Consistency Evaluation towards Enhancing the Conceptual Representation of Verbs in WordNet

Svetlozara Leseva and **Ivelina Stoyanova**
Department of Computational Linguistics
Institute for Bulgarian Language
Bulgarian Academy of Sciences
zarka@dcl.bas.bg, iva@dcl.bas.bg

Abstract

This paper outlines the process of enhancing the conceptual description of verb synsets in WordNet using FrameNet frames. On the one hand we expand the coverage of the mapping between WordNet and FrameNet, while on the other – we improve the quality of the mapping using a set of consistency checks and verification procedures. The procedures include an automatic identification of potential inconsistencies and imbalanced relations, as well as suggestions for a more precise frame assignment followed by manual validation. We perform an evaluation of the procedures in terms of the quality of the suggestions measured as the potential improvement in precision and coverage, the relevance of the result and the efficiency of the procedure.

Keywords: FrameNet, WordNet, Frame semantics, consistency evaluation

1. Introduction and Motivation

Our work aims at enhancing the conceptual description of verb synsets in WordNet through integrating frame semantics as represented in FrameNet. Below we briefly discuss the resources used in the study and the methodology we apply. We first overview existing alignments between WordNet and FrameNet which make an impact on the adopted methodology for the mapping of the two resources. The mapping proposed herein elaborates on and expands previous alignments by employing the inheritance of conceptual and lexical information (Section 2.2). In addition, we devise a set of consistency checks and frame suggestion procedures in order to further improve the quality and coverage of the resulting resource; these procedures rely on lexical and semantic properties, similarity and the relational structure of WordNet and FrameNet (Section 3). The evaluation of the procedures is based on a manually validated dataset (Section 4) and is presented in Section 5, which is followed by a brief discussion of the results in Section 6. The conclusions are summed up in Section 7.

1.1. Resources

We employ two lexical semantic resources – WordNet and FrameNet. WordNet (Miller, 1995; Fellbaum, 1998) is a large lexical database that represents comprehensively conceptual and lexical knowledge in the form of a network whose nodes denote sets of cognitive synonyms (synsets) interconnected through a number of conceptual-semantic and lexical relations such as synonymy, hypernymy, meronymy, etc. The main relation that determines WordNet's taxonomical structure is the relation of hypernymy.

FrameNet (Baker et al., 1998; Baker, 2008) represents lexical and conceptual knowledge couched in the apparatus of frame semantics. Frames are conceptual structures describing particular types of objects, situations, or events along with their components, called frame elements, or FEs (Baker et al., 1998; Baker and Ruppenhofer, 2002; Ruppenhofer et al., 2016). Depending on their semantic obligatoriness and contribution to the conceptual description, FEs may be core, peripheral or extrathematic(Ruppenhofer et al., 2016); core FEs are the most essential as their configuration makes a frame unique, which is why our focus is on them. Frames are instantiated by lexical units (LUs) which are

included as part of the description of the relevant frame. In addition, frames are related by means of frame-to-frame relations, some of which are discussed below.

The combination of the two resources is expected to strengthen their individual advantages, in particular the great lexical coverage and the branched and rich relational structure of WordNet with the detailed conceptual description of the combinatorial potential of lexical units supplied by FrameNet. The contribution of our research is directed to the expansion of the mapping of the two resources (with the prospect of integrating others, such as VerbNet, Propbank, etc.) to the end of overcoming the sparsity of the overlap between synset members (literals) and LUs in FrameNet.

1.2. Methodology

The work presented here is a continuation of our previous work on enhancing the conceptual description of verbs in WordNet that goes in two directions – expanding the coverage and improving the quality of the mappings (Leseva and Stoyanova, 2019; Stoyanova and Leseva, 2019). To this end, we propose a set of methods for expanding the mapping between the resources based on the relations of inheritance (cf. Section 2.2), which are further enhanced by means of automatic procedures for validation and improvement that focus on minimising manual work (cf. Section 3.1).

Most of the procedures that have been proposed rely on: (i) the notion of inheritance and the hierarchical relational structure of FrameNet and WordNet – the internal structure of the two resources is determined to a great extent by the notion of inheritance: in WordNet it is realised by the hypernymy relation, whereas in FrameNet it is represented mainly by the relations of *Inheritance* (strong inheritance) (Ruppenhofer et al., 2016), *Using* (weak inheritance) (Petruck, 2015) , as well as by relations such as *Subframe*, and *Perspective on*, although in a very limited way; (ii) semantic and lexical analysis of the components of the description in FrameNet and WordNet.

2. Mapping WordNet and FrameNet

2.1. Compilation of Existing Mappings between WordNet and FrameNet

Earlier research aimed at maximising the advantages and the richness of the conceptual and lexical information encoded in WordNet, FrameNet and other resources has led to a number of proposals, including the mapping of WordNet, FrameNet and VerbNet by Shi and Mihalcea (2005), the elaboration of WordFrameNet[1] by Laparra and Rigau (2010) and MapNet[2] by Tonelli and Pighin (2009), the implementation of other FrameNet-to-WordNet mappings, e.g. by Ferrandez et al. (2010). More enhanced proposals have been made too, such as Semlink[3] (Palmer, 2009), which unifies WordNet, FrameNet and VerbNet with PropBank, and its follow-up Semlink+ that brings in a mapping to Ontonotes (Palmer et al., 2014).

In the domain of verbs, out of 14,103 verb synsets, only 4,306 (30.5%) have been mapped through finding existing equivalents (using existing mappings) (Leseva and Stoyanova, 2019). A number of consistency checks were also implemented on the result of the initial mapping before the application of the frame assignment procedure described in 2.2. These checks led to improved connectivity between synsets, which in turn made the hypernym-to-hyponym frame assignment more efficient; part of the checks involved correction of already assigned frames so as to avoid the propagation of errors in the course of hypernym-to-hyponym frame assignment (Leseva and Stoyanova, 2019).

2.2. Inheritance-based mapping

After the implementation of the initial mapping and preliminary validation procedures mentioned above, we undertake expansion of the mapping by incorporating procedures aimed at 'digging up' non-explicit information about the frame membership of WordNet literals on the basis of the relational information in the two resources. Along these lines Burchardt et al. (2005) propose the expansion of the inter-resource coverage (mapping WordNet literals to FrameNet frames) by weighing the candidate frames evoked by literals related to a given target literal through certain semantic relations (synonymy, hypernymy,

[1] `http://adimen.si.ehu.es/web/WordFrameNet`
[2] `https://hlt-nlp.fbk.eu/technologies/mapnet`
[3] `https://verbs.colorado.edu/semlink/`

antonymy). Another strategy is to exploit the relational structure of the resources – particularly that of WordNet – by mapping frames to synsets on the basis of the inheritance of conceptual features in hypernym trees, i.e. by assigning frames from hypernyms to hyponyms (Leseva et al., 2018).

We adopt this latter approach and using the initial automatic mapping of 4,306 synsets (cf. Section 2.1), we implemented a procedure of transferring the hypernyms' frames to their hyponyms in the cases where the hyponyms were not directly mapped to FrameNet frames (through the initial mapping). In such a way, we obtained an extended coverage of 13,226 synsets (synsets with an assigned FrameNet frame) out of the total of 14,103 verb synsets (Leseva and Stoyanova, 2020).

The main drawback of the inheritance approach is that especially for deeper level WordNet synsets the inherited frames may be underspecified. Thus, a natural follow-up was to look for ways to discover appropriately specific frames which have already been defined in FrameNet. This is where the FrameNet's relational structure comes into play as it may point to where to look first for probable candidate frames. Most of the procedures that are proposed below rely on the information and the overall relational structure of FrameNet and WordNet, as well as the semantic and lexical analysis of the components of the description in FrameNet and WordNet.

3. Consistency Checks and Procedures to Verify and Enhance the Conceptual Description

The main idea of these procedures is to explore: (i) the lexical mapping of the target synset's literals to lexical units in FrameNet frames that are related (through frame-to-frame relations) to the frame assigned to the target synset from its hypernym (i.e. exploring the vicinity of the frame inherited from the hypernym); (ii) the lexical mapping of the target synset's literals to lexical units in FrameNet frames that are assigned to synsets in the vicinity of the target synset (its hyponyms and sister synsets in particular); (iii) similarity, e.g. the similarity between keywords in WordNet glosses and the definitions of FrameNet lexical units, cf. Section 3.1. The methodology and implementation have been described in detail in Leseva and Stoyanova (2019) and Stoyanova and Leseva (2019).

3.1. Procedures based on Lexical and Semantic Analysis Involving Hierarchical FrameNet Relations

The procedures involve several steps, as described below:

(1) Check whether any of the literals of the target synset appears as a LU in: (a) the frame assigned from the synset's hypernym (to confirm its validity); (b) more specific frames the frame under discussion is linked to by means of any of the considered frame-to-frame inheritance relations (so as to try to find a suitable more specialised frame); (c) the sister frames of the assigned frame (the frames sharing a parent with the one assigned from the synset's hypernym).

Example 1. Synset: eng-30-01900255-v {flutter:3}
Gloss: flap the wings rapidly or fly with flapping movements
Assigned frame from hypernym: Body_movement
Suggested from (1a): Body_movement (LU: flutter)

The synset in Example 1 is assigned the frame Body_movement directly from its hypernym {beat:8, flap:3} 'move with a thrashing motion', which in its own right is assigned this frame through one or more of the automatic mappings described in Section 2.1. The appropriateness of the assignment through inheritance is confirmed by means of Procedure (1a) as the single literal in this synset, *flutter*, is found as a LU in the frame Body_movement.

Example 2 illustrates Procedure (1b). The synset is originally assigned the frame Cause_change from its hypernym {change:1; alter:1; modify:3} 'cause to change; make different; cause a transformation'. A more specific and better matching frame Cause_change_of_strength is suggested by the procedure on the basis of: (i) the fact that the three literals in the synset are found as LUs in this frame; (ii) there is an inheritance relation between the frame assigned from the hypernym and the newly suggested frame.

Example 2. Synset: eng-30-00220869-v {strengthen:1; beef up:1; fortify:1}
Gloss: make strong or stronger

Assigned frame from hypernym: Cause_change
Suggested from (1b): Cause_change_of_strength (LU: strengthen; beef up; fortify)

Example 3 illustrates Procedure (1c). The synset {educate:3, school:2, train:5, cultivate:3, civilize:1, civilise:1} is assigned the frame Cause_to_make_progress from its hypernym, while three of the literals are found as LUs in its sister frame Education_teaching (both Cause_to_make_progress and Education_teaching inherit from the frame Intentionally_affect).

Example 3. Synset: eng-30-02388403-v {educate:3, school:2, train:5, cultivate:3, civilize:1, civilise:1}
Gloss: teach or refine to be discriminative in taste or judgment
Assigned frame from hypernym: Cause_to_make_progress
Suggested from (1c): Education_teaching (LUs: educate, school, train)

To sum up, Procedure 1 aims at establishing whether a synset's literal or literals may be found as a LU/LUs in a substructure of frames related through inheritance or sisterhood to the frame assigned from the hypernym, thus expanding the frame window explored for literal-to-LU lexical match.

(2) Check whether any of the target synset' literals appears as a LU in: (a) any of the frames assigned to its hyponyms; (b) any of the frames assigned to its sister synsets; and (c) any of the frames related to the frames in (a) and/or (b).

Example 4. Synset: eng-30-00097621-v {regenerate:9; revitalize:1}
Gloss: restore strength
Assigned frame from hypernym: Cause_to_make_progress
Suggested frame from (2a): Rejuvenation (LU: revitalize)

In Example 4 the synset {regenerate:9; revitalize:1} is assigned the frame Cause_to_make_progress from its hypernym {better:2, improve:1, amend:2, ameliorate:1, meliorate:1} 'to make better'. A more specific and accurate frame is suggested through Procedure (2a) by virtue of the fact that the frame Rejuvenation, which is assigned to the single hyponym of {regenerate:9; revitalize:1} – {rejuvenate:3} 'make younger or more youthful', contains a LU matching one of the literals of {regenerate:9; revitalize:1}: *revitalize*.

Example 5. Synset: eng-30-00080705-v {nurse:1}
Gloss: try to cure by special care of treatment, of an illness or injury
Assigned frame from existing mapping: Medical_professionals
Suggested frame from (2b): Cure (LU: nurse)

The synset {nurse:1} in Example 5 was originally assigned the frame Medical_professionals (from the initial automatic mapping), which includes nouns denoting medical workers. Through the application of Procedure (2b), we found out that some of the sisters of the synset are assigned the frame Cure (e.g. {massage:2} 'give a massage to' and {insufflate:2} 'treat by blowing a powder or vapor into a bodily cavity') and that the literal *nurse* corresponds to a LU in the same frame. These two facts in conjunction motivate the suggestion of the frame Cure for the synset {nurse:1}.

Example 6. Synset: eng-30-01013230-v {remonstrate:2; point out:3}
Gloss: present and urge reasons in opposition
Assigned frame from hypernym: Telling
Suggested frame from (2c): Judgment_communication (LU: remonstrate)

Example 6 illustrates the application of Procedure (2c) and the resulting suggestion of Judgment_communication as a possible replacement of the frame Telling originally assigned to {remonstrate:2; point out:3} from its hypernym. In this case, some of the sisters of the target synset, such as {announce:3; denote:3} 'make known; make an announcement', are assigned the frame Statement, which is a more general frame related through the relation *Inheritance* ('strong inheritance') to Telling, and through the relation *Using* ('weak inheritance') to Judgment_communication; in addition, the target synset's literal *remonstrate* is found as a LU in the latter frame. The frame assigned from the hypernym and the frame suggested through Procedure (2c) are 'step-sisters' as they are related to the same frame (Statement) through similar, but not identical relations.

The purpose of the procedures subsumed under Procedure 2 is to establish whether a target synset's literal(s) may be found as a LU/LUs in frames assigned to synsets in the WordNet substructure defined by the target synset's children (hyponyms) and/or sisters as well as in the vicinity of such frames.

(3) Check whether any of the synset literals appear as a LU in any other frame in FrameNet.

Example 7. Synset: eng-30-02217266-v {finance:1}
Gloss: obtain or provide money for
Assigned frame from hypernym: Commerce_pay
Suggested from (3): Funding (LU: finance)

The synset in Example 7 is assigned the frame Commerce_pay from its hypernym {pay:1} 'give money, usually in exchange for goods or services', which is close in meaning as financing involves paying. The more appropriate frame Funding is not related through any of the frame-to-frame relations to Commerce_pay and is suggested by virtue of the fact that the verb *finance* is a LU in this frame. As Procedure 3 is based primarily on the lexical correspondence between literals and LUs, it is more reliable for verbs with fewer senses or in cases where the suggested frame is additionally confirmed. The latter is true for the example under discussion: the proposed frame Funding receives support from the fact that two of the hyponyms of {finance:1} – {back:5} 'support financial backing for' and {fund:1} 'convert (short-term floating debt) into long-term debt that bears fixed interest and is represented by bonds' – are LUs in the same frame (hence their mapping to Funding is suggested by the same Procedure 3).

Procedure 3 may be the only available one when the assignment through inheritance has failed and hence no frame has been mapped to the synset from its hypernym (therefore procedures 1b and 1c are not applicable), consider 8 below.

Example 8. Synset: eng-30-02227741-v {abandon:4; give up:5}
Gloss: give up with the intent of never claiming again
Assigned frame from hypernym: none (ROOT synset)
Suggested from (3): Surrendering_possession (LU: give up); Abandonment (LU: abandon); Quitting_a_place (LU: abandon); Activity_stop (LU: abandon)

In this particular example four frames – Surrendering_possession, Abandonment, Quitting_a_place and Activity_stop – are suggested by Procedure 3 based on the literals (*give up, abandon*) found in them. Manual analysis is then performed to select the most appropriate frame by also considering frames related to the ones suggested.

(4) In this step we use keywords (words found in the name of a FrameNet frame, plus their derivatives collected from WordNet through the *eng_derivative* relation) and identify synsets with literals and/or glosses containing these keywords as candidates to be assigned the frame under discussion.

Example 9. Synset: eng-30-00768389-v {talk out of:1}
Gloss: persuade someone not to do something
Assigned frame: Suasion
Confirmed by (4): keyword:persuade (in gloss)

The synset {talk out of:1} is assigned the frame Suasion from its hypernym {dissuade:1, deter:2} 'turn away from by persuasion', which is the appropriate one. While Procedures (1-3) fail to suggest a frame, Procedure (4) confirms the assignment of Suasion through the keyword *persuade* in the gloss of {talk out of:1}, which is a LU in the frame Suasion.

3.2. Similarity-based Procedures

In addition to the procedures described in Section 3.1, we also introduce checks based on the similarity measures between synset glosses in WordNet and LU definitions in FrameNet. Similarity is measured as the degree of overlapping word roots (using stemming) where direct overlaps of words are given a higher score than the overlaps after stemming.

The procedures include:

(1) Direct similarity: In this step we identify candidate frames for a target synset by checking the similarity between its gloss and FrameNet LU definitions (even though there is no lexical correspondence between the synset's literals and the LUs).

Example 10. Synset: eng-30-01399821-v {beetle:3}
Gloss: beat with a beetle
Assigned frame from hypernym: Cause_ harm
Suggested from (1): Cause_ harm
Confirmed by: similarity of the WordNet gloss with LU definitions (bludgeon; cudgel; whip)

In this example {beetle:3} is assigned the frame Cause_ harm from its hypernym {beat:3} 'hit repeatedly'. This suggestion is confirmed by the similarity existing between the gloss of the synset and the definitions of the LUs *bludgeon*, *cudgel* and *whip* in the frame Cause_harm – 'beat with a bludgeon', 'beat with a cudgel', 'beat with a whip'.

(2) Indirect similarity: In this procedure we identify candidate frames for the target synset by checking the similarity between the glosses of the synsets that are derivationally related to it (as well as the glosses of their hypernyms which are their closest semantic generalisations) and FrameNet LU definitions.

In Example 11 the synset {solo:1} was initially mapped to the frame Shaped_part, an assignment originating from an error in an existing mapping between the synset {handle:4; palm:1} 'touch, lift, or hold with the hands' and this frame; the wrong assignment was then transferred in four steps down the tree from hypernyms to hyponyms to {solo:1}. This error was corrected after an appropriate frame, Operate_vehicle, was suggested using the indirect similarity procedure based on the lexical similarity between the gloss of the derivationally related noun {solo:3} and the definition of one of the LUs in the frame Operate_vehicle, *fly.v*. The similarity is calculated on the basis of matching words (in bold) excluding closed class lexemes and auxiliaries and taking into account the length of the glosses. Scores of over 1.0 are considered a strong indicator of similarity between the definitions.

Example 11. Synset: eng-30-01941987-v {solo:1}
Gloss: fly alone, without a co-pilot or passengers
Assigned frame from hypernym: Shaped_part
Derivationally related synset: eng-30-00304729-n {solo:3}
Gloss: a **flight** in which the **aircraft** pilot is unaccompanied
Suggested from (2): Operate_vehicle (1.11)
Confirmed by: similarity between the gloss of {solo:3} and the gloss of the LU fly.v
Gloss of LU fly.v: control the **flight** of (an **aircraft**)

3.3. Ranking the Suggestions

We consider separately the suggested frames that are related to the frame assigned from the hypernym as they are given higher priority over unrelated suggested frames. We give each lexical match a score based on the calculated similarity, then assign an overall cumulative score to each frame and rank the suggestions so that the more likely candidates are analysed first in order to optimise the manual work. In Example 12 the suggestion Motion_noise has been yielded by the LUs *crackle*, *squelch* and *hiss*, and this is why it is ranked higher than Fluidic_motion (yielded only by the LU *hiss*).

Example 12. Synset: eng-30-02069120-v {woosh:1, whoosh:1}
Gloss: move with a sibilant sound
Hypernym: eng-30-01850315-v {move:2, displace:4}
Assigned frame from hypernym: Cause_motion
Suggested frames from the procedures: Motion_noise:crackle (to move making soft sharp repeating sounds 1.17); Motion_noise:squelch (move with such a sound 1.2); Motion_noise:hiss (to move making a sibilant sound as of the letter s 1.12); Self_motion:wriggle (move with wiggling movements 1.2); Fluidic_motion:hiss ((for air) to move producing a sharp sibilant sound 1.29)
Assigned correct frame: Motion_noise

Future work will be directed to the development of a methodology for quantifying relevance and more precise ranking of the suggestions, so that manual work is minimised and an automatic procedure for filtering and selection of suggestions is implemented.

4. Manual Verification of Automatic Frame Suggestions

The output of all the applied procedures is produced as a list for the experts to analyse and possibly confirm the appropriate candidates. Consider Example 13 below. In this particular instance, each of the frames was suggested on the basis of a direct or an indirect derivational or semantic relation between the target synset and another synset – represented by the relevant literals – *terrifying*, *pleasing*, *loathing*, etc., which in their own right have been assigned the suggested frames on the basis of lexical mapping with LUs in the respective frames. A linguist needs to study the list of suggested frames and select an appropriate one if such is available.

Example 13. Synset: eng-30-01813668-v {exult:1, walk on air:1, be on cloud nine:1, jump for joy:1}
Gloss: feel extreme happiness or elation
Hypernym: eng-30-01813884-v {rejoice:1, joy:1}
Assigned frame from hypernym: Feeling
Suggested frames from the procedures: Stimulus_focus: terrifying (1.2); Stimulus_focus:pleasing (1.17); Stimulus_focus:exhilarating (1.17); Emotion_directed:agony (1.25); Emotion_directed:ecstatic (1.5); Emotion_directed:fury (1.25); Experiencer_focused_emotion:loathing (1.4)
Assigned correct frame: Experiencer_focused_emotion

The candidate frames produced via the assignment through inheritance and the devised procedures have been validated manually for approximately one third (4,522 out of 14,103) of the verb synsets. In the following Section 5 we present our findings on the performance of the various types of assignments.

5. Evaluation

We evaluate the development of the mappings between WordNet and FrameNet in terms of: (a) the improvement in the overall precision and coverage of the mapping; (b) the relevance of each procedure; and (c) the efficiency of each procedure. Precision is measured by counting, on the one hand, the number of mappings that we consider consistent, and on the other – the inconsistent mappings that have been identified and corrected, as well as any cases of missing elements (either in FrameNet or in WordNet, or both) that create imbalance and a skewed relation in any of the resources. Relevance takes into account the degree to which a frame suggested from the automatic procedures is directly related through a frame-to-frame relation to the manually validated frame (see details below). In addition, we pay attention to the efficiency of each procedure which we evaluate as the proportion of valid frame proposals out of all the suggestions obtained through a given procedure.

We perform a detailed analysis on each type of procedure and evaluate its results and efficiency with respect to the evaluation dataset containing 4,522 manually validated synsets (Section 4). We use the following data: (a) the initial mapping (baseline 1) – compiled from the evaluation dataset by applying only existing previous mappings before any extensions and consistency procedures are carried out (cf. Section 2.1); (b) the extended mapping (baseline 2) – compiled from the evaluation dataset baseline 1 by assigning a frame from a hypernym to its hyponyms in the cases where the hyponyms are not assigned a frame from the existing mappings (cf. Section 2.2); (c) the final mapping (manually validated) – after manual validation were carried out (cf. Section 4).

The baseline and the output mark the two extreme points on our evaluation scale. Our analysis takes into account the fact that the output is neither in its final state nor all the WordNet synsets are fully verified. The adopted level of detail in the classification of frames can vary for different resources and/or purposes, so we can always introduce more fine-grained frames, which will affect the evaluation of the procedures presented here with respect to the new output.

We introduce a detailed evaluation of the procedures involving not only precision and coverage but also relevance and efficiency of the result of each of the proposed procedures. Relevance measures the precision of the procedure itself and shows how close the output is to the desired result rather than whether it is precise since in the case of conceptual description assigning a more general frame to a synset is not considered wrong, although a more specific frame may be more informative.

The evaluation analysis tries to reflect the fact that the relevance of the assigned frames is not a

binary value (true/false). To this end, we introduce a scale from 0.00 to 1.00 to measure the relevance of a suggested assignment produced by the application of a given procedure with reference to the manually validated output: 1.00 is scored when a suggestion coincides with the final output; 0.50 – when a suggested frame is directly related to the final result via an inheritance relation; 0.25 – when a suggested frame is directly related to the final output via a different relation (e.g., Using, See also); 0.00 – when the two frames are not directly related.

The efficiency of each procedure is represented as the ratio between the number of changes undertaken and the total number of suggestions made using a particular procedure. The need to evaluate efficiency is related to the fact that manual verification is an expensive and time-consuming task, so the number of entities and suggestions to be manually checked needs to be optimised. Essentially, the efficiency measures the precision of the procedure itself. Procedures that require a lot of checks but identify very few relevant entries requiring changes are to be avoided unless essential for a particular task. The measure can also be used as a point of departure to optimising certain automatic procedures and consistency checks. One possible approach is the ranking of suggestions so that more likely ones appear first and thus, reduce the need to check lower ranking suggestions.

Table 5 shows the precision, coverage, relevance and efficiency when the discussed procedures are applied separately and in combination.

Procedure	**Precision**	**Coverage**	**Relevance**	**Efficiency**
BASELINE 1 (cf. sec. 2.1)	0.632	0.339	N/A	N/A
BASELINE 2 (cf. sec. 2.2)	0.782	0.774	N/A	N/A
Lexical & Semantic analysis (cf. sec. 3.1)	0.405	0.654	0.720	0.190
Similarity (cf. sec. 3.2)	0.334	0.590	0.691	0.117
All procedures (excl. BASELINES)	0.486	0.694	0.859	0.250
All procedures (with BASELINES)*	0.796	0.851	N/A	N/A

Table 1: Precision, coverage, relevance and efficiency of each of the procedures applied independently and in combination. *These results count the cases where either the baseline assignment was confirmed or the correct frame was suggested at least once by any of the procedures.

The reported results show that the semantic analysis and the similarity procedures as defined at present have limited contribution to improving the precision and coverage of the frame assignment. However, it is also evident that they complement the inheritance mapping and each other and that the improvement increases when all the procedures are applied in combination.

The observations on the relevance of the procedures show that we need to evaluate the results more broadly, not only in terms of the direct contribution to improving the precision and coverage but also as a contribution to the manual verification by facilitating expert decisions or giving helpful clues. The efficiency of the procedures is very low, which points to the need of narrowing down the possible suggestions in order to optimise manual work.

6. Discussion

While they are not conclusive (as all the data are not yet manually validated), the proposed expansion procedures and consistency checks are promising with respect to the task of frame-to-synset alignment.

6.1. Inheritance Mapping, Semantic Analysis, Similarity

Inheritance-based assignment proves to be the most powerful procedure in terms of its impact on both precision and coverage. Given that the frames assigned to synsets (where such assignments are available) are correct, the transfer of a synset's frame to its hyponyms must also be correct as hyponyms inherit an essential part of their semantic and lexical properties from their hypernyms, although the parent frame may be too underspecified, especially for deeper level synsets (ones assigned a frame from a distant hypernym).

As noted above, the procedures based on semantic analysis and on similarity have a marginal effect in terms of their contribution as compared with the inheritance-based mapping. The suggestions coming from these additional procedures, however, merit attention as they do yield appropriately specific frame candidates that cannot be discovered if inheritance assignment is used alone. Moreover, they offer suggestions which significantly narrow down the scope (the vicinity of related frames) where a more suitable frame can be found in the FrameNet structure and thus contribute to optimising manual work. This aligns with the considerable relevance score of the semantic analysis and the similarity-based procedures.

6.2. Towards the Definition of New Frames

Due to the discrepancy between the lexical coverage of FrameNet and WordNet, even with the application of inheritance mapping, there may not be a suitable enough parent frame to be assigned. In such cases, new frames need to be created in order to be able to achieve better coverage.

Our current efforts are directed towards defining frames that elaborate on more general ones in a way that is consistent with the formulation of already existing frames. As noted in Leseva et al. (2019), while frames have been created that describe changes in various attributes, such as *temperature* (Cause_temperature_change), *consistency* (Cause_change_of_consistency), *phase* (Cause_change_of_phase), *strength* (Cause_change_of_strength), among others, corresponding frames are missing for equally specific properties, such as *colour*, *taste*, *chemical composition*, etc. The definition of such new frames as undertaken in our work is modelled on already formulated ones. For instance, in defining the new frames Cause_chemical_reaction, Cause_change_colour, Cause_change_taste, we follow the definitions of Cause_change_of_consistency and Cause_change_of_phase with which they most closely correspond.

In addition, in a number of cases certain frames are predictable from the FrameNet structure but have not been implemented. A notable example is the lack of frame correspondences between causative and inchoative parts of the lexicon where either of the members may be missing. We take as a model pairs of frames, such as Cause_change and Undergo_change or Cause_change_of_position_on_a_scale and Change_of_position_on_a_scale, among many others, where the causative frame is related to the inchoative frame by means of the *Causative of* relation; we then proceed to define a new causative or inchoative frame where one must exist and link it to its counterpart by means of this relation. For instance, the frame Cause_change_of_strength assigned to {strengthen:1, beef up:1, fortify:1} 'make strong or stronger' does not have an inchoative counterpart that should be assigned to {strengthen:2} 'gain strength'. In a like manner, the causative frame may be missing: Change_direction, Motion_directional, Self_motion do not have causative correspondences although this distinction is made for their parent frame Motion (with its counterpart Cause_motion). Thus, for instance, {march:3} 'walk fast, with regular or measured steps; walk with a stride' is assigned the frame Self_motion, but there is no corresponding frame to account for {march:2} 'force to march' and other verbs describing self propelled motion of a person, animal, vehicle, etc. brought about through the action of another participant. One should either resort to the more general frame Cause_motion or define a new one Cause_self_motion. The same procedure of defining both the causative and the inchoative correspondence is carried out when defining new frames such as the ones described in the previous paragraph.

Finally, we intend to explore new frame suggestions made by teams working on framenets for different languages within the Global FrameNet project[4] and incorporate suitable ones.

7. Conclusions and future work

The work envisaged in the near future is aimed at providing further validation of the frame assignment to verb synsets in WordNet. A challenging prospective research will be to devise new frames that provide description of parts of the verb lexicon that have not yet been tackled in FrameNet as well as of parts of the Bulgarian verb lexicon that have no English counterparts.

A further goal is to employ the obtained linked resource in tasks such as semantic role labelling, event detection, syntactic parsing, machine translation, among others. The mapping between WordNet and FrameNet as well as the newly devised frames will be made available to the research community.

[4] https://www.globalframenet.org/

Acknowledgements

This study has been carried out as part of the project *Towards a Semantic Network Enriched with a Variety of Semantic Relations* funded by the National Science Fund of the Republic of Bulgaria under the Fundamental Scientific Research Programme (Grant Agreement 10/3 of 14.12.2016).

References

Baker, C. F. and Ruppenhofer, J. (2002). FrameNet's Frames vs. Levin's Verb Classes. In Larson, J. and Paster, M., Eds., *Proceedings of 28th Annual Meeting of the Berkeley Linguistics Society*, pages 27–38.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The Berkeley FrameNet Project. In *COLING-ACL '98: Proceedings of the Conference. Montreal, Canada*, pages 86–90.

Baker, C. F. (2008). FrameNet, present and future. In Webster, J., Ide, N., and Fang, A. C., Eds., *The First International Conference on Global Interoperability for Language Resources*, Hong Kong. City University, City University.

Burchardt, A., Erk, K., and Frank, A. (2005). A WordNet detour to FrameNet. In *Sprachtechnologie, mobile Kommunikation und linguistische Resourcen*, volume 8 of *Computer Studies in Language and Speech*. Lang, Frankfurt, Germany.

Fellbaum, C., Ed. (1998). *WordNet: an Electronic Lexical Database*. MIT Press, Cambridge, MA.

Ferrandez, O., Ellsworth, M., Munoz, R., and Baker, C. F. (2010). Aligning FrameNet and WordNet based on semantic neighborhoods. In *Proceedings of the 7th Conference on Language Resources and Evaluation (LREC 2010), May 17-23, Valletta, Malta*, pages 310 – 314.

Laparra, E. and Rigau, G. (2010). eXtended WordFrameNet. In *Proceedings of LREC 2010*, pages 1214–1219.

Leseva, S. and Stoyanova, I. (2019). Enhancing Conceptual Description through Resource Linking and Exploration of Semantic Relation. In *Proceedings of 10th Global WordNet Conference, 23 – 27 July 2019, Wroclaw, Poland*, pages 229–238.

Leseva, S. and Stoyanova, I. (2020). Beyond Lexical and Semantic Resources: Linking WordNet with FrameNet and Enhancing Synsets with Conceptual Frames. In *Towards a Semantic Network Enriched with a Variety of Semantic Relations*.

Leseva, S., Stoyanova, I., and Todorova, M. (2018). Classifying Verbs in WordNet by Harnessing Semantic Resources. In *Proceedings of CLIB 2018, Sofia, Bulgaria*, pages 115–125.

Leseva, S., Stoyanova, I., Todorova, M., and Kukova, H. (2019). Frame Specialisation Motivated by Inter-Frame Relations in FrameNet. In *Proceedings of the 14th International Conference on Linguistic Resources and Tools for Natural Language Processing, Cluj-Napoca, 18-20 November 2019, Editura Universităii Alexandru Ioan Cuza din Iasi*, pages 229–238.

Miller, G. A. (1995). WordNet: A Lexical Database for English. *Commun. ACM*, 38(11):39–41.

Palmer, M., Bonial, C., and McCarthy, D. (2014). SemLink+: FrameNet, VerbNet and Event Ontologies. In *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014), Baltimore, Maryland USA, June 27, 2014*, pages 13–17. Association for Computational Linguistics.

Palmer, M. (2009). Semlink: Linking PropBank, VerbNet and FrameNet. In *Proceedings of the Generative Lexicon Conference*. 9–15.

Petruck, M. R. (2015). *The Components of FrameNet*. `http://naacl.org/naacl-hlt-2015/tutorial-framenet-data/FNComponentsMRLP.pdf`.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R., Johnson, C. R., Baker, C. F., and Scheffczyk, J. (2016). *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California.

Shi, L. and Mihalcea, R. (2005). Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. In Gelbukh, A., Ed., *Computational Linguistics and Intelligent Text Processing. CICLing 2005. Lecture Notes in Computer Science*, volume 3406. Springe, Berlin, Heidelbergr.

Stoyanova, I. and Leseva, S. (2019). Structural Approach to Enhancing WordNet with Conceptual Frame Semantics. In Angelova, G., Mitkov, R., Nikolova, I., and Temnikova, I., Eds., *Natural Language Processing in a Deep Learning World. Proceedings of RANLP 2019, Varna, Bulgaria*, pages 630 – 638.

Tonelli, S. and Pighin, D. (2009). New Features for Framenet – Wordnet Mapping. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL'09), Boulder, USA*.