Unsupervised Techniques for Extracting and Clustering Complex Events in News

Delia Rusu *
Jožef Stefan Institute and
Jožef Stefan International
Postgraduate School
Ljubljana, Slovenia
`delia.rusu@ijs.si`

James Hodson, Anthony Kimball
Bloomberg Labs
New York, NY, USA
`{jhodson2,akimball2}`
`@bloomberg.net`

Abstract

Structured machine-readable representations of news articles can radically change the way we interact with information. One step towards obtaining these representations is event extraction - the identification of event triggers and arguments in text. With previous approaches mainly focusing on classifying events into a small set of predefined types, we analyze unsupervised techniques for complex event extraction. In addition to extracting event mentions in news articles, we aim at obtaining a more general representation by disambiguating to concepts defined in knowledge bases. These concepts are further used as features in a clustering application. Two evaluation settings highlight the advantages and shortcomings of the proposed approach.

1 Introduction

Event extraction is a key prerequisite for generating structured, machine-readable representations of natural language. Such representations can aid various tasks like a) question answering, by enabling systems to provide results for more complex queries, b) machine translation, by enhancing different translation models or c) novelty detection, as a basis for computing geometric distances or distributional similarities. Event extraction primarily requires identifying *what* has occurred and *who* or *what* was involved, as well as the *time interval* of the occurrence. Additional information related to the event mention may include its *location*. Moreover, the event mention can also be labeled as belonging to a certain *event type*. Generally speaking, the goal of event extraction is to identify the *event trigger*, i.e. the

words that most clearly define the event, and the *event arguments*. For example, the event mention {*Hurricane Katrina struck the coast of New Orleans in August 2005*} belonging to the *occurrence of natural disasters* type of events includes the location of the disaster - *New Orleans* and the time of occurrence - *August 2005*. The event trigger is the verb *struck* while the other words represent the arguments of this event. The generalized form of the event mention is {*natural disaster occurred at location on date*}. Another similar event mention is {*Hurricane Katrina hit New Orleans*}, having the generalized form {*natural disaster occurred at location*}. Both event mentions can be generalized to {*natural disaster occurred at location*}, with the first event mention providing additional details regarding the date of the occurrence.

Supervised approaches imply classifying extracted event mentions according to predefined event types (Hong et al., 2011; Li et al., 2013). Lexical databases such as FrameNet (Baker et al., 1998), VerbNet (Schuler, 2005) or PropBank (Kingsbury and Palmer, 2002) can serve as training data. However, the coverage of this data is still limited, especially for domain-specific applications, and acquiring more labeled data can be expensive. Unsupervised approaches, on the other hand, are usually used to extract large numbers of untyped events (Fader et al., 2011; Nakashole et al., 2012; Alfonseca et al., 2013; Lewis and Steedman, 2013). Despite the coverage of these techniques, some of the extracted events can suffer from reduced quality in terms of both precision and recall. Distant supervision aims at mitigating the disadvantages of both supervised and unsupervised techniques by leveraging events defined in knowledge bases (Mintz et al., 2009).

In this work we investigate unsupervised techniques for extracting and clustering complex events from news articles. For clustering events we are using their generalized representation ob-

---

Event pattern	Explanation	Event mention
$\{entity_1, verb, entity_2\}$	an event having two named entities as arguments; verb modifiers are also included	{**Obama**, <u>apologized</u> for problems with, **ACA** rollout}
$\{sub, verb\}$ $\{sub, verb, obj\}$	a sequence of inter-related events having as arguments a subject and an object	{**Obama**, <u>apologized</u>} {**Obama**, <u>offered</u>, fix}
$\{sub, entity_1, verb, obj, entity_2\}$	an event having a subject, an object and two named entities as arguments	{**Hurricane Katrina**, <u>struck</u>, coast, of **New Orleans**}

Table 1: Examples of extracted events from text, where the event triggers are underlined and named entities are marked in bold.

tained by disambiguating events to concepts defined in knowledge bases. We are primarily looking at Bloomberg news articles which have a particular writing style: complicated sentence structures and numerous dependencies between words. In such cases a first challenge is to correctly identify the event trigger and all event arguments. Moreover, an event is described in news in different ways. Therefore, a second challenge is to capture the relations between event mentions. Thirdly, Bloomberg news mainly focuses on financial news reporting. Lexical databases such as FrameNet are intended for the general domain and do not cover most of the events described in financial news.

2   General Approach

We propose the following pipeline for extracting and clustering complex events from news articles. Firstly, we identify events based on the output of a dependency parser. Parsers can capture dependencies between words belonging to different clauses, enabling the detection of sequences of inter-related events. Section 3 describes two complementary approaches to event extraction which leverage dependencies between verbs and shortest paths between entities. Secondly, we obtain more general representations of the events by annotating them with concepts defined in (multilingual) knowledge bases (see Section 4). We refer to such generalized events as *complex events*. The knowledge base structure allows us to experiment with different levels of generalization. As a final step we apply a data-driven clustering algorithm to group similar generalized events. Clustering can be seen as an alternative to labeling events with predefined event types. Details regarding the clus-

tering approach can be found in Section 5.

3   Event Extraction

Most of the previous unsupervised information extraction techniques have been developed for identifying semantic relations (Fader et al., 2011; Nakashole et al., 2012; Lewis and Steedman, 2013). These approaches extract binary relations following the pattern $\{arg_1, relation, arg_2\}$. An example of such a relation is {*EBX Group Co., founder, Eike Batista*}, with the arguments of the *founder* relation being *EBX Group Co.* and *Eike Batista*. Similar to relations, events also have *arguments* such as named entities or time expressions (Li et al., 2013). In addition to the arguments, events are also characterized by the presence of an *event trigger*. In this work we consider *verbs* as event triggers, and identify events following the pattern:

$$\{verb, arg_1, arg_2,...,arg_n\},$$

where $arg_1, arg_2,...,arg_n$ is the list of event arguments. Aside from *named entities* and *time expressions*, we find additional valid argument candidates to be the *subject* or *object* of the clause. Together with the verb we also include its modifiers. Table 1 lists a few examples of extracted events.

In order to extract the events, we use the output of a dependency parser. Dependency parsing has been widely used for relation and event extraction (Nakashole et al., 2012; Alfonseca et al., 2013; Lewis and Steedman, 2013). There are various publicly-available tools providing dependency parse at the sentence level. We use the output of ZPar (Zhang and Clark, 2011), which implements an incremental parsing process with the de-
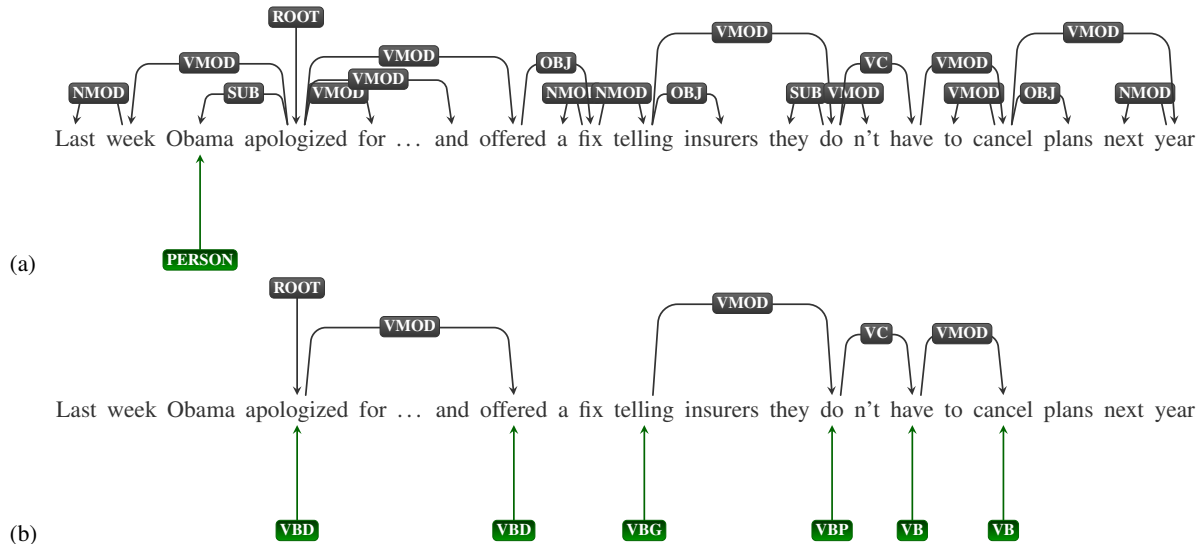
Figure 1: (a) Example sentence with highlighted word dependencies and named entities. (b) Example sentence marked with dependencies between verbs.

coding based on the Beam Search algorithm. The parser processes around 100 sentences per second at above 90% F-score.

The sentences that we are analyzing have a rather complex structure, with numerous dependencies between words. An example sentence is presented in Figure 1 (a). In this example there is a sequence of inter-related events which share the same subject: {*Obama apologized*} and {*Obama offered fix*}. Such events cannot be captured using only simple pattern matching techniques like the one implemented by REVERB (Fader et al., 2011). Other relations that are hard to identify are the lexically distant ones - this is the case with the dependence between the verb *apologized* and the verb *offered*. Consequently, we consider the following two complementary approaches to event extraction, both of them based on the output of the dependency parser:

1. Identifying verbs (including verb modifiers) and their arguments,

2. Identifying shortest paths between entities.

3.1 Identifying Verbs and Their Arguments

In order to identify inter-related events we extract dependency sub-trees for the verbs in the sentence. The verb sub-trees also allow us to extend the argument list with missing arguments. This is the case of the event mention {*Obama offered fix*}, where the subject *Obama* is missing.

The example sentence in Figure 1 (b) contains two verb sub-trees, the first one including the

nodes *apologized* and *offered* and the second one including the nodes *telling*, *do*, *have* and *cancel*. Once the sub-trees are identified, we can augment them with their corresponding arguments. For determining the arguments we use the REVERB relation pattern:

$$V|VP|VW^*P,$$

where V matches any verb, VP matches a verb followed by a preposition and VW^*P matches a verb followed by one or more nouns, adjectives, adverbs or pronouns and ending with a preposition.

3.2 Identifying Shortest Paths between Entities

Manual qualitative analysis of the events extracted using the approach described in Subsection 3.1 suggests that the *verbs and arguments* patterns do not cover all the events that are of interest to us. This is the case of events where two or more named entities are involved. For example, for the sentence in Figure 1 (a) we identify the event mentions {*Obama apologized*} and {*Obama offered fix*} using verb and argument patterns, but we cannot identify the event mention {*Obama apologized for problems with ACA rollout*} which includes two named entities: *Obama* and *ACA (Affordable Healthcare Act)*. We therefore expand our set of extracted events by identifying the shortest path connecting all identified entities. This is similar to the work of Bunescu and Mooney (2005) which
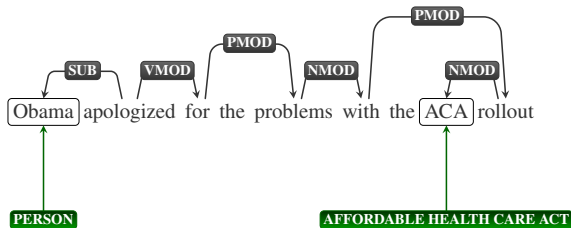
Figure 2: An event mention {*Obama apologized for problems with ACA rollout*} identified using the shortest path between entities approach.

Super-sense	Description
communi-cation.noun	communicative processes and contents
quantity.noun	quantities and units of measure
possession.noun	possession and transfer of possession
possession.verb motion.verb stative.verb	buying, selling, owning walking, flying, swimming being, having, spatial relations

Table 2: Example noun and verb super-sense labels and descriptions taken from WordNet.

build shortest path dependency kernels for relation extraction, where the shortest path connects two named entities in text.

We first use the Stanford Named Entity Recognizer (Finkel et al., 2005) to detect named entities and temporal expressions in the sentence. Next, we determine the shortest path in the dependency tree linking these entities. An example entity pattern discovered using this approach is shown in Figure 2.

4 Event Disambiguation

We disambiguate the events by annotating each word with WordNet (Fellbaum, 2005) super-senses and BabelNet (Navigli and Ponzetto, 2012) senses and hypernyms. WordNet super-senses offer the highest level of generalization for events, followed by BabelNet hypernyms and BabelNet senses. The choice of annotating with Word-Net concepts is motivated by its wide usage as a knowledge base covering the common English vocabulary. There are 41 WordNet super-sense classes defined for nouns and verbs. Table 2 depicts example WordNet super-senses with a short description.

Previous work on annotating text with WordNet super-senses mainly used supervised techniques. Ciaramita and Altun (2006) propose a sequential labeling approach and train a discriminative Hidden Markov Model. Lacking labeled data we investigate simple unsupervised techniques. Firstly, we take into account the first sense heuristic which chooses, from all the possible senses for a given word, the sense which is most frequent in a given corpus. The first sense heuristic has been used as a baseline in many evaluation settings, and it is hard to overcome for unsupervised disambiguation algorithms (Navigli, 2009). Secondly, we use a kernel to compute the similarity between the sentence and the super-sense definition. If x and y are

row vectors representing normalized counts of the words in the sentence and the words in the super-sense definition, respectively, the kernel is defined as:

$$k(x, y) = \frac{xy^T}{\|x\| \|y\|}$$

BabelNet is a multilingual knowledge base, mainly integrating concepts from WordNet and Wikipedia. The current version 2.0 contains 50 languages. We use the BabelNet 1.0.1 knowledge base and API to disambiguate words. As a starting point we consider the PageRank-based disambiguation algorithm provided by the API, but future work should investigate other graph-based algorithms.

5 Event Clustering

Events are clustered based on the features they have in common. We aim at obtaining clusters for the two types of extracted events: verbs and their arguments and shortest paths between entities in the dependency tree. The following two event patterns are considered for this experiment, for both event patterns: {*sub, verb, obj*} and {*sub, verb, obj, entities*}, where the verb and arguments can appear in the sentence in any order. Each event is described using a set of features. These features are extracted for the arguments of each event: the *sub*, *obj* and *entities*. The following feature combinations are used for each argument in the event argument list:

- WordNet super-senses,

- BabelNet senses,

- BabelNet hypernyms,

- WordNet super-senses, BabelNet senses and hypernyms.

For the WordNet experiments we include both disambiguation techniques - using the first sense heuristic and the kernel for determining the similarity between the sentence and the super-sense definition. Similar to the WordNet disambiguation approach we generate vectors for each event, where a vector x includes normalized counts of the argument features for the specific event. Thus we can determine the similarity between two events using the kernel defined in Section 4.

The Chinese Whispers algorithm (Biemann, 2006) presented in Algorithm 1 is used to cluster the events. We opted for this graph-clustering algorithm due to the fact that it is scalable and non-parametric. The highest rank class in the neighborhood of a given event e_i is the class of the event most similar to e_i.

Data: set of events E
Result: class labels for events in E

for $e_i \in E$ **do** class(e_i) = i;
while *not converged* **do**
    randomize order of events in E;
    for $e_i \in E$ **do**
        class(e_i) = highest ranked class in the neighborhood of e_i;
    end
end
Algorithm 1: Chinese Whispers Algorithm.

---

6 Evaluation

We evaluated the extracted events, as well as the clusters obtained for the disambiguated events. For each set of experiments we prepared a dataset by sampling Bloomberg news articles.

As there is no benchmark dataset for the news articles that we are analyzing, we propose to evaluate event extraction in terms of completeness. Clustering evaluation is done based on the model itself, and for different feature combinations. In what follows we describe the evaluation setting in more detail.

6.1 Event Extraction Evaluation

The evaluation dataset consists of a sample of 23 stories belonging to the *MEDICARE* topic, con-taining a total of 1088 sentences. The event extraction algorithms yields 229 entity paths and 515 verb and argument events. Each event is assessed in terms of completeness; an event is deemed to be *complete* if all event elements (the event trigger and the arguments) are correctly identified. We only analyze two event patterns: {*sub, verb, obj*} and {*sub, verb, obj, entities*}, as events belonging to other patterns are rather noisy. Two annotators independently rate each event with 1 if all event elements are correctly identified, and 0 otherwise. Note that incomplete events receive a 0 score. Cohen's kappa coefficient (Cohen, 1960) of inter-annotator agreement for this experiment was 0.70. The entity path approach correctly identified 78.6% of the entities while the verb arguments approach identified 69.1% of the events. Events obtained using entity paths tend to have a higher number of arguments compared to the verb arguments approach; this explains the higher score obtained by this technique.

6.2 Clustering Evaluation

As we do not know the cluster labels a priori, we opt for evaluating the clusters using the model itself. To this end, we use the *Silhouette Coefficient* (Kaufman and Rousseeuw, 1990); we plan to investigate other clustering evaluation metrics in future work. The Silhouette Coefficient is defined for each sample, and it incorporates two scores:

$$s = \frac{b - a}{max(a, b)},$$

where a is the mean distance between a sample and all other points within the same class whereas b is the mean distance to all other points in the next nearest class. To determine the coefficient for a set of samples one needs to find the mean of the coefficient for each sample. A higher coefficient score is associated with a model having better defined clusters. The best clustering model will obtain a Silhouette coefficient of 1, while the worst one will obtain a -1 score. Values close to 0 imply overlapping clusters. Negative values signify that the model assigned samples to the wrong cluster, as a different cluster is more similar.

The evaluation dataset comprises 325 *MEDICARE* news articles and 16,450 sentences. In this dataset we identify 7,491 verb and argument events and 2,046 shortest path events. Table 3 shows example events belonging to two event
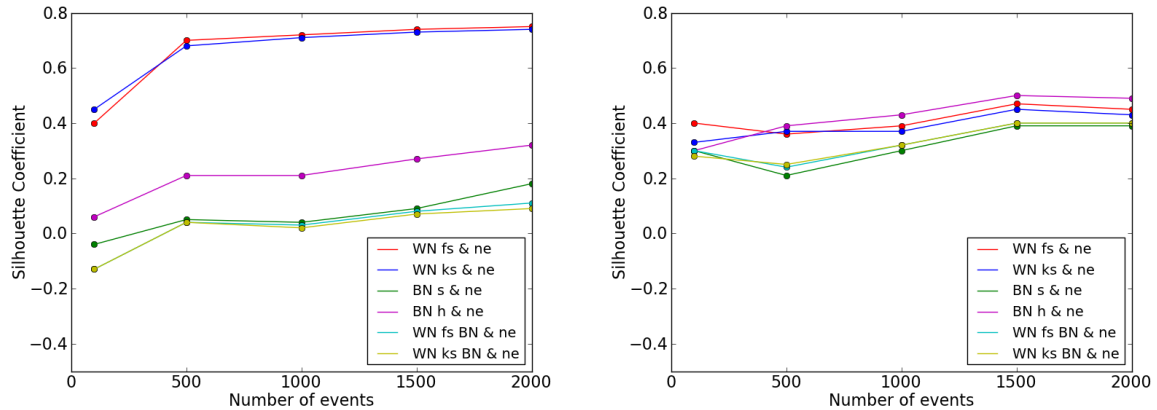
Figure 3: Clustering evaluation results for verbs and arguments (left) and shortest paths between entities (right) events, using different feature combinations.

clusters. The first cluster is obtained by extracting verb argument events while the second cluster is composed of shortest entity path events.

In Figure 3 we show clustering evaluation results for the (a) verbs and arguments and (b) shortest paths between entities, using different feature combinations. As expected, the best results are obtained in the case of the WordNet super-senses, which are the most generic senses assigned to the events. There is less overlap among the BabelNet senses and hypernyms, although results improve as more data is available. The results also mark the difference between the two types of events: verbs and arguments versus shortest paths between entities. Events extracted using the entity path approach tend to have a higher number of arguments, which in turn implies a richer set of features. This explains the higher scores obtained in the case of shortest path events compared to verb argument events.

7 Related Work

The event extraction task have received a lot of attention in recent years, and numerous approaches, both supervised and unsupervised, have been proposed. This section attempts to summarize the main findings.

Supervised approaches. These approaches classify events based on a number of predefined event types. A popular dataset is the NIST Automatic Content Extraction (ACE) corpora (Doddington et al., 2004) which consists of labeled relations and events in text. State-of-the-art approaches mainly use sequential pipelines to separately identify the event trigger and the arguments (Hong et al., 2011). More recently Li et al. (2013) propose a joint framework which considers event triggers and arguments together. Their model is based on structured perceptron with Beam Search. In another line of work (Alfonseca et al., 2013) events extracted in an unsupervised manner from the output of a dependency parser are the building blocks of a Noisy-OR model for headline generation. Tannier and Moriceau (2013) identify event threads in news, i.e. a succession of events in a story, using a cascade of classifiers.

Mintz et al. (2009) propose a distant supervision approach. They use Freebase relations and find sentences which contain entities appearing in these relations. From the sentences the authors extract a number of textual features which are used for relation classification. Dependency parsing features are used to identify relations that are lexically distant.

Unsupervised approaches. Most unsupervised approaches have been tailored to identifying relations in text. Fader et al. (2011) extract relations and their arguments based on part-of-speech patterns. However, such patterns fail to detect lexically distant relations between words. Therefore, most state-of-the-art unsupervised approaches also rely on sentence parsing. For example, Lewis and Steedman (2013) extract cross-lingual semantic relations from the English and French parses of sentences. Relational patterns extracted from the sentence parse tree have also been generalized to syntactic-ontologic-lexical patterns using a frequent itemset mining approach (Nakashole et al., 2012). Poon and Domingos (2009)

31

Event	Features
{owners are being incentivized to drop their health insurance coverage}	noun.person noun.possession
{analysts are not permitted receive compensation directly}	noun.person noun.possession
{HHS General issued report in July 2013}	noun.person noun.group noun.time
{lawmakers asked Kathleen Sebelius to respond by December 6}	noun.person noun.time

Table 3: Example events belonging to two event clusters. Each event is assigned WordNet supersense features.

learn a semantic parser using Markov logic by converting dependency trees into quasi-logical forms which are clustered.

DIRT (Lin and Pantel, 2001) is an unsupervised method for discovering inference rules from text. The authors leverage the dependency parse of a sentence in order to extract indirect semantic relations of the form "X $relation$ Y" between two words X and Y. Inference rules such as "X $relation_1$ $Y \approx X$ $relation_2$ Y" are determined based on the similarity of the relations.

ALICE (Banko and Etzioni, 2007) is a system that iteratively discovers concepts, relations and their generalizations from the Web. The system uses a data-driven approach to expand the core concepts defined by the WordNet lexical database with instances from its Web corpus. These instances are identified by applying predefined extraction patterns. The relations extracted using TextRunner (Banko et al., 2007) are generalized using a clustering-based approach.

Our aim is to identify events rather than any relation between two concepts. We therefore propose different extraction patterns based on the dependency parse of a sentence which allow us to detect event triggers and event arguments that can be lexically distant. Events are generalized by mapping them to concepts from two different knowledge bases (WordNet and BabelNet), allowing us to experiment with multiple levels of generalization.

8 Conclusions and Future Work

In this work we investigated different unsupervised techniques for extracting and clustering complex events from news articles. As a first step we proposed two complementary event extraction algorithms, based on identifying verbs and their arguments and shortest paths between entities, respectively. Next, we obtained more general representations of the event mentions by annotating the event trigger and arguments with concepts from knowledge bases. The generalized arguments were used as features for a clustering approach, thus determining related events.

As future work on the event extraction side, we plan to improve event quality by learning a model for filtering out noisy events. In the case of event disambiguation we are looking into different graph-based disambiguation algorithms to enhance concept annotations.

Acknowledgments

References

[Alfonseca et al.2013] Enrique Alfonseca, Daniele Pighin, and Guillermo Garrido. 2013. Heady: News headline abstraction through event pattern clustering. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1243–1253.

[Baker et al.1998] Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The Berkeley Framenet Project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1*, pages 86–90. Association for Computational Linguistics.

[Banko and Etzioni2007] Michele Banko and Oren Etzioni. 2007. Strategies for lifelong knowledge extraction from the web. In *Proceedings of the 4th international conference on Knowledge capture*, pages 95–102. ACM.

[Banko et al.2007] Michele Banko, Michael J Cafarella, Stephen Soderland, Matthew Broadhead, and Oren Etzioni. 2007. Open information extraction for the web. In *IJCAI*, volume 7, pages 2670–2676.

[Biemann2006] Chris Biemann. 2006. Chinese whispers: an efficient graph clustering algorithm and its application to natural language processing problems. In *Proceedings of the first workshop on graph based methods for natural language processing*, pages 73–80. Association for Computational Linguistics.

[Bunescu and Mooney2005] Razvan C Bunescu and Raymond J Mooney. 2005. A shortest path dependency kernel for relation extraction. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 724–731. Association for Computational Linguistics.

[Ciaramita and Altun2006] Massimiliano Ciaramita and Yasemin Altun. 2006. Broad-coverage sense disambiguation and information extraction with a supersense sequence tagger. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 594–602. Association for Computational Linguistics.

[Cohen1960] Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

[Doddington et al.2004] George R Doddington, Alexis Mitchell, Mark A Przybocki, Lance A Ramshaw, Stephanie Strassel, and Ralph M Weischedel. 2004. The automatic content extraction (ace) program-tasks, data, and evaluation. In *LREC*.

[Fader et al.2011] Anthony Fader, Stephen Soderland, and Oren Etzioni. 2011. Identifying relations for open information extraction. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1535–1545. Association for Computational Linguistics.

[Fellbaum2005] Christiane Fellbaum. 2005. Wordnet and wordnets. In Keith et al. Brown, editor, *Encyclopedia of Language and Linguistics*, pages 665–670. Oxford: Elsevier, second edition.

[Finkel et al.2005] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 363–370. Association for Computational Linguistics.

[Hong et al.2011] Yu Hong, Jianfeng Zhang, Bin Ma, Jianmin Yao, Guodong Zhou, and Qiaoming Zhu. 2011. Using cross-entity inference to improve event extraction. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 1127–1136. Association for Computational Linguistics.

[Kaufman and Rousseeuw1990] Leonard Kaufman and Peter J Rousseeuw. 1990. *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

[Kingsbury and Palmer2002] Paul Kingsbury and Martha Palmer. 2002. From treebank to propbank. In *Proceedings of the International Conference on Language Resources and Evaluation LREC*.

[Lewis and Steedman2013] Mike Lewis and Mark Steedman. 2013. Unsupervised induction of cross-lingual semantic relations. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing EMNLP*, pages 681–692.

[Li et al.2013] Qi Li, Heng Ji, and Liang Huang. 2013. Joint event extraction via structured prediction with global features. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*.

[Lin and Pantel2001] Dekang Lin and Patrick Pantel. 2001. Dirt - discovery of inference rules from text. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 323–328. ACM.

[Mintz et al.2009] Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. Distant supervision for relation extraction without labeled data. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2*, pages 1003–1011. Association for Computational Linguistics.

[Nakashole et al.2012] Ndapandula Nakashole, Gerhard Weikum, and Fabian Suchanek. 2012. Patty: a taxonomy of relational patterns with semantic types. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 1135–1145. Association for Computational Linguistics.

[Navigli and Ponzetto2012] Roberto Navigli and Simone Paolo Ponzetto. 2012. Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.

[Navigli2009] Roberto Navigli. 2009. Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR)*, 41(2):10.

[Poon and Domingos2009] Hoifung Poon and Pedro Domingos. 2009. Unsupervised semantic parsing. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pages 1–10. Association for Computational Linguistics.

[Schuler2005] Karin Kipper Schuler. 2005. Verbnet: A broad-coverage, comprehensive verb lexicon.

[Tannier and Moriceau2013] Xavier Tannier and Véronique Moriceau. 2013. Building event threads out of multiple news articles. Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing EMNLP.

33

[Zhang and Clark2011] Yue Zhang and Stephen Clark. 2011. Syntactic processing using the generalized perceptron and beam search. *Computational Linguistics*, 37(1):105–151.