Cross-Scenario Inference Based Event-Event Relation Detection

Yu Hong, Jingli Zhang, Rui Song, and Jianmin Yao(✉)

Computer Information Processing Technology of Soochow University, Suzhou 215006, Jiangsu, China
tianxianer@gmail.com, jlzhang05@gmail.com, cnsr27@gmail.com, jyao@suda.edu.cn

Abstract. Event-Event Relation Detection (RD_{2e}) aims to detect the relations between a pair of news events, such as `Causal` relation between `Criminal` and `Penal` events. In general, RD_{2e} is a challenging task due to the lack of explicit linguistic feature signaling the relations. We propose a cross-scenario inference method for RD_{2e}. By utilizing conceptualized scenario expression and graph-based semantic distance perception, we retrieve semantically similar historical events from Gigaword. Based on explicit relations of historical events, we infer implicit relations of target events by means of transfer learning. Experiments on 10 relation types show that our method outperforms the supervised models.

Keywords: Relation detection · Cross scenario · Semantic distance

1 Introduction

Event relation refers to the way in which an event exerts an influence on the other, such as `Conditionality`, `Causality`, `Concession`, etc. For example, the event "*Snowden was trained as a secret agent*" is the necessary `condition` of the event "*he successfully escaped scrutiny*".

The goal of a RD_{2e} system is to automatically detect the implicit relations between event mentions, some of which occur in a single document while others different documents (cross-document relations). In this paper, we limit our discussion to the ground-truth event mentions that have been manually extracted from news articles. As defined in ACE and KBP (Ji and Grishman, 2008; Liao and Grishman, 2010), an event mention is a text span that contains a trigger and the closely related arguments.

In this paper, we propose a cross-scenario inference approach, which performs with minimal supervision. The fundamental behind the inference approach is that if there are some historical events similar to the target events, and the relations between the historical events are explicit, thus the implicit relations of the target can be inferred accordingly using those explicit relations. For example, we can determine the unknown relation in (1) as conditionality, because such a relation is held between the similar historical events in (2) and has been explicitly signalled by the conjunction "*thus*".

M. Zhang et al. (Eds.): NLPCC 2018, LNAI 11109, pp. 257–267, 2018.
https://doi.org/10.1007/978-3-319-99501-4_22

(1) **Target:** *Edward Snowden was trained as a secret agent. The certification would have given him some of the skills he needed to escape scrutiny.* (in 2013)

T_1 : Snowden was trained as a secret agent. T_2 : Snowden escaped scrutiny.

S_{T1} : Person AND Education-teaching AND People-by-vocation

S_{T2} : Person AND Avoiding AND Scrutiny

(2) **Historical:** *Edward Howard, a CIA case officer, was trained as a spy and **thus** eluded FBI surveillance.* (in 1993)

H_1 : Howard was trained as a spy. H_2 : Howard eluded surveillance.

S_{H1} : Person AND Education-teaching AND People-by-vocation

S_{H2} : Person AND Avoiding AND Scrutiny

Methodologically, we leverage FrameNet, a frame-level semantic dictionary, to the description of event scenarios, transforming the words in an event mention to the semantic frame tags by looking-up. This allows the similarity computing at the level of semantics, and makes it easy to match semantically similar event scenarios (see the frame tags in $S_{T1\&T2}$ and $S_{H1\&H2}$). In addition, we utilize the labelled conjunctions (e.g., "*thus*") in the corpus of Penn Discourse Tree Bank (PDTB), so as to reinforce the explicit-to-implicit relation inference.

The rest of the paper is organized as below: Sect. 2 overview the related work; Sect. 3 presents the inference approach; Sect. 4 gives a refined scenario model; Sect. 5 provides the test results; we conclude the paper in Sect. 6.

2 Related Work

Girju et al. [8] use lexico-syntactic patterns of noun phrases (NPs) and verbs, <`cause`-NP, `causal` verb, `effect`-NP>, to automatically extract the `Causal` relation within one sentence. Soon thereafter, Chang and Choi [5] revise Girju et al. (2002)'s pattern, using lexical pair (LP) and cue phrase (e.g., *due to*) to generate the new version, <LP, cue, LP>. The LP and cue-phrase probabilities are estimated in raw corpus by EM procedure, and jointly used in a naive Bayes classifier. Abe et al. [1] go further by using co-occurrence probability.

Sufficient attentions have been given to the optimization of `causal` relation recognition from different aspects: fine-grained `causal` relation classification [10], feature selection [3], syntactic, graphical and sophisticated patterns [11], use of discourse structure [6], rule generation [21], and domain adaptation [22].

The pilot study on event-oriented temporal relation derives from Mani et al. [14] and Lapata and Lascarides [12], both of whom focus on machine learning of temporal relations. In the past decade, the SemEval [20] has promoted a great deal of experimental study, including on the grammatical, syntactic, semantic and ordering feature-based temporal relation classification [4], and the validate of sequence labeling and Markov Logic [25]. SemEval-2015 [16] defines cross-sentence and cross-document event ordering tasks, both of which go a step further than the previous challenges. Most recently, deep neural networks have shown promising results. Santo et al. [24] proposed the Ranking CNN(CR-CNN) model with a class embedding matrix. Miwa and Bansal [17] similarly observed that LSTM-based RNNs are outperformed by models using CNN due to limited linguistic structure captured in the network architecture.

3 Cross-Scenario Inference

3.1 Scenario Expression and Conceptualization

We model the scenario as a vector of scene elements. Given an event mention, we use the content words (nouns, verbs, adjectives and adverbs) in the mention as the representation of the scene elements. They are able to cover most types of scene elements, such as objects (e.g., participants and attributes), activities (event triggers) and status (time, locations, surroundings and other conditions): **Mention-** *Edward fled in 1983.* **Scenario-** <*Edward*, *fled*, *1983*>; **Concept-** <`PER`, `Fleeing`, 19830000>.

We conceptualize a scenario by transforming scene elements (content words) into their semantic frames. We brief the definition of frame semantics in the next section. In addition, we conceptualize named entities with the labels of entity types, including `Person`, `Location` and `Organization`. We generalize time expressions with the pattern `yyyy\mm\dd`. The granularity is set as `day`. See a conceptualized scenario in the above example.

3.2 Frame Semantics (FrameNet)

Semantic frame is defined as the concept of lexical units (words or phrases) that share similar semantic context [7]. It helps to identify the homogeneous lexical units. It is noteworthy that such units aren't definitely synonymous. For example, both the word *boat* and *plane* comply with the frame `Vehicle` but are different in sense.

Scenario conceptualization by semantic frame facilitates the discovery of semantically similar event mentions. See the following examples, which adhere to the same scenario at the level of semantics: **Mention1-** *He took a plane and ran away.* **Mention2-** *He fled on a boat.* **Concept-** <`PER`, `Fleeing`, `Vehicle`>.

We obtain the frames of words from FrameNet. FrameNet is a machine-readable lexical database of English [23]. It alphabetically indexes more than 10,000 frames. Each frame tag (`Nam`) corresponds to a cluster of lexical units (`Lus`) that evoke the frame, along with the related frames (also named as frame-to-frame relation, abbr., `Ffr`). Listed in Table 1 are the components of the frame `Avoiding`. The key behavioral elements in (1) and (2), i.e., *escape* and *elude*, both are the `Lus` of `Avoiding`.

3.3 Explicit-To-Implicit (ETI) Relation Assignment

Let the scenarios of a pair of target events be $< e_{t1}, e_{t2} >$. We assume that the relation of the events is determined by their scenarios, i.e., $R^t_{imp} \Leftarrow < e_{t1}, e_{t2} >$,

Table 1. Semantic frame (*e.g., scrutiny).*

Nam	Avoiding
Def	An agent avoids an undesirable situation under certain circumstances
Ffr	Inherits from Intentionally_act; Is inherited by Dodging and Evading
Lus	*avoid* (v), *avoidance* (n), ***escape* (v)**, ***elude* (v)**, *keep away* (v), etc.

where R^t_{imp} denotes an unknown (implicit) relation. For a pair of historical events that contain an known (explicit) relation R^h_{exp}, similarly, we have that $R^h_{exp} \Leftarrow < e_{h1}, e_{h2} >$. If the scenarios of the historical events are conceptually similar to that of the target, we can conclude that R^t_{imp} is very likely to be the same as R^h_{exp}. So, we produce the ETI relation assignment:

$$R^t_{imp} \Leftarrow R^h_{exp}, \text{if} < \mathrm{e_{t1}}, \mathrm{e_{t2}} > \Leftrightarrow < \mathrm{e_{h1}}, \mathrm{e_{h2}} >$$

We name the above assignment process as cross-scenario inference. Figure 1 shows an inference process from the historical defection events H_1 and H_2 to the targets T_1 and T_2 (The full contents of the events can be accessed in (1)).

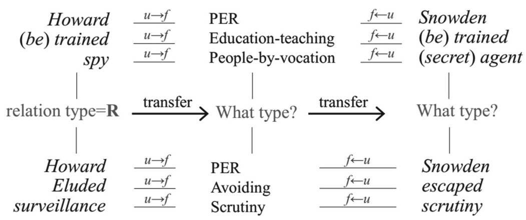

Fig. 1. Cross-scenario relation inference. Either "u→f" or "f←u" refers to a translation process from a lexical unit (word or phrase) to its semantic frame in a specific context. The left column shows historical events, the middle column shows conceptual scenario and the right column shows the targets.

It is strictly required by cross-scenario inference to obtain historical events that contain an explicit relation. The relation is used as prior knowledge to support ETI relation assignment, assigning explicit relation of reliable homogeneous historical events to the target.

3.4 Connective-Relation (C-R) Alignment

If two event mentions is connected by a connective, or included in the scope of the connective [27], they contain an explicit relation. We detect the explicit relation type by a connective-relation alignment approach. The connective "*thus*", for example, aligns with (i.e., signals) the `conditional` relation.

Table 2. Partial relation-connective mapping table

`Causal`- *because*	`Competitive`- *unlike*
`Temporal`- *before*	`Concessive`- *however*
`Subevent`- *involving*	`Disjunctive`- *except*
`Conjunctive`- *in addition*	`Conditional`- *in order to*

We aligns a connective with its relation type by searching it in a one-to-many relation-connective mapping table (See a part of the table in Table 2).

We initialize the table based on the ground-truth connectives and relation samples in PDTB 2.0 corpus[1] [19]. We filter the ambiguous connectives. For example, the connective "*since*" is ambiguous because it may signal a `causal` relation (meaning"*in view of the fact that*") or `temporal` ("*from a past time until now*"). We collected all pairs of events that contain an explicit relation beforehand from 8.2M documents in the 2003 English Gigaword[2]. They're employed as the only data available for the acquisition of the desired similar historical events.

3.5 Scenario Similarity Calculation

We acquire similar historical events based on scenario similarity to the target. We use a frame-based vector space to model event scenarios, in which each dimension indicates a semantic frame. We instantiate the vector by weighting the frames with 1 and 0. Given a frame, it will be weighted by 1 if the word of the frame occurs in the event mention, otherwise 0. We measure the similarity by the cosine metric. Then, we measure the joint similarity between a pair of historical events and the target based on maximum likelihood principle:

$$tuple(e_{hi}, e_{tj}) = argmax S(e_{hi}, e_{tj}), i, j \in 1, 2 \tag{1}$$

$$S(e_{h1}, e_{h2}, e_{t1}, e_{t2}) = \frac{S(e_{hi}, e_{tj})S(\overline{e_{hi}}, \overline{e_{tj}})}{exp(S(e_{hi}, e_{tj}) - S(\overline{e_{hi}}, \overline{e_{tj}}))} \tag{2}$$

where, e_{hi} and e_{tj} are, respectively, the historical and target events that yield the maximum S, while $\overline{e_{hi}}$ and $\overline{e_{tj}}$ are the other two. The exponential function, i.e., $exp(*)$, ensures that the value of the denominator is nonzero.

3.6 Confidence Level Computing

We search similar events from the data source mentioned in Sect. 3.4, and rank them by scenario similarity to the target. We use top n search results as eligible **R**eference **S**ource (RS, i.e., set of similar historical events) for cross-scenario inference.

[1] https://www.seas.upenn.edu/~pdtb/.
[2] https://catalog.ldc.upenn.edu/LDC2003T05.

For the case that there is only one type of relation R_{exp} occurred in RS, we directly assign it to R_{imp} of the target events as the RD_{2e} result. If there are multiple types occurred, we will evaluate their confidence, and select the most confident one as the RD_{2e} result. Given a type of relation, we measure its confidence level based on the probability it occurred in RS and the ratings:

$$C(\Re) = \frac{f(\Re)}{\prod_{i\in n} rank(\Re)} \tag{3}$$

where, $\Re$ is a relation type, $f(\Re)$ denotes the frequency of $\Re$ occurred in RS. For a reference sample that contains $\Re$, $rank(\Re)$ denotes its rating in the similarity-based ranking list of RS.

4 Graph Based Scenario Modeling

It was found that some lexical units's senses adhere to a semantic frame at different levels, some strictly, others loosely. In the cases, the scenario modeling mentioned in section 2.5 fails to identify the difference, assigning the same weight to the units at the dimension of the frame, either 1 if they occur in event mentions, otherwise 0. This causes biases in similarity calculation among scenarios. In this section, we use a lexicon-frame graph to improve scenario modeling.

4.1 Hybrid Lexicon-Frame Graph

Given a lexical unit u, if its sense conforms to a frame f and the semantic distance is d, the weight of f can be measured by the reciprocal of d, i.e., $w_u(f) = 1/d$.

In order to facilitate the measurement of semantic distance, we built a hybrid lexicon-frame graph based on `Ffr`. `Ffr` is a connection between frames, indicating a semantic relation. There are totally 8 `Ffr` types defined in FrameNet, including `Inheritance`, `Subframe`, `Precedence`, `Perspective`, `Inchoation`, etc. By tracing along the `Ffr`s in an end-to-end manner, we can find out new semantically related unit-frame, unit-unit or frame-frame pairs. They are useful for the generation of a hybrid directed lexicon-frame graph. In practice, we only obtained a collection of subgraphs because the manufactured `Ffr` edges are incompleted. We brought the subgraphs into use in our experiments. See an example in Fig. 2.

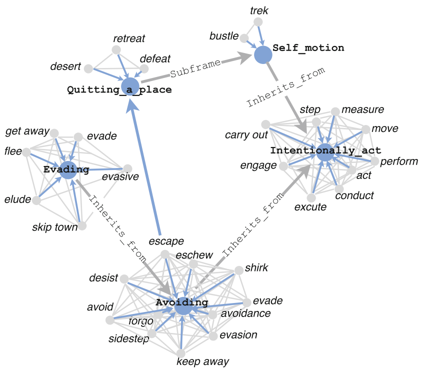

Fig. 2. Hybrid lexicon-frame subgraph

For two nodes in the graph, we specify the length of the shortest path between the nodes as their semantic distance. The length is measured by the number of edges between the nodes. We take into consideration all frames that are directly or indirectly connected with a lexical unit in the graph, and use them as semantically related to the unit. Based on the semantic distance, the refined scenario model will assign a weight $w_u(f)$ to every frame f related to a unit u, for every unit occurred in the event mention. Still, it assigns 0 to other frames. *escape* → `Avoiding` → `Intentionally_act`.

4.2 Probabilistic Scenario Model

We employ a probabilistic matrix as a substitute for VSM. Each matrix row is an uneven projection on all semantic frames, denoting the probability distribution of a specific lexical unit over the frames. The probability is estimated by semantic distance in the same way as the weighting method $w_u(f)$, except that it is normalized by the sum of the weights:

$$p_u(f) = \frac{w_u(f)}{\sum_f w_u(f)} \tag{4}$$

Accordingly an event scenario is represented as a lexicon-frame matrix, where each row indicates the probability model $p_u(f)$ of a lexical unit u over all frames. If a unit doesn't occur in the mention, the probability $p_u(f)$ over the frames in the corresponding row will all be set as 0. Thus, the similarity between events can be estimated by the agreement of their lexicon-frame probability matrixes.

Technically we use the Kullback-Liebler (KL) divergence [15] as a measure of the agreement $A(*,*)$. A significant KL divergence is equivalent to little agreement. The agreement between matrixes is fully determined by that between all their rows. In addition, we involve the standard deviation in the measurement of partial agreement between a row in a matrix and all in another. The deviation is employed to reduce the effect of general rows on the measure of agreement. A general row corresponds to a unit who has similar $p_u(f)$ over the usual frames. We measure the full agreement between scenarios by:

$$S(e_h, e_t) \propto A(P_h, P_t) = \frac{G}{D_{KL}(P_h, P_t)}, \qquad g(u) = \sqrt{\frac{\sum_{f \in F}(p(f) - \overline{p(f)})^2}{n-1}} \tag{5}$$

$$D_{KL}(P_h, P_t) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{f \in F} p_{u_i,h}(f)\; log\frac{p_{u_i,h}(f)}{p_{u_j,t}(f)}, \qquad G = \sum_{i=1}^{n}\sum_{j=1}^{n} g_h(u_i)g_t(u_j) \tag{6}$$

5 Experiments

5.1 Corpus and Evaluation Metric

We follow Yu et al. [9] relation schema (see Table 3). We collected 828 event mentions from 24 expository texts in the American National Corpus (ANC)[3].

[3] https://framenet.icsi.berkeley.edu/fndrupal/fulltextIndex.

There are 968 pairs of events found related to each other, in which 569 are structurally adjacent, 330 cross-sentence and 69 cross-document. Three students who major in computational linguistics were employed to label the relations. The agreement among them is fairly good (kappa = 0.69 at worst).

In addition, we annotate the relations for 303 pairs of related events in 159 ACE2005 news stories . They were used for the purpose of evaluating the cross-domain stability of our RD_{2e} system. We employ Precision (P), Recall (R) and F-score (F) as the evaluation metrics.

5.2 Main Results

We test our basic RD_{2e} on ANC. The frame-based VSM was used. Table 4 shows the performance for each existing relation type in ANC. The performance for the two temporal relation types (*Bef.* & *Eql.*) is different from the true state. Annotators were unable to detect temporal relations for most samples in ANC due to lack of visible evidence, except the 80 ones shown in Table 3. The scores of *Bef.* & *Eql.*, hence, are yielded only for them not all. For uncommon relation types (Table 3), our RD_{2e} shows poor recall (Table 4). We suggest that if a relation type is uncommon, there is little explicit and definitive evidence in neither our mind nor an electronic database available for inferring it.

Table 3. Relational Schema

Relation(*REL.*)	Num
Before(*Bef.*)	29
Equal(*Eql.*)	51
Opposite(*Opp.*)	91
Variance(*Var.*)	31
Cause(*Cas.*)	131
Condition(*Con.*)	197
Conjunction(*Coj.*)	203
Concession(*Coc.*)	43
Coreference(*Cor.*)	89
Sub-event(*Sub.*)	103

Table 4. Performance for all types

REL.	# P	# R	# F
Bef.	0.61	0.31	0.41
Eql.	0.82	0.25	0.38
Opp.	0.51	0.67	0.58
Var.	0.65	0.30	0.41
Cas.	0.79	0.66	0.72
Con.	0.61	0.43	0.50
Coj.	0.39	0.73	0.51
Coc.	0.85	0.16	0.27
Cor.	0.57	0.29	0.38
Sub.	0.31	0.69	0.43

We built RD_{2e} systems by using different models proposed in the paper, including word-level and frame-level scenario description along with VSM and probability (PRO) matrix models. We evaluate them on both ANC and ACE. See performance in Table 5. It is illustrated that the frame based scenario conceptualization and the generalized probability model are both conducive to cross-scenario inference. Moreover, they show better stability than the word-level model in different domains.

Table 5. Performance (F) on ANC and ACE

System construction	# ANC	# ACE
Lexical Unit + VSM	0.45	0.41
Frame + VSM	0.51	0.50
Frame + PRO	0.56	0.53

We reproduce some state-of-the-art relation detection methods. Some of them were employed in determining discourse-oriented structural relations between adjacent text spans, such as the supervised relation classification (CLA) based on syntactic constituent and dependency [13] or multiple types of linguistic features [18], as well as language model (LG) based connective and relation prediction [26]. Others focus on event-event relation problems, such as the use of coocurrence (Coo) information of pairwise phrases or pattens for relation analysis [2]. We evaluate the methods on the samples of different structures in ANC.

Table 6. F-measure for event pairs in adjacent (ADJ), CroS and CroD structures

System construction	# ADJ	# CroS	# CroD
Syntactic Features (CLA)	0.50	N/A	N/A
All features (CLA)	0.61	N/A	N/A
Connective (LG)	0.51	0.30	0.33
Phrase (Coo)	0.49	0.39	0.35
Pattern (Coo)	0.56	0.41	0.40
Our best	**0.55**	**0.59**	**0.53**

Table 6 shows their performance and our best. The classifiers didn't yield any performance for both the structurally cross-sentence (CroS) and cross-document (CroD) test samples, as in the cases, the syntactic features are non-existent and therefore unavailable. Beyond that, we can see that our unsupervised cross-scenario inference method yields insignificant performance loss for the cross-sentence and cross-document samples. It means the method is more credible than others when provided few available structural-level linguistic features.

6 Conclusion

We propose a cross-scenario inference method for event-event relation detection. It outperforms the state of the art in low resource settings. This can facilitate the fast detection of relations for the newly occurred related events, although the detection results are not quite exact (pseudo relations). Encouraged by this observation, we will focus on the refinement of pseudo relations by using semi-supervised learning method.

Acknowledgements. This work was supported by the national Natural Science Foundation of China via Nos. 2017YFB1002104, 61672368 and 61672367.

References

1. Abe, S., Inui, K., Matsumoto, Y.: Acquiring event relation knowledge by learning cooccurrence patterns and fertilizing cooccurrence samples with verbal nouns. In: IJCNLP, pp. 497–504 (2008)
2. Abe, S., Inui, K., Matsumoto, Y.: Two-phased event relation acquisition: coupling the relation-oriented and argument-oriented approaches. In: COLING, pp. 1–8 (2008)
3. Blanco, E., Castell, N., Moldovan, D.I.: Causal relation extraction. In: LREC (2008)
4. Caselli, T., Fokkens, A., Morante, R., Vossen, P.: Spinoza vu: an NLP pipeline for cross document timelines. In: SemEval-2015 p. 787 (2015)
5. Chang, D.-S., Choi, K.-S.: Causal relation extraction using cue phrase and lexical pair probabilities. In: Su, K.-Y., Tsujii, J., Lee, J.-H., Kwong, O.Y. (eds.) IJCNLP 2004. LNCS (LNAI), vol. 3248, pp. 61–70. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-30211-7_7
6. Do, Q.X., Chan, Y.S., Roth, D.: Minimally supervised event causality identification. In: EMNLP, pp. 294–303 (2011)
7. Gildea, D., Palmer, M.: The necessity of parsing for predicate argument recognition. In: ACL, pp. 239–246 (2002)
8. Girju, R., Moldovan, D.I., et al.: Text mining for causal relations. In: FLAIRS, pp. 360–364 (2002)
9. Hong, Y., Zhang, T., O'Gorman, T., Horowit-Hendler, S., Ji, H., Palmer, M.: Building a cross-document event-event relation corpus. In: LAW X, p. 1 (2016)
10. Inui, T., Inui, K., Matsumoto, Y.: Acquiring causal knowledge from text using the connective marker tame. TALIP **4**(4), 435–474 (2005)
11. Ittoo, A., Bouma, G.: Extracting explicit and implicit causal relations from sparse, domain-specific texts. In: Muñoz, R., Montoyo, A., Métais, E. (eds.) NLDB 2011. LNCS, vol. 6716, pp. 52–63. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22327-3_6
12. Lapata, M., Lascarides, A.: Learning sentence-internal temporal relations. J. Artif. Intell. Res. (JAIR) **27**, 85–117 (2006)
13. Lin, Z., Ng, H.T., Kan, M.Y.: A PDTB-styled end-to-end discourse parser. Nat. Lang. Eng. **20**(02), 151–184 (2014)
14. Mani, I., Verhagen, M., Wellner, B., Lee, C.M., Pustejovsky, J.: Machine learning of temporal relations. In: COLING and ACL, pp. 753–760 (2006)
15. Migon, H.S., Gamerman, D., Louzada, F.: Statistical inference: an integrated approach. CRC Press, Boca Raton (2014)
16. Minard, A.L., et al.: Semeval-2015 task 4: Timeline: Cross-document event ordering. In: SemEval, pp. 778–786 (2015)
17. Miwa, M., Bansal, M.: End-to-end relation extraction using lstms on sequences and tree structures (2016). arXiv preprint. arXiv:1601.00770
18. Pitler, E., Louis, A., Nenkova, A.: Automatic sense prediction for implicit discourse relations in text. In: ACL and AFNLP, pp. 683–691 (2009)
19. Prasad, R., et al.: The penn discourse treebank 2.0. In: LREC (2008)

20. Pustejovsky, J., Verhagen, M.: Semeval-2010 task 13: evaluating events, time expressions, and temporal relations (tempeval-2). In: Workshop on Semantic Evaluations, pp. 112–116 (2009)
21. Radinsky, K., Davidovich, S., Markovitch, S.: Learning causality for news events prediction. In: WWW, pp. 909–918 (2012)
22. Raja, K., Subramani, S., Natarajan, J.: Ppinterfinder–a mining tool for extracting causal relations on human proteins from literature. Database 2013, bas052 (2013)
23. Rugemalira, J.M.: What is a symmetrical language? multiple object constructions in Bantu. In: Berkeley Linguistics Society, vol. 17 (2012)
24. Santos, C.N.d., Xiang, B., Zhou, B.: Classifying relations by ranking with convolutional neural networks (2015). arXiv preprint arXiv:1504.06580
25. Velupillai, S., Mowery, D.L., Abdelrahman, S., Christensen, L., Chapman, W.W.: Blulab: Temporal information extraction for the 2015 clinical tempeval challenge. In: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pp. 815-819 (2015)
26. Zhou, Z.M., Xu, Y., Niu, Z.Y., Lan, M., Su, J., Tan, C.L.: Predicting discourse connectives for implicit discourse relation recognition. In: COLING, pp. 1507–1514 (2010)
27. Zou, B., Zhou, G., Zhu, Q.: Tree kernel-based negation and speculation scope detection with structured syntactic parse features. In: EMNLP, pp. 968–976 (2013)