Using Ontologies for Semi-automatic Linking VerbaLex with FrameNet

Jiří Materna, Karel Pala

Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic
{xmaterna,pala}@fi.muni.cz

Abstract

This work presents a method of linking verbs and their valency frames in VerbaLex database developed at the Centre for NLP at the Faculty of Informatics Masaryk University to the frames in Berkeley FrameNet. While completely manual work may take a long time, the proposed semi-automatic approach requires a smaller amount of human effort to reach sufficient results. The method of linking VerbaLex frames to FrameNet frames consists of two phases. The goal of the first one is to find an appropriate FrameNet frame for each frame in VerbaLex. The second phase includes assigning FrameNet frame elements to the deep semantic roles in VerbaLex. In this work main emphasis is put on the exploitation of ontologies behind VerbaLex and FrameNet. Especially, the method of linking FrameNet frame elements with VerbaLex semantic roles is built using the information provided by the ontology of semantic types in FrameNet. Based on the proposed technique, a semi-automatic linking tool has been developed. By linking FrameNet to VerbaLex, we are able to find a non-trivial subset of the interlingual FrameNet frames (including their frame-to-frame relations), which could be used as a core for building FrameNet in Czech.

1. Introduction

The method of linking Czech verb valency frames contained in lexical database VerbaLex developed at the Centre for NLP at the Faculty of Informatics, Masaryk University with the frames in Berkeley FrameNet is aimed at building the core of Czech FrameNet. The process of linking the VerbaLex frames with the FrameNet frames consists of two phases. In the first phase we find an appropriate FrameNet frame for each frame in VerbaLex. The second phase includes assigning FrameNet frame elements to the deep semantic roles in VerbaLex.

First, a few words should be said about FrameNet (Ruppenhofer et al., 2006). Main goal of this project is to build a large lexical resource of English based on frame semantics and supported by corpus evidence. Frame semantics has been proposed and worked out by Charles J. Fillmore, see e.g. (Fillmore, 1976) or (Fillmore, 1982). Throughout the world, there are several projects built on the same idea as the Berkeley FrameNet – Saarbrücken team has been developing German frame-based electronic lexicon SALSA (Burchardt et al., 2006b), Spanish team has been working on Spanish FrameNet (Subirats and Petruck, 2003) etc. For Czech some experiments have been carried out by the J. Hajič's group in Prague (Benešová et al., 2008), taking advantage of the lexical database Vallex built in this group. However, at present, Czech FrameNet as such has not been worked out yet. It should be remarked that one of the pitfalls of FrameNet is its relatively low recall and the fact that its building reminds a never ending story. If we start with a reasonably determined list of verbs such as VerbaLex the task can get a more distinct outline.

To cover this gap at least partially we present a method of linking Czech verb valency database VerbaLex (Hlaváčková and Horák, 2006) frames to FrameNet frames in order to build a core of Czech FrameNet. While lexical units in FrameNet consist of verbs, nouns, adjectives, adverbs and prepositions, VerbaLex captures verb valencies of approximately 10,500 Czech verbs. VerbaLex takes advantage of the Czech and Princeton WordNet (CZWN and PWN) notation (Pala and Smrž, 2004), working with synsets and using the same verb sense labeling. Also hierarchical structure (hypero/hyponymy trees) of both WordNets is exploited. By linking VerbaLex valency frames with the FrameNet ones we obtain a mapping of the frame elements in FrameNet to the semantic roles in VerbaLex. The mapping is based on the similarity between the set of frame element realisations in FrameNet and the hypernyms of possible verb valency realisations in VerbaLex, as well as on the utilization of the semantic type hierarchy in FrameNet.

One of the aims of this work is to reduce manual work that is expensive and time consuming using the proposed semi-automatic approach, which requires a smaller amount of human effort. By linking FrameNet with VerbaLex, we are able to find a non-trivial subset of the interlingual FrameNet frames (including their frame-to-frame relations), which we are going to use as a core for building Czech FrameNet.

2. Frame semantics and FrameNet

The central idea of the frame semantics (Fillmore, 1982) is that word meaning is described in a relation to *semantic frame*, which consists of a target *lexical unit* (pairing of a word with a sense), *frame elements* (its semantic arguments) and relations between them.

FrameNet is a project in which the information about the linked semantic and syntactic properties of the English words is extracted from a large electronic text corpora, using both manual and automatic procedures. The information about words and their properties is stored in an electronic lexical database. Possible syntactic realizations of the semantic roles associated with a frame are exemplified in the annotated FrameNet corpus.

2.1. Semantic Frames

A semantic frame is defined as a script-like conceptual structure that describes a particular type of situation, object or event together with its participants and properties

(Fillmore, 1977). The semantic links of the lexical unit are expressed in terms of the kinds of entities that can participate in frames of the type evoked by the lexical unit.
Lexical unit is a pairing of a word with a meaning. Typically, each sense of a polysemous word belongs to a different semantic frame. For example, the *Commerce_sell* frame describes a situation in which a `seller` sells some `goods` to a `buyer`, and is evoked by lexical units such as *auction*, *retail*, *retailer*, *sale*, *sell*, etc. The semantic participants are called frame elements.

2.2. Frame Elements

Frame elements bear some resemblance to the argument variables used in first-order predicate logic, but display important differences coming from the fact that frames are much more complex than logical predicates (Fillmore et al., 2003). In the example above, the frame elements include `Seller`, `Goods`, `Buyer`, etc.
FrameNet distinguishes three types of frame elements – *core* FEs (the presence of such FEs is necessary to satisfy a semantic link of the given frame), *peripheral* FEs (they are not unique for a given frame and can usually occur in any frame, typically expressions of time, place, manner, purpose, attitude, etc.) and *extra-thematic* FEs which have no direct relation to the situation identified by the frame, but add new information, often showing how the event represented by one frame is a part of the event involving another frame.

2.3. FrameNet relations

In order to make FrameNet more comprehensive several frame-to-frame relations are introduced. Each frame relation in the FrameNet data is an asymmetric relation between two frames, where one frame (the more abstract) can be called the *super-frame* and another (the less abstract) can be called the *sub-frame*. The set of the most important relations comprises the following ones:

- **Inheritance** – the strongest relation between frames, corresponding to IS-A relation in ontologies. Each sub-frame must inherit all properties of its super-frame and share all FEs.
- **Subframe** – some frames are complex in the sense that they refer to sequences of states and transitions, each of them can be separately described as an individual frame. The separate frames (sub-frames) are related to the complex frames via the Subframe relation.
- **Perspective_on** – this relation indicates the presence of at least two different points-of-view. For example, a neutral *Commerce_goods_transfer* frame has two points-of-view – *Commerce_buy* and *Commerce_sell*. The neutral frame is usually non-lexical (Ruppenhofer et al., 2006).
- **Using** – the super-frame constitutes the background for its sub-frames. Not all attributes of the super-frame must be inherited by the sub-frames. For example, Volubility uses the Communication frame, since Volubility describes a quantification of communication events.

3. VerbaLex

VerbaLex is a lexical database consisting of the valency frames of Czech verbs. It has been developed in the Centre for Natural Language Processing at the Faculty of Informatics Masaryk University. Basic units (entries) in VerbaLex consist of verb lemmata grouped in synsets together with their sense numbers in the standard Princeton WordNet (PWN) notation. For the description of verb valencies we use two level notation consisting of the semantic labels corresponding to semantic roles (semantic values of the verb arguments) and the surface level comprising the information about morphosyntactic properties (cases) of Czech verbs. A valency frame also contains the information about obligatoriness and optionality of its constituents (noun, prepositional and adverbial phrases, dependent clauses). Further, valency frames contain additional information including:

- definition of verb meaning;
- verb ability to create passive form;
- number of meanings for homonymous verbs;
- semantic class a verb belongs to;
- aspect (perfective or imperfective);
- example of verb use;
- types of reflexivity for reflexive verbs.

Combining all this information we obtain Complex Valency Frames (CVFs).
Information about the verb arguments is represented by the complex *semantic roles* whose first part are main labels primarily based on the EuroWordNet (Vossen et al., 1998) first-order and second-order Top Ontology entities arranged in a hierarchical structure. The inventory of these labels is closed and currently contains 33 items (concept labels). Second part includes the collection of the selected lexical units (literals) from the set of EuroWordNet and Balkanet Base Concepts (Pala and Smrž, 2004) with their respective sense numbers. This list of labels is open and currently contains about 1,200 literals. We can view them as quite detailed sub-categorization features. Since they represent nodes in the WordNet hypero/hyponymy trees this notation can be characterized as endogenous. An example of a VerbaLex frame corresponding to the synset OPUSTIT:6 (OPOUŠTĚT), NECHAT:9 (to leave) is shown in figure 1.

4. Method of linking

The process of linking frames is asymmetric, which means that at most one frame from FrameNet is assigned to a frame in VerbaLex. There are several ways of looking for the appropriate frame. The most straightforward solution is to translate verbs belonging to a given synset in VerbaLex from Czech to English and to find their equivalents among lexical units in FrameNet. In this approach, each VerbaLex

```
(1) OPUSTIT:6(OPOUŠTĚT), NECHAT:9
(2) pf: opustit:6 impf: popuštět pf nechat:9
(3) -frame: AG(kdo1;<person:1>;obl)+++VERB+++PAT(koho4;<(person:1)|(animal:1)>;obl)
(4) -synon:
(5) -example: dok: opustila své děti (she left  her children)
(5) -example: dok: opustil svého psa (he left his dog)
(6) -use: prim
(7) -class: leave
```

Figure 1: Example of a VerbaLex frame.

frame is represented by a set of translations of the verb literals from the given synset. As a translation dictionary the electronic Czech-English dictionary, developed at the University of West Bohemia [1], which can be freely downloaded under the GNU/FDL licence, has been used. In order to compare two frames, the Jaccard coefficient [2] is used. The method based on verb translations is shown in figure 2.

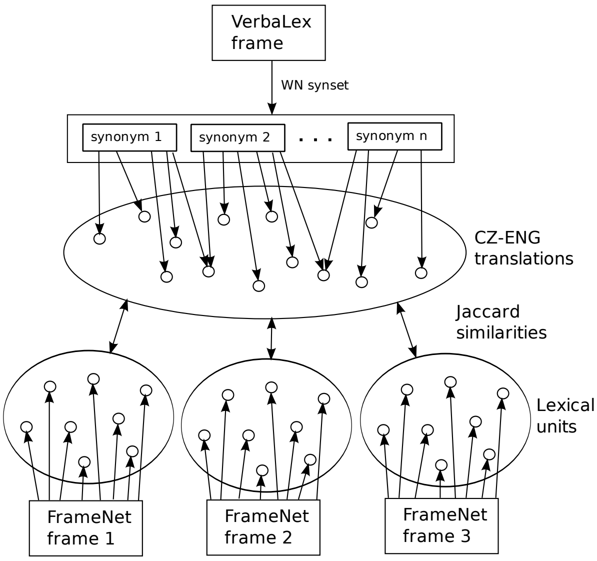

Figure 2: Verb translation approach.

However, there is a problem of finding a correct translation. Another possible solution is to exploit direct links between VerbaLex and Princeton WordNet (Pala and Smrž, 2004; Fellbaum, 1998) and to use WordNet as an inter-language between VerbaLex and FrameNet. To ensure maximum reliability, the linking of PWN verbs onto FrameNet lexical units is based on a WordNet verb–FrameNet dictionary, which has been built manually as a part of the work presented in (Shi and Mihalcea, 2005). This manually built dictionary has high precision but very low recall – it covers 2,651 from 10,500 verbs. This sparseness of the dictionary function can be overcame by the usage of some additional automatic method. In this work, a rule-based system for mediating frame assignment by a “Detour via WordNet” developed by Burchardt et al. (Burchardt et al., 2006a) has been used. The process of linking VerbaLex frames with FrameNet frames based on the inter-language is described in figure 3. Note that multiple frames can be assigned to a single synset if all English translations do not share the same FrameNet frame.

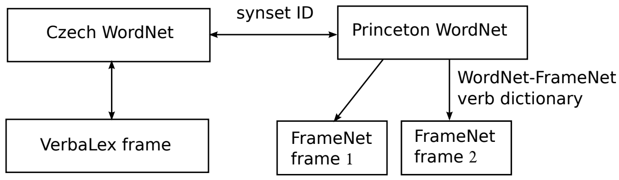

Figure 3: Using WordNet as an inter-language between VerbaLex and FrameNet.

4.1. Assigning verb arguments

An assumption of this task is that the investigated VerbaLex frame is already connected with exactly one appropriate FrameNet frame. Once each VerbaLex frame is connected with a semantic frame from FrameNet, a mapping between semantic roles in VerbaLex and frame elements in FrameNet can be easily identified. Because of the nature of VerbaLex arguments (reflecting combinatory possibilities of verbs), in the case of FrameNet frame, only the *core* frame elements are used.

The linking of semantic roles in VerbaLex with frame elements from FrameNet is based on the most probable pairing. The FrameNet corpus provides a wide range of annotated frame element realizations. Each frame element of the investigated frame is represented by the set of all its realisations in the FrameNet corpus. If the realisation comprises more than one word form all of them are used. In addition, the constituents are normalized and lemmatized using the TreeTagger (Schmid, 1994), developed at the Institute for Natural Language Processing, Stuttgart University.

In this method, semantic roles in VerbaLex are not exemplified sufficiently but a large enough set of examples can be obtained by taking into account a hyponym relation in CZWN. Formally, each verb argument of the investigated frame is represented by the associated synset from PWN and all its hyponyms. The similarity of a frame element and a verb argument in VerbaLex is defined as a Jaccard measure between their representing sets. The method of linking VerbaLex semantic roles with FrameNet frame elements is shown in figure 4.

The goal then is to find the best pairing in terms of the great-

[1] http://slovnik.zcu.cz/

[2] The Jaccard coefficient measures similarity between sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets.

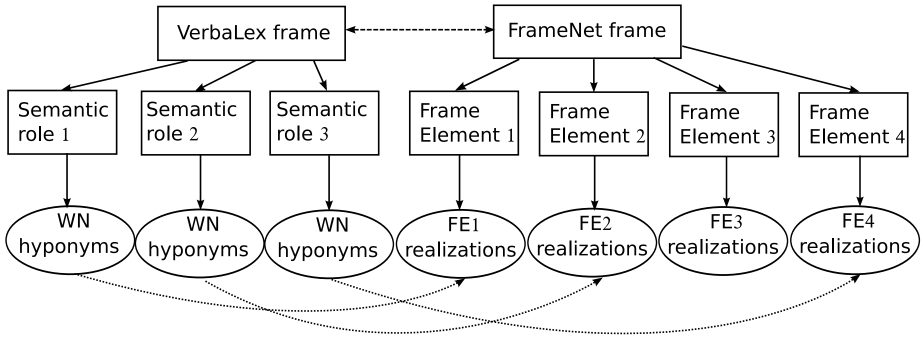

Figure 4: Linking VerbaLex semantic roles with FrameNet frame elements

est sum of their similarities. The key issue of this approach is a definition of the „best pairing". The two following requirements on the pairing are formulated:

- sum of the similarities of all the participated pairs must be maximal;
- each frame element (verb argument) can be connected at most to one verb argument (frame element), i. e. a particular node may not share more than one pair.

The task can be modeled in the graph theory as the most expensive pairing problem of the bipartite graph $G = (E, V_1, V_2, f)$, where V_1 is the set of all frame elements in the FrameNet frame, V_2 is the set of all verb arguments in the VerbaLex frame, $E \subseteq V_1 \times V_2$ is the set of edges, $s : V_1 \cup V_2 \rightarrow 2^W$ is the set of lexical units associated with the vertex $v \in V_1 \cup V_2$, W is the set of all possible lexical units and $f : E \rightarrow \mathbb{R}$ is the weight of an edge $(v_1, v_2) \in E$ and is defined as

$$f((v_1, v_2)) = \frac{|s(v_1) \cap s(v_2)|}{|s(v_1) \cup s(v_2)|}$$

Even though several algorithms performing in polynomial time complexity exist (Cook et al., 1997), the simplest brutal-force algorithm, which searches all possible pairings, has been used. The cardinality of a frame element set is usually up to 5 and the cardinality of a VerbaLex argument set up to 3. For such small graphs, even the brutal-force exponential algorithm can overcome some sophisticated polynomial algorithm.
For a better insight into the problem, an example of linking VerbaLex frames with FrameNet ones is given. Let us consider following synset from CZWN: *prodat:n9, střelit:n2* (to sell something) and a shortened list of simplified VerbaLex frames corresponding to the synset: (a) ***AG*** *(institution:1)* ***VERB ART*** *(goods:1)* ***REC*** *(institution:1)*, (b) ***AG*** *(person:1)* ***VERB ART*** *(goods:1)*, (c) ***AG*** *(person:1)* ***VERB ART*** *(goods:1)* ***REC*** *(person:1)*. All listed frames should be linked to the *Commerce_sell* frame from FrameNet. The core frame elements of the FrameNet frame include *Seller* (linked to *AG (institution:1)* in (a) and to *AG (person:1)* in (b), (c)), *Goods* (linked to *ART (goods:1)* in (a), (b) and (c)) and *Byuer* (linked to *REC (institution:1)* in (a) and to *REC (person:1)* in (c)). In frame (b), there is no equivalent for *Buyer*.

4.2. Ontologies behind FrameNet and VerbaLex

We can say that FrameNet lexical resource comprises two independent ontologies. First of them is the hierarchy of FrameNet frames arranged according to the frame-to-frame relations, using especially the inheritance relation. The second one is the hierarchy of semantic types (e. g. Sentient for the Cognizer FE), which are connected with some general frame elements. While the frame-to-frame relations were exploited in the methods identifying links between FrameNet frames and WordNet synsets (Shi and Mihalcea, 2005; Burchardt et al., 2006a), and subsequently in linking FrameNet frames with VerbaLex frames, the relation between frame elements stood apart from our interest so far.
One of the fundamental differences between FrameNet and VerbaLex consists in the fact that while VerbaLex uses as labels of the verb arguments a fixed set of Complex Semantic Roles (CSRs) for all valency frames, FrameNet defines a unique frame element set for each frame. Nevertheless, there are frame elements shared by two or more frames (e. g. Time). On the other hand, some frame elements of the same name can have different meaning in different frames. To extend information value of the frame elements a hierarchy of 72 semantic types has been introduced. Each frame element may be connected with one or more semantic types. Unfortunately, only approximately 20 % of all frame elements (usually very general) is connected with any of them. Moreover, the hierarchy of semantic types does not correspond to any existing ontology (although it reminds the Top Ontology in EuroWordNet). However, we can mention a project (Scheffczyk et al., 2006) in which the authors linked semantic types and frame elements to SUMO ontology (Niles and Pease, 2001).

4.3. Exploitation of Ontologies in Linking FrameNet Frame Elements with Semantic Roles from VerbaLex

The most problematic part of the process of linking VerbaLex with FrameNet is the connecting VerbaLex verb arguments with appropriate frame elements from FrameNet. The methods of linking frames described above use statistical technique based on comparison of a typical frame element and semantic role fillers, but it can be improved by using the links between semantic types (or frame elements) and SUMO concepts. Once a SUMO concept is linked to each frame element (or most of them), we can use a SUMO-

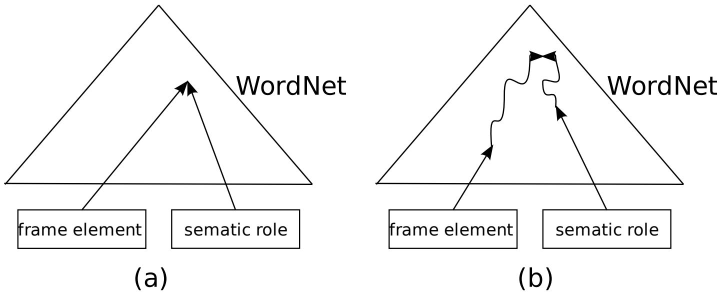

Figure 5: Distance of a VerbaLex semantic role and a FrameNet frame element.

WordNet mapping described in (Niles and Pease, 2003) and get the corresponding WordNet synset. The distance of the frame element and semantic role from VerbaLex is then defined in the following way:

- The linked synset is equal to the synset associated with the VerbaLex semantic role (see fig. 5.a). In this case, the distance is zero.
- The linked synset is not equal to the synset associated with the VerbaLex semantic role (see fig. 5.b). In this case, the distance is defined as the length of the path from one synset to the other using the hypernym relation in WordNet.

The ontology-based method of identifying the nearest FrameNet frame element for a semantic role from VerbaLex is very straightforward solution but suffers from many errors in the SUMO-WordNet mapping as well as the lack of frame elements annotated by a semantic type. That is why the method is used only as a correction of the statistical approach described in section 4.1. The bottleneck of the statistical approach consists especially in ambiguity. The problem arises when two or more edges in the pairing graph have (nearly) equal costs. In that case (on condition that the ontology-based cost is defined), the ontology-based instead of the statistical-based distance is used to decide which edge should be selected.
The combination of the statistical and the ontology-based method increases the coverage as well as accuracy of the system.

5. Evaluation

The first phase describing the mapping of the frames from FrameNet onto VerbaLex is proposed using two different approaches. Both methods return a list of frames with a probability. For the purposes of the future applications and the evaluation, the methods are combined in the following way.
Let $L^A(vf) = (f_1^A, p_1^A) \dots (f_m^A, p_m^A)$ be a list of potential (*frame*, *probability*) pairs for the VerbaLex frame *vf*, acquired by the method based on verb translations (A) and sorted in the increasing order according to their probabilities. Similarly, $L^B(vf) = (f_1^B, p_1^B) \dots (f_n^B, p_n^B)$ be an output of the method based on WordNet (B). The final list of potential (*frame*, *probability*) pairs is computed by merging L^A and L^B. If any frame is a member of both lists, the arithmetic mean of their probabilities is used.

The whole output list can be useful in some application but for the evaluation only the first, most probable item, is used. The method has been evaluated manually by the human annotator on the randomly chosen set of 200 VerbaLex frames.
Main problem of this task are translation errors of two types – translation between Czech WordNet and Princeton WordNet and translation between PWN and FrameNet. In spite of the fact that the sense of a word can relatively easily change, acquired precision of the linking between FrameNet frames and VerbaLex ones is about 84 %. The errors include cases where no frame was assigned at all. The frame is not assigned if either the corresponding frame does not exist or the PWN–FrameNet mapping is incomplete.
The second phase deals with the linking between verb arguments in VerbaLex and frame elements in FrameNet. This task is, in the principle, more complicated than the previous one. Regardless of the wrongly assigned frame elements, there is a problem with ambiguity. For instance, if two or more arguments of a verb in VerbaLex represent persons, they all are usually connected with the same literal from WordNet (*person:1*). The similarities of these arguments with frame elements are then the same and the assignment process is ambiguous (see frames (b) and (c) in examples from section 4.1). The experiments have shown that the error rate of the system is about 24 % and the ambiguity rate about 12 %. The former version of the system without corrections by the ontology-based algorithm has performed with the error rate about 26 % and the ambiguity rate about 19 %.
Two main problems were identified – first, some characteristic sets of the lexical units are not representative enough and second, there is a problem with verb arguments for which the corresponding frame element in FrameNet does not exist. The relatively low accuracy is also caused by the high number of combinations for each frame pair – the average number of arguments in VerbaLex is 2.7 and the average number of frame elements in FrameNet is 14.6.

6. VerbaLex-FrameNet linking tool

Based on the described methods, a semi-automatic VerbaLex–FrameNet linking tool has been developed. The tool is a client-server application, which enables different users to link FrameNet frames and their frame elements to VerbaLex frames simultaneously.

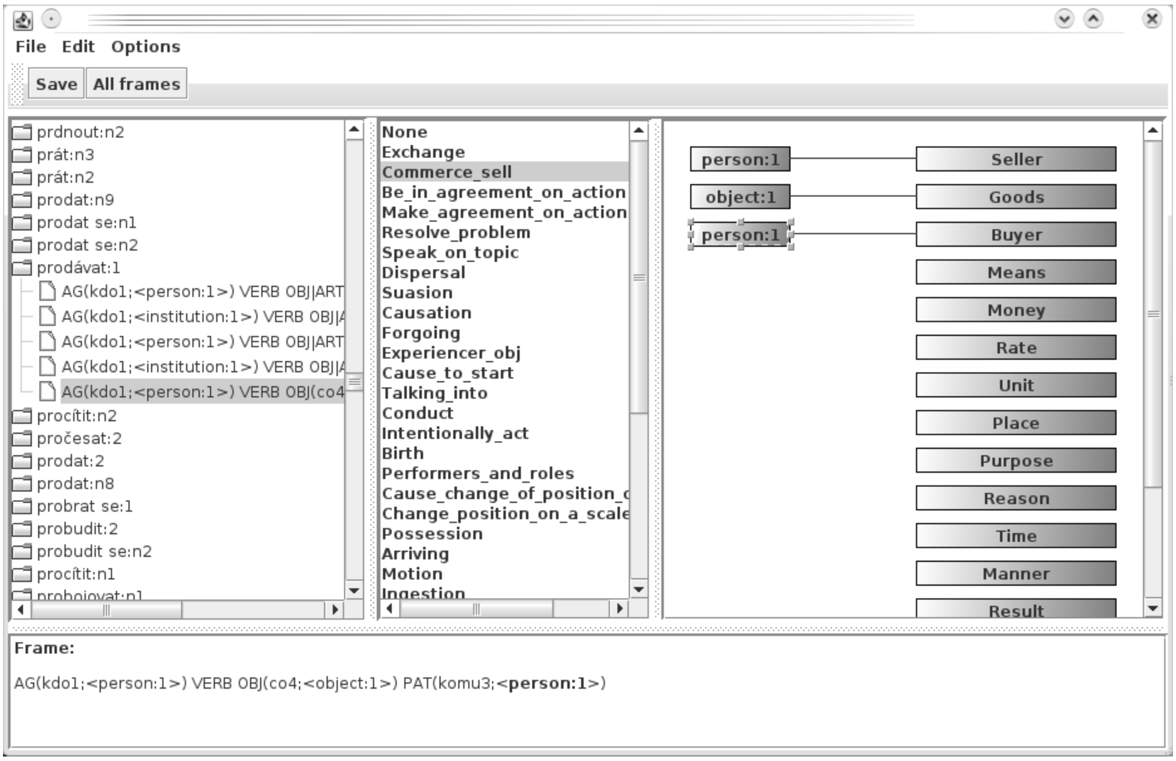

Figure 6: VerbaLex–FrameNet linking tool

Users are first asked by the system to select the most appropriate FrameNet frame for a given VerbaLex Frame. The possible frames are ordered from the most probable to the least probable one using the algorithm from section 4. After selecting it, based on the method described in sections 4.1 and 4.3, a graph representing linking between VerbaLex arguments and FrameNet frame elements is generated. If the user does not agree with the automatically generated pairing, he or she can make it better simply using the mouse. During the whole process one can see a detailed description of any object (frame, synset, frame element, etc.) in the bottom panel by hovering mouse over the it.

A screenshot of the system is shown in figure 6. First column in the screenshot contains frames from VerbaLex, second one contains frames from FrameNet and the third one shows linking between VerbaLex verb arguments and FrameNet frame elements. The figure illustrates annotation of one of the frames belonging to CZWN synset `prodávat:1, obchodovat:2, distribuovat:n3` (to sell something). The VerbaLex frame (placed on the bottom panel) with verb arguments `person:1, object:1` and `person:1` is linked to `Commerce_sell` frame from FrameNet.

7. Conclusions

The presented work describes the development of a semi-automatic tool for linking FrameNet frames to Czech verb valency frames. The method is based on a searching for an appropriate linking between Czech verb valency lexicon VerbaLex and FrameNet. The process of linking consists of two phases – connecting VerbaLex and FrameNet frames and linking verb arguments to frame elements. The described method was evaluated on a randomly chosen set of sample frames. The results showed that the algorithm of linking frames works with the error rate about 16 %, and the verb arguments to frame elements linking algorithm performs with the error rate about 24 % and ambiguity rate about 12 %. These results appear to be acceptable for semi-automatic methods used in the linking tool, which has been developed during the project.

For the future work, the goal is to build a core of Czech FrameNet by means of the developed tool. Such FrameNet based lexicon can be used for information retrieval and searching semantic relations in texts. Also other challenging tasks come into consideration, namely in the area of the Semantic Web.

Acknowledgements

This work has been partly supported by the Ministry of Education of Czech Republic under the projects LC526, NPV II 2C06009 and by the Czech Grant Agency under the project GA 407/07/0679 and project GA P401/10072.

8. References

V. Benešová, M. Lopatková, and K. Hrstková. 2008. Enhancing Czech Valency Lexicon with Semantic Information from FrameNet: The Case of Communication Verbs. In *Proceedings of The First International Conference on Global Interoperability for Language Resources*, pages 18 – 25. Hong Kong.

Aljoscha Burchardt, Katrin Erk, and Anette Frank. 2006a. A WordNet Detour to FrameNet. Technical report, Saarland University and DFKI Saarbrucken.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. 2006b. The Salsa Corpus: A German Corpus Resource for Lexical Semantics. In *The International Conference on Language Resources and Evaluation (LREC)*. Genoa, Italy.

William J. Cook, William H. Cunningham, William R. Pulleyblank, and Alexander Schrijver. 1997. *Combinatorial Optimization*. Wiley-Interscience.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. The MIT Press.

Charles J Fillmore, Christopher R Johnson, and Miriam R L Petruck. 2003. Background to FrameNet. In *International Journal of Lexicography*, pages 235 – 250. Oxford University Press.

Charles J. Fillmore. 1976. Frame Semantics and the Nature of Language. In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, volume 280, pages 20 – 32.

Charles J. Fillmore. 1977. Scenes-and-frames Semantics. In *A. Zampolli, ed. Linguistic Structures Processing*, pages 55 – 81. Amsterdam, North-Holland.

Charles J. Fillmore. 1982. Frame Semantics. In *Linguistics in the Morning Calm*, pages 111 – 137. Seoul, Hanshin Publishing Co.

Dana Hlaváčková and Aleš Horák. 2006. VerbaLex – New Comprehensive Lexicon of Verb Valencies for Czech. In *Computer Treatment of Slavic and East European, Languages Bratislava*, pages 107 – 115. Slovakia: Slovenský národný korpus.

I. Niles and A. Pease. 2001. Towards a Standard Upper Ontology. In *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems*, pages 2 – 9.

I. Niles and A. Pease. 2003. Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In *Proceedings of the 2003 International Conference on Information and Knowledge Engineering*, pages 23 – 26. Las Vegas.

Karel Pala and Pavel Smrž. 2004. Building Czech WordNet. In *Romanian Journal of Information Science and Technology*, pages 79 – 88. Romanian Academy.

J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, C. R. Johnson, and J. Scheffczyk. 2006. FrameNet II: Extended Theory and Practice. http://www.icsi.berkeley.edu/framenet.

Jan Scheffczyk, Adam Pease, and Michael Ellsworth. 2006. Linking FrameNet to the Suggested Upper Merged Ontology. In *Proceeding of the 2006 conference on Formal Ontology in Information Systems*, pages 289 – 300. Baltimore, Maryland.

Helmut Schmid. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *International Conference on New Methods in Language Processing*, pages 44 – 49. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Lei Shi and Rada Mihalcea. 2005. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robus Semantic Parsing. In *Proceedings of Cicling*. Mexico City, Mexico.

Carlos Subirats and Miriam Petruck. 2003. Surprise: Spanish Framenet. In *International Congress of Linguists. Workshop on Frame Semantics*. Prague, Matfyzpress.

Piek Vossen, Graeme Hirst, Horacio Rodriguez, Salvador Climent, Nicoletta Calzolari, Adriana Roventini, Francesca Bertagna, Antonietta Alonge, and Wim Peters. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic*. Kluwer Academic Publisher.