european lexicographic infrastructure

D3.2

Multilingual Word Sense Disambiguation

Author(s): Rexhina Blloshmi, Federico Martelli, Ivan Yevtushenko, Paola Velardi, Giovanni Stilo, Paolo Spadoni, Roberto Navigli.

Date: 31.01.2020

H2020-INFRAIA-2016-2017

Grant Agreement No. 731015

ELEXIS - European Lexicographic Infrastructure

D3.2 Multilingual Word Sense Disambiguation and Entity Linking algorithms - initial report

Deliverable Number: 3.2

Dissemination Level: PUBLIC

Delivery Date: 31.01.2020

Version: 1.0

Author(s):

Rexhina Blloshmi

Federico Martelli

Ivan Yevtushenko

Paola Velardi

Giovanni Stilo

Paolo Spadoni

Roberto Navigli

Project Acronym:	ELEXIS
Project Full Title:	European Lexicographic Infrastructure
Grant Agreement No.:	731015

Deliverable/Document Information

Project Acronym:	ELEXIS
Project Full Title:	European Lexicographic Infrastructure
Grant Agreement No.:	731015

Document History

Version Date	Changes/Approval	Author(s)/Approved by

Table of Contents

1 INTRODUCTION

The goal of Task 3.1, Word Sense Disambiguation (WSD) and Entity Linking (EL), is to provide novel algorithms and resources which enable disambiguation and entity linking in dozens of languages. This deliverable, D3.2, provides a new lexical-semantic resource and a knowledge-based algorithm for enabling high-quality disambiguation in multiple languages.

Most common approaches towards WSD can be classified into two categories: supervised and knowledge-based. Despite the promising performance of supervised approaches, which nowadays are characterized by deep neural network architectures, the need for large quantities of sense-annotated training data comes as a disadvantage when aiming to cover most of the European languages. Therefore, we put forward novel knowledge-based approaches, which most importantly drop the requirement of large amounts of data typically needed by neural networks.

This deliverable is organized as follows. We first describe the task and the main characteristics of the approaches employed to tackle it. Secondly, we explain our approach to the problem and illustrate the algorithm employed. Then we report a quantitative evaluation on the standard multilingual datasets in the field, in particular from the international standard competitions in WSD, i.e. Senseval and SemEval.

2 TASK DESCRIPTION

Word Sense Disambiguation and Entity Linking are the computational tasks of automatically determining the meaning of words in context [Navigli, 2009]. Moro et al [2014] were the first to merge these tasks together, demonstrating that a knowledge-based approach can be extended to perform both WSD and EL. The key assumption was that they can be complementary to each other such that the lexicographic knowledge used in WSD is also useful for tackling the EL task, and vice versa the encyclopedic information utilized in EL helps disambiguate nominal mentions in a WSD setting.

To perform WSD and EL, supervised approaches train a classifier on sense-annotated datasets and come with a better performance than knowledge-based ones. However, their application is limited to those languages for which there exist large sense-annotated datasets which enable the needed supervision, thus making the task unfeasible in many languages, especially under-resourced ones. Knowledge-based (KB) systems, instead, relieve

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

the burden of training data and have a straightforward application on multiple languages by relying on rich Lexical Knowledge Bases (LKBs) such as WordNet [Fellbaum, 1998] and BabelNet [Navigli and Ponzetto, 2012]. Therefore, the performance of KB approaches heavily relies on the nature of the relations offered by the underlying LKBs. In this deliverable, we present an algorithm which draws on an LKB augmented with a novel lexical-semantic resource, to carry out high-performance WSD and EL in multiple languages.

3 APPROACH

In this section we enable multilingual WSD and EL by providing two key components: SyntagNet [Maru et al, 2019], a novel lexical-semantic combination knowledge resource and SyntagRank, a graph-based algorithm employed to perform disambiguation.

3.1 SyntagNet

The performance of knowledge-based approaches to WSD highly depends on the type of relations present in the underlying Lexical Knowledge Base (LKB). Most commonly, the LKBs focus on paradigmatic relations between concepts while leaving uncovered syntagmatic relations, i.e., relations that exist between two words which co-occur in the same context frequently. SyntagNet addresses this gap and provides a semi-automatic large-scale lexical-semantic combination resource which associates pairs of co-occurring words with their respective meanings.

Constructing SyntagNet is a procedure of two main steps which include lexical combinations extraction and manual disambiguation. To this end, the lexical combinations were extracted from the English Wikipedia and the British National Corpus (BNC) [Leech, 1992]. More specifically, the Stanford CoreNLP pipeline was used to extract the dependency trees for all the sentences of these two corpora from which only the pairs of POS-tagged and lemmatized words co-occurring within a slide window of 3 words were considered. The candidate pairs were then associated with a strength of correlation score exploiting the Dice's coefficient formula, used to determine their relevance. Once all the possible candidate pairs are extracted and ranked according to their relevance score, 8 annotators were asked to manually disambiguate a total of 78 000 top-ranking lexical combinations. Therefore, the annotators were assigned the task of associating the most appropriate sense to each word in a lexical pair according to WordNet sense inventory. To ensure quality, they were also asked to filter out lexical combinations which contained mistakes due to the automatic extraction process, pairs which were not associated with any of

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

the senses in WordNet, those reflecting idiomatic expressions and finally those pairs which were multi-word Named Entities.

Below we report examples of lexical (left) and semantic (right) combinations:

$word_1$	$word_2$	$sense_1$	$sense_2$
$miss_v$	$home_n$	$miss_v^2$ (feel or suffer from the lack of)	$home_n^7$ (an environment offering affection and security)
$miss_v$	$concert_n$	$miss_v^3$ (fail to attend an event or activity)	$concert_n^1$ (a performance of music by players or singers not involving theatrical staging)
$miss_v$	$train_n$	$miss_v^5$ (fail to reach or get to)	$train_n^1$ (public transport provided by a line of railway cars coupled together and drawn by a locomotive)

Table 1: Excerpt of SyntagNet relations (senses from WordNet)

3.1.1 Statistics

SyntagNet covers 78,000 lexical combinations (10,218 unique nouns and 3,786 unique verbs for 52,432 noun-verb relations and 25,568 noun-noun relations) which, once disambiguated, make up 88,019 semantic combinations (61,249 noun-verb and 26,770 noun-noun semantic relations) linking 20,626 WordNet 3.0 nodes, i.e., unique synsets (14,204 noun synsets and 6,422 verb synsets), with a relation edge.

3.2 SyntagRank

In order to leverage the syntagmatic relations provided by SyntagNet, we developed a new graph-based algorithm for WSD, called SyntagRank, which is based on Personalized PageRank (PPR) [Haveliwala, 2002] to choose the correct meaning of words according to the context they appear in. This approach has been already shown to be successful [Agirre et al, 2014], however without exploiting explicit syntagmatic information. PPR

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

relies on a semantic network where the nodes represent concepts and edges represent semantic relations between them. Our algorithm consists in a word-to-word version of the Personalized PageRank algorithm (PPR_{w2w}), where a separate PPR is run from each target word in context, initializing the PPR vector using only the context and excluding the target word itself. This choice is motivated in that it enables the context to decide which concept is more relevant to the target word, without having the target affect this decision. The semantic network used by SyntagRank is composed of WordNet, the semantic relations in the Princeton WordNet Gloss Corpus[1] and the syntagmatic information provided by SyntagNet.

4 SYNTAGRANK IMPLEMENTATION

SyntagRank performs WSD in three steps:

1. **Natural Language Processing Pipeline**

   SyntagRank integrates a multilingual NLP pipeline which supports 5 languages, i.e., English, French, German, Italian and Spanish. Specifically, we use Stanford CoreNLP[2] pipeline [Manning et al., 2014] for the tokenization, sentence splitting, Part-of-Speech (POS) tagging and lemmatization. It fully supports the English language, while for the French, German and Spanish language we used its available models for POS-tagging and its tree-tagger for the lemmatization. For the Italian language instead, we use TINT[3], an implementation for the Italian language only, which offers high quality models for the pipeline in the same format as Stanford CoreNLP.

2. **Personalized PageRank Computation**

   Initially, the PPR vector is calculated for all the nodes in the graph, i.e., as a preprocessing step, and the vectors are serialized. Notice that this step is performed only once, therefore not affecting the overall response time.

[1] https://wordnetcode.princeton.edu/glosstag.shtml
[2] https://stanfordnlp.github.io/CoreNLP/
[3] http://tint.fbk.eu/

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

3. Disambiguation Procedure

At this step, the input is lemmatized and POS-tagged. Once the sentence is parsed, the correct sense for any target word is chosen as follows:

1: Gather all the candidate senses of the content words in the sentence, excluding the target word.

2: Gather all the PPR vectors of the candidates from the pre-calculated vectors, and calculate their weighted average.

3: Gather the candidate senses of the target word and their values in the average vector calculated in Step 2.

4: Return the sense with the highest value, i.e., as the most probable sense of the target word.

5 EVALUATION SETUP

In order to assess the effectiveness of SyntagNet relations when employed in a knowledge-based WSD system, i.e., SyntagRank, we conduct a set of experiments and compare with other knowledge-based and supervised systems on several English and multilingual WSD tasks.

5.1 Evaluation Benchmark

We use the evaluation framework made available by Raganato et al. [2017a] which comprises five standard test sets for WSD, i.e. Senseval-2 (SE2) [Edmonds and Cotton, 2001], Senseval-3 (SE3) [Synder and Palmer, 2004], SemEval-07 (SE7) [Pradhan et al., 2007], SemEval-13 (SE13) [Navigli et al., 2013], SemEval-15 (SE15) [Moro and Navigli, 2015]. To run experiments on multilingual WSD, we used the last two of the foregoing datasets, which also include German, Spanish, French and Italian, employing, as sense inventory, the synset lexicalizations provided in BabelNet 4.0[4]. We computed precision, recall and F1, which in our case are equal, since SyntagRank always outputs a sense for each target word.

[4] https://babelnet.org/

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

5.2 Comparison Systems

To evaluate the impact of SyntagNet relations on the disambiguation performance, we compare the results obtained by the disambiguation algorithm, using different LKBs. We employed WordNet + Princeton WordNet Glosses (WNG) as baseline. We measure the performance of the algorithm when integrating SyntagNet on top of the baseline. We also evaluated the following LKBs when integrated separately on top of WNG: 1) KnowNet (KnowNet20) [Cuadros and Rigau, 2008] and deep KnowNet (deepKnowNet95d) [Cuadros et al., 2012], 2) the subgraph of BabelNet 4.0 [Navigli and Ponzetto, 2012] induced by WordNet 3.0, 3) eXtended WordNet [Mihalcea and Moldovan, 2001], 4) ColWordNet [Espinosa-Anke et al., 2016]. In the table below, we give details for the resources used together with WNG and their respective number of lexical-semantic relations.

Resource	Number of Relations
WNG	671,779
WNG+KnowNet20[5]	520,682
WNG+deepKnowNet95d[6]	522,880
WNG+BabelNet 4.0[7]	9,447,341
WNG+eXtended WordNet[8]	551,551
WNG+ColWordNet[9]	8,424
WNG+SyntagNet	88,019

Table 2: Lexical-Semantic relations per LKB.

Moreover, we also perform a comparison of SyntagRank against state-of-the-art supervised approaches to observe the impact of SyntagNet in closing the performance gap between knowledge-based and supervised systems.

6 RESULTS

As mentioned in the previous section, we compare the performance of the disambiguation algorithm when using different LKBs.

[5] http://adimen.si.ehu.es/web/KnowNet
[6] http://adimen.si.ehu.es/web/deepKnowNet
[7] the subgraph of BabelNet 4.0 induced by WordNet 3.0
[8] http://www.hlt.utdallas.edu/~xwn/
[9] http://bitbucket.org/luisespinosa/cwn/

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

	ENGLISH						MULTILINGUAL						
							SemEval-13				SemEval-15		
	SE2	SE3	SE7	SE13	SE15	ALL	IT	ES	DE	FR	IT	ES	ALL
WNG	69.2	65.9	54.9	66.8	70.7	67.1	71.4	71.2	68.0	69.6	62.2	58.1	67.2
WNG+KnowNet20	67.2	65.8	53.8	67.3	71.5	66.6	71.6	73.1	68.3	70.4	61.4	59.9	67.9
WNG+deepKnowNet95d	66.9	64.9	53.6	66.9	71.6	66.2	71.4	71.9	67.7	70.5	62.4	58.7	67.5
WNG+BabelNet 4.0	67.5	64.1	53.0	67.6	66.9	65.6	73.8	71.6	**69.9**	67.1	62.4	57.8	67.6
WNG+eXtended WordNet	67.7	65.7	52.3	67.6	71.0	66.7	72.4	71.8	68.5	69.3	62.4	58.9	67.7
WNG+ColWordNet	69.2	65.9	54.1	66.7	70.7	67.1	71.4	71.0	68.0	69.3	61.9	57.8	67.0
SyntagRank (WNG+SyntagNet)	**71.2**	**71.6**	**59.6**	**72.4**	**75.6**	**71.5**	**74.2**	**73.4**	66.9	**72.7**	**65.0**	**61.2**	**69.3**

Table 3: Knowledge-based approaches comparison. Performance is in terms of F1 scores.

Table 3 shows the performance of all the knowledge-based systems relying on the different LKBs in terms of F1 scores. As one can see, SyntagRank achieved the best results in the English all-words disambiguation tasks, attaining 4.4 points above the WNG baseline overall, i.e., in ALL dataset, a concatenation of all English evaluation datasets. Considering the lower improvement or degradation in performance of the baseline when combined with the other LKBs, we conclude that the syntagmatic nature of SyntagNet provides an advantage as opposed to the noiser character of other LKBs.

We conducted a deeper analysis of the nature of the relations found in each LKB by sampling 500 relations of each LKB and classifying the edges as syntagmatic or paradigmatic and found out that the highest percentage of syntagmatic relations among comparison LKBs was found in eXtended WordNet, i.e., 54%. In contrast, the fully syntagmatic nature of SyntagNet relations proved to improve the performance significantly. Similarly, SyntagRank outperforms the other systems in the multilingual ALL dataset, and in all the languages but one, thus enforcing the benefit of SyntagNet resource.

system	SE2	SE3	SE7	SE13	SE15	ALL
LSTMMLP	**73.8**	71.8	**63.5**	69.5	72.6	**71.5**
$IMSC2V_{+PR}$	**73.8**	71.9	63.3	68.2	72.8	71.2
fastSense	73.5	**73.5**	62.4	66.2	73.2	71.1

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

SyntagRank	71.2	71.6	59.6	**72.4**	**75.6**	**71.5**

Table 4: F1 scores for English all-words WSD.

Table 4 compares SyntagRank against several supervised English WSD systems such as LSTMMLP [Yuan et al., 2016], IMSC2V$_{+PR}$ [Melacci et al., 2018] and fastSense [Uslu et al., 2018]. We remark that the differences in ALL dataset are not statistically significant, according to a χ^2 test ($p<0.01$), which leads to the conclusion that SyntagRank, despite being a knowledge-based system, remains in the same ballpark with state-of-the-art supervised approaches for English WSD, proving the effectiveness of syntagmatic relations it provides.

	SemEval-13				SemEval-15		
system	IT	ES	DE	FR	IT	ES	ALL
BILSTM	62.0	66.4	69.2	55.5	-	-	-
UMCC-DLSI	65.8	71.0	62.1	60.5	-	-	-
T-O-M	68.2	66.9	63.2	60.5	-	-	-
SUDOKU-RUN1	-	-	-	-	59.9	56.0	-
SUDOKU-RUN2	-	-	-	-	56.9	57.1	-
Best System	68.2	71.0	**69.2**	60.5	59.9	57.1	64.7
SyntagRank	**74.2**	**73.4**	66.9	**72.7**	**65.0**	**61.2**	**69.3**

Table 5: F1 scores for multilingual WSD.

Not only does SyntagRank achieve similar results to the best supervised approaches in English WSD, but it can easily scale across multiple languages. In table 5, we show the results in multilingual WSD evaluation datasets attained by several supervised systems, namely BILSTM [Raganato et al., 2017b], UMMCC-DLSI [Gutierrez Vazquez et al., 2010], T-O-M [Pasini and Navigli, 2017], SUDOKU RUN1 and SUDOKU RUN2 [Manion, 2015] and SyntagRank. Moreover, for ease of comparison, we report the results obtained by aggregating the outputs of the best performing system for each dataset, i.e., Best System. As one can see, SyntagRank achieves state-of-the-art results in five out of six datasets, with a significant boost in performance with 4.6 F1 points on overall over the Best System.

7 RESTful SERVICE

SyntagRank enables interaction through a REST interface. The system is queried by specifying the text to be disambiguated and its written language. The interface is reachable from two endpoint addresses. The output to the requests is a JSON object which contains the predicted senses for each token in the text and the anchor of each token specifying their positional indices in the input text.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

More specifically, the tables below show details on the RESTful API documentation:

Title	Disambiguate Text
URL	http://api.syntagnet.org/disambiguate?lang=*lang*&text=*text*
Method	POST
URL Params	Required: text = "string" lang = "string" **text**: *the sentence or text to be disambiguated (max length = 1500 characters)* **lang**: *the language of input text (to be chosen from: DE, EN, ES, FR, IT)* Example: text=this is a text lang=EN

Success Response	Code: 200
	Content:
	{
	  "language": "EN",
	  "tokens": [
	    {
	      "senseID": "wn:02604760v",
	      "position": {
	        "charOffsetBegin": 5,
	        "charOffsetEnd": 7
	      }
	    },
	    {
	      "senseID": "wn:06387980n",
	      "position": {
	        "charOffsetBegin": 10,

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

	"charOffsetEnd": 14
	        }
	      }
	   ]
	}

Table 6: RESTful request example 1.

Title	Disambiguate Text
URL	http://api.syntagnet.org/disambiguate_tokens
Method	POST

Data Params	Required data as JSON Object:
	data = {
	   "lang": "string"
	   "words": [
	      {
	         "id": "string",
	         "word": "string",
	         "lemma": "string",
	         "pos": "string",
	         "isTargetWord": bool
	      }
	   ]
	}
	
	lang: *the language of input text (to be chosen from: DE, EN, ES, FR, IT)*
	words: *list of words in a sentence where each word is an object containing the following fields:*
	- **id**: *the index of the word in the sentence*
	- **word**: *the inflicted form of the word as it appears in the sentence*
	- **lemma:** *lemma corresponding to the word*
	- **pos**: *the part of speech tag from the Penn Treebank tag set.*

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

<table>
<tr>
<td></td>
<td>

- **isTargetWord**: *a boolean variable denoting if the word should be disambiguated*

Example:


```
data = {
  "lang": "EN",
  "words": [
    {
      "id": "0",
      "word": "this",
      "lemma": "this",
      "pos": "X",
      "isTargetWord": false
    },
    {
      "id": "1",
      "word": "is",
      "lemma": "be",
      "pos": "VERB",
      "isTargetWord": true
    },
    {
      "id": "2",
      "word": "a",
      "lemma": "a",
      "pos": "X",
      "isTargetWord": false
    },
    {
      "id": "3",
      "word": "text",
      "lemma": "text",
      "pos": "NOUN",
      "isTargetWord": true
    }
  ]
}
```


</td>
</tr>
</table>

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

Success Response	Code: 200
	Content:
	
	{
	"result": [
	{
	"id": "3",
	"synset": "wn:06387980n"
	},
	{
	"id": "1",
	"synset": "wn:02604760v"
	}
	]
	}

Table 7: RESTful request example 2.

8 WEB INTERFACE

SyntagRank provides a web interface for looking up terms and their lexical-semantic combinations according to the SyntagNet resource. SyntagRank can be accessed at http://syntagnet.org. Using the web interface, a user can look up syntagmatic relations connecting meanings of specific terms. An example is shown below for verb *miss*:

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

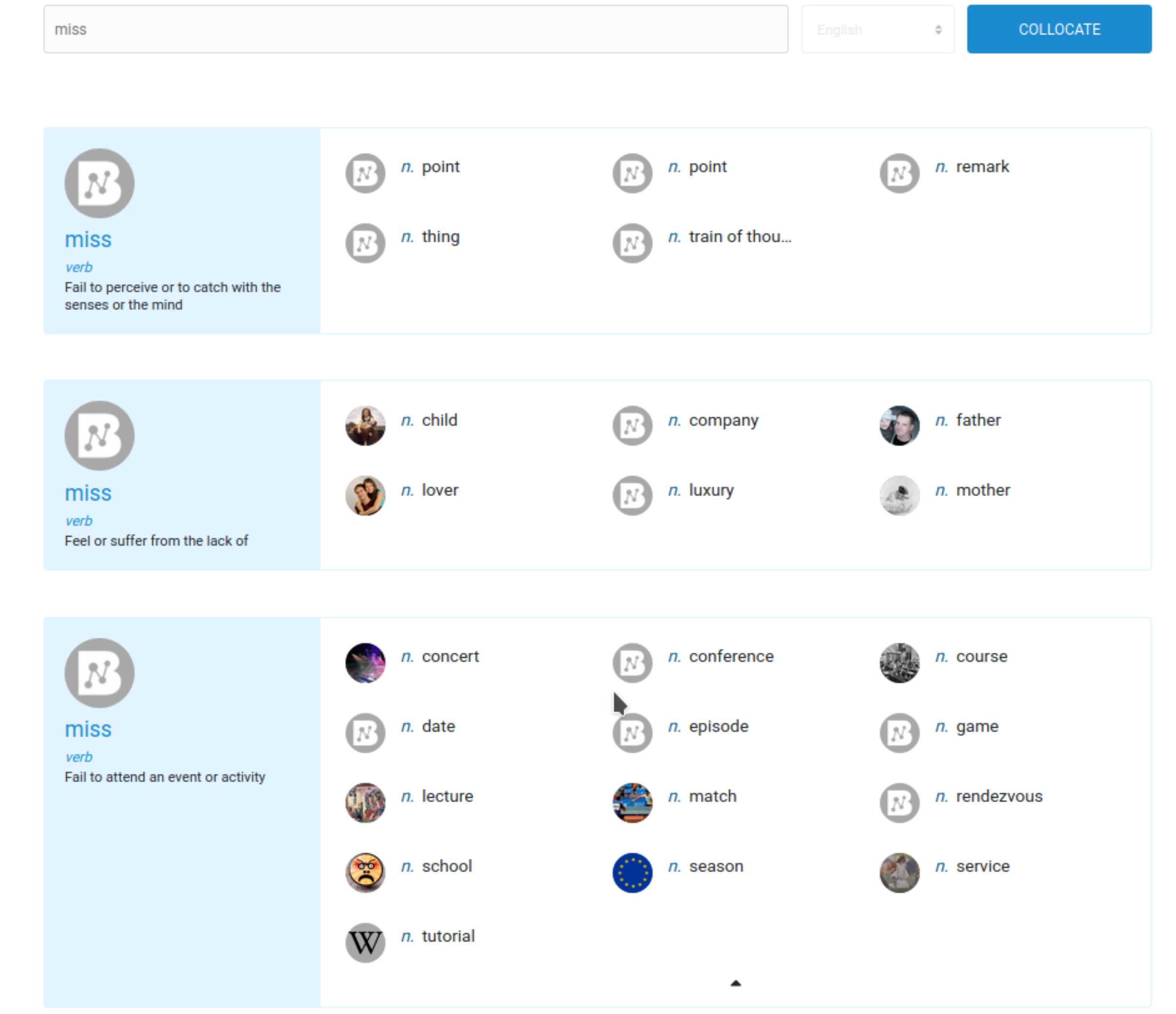

Figure 1: Lexical-Semantic combinations result for the verb *miss*.

As shown in Figure 1 the collocations of a word are specified for each of the different meanings of the word, which in addition are associated with the respective gloss from WordNet.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

Moreover a user can disambiguate phrases by specifying the language and the sentence in natural language. One feature of SyntagRank is that it identifies the Named Entities and Concepts. Each concept is connected by a link to BabelNet synsets. Figure 2 and 3 show examples of disambiguating sentences in different languages:

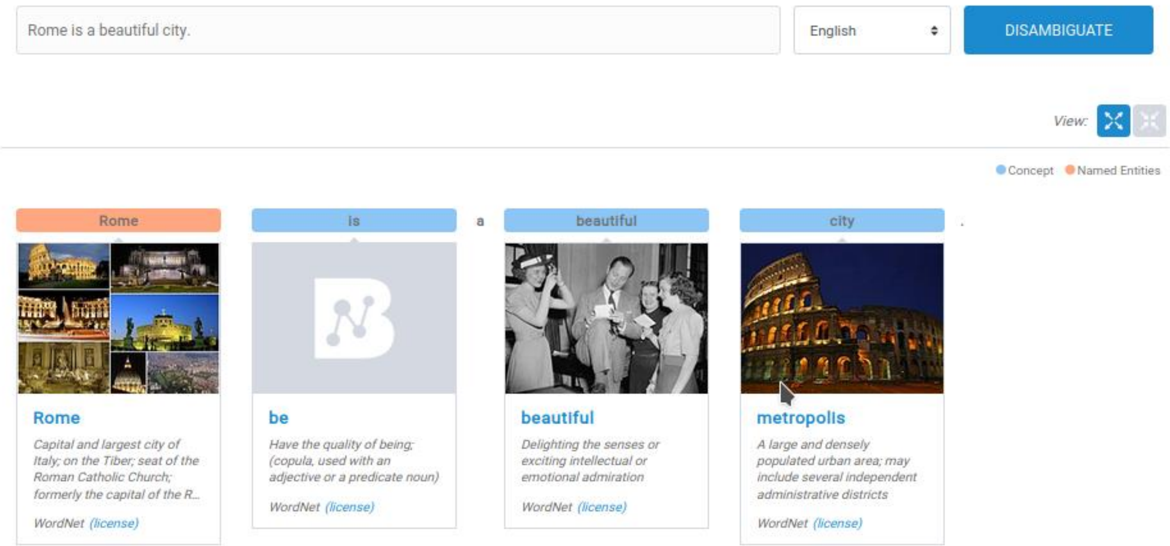

Figure 2: Disambiguation result of an English sentence.

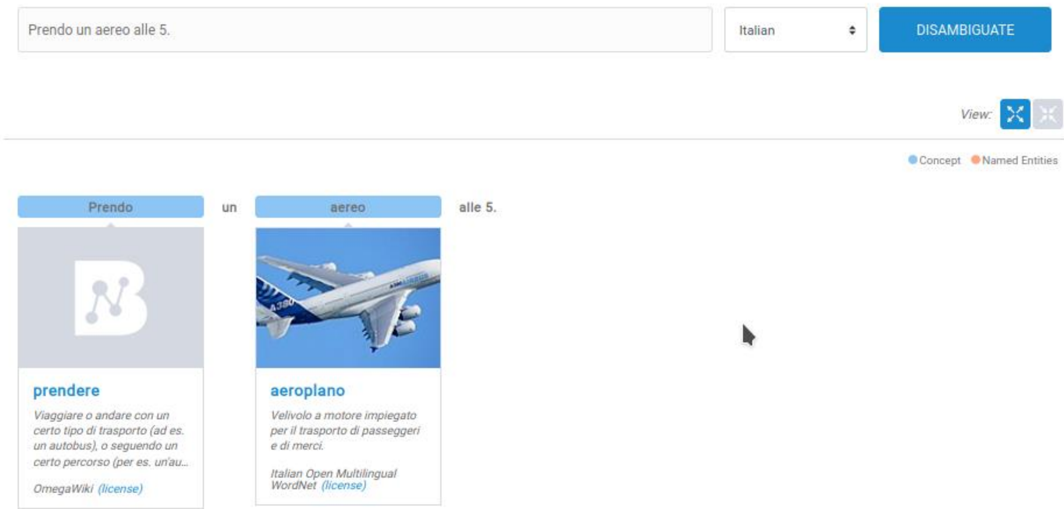

Figure 3: Disambiguation result of an Italian sentence.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

9 REFERENCES

Eneko Agirre, Oier Lopez de Lacalle, and Aitor Soroa. 2014. *Random Walks for Knowledge-Based Word Sense Disambiguation*. Computational Linguistics, 40(1):57–84.

Montse Cuadros, Lluis Padro, and German Rigau. 2012. *Highlighting relevant concepts from topic signatures.* In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012), pages 3841–3848, Istanbul, Turkey.

Montse Cuadros and German Rigau. 2008. *KnowNet: Building a Large Net of Knowledge from the Web*. In Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, pages 161–168, Manchester, United Kingdom.

Philip Edmonds and Scott Cotton. 2001. *SENSEVAL-2: Overview*. In Proceedings of SENSEVAL-2 Second International Workshop on Evaluating Word Sense Disambiguation Systems, pages 1–5,Toulouse, France.

Luis Espinosa-Anke, Jose Camacho-Collados, SaraRodriguez Fernandez, Horacio Saggion, and Leo Wanner. 2016. *Extending WordNet with Fine-Grained Collocational Information via Supervised Distributional Learning.* In Proceedings of COLING 2016: Technical Papers. The 26th InternationalConference on Computational Linguistics, pages 900–910, Osaka, Japan.

Christiane Fellbaum, editor. 1998. *WordNet: an electronic lexical database.* MIT Press, Cambridge, MA, USA.

Yoan Gutierrez Vazquez, Antonio Fernandez Orquin, Andres Montoyo Guijarro, and Sonia Vazquez Perez. 2010. *UMCC-DLSI: Integrative Resource for Disambiguation Task.* InProceedings of the 5th International Workshop on Semantic Evaluation, pages 427–432, Uppsala, Sweden.

Taher H. Haveliwala. 2002. *Topic-Sensitive PageRank*. In Proceedings of the 11th international conference on World Wide Web, pages 517–526, Honolulu, HI, USA.

Geoffrey Neil Leech. 1992. *100 million words of English: the British National Corpus (BNC)*. Language Research.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

Steve L. Manion. 2015. *SUDOKU: Treating Word Sense Disambiguation & Entity Linking as a Deterministic Problem via an Unsupervised & Iterative Approach*. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval-2015), pages 365–369, Denver, CO, USA.

Christopher Manning, Mihai Surdeanu, John Bauer,Jenny Finkel, Steven Bethard, and David McClosky. 2014. *The Stanford CoreNLP natural language processing toolkit.* In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 55–60, Baltimore, MD, USA.

Marco Maru, Federico Scozzafava, Federico Martelli, and Roberto Navigli. 2019. *SyntagNet: Challenging Supervised Word Sense Disambiguation with Lexical-Semantic Combinations.* In Proceedings of EMNLP-IJCNLP 2019.

Stefano Melacci, Achille Globo, and Leonardo Rigutini. 2018. *Enhancing Modern Supervised Word Sense Disambiguation Models by Semantic Lexical Resources.* In Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, May 7-12, 2018, Miyazaki, Japan.

Rada Mihalcea and Dan Moldovan. 2001. *eXtendedWordNet: Progress Report*. In Proc. of the NAACL Workshop on WordNet and Other Lexical Resources, pages 95–100, Pittsburgh, PA, USA.

Andrea Moro, Alessandro Raganato and Roberto Navigli. *Entity Linking meets Word Sense Disambiguation: a Unified Approach.* In *Transactions of the Association for Computational Linguistics* 2 (2014): 231-244.

Andrea Moro and Roberto Navigli. 2015. *SemEval-2015 task 13: Multilingual all-words sense disambiguation and entity linking.* In Proceedings of the9th International Workshop on Semantic Evaluation(SemEval 2015), pages 288–297, Denver, CO, USA.

Roberto Navigli. 2009. *Word sense disambiguation: A survey*. ACM Comput. Surv., 41, 10:1-10:69.

Roberto Navigli, David Jurgens, and Daniele Vannella.2013. *SemEval-2013 task 12: Multilingual word sense disambiguation.* In Second Joint Conference 3540 on Lexical and Computational Semantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), pages 222–231, Atlanta, GA, USA.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.

Roberto Navigli and Simone P. Ponzetto. 2012. *BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage MultilingualSemantic Network*. Artificial Intelligence Journal,193:217–250.

Tommaso Pasini and Roberto Navigli. 2017. *Train-O-Matic: Large-Scale Supervised Word Sense Disambiguation in Multiple Languages without Manual Training Data*. In Proc. of EMNLP: System Demonstrations, pages 78–88, Copenhagen, Denmark.

Sameer S. Pradhan, Edward Loper, Dmitriy Dligach,and Martha Palmer. 2007. *SemEval-2007 Task 17: English Lexical Sample, SRL and All Words*. In Proc. of SemEval-2007, pages 87–92, Stroudsburg,PA, USA.

Alessandro Raganato, Jose Camacho-Collados, and Roberto Navigli. 2017a. *Word Sense Disambiguation: A Unified Evaluation Framework and Empirical Comparison*. In Proc. of EACL, pages 99–110, Valencia, Spain.

Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli. 2017b. *Neural Sequence Learning Models for Word Sense Disambiguation*. In Proc. of EMNLP, pages 1167–1178, Copenhagen, Denmark.

Benjamin Snyder and Martha Palmer. 2004. *The English all-words task.* In Proceedings of the 3rd International Workshop on the Evaluation of Systems for the Semantic Analysis of Text (SENSEVAL-3), pages 41–43, Barcelona, Spain.

Tolga Uslu, Alexander Mehler, Daniel Baumartz, Alexander Henlein, and Wahed Hemati. 2018. *Fast-Sense: An Efficient Word Sense Disambiguation Classifier.* In Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, May 7-12, 2018, Miyazaki,Japan.

Dayu Yuan, Ryan Doherty, Julian Richardson, ColinEvans, and Eric Altendorf. 2016. *Semi-supervised Word Sense Disambiguation with Neural Models*. In Proc. of COLING, pages 1374–1385, Osaka,Japan.

This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Union.