Integrating Semantic Frames from Multiple Sources

Namhee Kwon and Eduard Hovy

Information Sciences Institute
University of Southern California
Marina del Rey, CA 90292, USA
{nkwon,hovy}@isi.edu

Abstract. Semantic resources of predicate-argument structure have high potential in language understanding. Several alternative frame collections exist, but they cover different sets of predicates and use different role sets. We integrate semantic frame information given a predicate verb using three available collections: FrameNet, PropBank, and LCS database. For each word sense in WordNet, we assign the corresponding FrameNet frame and align frame roles between FrameNet and PropBank frames and between FrameNet and LCS frames.

1 Introduction

By providing more accurate semantic analysis, we should obtain higher performance in many applications such as machine translation, question answering, and summarization. Thanks to the release of annotated corpora with semantic argument structures and manually constructed lexical-semantic information such as FrameNet [1], PropBank [10], LCS database [3], and VerbNet [11], many models inducing semantic frames have been developed. ([7], [6], [13], [17], [18]).

Such data collections cover different sets of predicates, no collection covers all (or most) of the (English) predicates, and the roles and other definitional aspects of the collections differ. Due to these differences, most approaches to semantic analysis using these available resources (semantic role tagging) are specific to only one of these resources and their results are not comparable and usable over other resources.

We believe that we can build a broader and consistent semantic resource by integrating all semantic frame information of disparate collections. The value of the integrated resource is apparent at many levels. First, as a theoretical device to highlight differences and generate further refinements in lexical semantic theory, second, as a practical resource that can be used by semantic analysis and other applications. And third, an automatic aligning method between different resources can be applied to more general integration of lexical information. As more such annotated collections are created in the future, the value of (semi)automatic alignment/integration processes will increase.

In this paper, we provide a method to combine individual semantic resources - FrameNet, PropBank, and LCS [1] - and investigate similarities and differences among these collections. We align frames including semantic roles for each lexical sense defined in WordNet 2.0 [4]. We currently work on verbs since PropBank and LCS define only verb predicates.

Shi and Mihalcea have attempted to combine FrameNet frames and VerbNet verb classes using WordNet semantic networks in [15] and generate a rule-based semantic parser in [14]. Their rule-based parser is a good example of utilizing an integrated resource. Different from their work, we also combine PropBank argument structures and provide automatic method to assign FrameNet frame to each WordNet sense with a careful evaluation/correction process.

The rest of the paper is organized as follows: Section 2 describes frame resources in detail, Section 3 explains the automatic process for frame and role alignment, Section 4 provides the system evaluation and results, Section 5 discusses an expanding process in ontology, and finally, Section 6 concludes.

2 Lexical Resources

We gather frame information from three different collections (FrameNet, PropBank, and LCS), and produce an integrated frame alignment, at the word sense level represented in the WordNet. Each resource has different coverage of lexicons, and provides different frame definitions and different generalization levels in role naming.

FrameNet The FrameNet [1] defines semantic frames which are semantic representations of situations involving various participants, properties, and other conceptual roles. For each frame, corresponding predicating words are associated. It uses particular role names specific to a given situation such as *Communicator, Speaker*, and *Traveler*. FrameNet release 1.1 (Jan. 2004) defines 487 distinct frames with 696 distinct roles, and 6,743 predicates (2,300 verbs, 3,103 nouns, 1,264 adjectives, etc.) are associated with the frames.[2]

PropBank The Proposition Bank project adds semantic information to Penn Treebank [10]. It defines a predicate-argument structure on a per-predicate basis, not at the frame level as the FrameNet. The core arguments of each predicate are simply numbered while keeping general role names such as "temporal" or "locative". The numbered arguments (*Arg0, Arg1,.., Arg5*) are specific to a predicate verb, in other words, *Arg1* of one verb and *Arg1* of another verb do not necessarily mean the same role. PropBank (Feb. 2004) covers 3,323 predicate verbs and 4,659 framesets.

[1] Both semantic structures of LCS (Lexical Conceptual Structure) database and VerbNet are developed based on Levin's verb classes [12]. In this work, we integrated LCS (rather than VerbNet) that has more subdivided classes.

[2] Data release 1.2 became available (Jul. 2005) after we had started this research.

LCS database The Lexical Conceptual Structure [3] abstracts the language-independent conceptual information including semantic structures. The semantic structure of a verb is based on Levin's English verb classes [12] which assumes syntactic verb-argument frames represent inherent semantics. The semantic structure uses 15 thematic roles such as *Agent, Theme*, and *Predicate*. The database contains 4,432 verbs in 492 classes.

WordNet WordNet [4] is a lexical network covering most of English words (11,306 verbs, 114,648 nouns, 21,436 adjectives, and 4,669 adverbs) including semantic relations between lexical units. Several terms in this paper are from WordNet, and the following is a short description excerpted:
Sense: a meaning of a word in WordNet.
Synset: a synonym set; a set of words that are interchangeable in some context.
Hypernym: the generic term used to designate a whole class of specific instances. Y is a hypernym of X if X is a (kind of) Y.
Hyponym: The specific term used to designate a member of a class. X is a hyponym of Y if X is a (kind of) Y.

3 Integrating Process

A single lexical item can have more than one frame depending on its context, so we align frames for the item's WordNet sense. Given a WordNet sense, the corresponding frames from each source are assigned and the result will show the alignment between frames. The PropBank and LCS frames have been previously assigned to each sense by hand. Here, we associate word senses with FrameNet frames, and align frame roles between frames for one sense. We split the process into two steps: one is to assign a FrameNet frame to each sense, and the other is to map roles between FrameNet and PropBank frames, and between FrameNet and LCS frames.

3.1 Frame Alignment

FrameNet defines frames containing a frame name and a set of conceptual roles, and also defines corresponding predicate lexemes for each frame. Given a predicate, we search all senses in WordNet 2.0 and assign frames to a sense if there is enough evidence connecting the frame and the sense. The two steps of the process are:

Step I: Given a verb lexical unit from FrameNet, we search all WordNet senses and compute scores for each mapping between a sense and a frame. When the score is higher than an empirically found threshold[3], we assign the frame to the sense. We build scores mainly from the definition of frames and word senses,

[3] All thresholds used in frame and role alignment are determined by testing on 100 verbs.

Input: synset c, frame f	
Output: scores of relation between c and f	
1	The frequency of the pair of *(c,f)* over all synsets associated with the frame *f*.
2	The frequency of the pair of *(c,f)* over all frames associated with the synset *c*.
3	The ratio of the lexemes associated to *f* in *c* to all lexemes in the synset *c*.
4	The ratio of the number of matching words between sense description and FrameNet lexunit definition to the number of words in FrameNet lexunit definition.
5	The number of matching words between sense description and FrameNet lexunit definition.
6	The number of matching stem words between FrameNet frame definition and sense description.
7	The frequency of the pair of *(immediate hypernym of c, f)* over all senses whose immediate hyponyms are associated with the frame *f*.
8	The frequency of the pair of *(immediate hypernym of c, f)* over all frames associated with *c*s immediate hypernym.
9	The number of matching role names between FrameNet and LCS over the number of FrameNet frame element roles.
10	The number of matching role names between FrameNet and LCS over the number of LCS theta roles.
11	The number of matching words between FrameNet frame name and PropBank structure description.
12	The frequency of the pair of *(Lexicographer file name of c, f)* over the Lexicographer file names of all senses associated with the frame *f*. Verb synsets in WordNet are organized into 15 lexicographer files based on syntactic category and logical groupings such as *verb.body, verb.change*, and *verb.cognition*.
13	The frequency of the pair of *(Lexicographer file name of c, f)* over all frames associated with the Lexicographer file name of *c*.

Table 1. The scores used for computing the degree of association between a frame and a synset.

and also from frame occurrence distribution over senses. We also utilize the WordNet relations between senses, and pre-assigned PropBank and LCS frames. Since all verbs in a synset share a definition and WordNet relations, the scores are computed between a frame and a synset. The detailed description of each score is in Table 1.

Step II: After assigning initial frame-sense mappings, some lexemes have no senses connected to the given frame although the lexeme is associated with the frame in FrameNet. However, if a lexeme is associated with the frame, it means at least one of the senses of the given lexeme has to be connected to the frame. Based on this assumption, we assign the frame to the sense having the highest score if none of the senses has a frame after Step I.

We associated 2,679 senses with FrameNet frames from 8,269 senses corresponding to the lexemes defined in FrameNet.

3.2 Role Alignment

After assigning frames to each word sense, we align roles among all FrameNet, Propbank, and LCS frames associated to a given sense. We find the mapping between FrameNet and LCS frames, and between FrameNet and PropBank frames. Given a FrameNet frame role[4], we map a proper role from corresponding LCS and PropBank frames, since the FrameNet frame has a set of possible roles rather than an exhaustive set of roles for a sentence. For example, the FrameNet frame "Performer and roles" has the core roles: *Audience, Medium, Performance, role, Performer, Performer1, Performer2, score, script*, and *type*, while PropBank defines two roles, *Arg0(Actor)* and *Arg1(role)*, in "act.01" or "play.02" structure. When there are two performers in a sentence, they are called *Performer1* and *Performer2* respectively in FrameNet, but all match *Arg0 (Actor)* in PropBank. In other words, different roles from a FrameNet frame can have the same corresponding roles from a PropBank or LCS frame.

Mainly we analyze FrameNet definitions and compare the example sentence patterns with the patterns of LCS frames and PropBank argument structures.

FrameNet vs. LCS: FrameNet uses 392 distinct role names for frames of verb predicate, specific to a frame. To map those roles to more general theta roles in LCS, we use hierarchical frame information in FrameNet as a feature. We traverse up the role inheritance relations, and check if a role name matches the LCS role name. For example, *Deformer* in the "Reshaping" frame is inherited from *Agent* in the "Damaging" frame, so we assign points to the mapping from *Deformer* to *Agent* in the LCS frame. However, only 18% of the frame roles (including ancestor's name) use the same name as in LCS.

To cover many frame roles, we use example sentences from FrameNet frame. Figure 1 explains the mapping procedure with an example. From a sentence annotated with semantic roles, we delete parts not annotated with core roles, and substitute role names for real constituents. When a frame element starts with a preposition, we leave a preposition in front of the role name. With these simplified patterns, we compare LCS theta roles expressed in the same way. We count the frequency of matching roles between FrameNet and LCS.

With these features, we compute the sum of scores for a pair of a FrameNet role and a LCS role, and finally align the roles if the score is greater than a threshold.

FrameNet vs. PropBank: As in FrameNet vs. LCS, we check role names and role occurrence patterns in a sentence. PropBank uses numbered argument

[4] FrameNet frames have three types of roles: core, peripheral, and extra-thematic. For the role alignment, we only consider core roles that are required in a given frame.

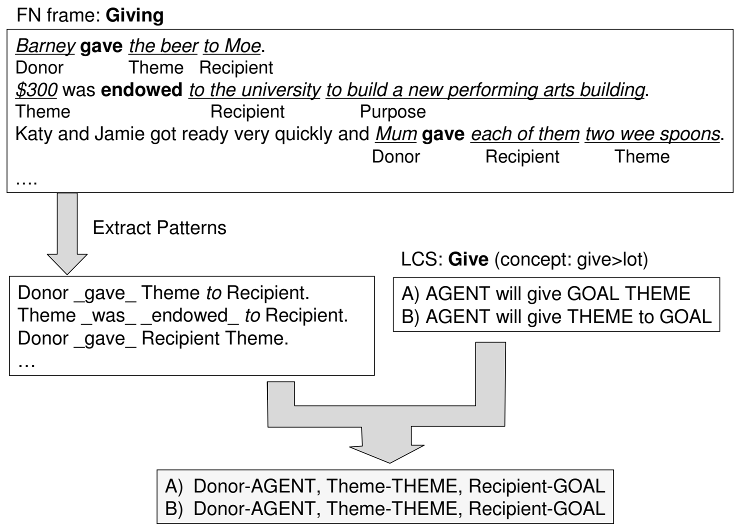

Fig. 1. Role mapping between FrameNet and LCS using annotated sentences.

names from *Arg0* to *Arg5*, and it includes short descriptions for the roles. We check the role description if it contains the FrameNet role name. In addition, we extract and compare role occurrence patterns in a sentence using example sentences from FrameNet and PropBank. The process is same as in FrameNet vs. LCS.

Since we already have the mappings between FrameNet frames and LCS frames, we use these mappings. PropBank definition partially covers the mapping to VerbNet [11] verb classes, and 27% of total roles have corresponding VerbNet theta roles. Since VerbNet semantic structure is also based on Levin's verb classes [12], the role definition is similar to roles in LCS. For the arguments having no corresponding theta roles in PropBank, we generalize an argument description to the patterns in Table 2. For instance, *Arg0* in "abandon.03" is described as "entity abandoning something" and this is generalized to "thing|entity -ing". We compute the probability of theta roles given one of these patterns.

From this probability given a generalized role description and from the example sentences, we compute the score of the relationship between FrameNet role name and PropBank argument, and apply a threshold again to decide. Table 3 shows the number of aligned roles after this process. There are 1,724 senses having FrameNet and LCS frame. 1,412 of them have role mapping information, and there are 9,635 role pairs between frames.

-er
-ed
thing|entity [being] -ed
thing|entity -ing
-ed thing|entity
-ing thing|entity
-ing from|to|into...(preposition)
-ed from|to|into...(preposition)

Table 2. The generalized patterns of role description of PropBank.

Automatically Aligned	FrameNet vs. LCS	FrameNet vs. PropBank
Senses having frame mappings	1,724	1,714
Senses having role mappings	1,412	1,350
Sense-Frame pairs	3,897	1,601
Role pairs	9,635	4,514

Table 3. The number of aligned roles.

4 Evaluation

The evaluation is performed by checking the system output manually. Not only to evaluate the result but also to use the integrated data for a future resource, we checked and corrected all the mapping output. Given the system output, an annotator first checked the frame alignment for each WordNet sense, and second checked the role mapping between FrameNet and LCS and between FrameNet and PropBank. In order to focus on the differentiation between frames given a verb, the interface was designed to link a frame and a sense given a verb, and to align roles between aligned frames. We hired two annotators for parallel checking. They checked 2,300 verbs defined in FrameNet.

4.1 Agreement between Annotators

In FrameNet, a predicate is often connected to multiple frames corresponding to its context, but it does not exactly match the WordNet sense, which results in disagreements between annotators.

Table 4 shows the agreement between two annotators for frame alignment and role alignment respectively. Agreement on frame alignment means if they provide the same answer (yes or no) for the association between the given frame and sense. For the frames that both annotators associate with a given sense, we check if they also agree on each role mapping.

We also compute the Kappa value for each frame alignment and role alignment. The Kappa statistic [16] is an agreement measure between judges' assessment considering chance agreement. An assessment with $K \geq 0.8$ is generally viewed as reliable, and $0.67 < K < 0.8$ allows tentative conclusions. In our experiment, the Kappa value is 0.71 for frame alignment and 0.75 for role alignment.

Annotation	Frame	Role
Agreement	0.85	0.82
Kappa	0.71	0.75

Table 4. The agreement between annotators.

Based on the assumption that all verbs in a synset share the same frame information, we checked for inconsistencies within a synset and fixed them manually. Finally we asked a third person to check the mismatch between annotators, and we used the answers where two of the three agreed.

4.2 Automatic Alignment Accuracy

Figure 2 shows the performance of the automatic process when we assume the hand-corrected data is correct. *Precision* is a measure representing how many senses are actually associated with the automatically assigned frames, *Recall* is how many sense-frame relations are detected by the process, and Accuracy is a measure how many of considered sense-frame pairs are correctly answered, including saying "no relation". The system shows 76% accuracy for frame alignment and 70% accuracy for role mapping. The precision is pretty high but the recall is not high enough because we don't have enough information for some frames. Especially when the frame or verb is not common, we do not find adequate generalized information, or when the role definition is short and includes no words that provide clues. In other words, when there is enough evidence, we can get good matches, but if not, we cannot say anything.

The baseline is to assign all frames to all senses of the given verb when the verb is associated with the frame in FrameNet. Since we are considering only verbs defined in FrameNet, the baseline recall is 1. As a role mapping baseline, we match a role with a role having the same name or description and match *Agent* and *Arg0*. High accuracy results in the high speed hand-correction since annotators do not need to change a lot. To help annotators, we update the output of the automatic process (input of hand-checking process) using already corrected results. After annotating 1,000 senses, we update the remaining 7,295 senses when their synonyms already have frames that both annotators have agreed on. In Figure 2, System1 represents the original automatic process and System2 represents the modified system. It shows much progress for System2 (76% to 82% in frame alignment, and 70% to 78% in role mapping) and it implies more improvement as we add more resources to modify remaining annotations. As we obtain more aligned data, we will speed up the next manual annotation.

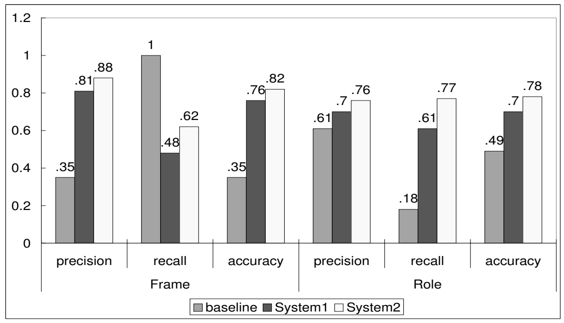

Fig. 2. Automatic process evaluation.

4.3 Results

By the integrating process, 4,240 senses are associated with FrameNet frames. In Table 5, we show the total number of senses associated with frames. 1,757 senses are associated frames from FrameNet, PropBank, and LCS, and 559 senses have only FrameNet frames. Figure 3 shows the system output with an example.

Associated Frames	Number of Senses
FrameNet	559
FrameNet & LCS	674
FrameNet & PropBank	1,250
FrameNet & PropBank & LCS	1,757

Table 5. WordNet senses associated with frames.

5 Expanding Frame Alignments

The verbs defined in FrameNet have been associated with the corresponding frames for each sense, but many senses still have no frame alignment. We attempt to expand the aligned frame information to other verb senses (not defined in FrameNet) using relations in WordNet.

LCS defines verb classes sharing semantics although it does not always match semantics (frames) in FrameNet [2]. We check the WordNet sense hierarchy within a verb class, and induce the FrameNet frame for the verbs defined in LCS. First, we check the immediate hyponyms of the word sense, and if most (90%) of the senses associated with a frame are connected to the same frame, and then we assign the frame to the sense (bottom-up approach). Second, given

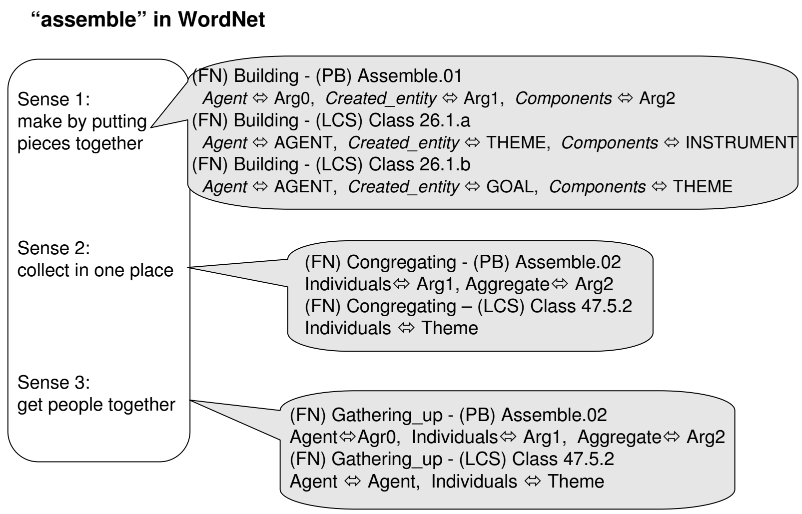

Fig. 3. The system output with an example. Given a verb "assemble", our systems finds FrameNet frames (Gathering_up, Congregating, or Building) corresponding to each sense, and aligns roles between frames.

a sense, we find the frames associated with the immediate hypernym of the sense, and if the hyponyms of the hypernym do not have much conflict (if there is less than 10% hyponyms mapped to other frames), we assign the frames to the sense (top-down approach).

We randomly chose 170 senses from the output of bottom-up and top-down approach, and computed the precision to see how many frames were correctly assigned. Table 6 shows the precision for each step.

Step	Frame-assigned Senses	Precision
I. Bottom-up	66	73%
II. top-down	486	67%
Overall Output	552	63%

Table 6. The result of expanding alignment. 791 senses are newly associated with frames from 3,207 of senses defined in LCS (not in FrameNet).

When verbs are not in FrameNet, it probably means that the appropriate FrameNet frames are not defined yet. However, we see the output provides an approximate for the real (undefined) frames with reasonable precision. Since we propagate the frames traversing the relations between senses within LCS verb classes, our system does not assign totally different frames but often mistakes fine-grained frame differences. For example, our system confuses *Attempt_suasion, Suasion*, and *Talking_into.* All three frames are about the situation that "the speaker expresses through the language his wish to get the addressee to act". *Attempt_suasion* "does not imply that the addressee forms an intention to act, and let alone acts", *Suasion* is for the situation "as a result, the addressee forms an intention to do so, but no implication that the addressee actually acts upon the intention", and *Talking_into* implies "the addressee forms an intention to act and does so". Even though they are different in detailed meaning, they all have semantic roles "Speaker", "Addressee", and "Content", and this error may be acceptable in some applications.

6 Conclusion

We have described the process of aligning frames to each WordNet sense. For each verb sense, we associated FrameNet frame and aligned frame slots (semantic roles) between FrameNet and PropBank and between FrameNet and LCS database. We aligned about 4,200 senses in FrameNet frame, and expanded the result to around 800 words not defined in FrameNet and defined in LCS. After obtaining valuable amount of data, we could infer the frames within the same verb class in LCS.

The alignment is performed for each word sense, and we could find the consistencies within a FrameNet frame. PropBank argument structure is defined per predicate and the same argument number does not represent the same role. The role mapping between FrameNet and PropBank within a FrameNet frame was consistent by 93%. For example, given a frame "Adorning", *Location* maps to *Arg2* and *Theme* maps to *Arg1* in the PropBank structure "encircle.01", but *Location* maps to *Arg1* and *Theme* maps to *Arg2* in "festoon.01".

As a practical resource for the application, this resource can be used in diverse ways. By integrating and expanding the frame information, semantic role labeling can utilize this data directly. In addition, the frames assigned to each word sense can be understood as coarse-grained sense disambiguation information and it can be applied to computing similarity measures between senses in WordNet. We leave the application using this resource for the future work.

The method we showed in this paper is not restricted to the frame alignment, and this could be applied to other applications investigating similarities and differences between heterogeneous resources. We are encouraged by the good agreement levels exhibited by our algorithm and plan to explore theoretical verifications further.

References

1. Baker, C.F., Fillmore, C.J., Lowe, J.B.: The Berkeley FrameNet project. Proceedings of COLING-ACL, Monteral, Canada (1998)
2. Baker, C.F., Ruppenhofer, J.: FrameNet's Frames vs. Levin's Verb Classes. Proceedings of 28th Annual Meeting of the Berkeley Linguistics Society, (2002) 27–38
3. Dorr, B.J.: Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Translation. Machine Translation 12(4). (1997) 271–322
4. Fellbaum, C.: Wordnet: An Electronic Lexical Database. The MIT press (1998)
5. Fillmore, C.J.: Frame Semantics and the Nature of Language. Annals of the New York Academy of Science Conference on the Origin and Development of Language and Speech, 280 (1976) 20–32
6. Fleischman, M.B., Kwon, N., Hovy,E.: Maximum Entropy Models for FrameNet Classification. Proceedings of EMNLP-03, Sapporo, Japan. (2003)
7. Gildea D., Jurafsky D.: Automatic Labeling of Semantic Roles. Computational Linguistics, 28(3) (2002) 245–288
8. Green, R., Dorr, B.J., Resnik, P.: Inducing Frame Semantic Verb Classes from WordNet and LDOCE. Proceedings of ACL-04, Barcelona, Spain. (2004)
9. Green, R., Pearl, L., Dorr, B.J., Resnik, P.: Mapping Lexical Entired in a Verbs Database to WordNet Senses. Proceedings of ACL-01, Toulouse, France. (2001)
10. Kingsbury, P., Palmer,M.: Adding Semantic Annotation to the Penn Treebank. Proceedings of HLT-2002, San Diego, CA. (2002)
11. Kipper, K., Dang, H.T., Palmer,M.: Class-Based Construction of a Verb Lexicon. Proceedings of AAAI-2000, Austin, TX, USA. (2000)
12. Levin, B.: English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago press (1993)
13. Pradhan, S., Ward, W., Hacioglu, K., Martin, J.H., Jurafsky, D.: Shallow Semantic Parsing Using Support Vector Machines. Proceedings of HLT/NAACL-04, Boston, MA. (2004)
14. Shi, L., Mihalcea, R.: An Algorithm Open Text Semantic Parsing. Proceedings of the workshop on Robust Methods in Analysis of Natural Language Data, Geneva, Switzerland. (2004)
15. Shi, L., Mihalcea, R.: Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. Proceedings of the 6th international Conference on Intelligent Text Processing and Computational Linguistics, Mexico city, Mexico. (2005)
16. Siegel, S., Castellan Jr., N.J.: Nonparametric Statistics for the Behavioural Sciences. McGraw-Hill (1988)
17. Surdeanu, M., Harabagiu, S., Williams, J., Aarseth, P.: Using Predicate-Argument Structure for Information Extraction. Proceedings of ACL-03, Sapporo, Japan. (2003)
18. Swier, R.S., Stevenson, S.: Unsupervised Semantic Role Labelling. Proceedings of EMNLP-04, Barcelona, Spain. (2004)