTool* or *WordTies* (Pedersen et al., 2012). *GernEdiT* (Henrich and Hinrichs, 2010) offers visualisation of the wordnet structure in the range selected by the user, but it is hierarchical and focused mainly on hypernymy. Moreover the visual presentation does not allow for direct editing of the structures. *WordnetLoom* introduced elaborated presentation of the relation graph and direct visual editing (Piasecki et al., 2013). As it is an open tool, it was used as a basis for the solution presented in this paper.

3 Basic Assumptions

WordnetLoom 1.0 has been used for plWordNet development since 2005 and proved to be a generally useful system. Thus, although software architecture has been reconstructed, the main philosophy of the system was preserved.

In order to avoid errors in the representation format, all editing actions should be done only via GUI client application and the results are stored in the central database. The XML-based format is secondary in relation to the database. WordnetLoom supports distributed group work by a group of linguists on the central database.

plWordNet construction has been following corpus-based wordnet development paradigm. Each iteration starts with the extraction of the most frequent lemmas from a large corpus together with the automated extraction of their semantic description, e.g. as a measure of semantic similarity. New lemmas are divided into *packages* on the basis of similarity-based clustering. The packages are assigned to linguists as work assignments and presented in WordnetLoom.

Substitution tests[2] are an intrinsic part of the relation definitions. Test templates are kept together with the relation definitions in the database. Before every editing decision is made, a test for a relation considered by the linguist is presented in a pop-up window and instantiated with the lemmas from the two synsets to be linked.

A wordnet is a network of lexico-semantic relation, and a graph is the basic means for both browsing and editing the wordnet structure. A network of synsets linked by synset relations is visually presented on the screen as a graph. The user can freely browse the network by clicking on synsets and unfolding as many levels of relations as needed, see Fig. 2. Every link can be added or removed directly on the graph presentation. This facilitates better comprehension of the wordnet structure, shorter connection between the editing intention and the resulting change in the wordnet structure, as well as a better understanding of the consequences of the intended and/or performed action to the wordnet structure beyond the local connections of the edited synset.

The same system and the same presentation means should also support the construction of the mappings between wordnets. Thus wordnets for different languages should be presented simultaneously on the screen as graphs that are connected by inter-lingual relations which are also visually presented on the screen. The editing of the mapping is performed in a way similar to monolingual editing by linking synsets or deleting links selected on the screen with the mouse.

Every wordnet includes also elements of the description that are not relations but attributes, e.g.: glosses, usage examples, and different attributes, e.g. stylistic register, sentiment polarity etc. As this kind of information is getting richer with the subsequent versions of plWordNet, we need also to introduce different perspectives on wordnet, not only graph-based, but also more dictionary-oriented. It is not also possible to fit everything into one single screen graph-presentation – the graph would be too cluttered. Attributes for a synset in focus are presented in side panels. Word-

[2] Each consists of one or more test sentences with slots for the tested lemmas.

netLoom offers three main perspectives on data: the *perspective of lexical units*, *visualisation* and *synsets*. The perspective of lexical units presents the wordnet as a dictionary. The searching is focused on lexical units (henceforth, LUs) and their relations, for a selected LU all synsets which it belongs to are listed. In addition the complete description of its attributes and lexical relations is shown. The synset perspective is organised in similar way, but around synsets as basic elements, and the visualisation perspective presents visually wordnet as a network of synsets. For a synset in focus its LUs are presented in the side panels together with their lexical relations.

4 Graph-based Presentation

A wordnet is intrinsically a graph. Lexical meanings are described by subgraphs of lexico-semantic relations. Thus a visual presentation of the wordnet graph should be a basis for a wordnet editing system.

From a formal point of view, there are not many restrictions on the shape of the wordnet graphs. However, the semantics of the relations reveals two basic groups of wordnet relations: relations expressing some aspects of hierarchy (e.g. hypernymy/hyponymy, type/instance) and other relations (e.g. holo/meronymy). The former defines some levels: synsets located at the upper levels are more general, those on the lower – more specific. The latter group does not show any preference concerning the location of elements belonging to one link (a graph arc) on the screen.

In many systems, a wordnet graph is visualised in way following the radial scheme, i.e. for a synset in focus its nearest neighbours are presented around it in equal distance, e.g. (Fišer and Novak, 2011, Pedersen et al., 2012) or the system tries to cover equally the whole area of the screen. In both cases, the important characteristic features of the hierarchical relations are lost together with the information about the hypernymic paths and top synsets which is crucial for the wordnet editors. The wordnet graph cannot be also presented as a tree, because, firstly, the majority of its relations do not form a tree, and secondly, truly hierarchical relations would be visually lost in such a presentation with a significant loss the information for the editors. In order to avoid drawbacks of both basic presentation paradigms, an unique combination of the radial and tree-like presentation was proposed for WordnetLoom. Structure relations are presented along the vertical dimension, while other relations are presented radially around synsets, but in a way limited to horizontal zone of limited height centred on a given synset (i.e. only two sectors are used for radial presentation for each synset). The proposed visualisation scheme is illustrated in Fig. 1.

In Fig. 1, the octagonals represent synsets, *P 2.3* and *E 3.1* labels – wordnets, navy blue triangles can be clicked to unfold hidden branches, red to fold those shown. If a very large number of links for a synset and presentation direction (top/down, left/right), exceeds a threshold, then the rest is hidden in the green circle symbol and can be 'taken out' by user clicking it. The threshold, categorisation of particular relation types as vertical or horizontal, as well as link labels and colours used are defined in the WordnetLoom set-up file.

Division of relations into synset and lexical relations is orthogonal to the previous one. Moreover, lexical relations are linked directly to LUs as graph nodes. In order to visualise lexical relations and synset relations on the same screen, it would be necessary to present two inter-connected graphs, in fact, namely, the graphs of synsets and LUs. What is worse, a synset can be connected to a number of LUs on average. Thus, it would be too much information for one screen to present both graphs in the same time. Such a design of the screen was evaluated by linguists as too much cluttered to be useful. Thus, only synset graph is visually presented, and for a synset in focus its LUs are presented in the middle-right panel, see Fig. 1, and the relations of the selected LU are textually presented in the bottom-right panel.

The largest synsets can include even more than 20 LUs, but the average size is much smaller, e.g. less than 2 in plWordNet. However, the initial tests of the visualisation showed that when the number of the presented synsets on the screen approaches 10, it starts to be perceived as cluttered, when all synset members are visible inside the synset symbols. A kind of dynamic adaptation of the number of synset members presented would be an unnecessary complication (it depends also on synset sizes). So, finally, only one synset member, the first LU from a synset, is presented as its representative, the rest is presented in the middle-right panel. Its different sub-panels give access to the attributes of the given synset. For a LU selected

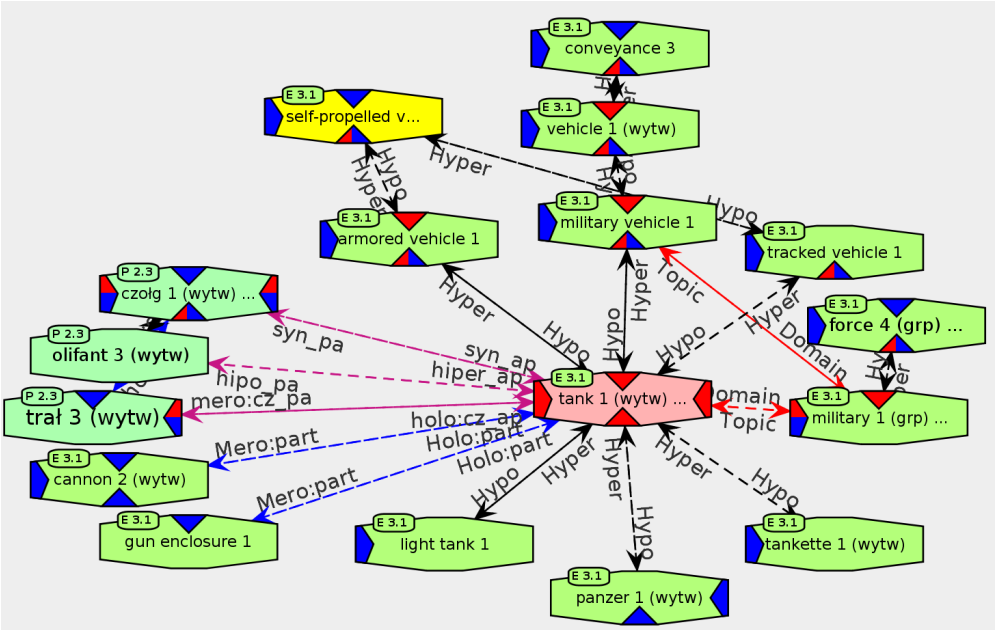

Figure 1: An example of the visualisation scheme.

among the synset members, its list of the lexical relations in shown in the bottom-right panel. Attributes of the synset LUs can be inspected with the context menu (right click).

Users can dynamically open several visualisation panels, e.g. to present search results, compare structures, create links between distant sub-trees or synsets from two wordnets etc.

5 Architecture

Contrary to several other systems, e.g. *DEBVisDic*, WordnetLoom 1.0 was written in Java, and the same approach was followed in the version 2.0. Since Java was quite stable, WordnetLoom 1.0 could be easily installed by non-technical users (by simply unpacking files including the *jar* file). The Java-based solution was free of the problems related to the changing versions of web-browsers (like plug-ins do have), and Java provided more flexibility in the implementation in contrast to the script languages used in the Web.

The construction of WordnetLoom 1.0 was initiated in 2005 as a client-server application with a direct connection to the database. The applied trigger mechanism allowed to encapsulate the whole system with a change control mechanism on selected database tables. WordnetLoom 1.0 supported the Polish language only. The lack of dynamic dictionaries made it difficult to expand and every change in the database required redistribution of the application. At the same time, several supporting tools based on the central wordnet database such as the monitoring system and statistics, API REST, mobile application and web application were created.

In order to adapt the system to new functionalities and other wordnets, we changed its architecture and enhanced the role of the central module which is shared between peripheral applications to increase maintainability of the whole system. In the new three-layer architecture, presented in Fig. 3 an additional, intermediate layer – a *service layer* – was introduced. This layer encompasses now the entire business logic code responsible for CRUD[3] operations and validation. The trigger mechanism has been replaced by the *Hibernate Envars*[4] module allowing the easier undoing of changes. The new schema migration mechanism has been introduced and the application has been secured by mechanisms provided with the *Wildfly Server*. The database schema itself has been rebuilt to be similar to an UBY-LMF structure (Gurevych et al., 2012), new tables have been added, allowing for dynamic construction of dictionaries and a localization mechanism.

WordnetLoom's thick client was transformed into a *thin client* model where all business logic

[3]The CRUD cycle describes the elemental functions of a persistent database. CRUD stands for Create, Read, Update and Delete

[4]`http://hibernate.org/orm/envers`

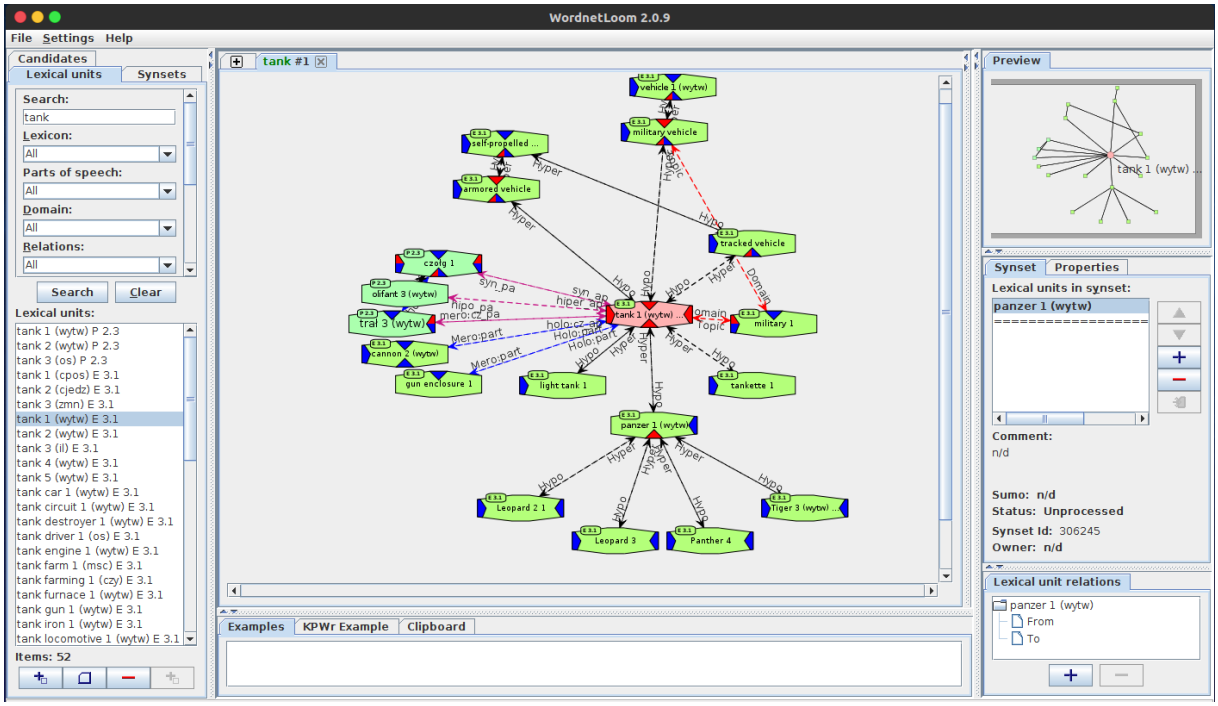

Figure 2: An example of the general layout of the visualisation perspective of WordnetLoom 2.0.

was transferred to the EJB[5] service module. Some of the UI perspectives for data presentation in tabular manner were removed from the client, while the graphical visualization perspective became the main workspace with all functionality accessible, especially the searching facilities have been expanded in this perspective to fully encompass both synsets and LUs[6]. The language of UI can be changed in any moment. The user can also choose which lexicons, mostly wordnets, they want to work with in the given moment. In the previous version the user was working with all lexicons at the same time. The constantly increasing number of relations, both within individual lexicons and across them obscured the graph visualization window. In order to make work more efficient, a functionality to hide selected lexicons was introduced.

The key business logic module has been implemented as an *EJB module* so that it can be a common unitary element. It offers API access to the data layer and, at the same time, has a common extensible validation module aimed at preventing the establishment of wrong relations and thus forcing the correctness of data input. The key tables are audited and, in addition, a special table contains a register of all operations carried out. A very important feature of our new version of the system is the fact that each element belongs to a certain lexicon, that gives the possibility to expand with new collections of elements.

The server is based on *MySql 5.7* database management system [7] and *Wildfly 10.1.0*[8]. For the construction of the system we have used also the following frameworks: *Java EE 7* enterprise edition platform, *JPA 2.+* [9] (*Hibernate 5+* [10]), *EJB 3.1*[11], *JAX-RS* [12], *JSF* [13] (*PrimeFaces*[14]), *Java Swing* [15] and *JUNG 2* [16].

[5] EJB is a server-side software component that encapsulates business logic of an application. The EJB specification is a subset of the Java EE specification

[6] In the visualisation perspective of WordnetLoom 1.0 searching was possible only for LUs and lexical relations.

[7] https://www.mysql.com

[8] Used as EJB web container which provide a runtime environment for web related components, e.g. computer security, Java servlet lifecycle management, transaction processing, and other web services. http://wildfly.org

[9] Java Persistence API. Part of Java EE 7 Specification

[10] http://hibernate.org

[11] http://download.oracle.com/otndocs/jcp/ejb-3.1-pfd-oth-JSpec

[12] JAX-RS: Java API for RESTful Web Services is a Java programming specification that provides support in creating web services according to the Representational State Transfer (REST) architectural pattern. Part of Java EE 7 Specification

[13] JavaServer Faces (JSF) is a Java specification for building component-based user interfaces for web applications

[14] https://www.primefaces.org

[15] GUI widget toolkit for Java provides API for building user interfaces

[16] JUNG — the Java Universal Network/Graph Framework, http://jung.sourceforge.net

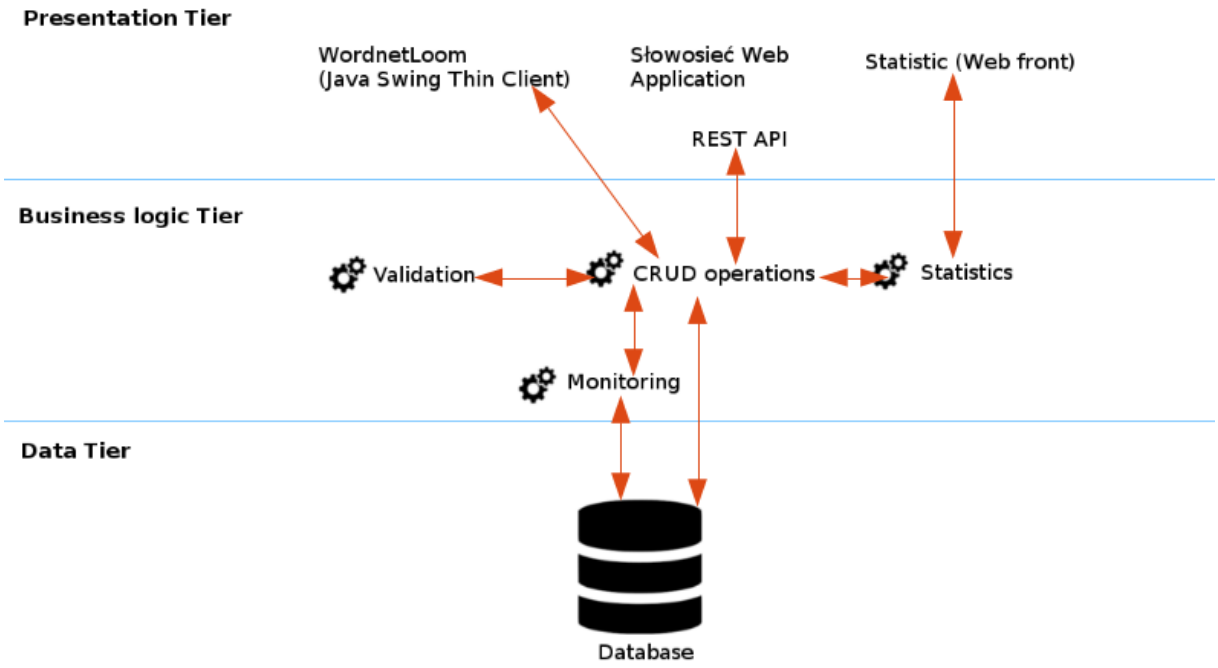

Figure 3: WordnetLoom 2.0 architecture.

6 Extensions and Applications

The architecture of the version 2.0 has been significantly improved in comparison to the previous one, but WordnetLoom has been used for more than 10 years for plWordNet editing (resulting in ≈200k lemmas, ≈300k LUs and ≈200k inter-lingual mappings processed), as well the new version has become a basis for system's adaptations to other wordnets, e.g. a Portuguese wordnet.

6.1 plWordNet Development

As inter-lingual relations are synset relations, but between synsets in different languages, subgraphs for plWordNet and Princeton WordNet should be presented on the same screen. In the new version we added possibility to work on any number of wordnets for any number of languages. Thus it became necessary to introduce labels representing wordnets (defined in set-up) that are attached to synset symbols. Moreover, searching can be limited to elements of a specified wordnet.

Many improvements requested by users were introduced. In the visualisation perspective, in the bottom-right panel of lexical relations double click on the target of relation, a LU, opens a new graph panel with the synset which this LU belongs to.

Every LU and synset is described by additional, meta-attribute of status with the following values: *not processed* (default value), *error*, *verified*, *new*, *partially processed* and *added sense*. Editors can also provide comments to the status, especially important for *error* and *partially processed* statuses, as an explanation of the error, or missing actions, respectively. The status *not processed* marks the material introduced earlier, while *new* signals newly added element especially requiring verification. According to the plWordNet work procedure editors are assigned packages of lemmas, cf Sec. 4, and are obliged to identify and add all LUs for each lemma. However, during their work they may need introduction of LUs for other lemmas than assigned to them, e.g., to add a relation link describing one of the assigned lemmas. In such cases a linguist marks the introduced new LUs and synsets with the *added sense* status that means that some other senses of the same lemma may be lacking. The system of statuses is defined in the database, can be further expanded and supports the management of the linguistic team.

In WordnetLoom 1.0, verb aspect was implicitly expressed by the aspectual relations. In order to facilitate searching and diagnostic procedures, aspect attribute has been added to verbs. Search function was also expanded to cover all attributes, e.g., synset identifiers that are automatically assigned and are not manually edited, but visible in the results of WSD. The search results

can be downloaded in CSV format useful for co-ordinators and plWordNet users.

Diagnostics was also improved by adding PoS tags to variables in substitution tests in the relation definitions stored in the database[17]. This PoS specification allows for automated controlling of the correctness of the links that are considered to be added, but also already present in the database.

The introduced easier expansion of the database and UI allows for adding new types of lexicographic files and annotation with semantic domains. The former facilitates wordnet editing (e.g. the extension includes verb classes used in plWordNet), while the latter supports applications. The domains are based on WordNet Domains (Bentivogli et al., 2004), but we plan to manually modify and expand this classification.

6.2 Portuguese Wordnet

As WordnetLoom is getting consolidated, it can be used to help the construction of wordnets other than just plWordNet. This is what is happening with the MultiWordnet of Portuguese, a quality wordnet for Portuguese (Branco et al., 2009).

This Portuguese wordnet is a project started in 2004 as a branch of Multi-WordNet (Pianta et al., 2002), which until now gathered seven different languages (English, Hebrew, Italian, Latin, Portuguese, Romanian and Spanish), and was one of the first consistent initiatives pursuing the goal of establishing a multilingual wordnet that remains open for further languages. The wordnets in these languages, were transitively aligned with each other by resorting to its alignment to Princeton WordNet, whose format all are following, and thus having English as the pivot language.

This pilot application of WordnetLoom to a different wordnet is providing an important testbed to assess its generality, to find aspects where it can be enhanced, and also to check its technical fitness. For instance, there have been a number of usability enhancements whose need emerged by having new users effectively using this application under different conditions and for a different language, thus stretching its usability requirements. A number of technical improvements have been also motivated in this context of extending the cooperative usage of WordnetLoom to further users.

The outcome of this process and key lessons learned with it are reported in this section.

6.2.1 Enhancing WordNet Content

When creating a quality WordNet for a given language, differences among its language variants should be taken into account and be duly recorded. The differences to be registered can be just superficial: the same word may have different spellings in different variants. Or they may be more substantial: a given concept may be expressed by the same words in different variants, or different variants may resort to different words.

Portuguese is the third European language in number of speakers worldwide. It is the official or national language of several countries and territories in four continents, including Portugal and Brazil. While all speakers of Portuguese can easily communicate, this language have a number of variants. In this context, the Portuguese wordnet has synsets that includes words that belong to only one variant. A word in a synset that belongs to all language variants receive no special marking. A word that belongs to one variant but not to others should be registered as expressing that concept *in that variant* (in addition to being included in that synset). Currently, the Portuguese WordNet covers both the European (spoken in Portugal) and American (spoken in Brazil) variant.

This need resulted in a contribution to enhance wordnet's content with which WordnetLoom can cope. There is now a new field by means of which word forms can be associated to one or other variant, or to none, in which that indicates that a word form is common to all variants.

Portuguese WordNet includes the mapping of synsets into concepts in SUMO ontology (Niles and Pease, 2001). A new field in the WordnetLoom database was introduced in order to represent this type of information, that can be also useful for plWordNet for which its mapping to SUMO was stored so far as a separate resource.

6.2.2 Enhancing Lexicographers Work

The quality Portuguese WordNet is being constructed under the semi-automatic methodology of MultiWordnet. After a first projection of tentative synsets and their relations obtained on the basis of Princeton WordNet and bilingual dictionaries, these synsets are adjusted and confirmed by human lexicographers.

In the initial version of WordnetLoom which the Portuguese WordNet started being edited with,

[17] A dedicated window for editing the definitions is accessible only for the co-ordinators of the linguistic team.

there were just a few search options, namely by word or POS. As the lexicographic labour was proceeding, we realized that it would be faster and easier, if it would be possible to keep track of synsets and senses that have been already checked before, to not check them again, wasting useless effort by the lexicographers. This could be done if there was an identifier for a sense or a synset status, indicating whether it had been checked.

As we discussed in Sec. 6.1, this need resulted in another contribution to enhance the versatility of WordnetLoom to support lexicographers work. In its current version, the users are provided with additional search options based on these statuses, so that they can retrieve only synsets that are yet to be checked or synsets whose edition are finalised.

6.2.3 Enhancing Format Compatibility

There is a main difference between the format of Princeton WordNet and the wordnet designed and developed for plWordNet. The latter is sense-based while the former is synset-based. This creates the need for new information (i.e. data-types and data-relations) in the database. Some instances are "sense relations" and "sense to synset connections". WordnetLoom was originally designed to be compatible with the Polish wordnet. Hence, before it could be employed, the data of the Portuguese WordNet – in Princeton WordNet format – had to be migrated to the plWordNet format. A converter[18] from the Princeton WordNet format to the WordnetLoom format was developed by the Portuguese team. It can now be reused to convert any wordnet in a format compatible, or convertible to the Princeton WordNet format (a *de facto* standard), into the WordnetLoom format, thus greatly enlarging the number of possible wordnets that now can be uploaded into and edited by WordnetLoom.

This step was rather challenging and demanding as there are substantial differences in the organisation of both representations, although facilitated by higher expressiveness of the plWordNet format (e.g. it allows for assigning a set of attributes to both: synsets and lexical units).

The fact that WordnetLoom is under continuous improvement is a positive aspect as teams can ask for changes according to their needs. These changes might be kept as useful suggestions for the final version of WordnetLoom or could be kept local for that specific team.

[18] the link temporarily anonymized for submissions

6.2.4 Technical Enhancements

One very important step in developing any system is its testing and debugging. The work on the Portuguese wordnet is part of the former, with the reporting to the central development team about the issues encountered while working with WordnetLoom, thus being contributing to its technical enhancement.

Three examples of more salient issues that were reported, and that were then solved, are indicated here. (1) Problems with multiple senses of a word. This problem occurred for ambiguous words where one of their senses already existed in WordnetLoom database. When adding a new sense, the UI raised a warning about repetitive entry even though it was actually the same word but in a new synset. (2) Some dis-functionality in the UI. There were cases that the buttons did not function correctly or clicking them caused exceptions that forced to restart the client. (3) Difficulties with setting up the server and client. Problems can be categorized into (i) incompatibility of Java versions and Java basic set-ups; (ii) local settings for both the server and each of the clients; and (iii) issues with running Java-Web-Start. The first two of these types of problems are already solved and the resolution of the third category is under progress.

7 Conclusions and Further Works

WordnetLoom incorporates more than 10 years of experience in the development of a very large wordnet by many linguists on daily basis and this rich experience has become a good basis for the development of new version improved with respect to both: technology and functionality. The system is open[19]. Its most unique feature is direct work on the visually presented wordnet graph, as well as support for simultaneous editing and interlinking of many wordnets (editors see the multilingual structures they are going to map).

WordnetLoom adaptation to the needs of the Portuguese Wordnet developed according to completely different method than plWordNet showed system's potential, and paved way for its adaptations to other resources and tasks. We plan to integrate both variants and continue collaborative development of the system.

[19] `https://github.com/CLARIN-PL/WordnetLoom`

Acknowledgment

Work partially supported by the Polish Ministry of Education and Science, Project CLARIN-PL, and the Portuguese Ministry of Higher Education, Science and technology, by the Infrastructure for the Science and Technology of the Portuguese Language (CLARIN Língua Portuguesa), and by the ANI/3279/2016 grant.

References

Luisa Bentivogli, Pamela Forner, Bernardo Magnini, and Emanuele Pianta. Revising wordnet domains hierarchy: Semantics, coverage, and balancing. In *COLING 2004 Workshop on "Multilingual Linguistic Resources", Geneva, Switzerland, August 28*, pages 101–108. ACL, 2004. URL http://wndomains.fbk.eu/publications/Coling-04-ws-WDH.pdf.

António Branco, Francisco Costa, Eduardo Ferreira, Pedro Martins, Filipe Nunes, Joao Silva, and Sara Silveira. Lx-center: a center of online linguistic services. In *Proceedings of the ACL-IJCNLP 2009 Software Demonstrations*, pages 5–8. Association for Computational Linguistics, 2009.

Nicoletta Calzolari et al., editor. *Proc. Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, 2012. European Language Resources Association. ISBN 978-2-9517408-7-7.

Christiane Fellbaum, editor. *WordNet – An Electronic Lexical Database*. The MIT Press, 1998.

Darja Fišer and Benoît Sagot. Constructing a poor man's wordnet in a resource-rich world. *Language Resource and Evaluation*, 49(3): 601–635, September 2015. ISSN 1574-020X. doi: 10.1007/s10579-015-9295-6. URL http://dx.doi.org/10.1007/s10579-015-9295-6.

Darja Fišer and Jernej Novak. Visualizing sloWNet. In *Proceedings of eLex*, pages 76–82, 2011. URL http://elex2011.trojina.si/Vsebine/proceedings/eLex2011-9.pdf.

I. Gurevych, J. Eckle-Kohler, S. Hartmann, M. Matuschek, Ch.M. Meyer, and Ch. Wirth. UBY – a large-scale unified lexical-semantic resource based on LMF. In *Proceedings of EACL 2012*. ACL, 2012.

Verena Henrich and Erhard Hinrichs. GernEdiT – the GermaNet editing tool. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May 2010. European Language Resources Association (ELRA). ISBN 2-9517408-6-7.

Ales Horak, Adam Rambousek, and Piek Vossen. A distributed database system for developing ontological and lexical resources in harmony. In *In 9th International Conference on Intelligent Text Processing and Computational Linguistics*, pages 1–15, Berlin, 2008. Springer.

Aleš Horák and Pavel Smrž. New features of wordnet editor VisDic. *Romanian Journal of Information Science and Technology*, 7(1–2): 201–213, 2004.

Aleš Horák, Karel Pala, Adam Rambousek, and Martin Povolný. DEBVisDic — first version of new client-server wordnet browsing and editing tool. In *Proceedings of the Third International WordNet Conference — GWC 2006*, pages 325–328. Masaryk University, 2006.

Hitoshi Isahara and Kyoko Kanzaki, editors. *Advances in Natural Language Processing: Proc. 8th International Conference on NLP, JapTAL*, volume 7614 of *Lecture Notes in Artificial Intelligence*, 2012. Springer-Verlag.

Marek Maziarz, Maciej Piasecki, Ewa Rudnicka, and Stan Szpakowicz. Beyond the transfer-and-merge wordnet construction: plWordNet and a comparison with WordNet. In Ruslan Mitkov, Galia Angelova, and Kalina Boncheva, editors, *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, pages 443–452, Hissar, Bulgaria, September 2013. INCOMA Ltd. Shoumen, BULGARIA. URL http://aclweb.org/anthology/R13-1058. **ACL Anthology**.

Luís Morgado da Costa and Francis Bond. Omwedit - the integrated open multilingual wordnet editing system. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pages 73–78. Association for Computational

Linguistics and The Asian Federation of Natural Language Processing, 2015. doi: 10.3115/v1/P15-4013. URL http://aclanthology.coli.uni-saarland.de/pdf/P/P15/P15-4013.pdf.

Ian Niles and Adam Pease. Towards a standard upper ontology. In *Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001*, pages 2–9. ACM, 2001.

B.S. Pedersen, L. Borin, M. Forsberg, K. Lindén, H. Orav, and E. Rögnvalssson. Linking and validating nordic and baltic wordnets – a multilingual action in META-NORD. In *Proceedings of 6th International Global Wordnet Conference*, pages 254–260., Matsue, Japan., 2012.

Emanuele Pianta, Luisa Bentivogli, and Christian Girardi. Multiwordnet: developing an aligned multilingual database. 1st gwc. *India, January*, 2002.

Maciej Piasecki, Michał Marcińczuk, Radosław Ramocki, and Marek Maziarz. WordNetLoom: a WordNet development system integrating form-based and graph-based perspectives. *International Journal of Data Mining, Modelling and Management*, 5(3):210–232, 2013. doi: 10.1504/IJDMMM.2013.055861.

Borislav Rizov. Hydra: A software system for wordnet. In Heili Orav, Christiane Fellbaum, and Piek Vossen, editors, *Proceedings of the Seventh Global Wordnet Conference*, pages 142–147, Tartu, Estonia, 2014. URL http://www.aclweb.org/anthology/W14-0119.