Finding the Causality of an Event in News Articles

Pattabhi RK Rao, Sobha Lalitha Devi
AU-KBC Research Centre
MIT Campus of Anna University, Chennai-600044
sobhanair@yahoo.com

Abstract

This paper discusses about the finding of causality of an event in newspaper articles. The analysis of causality , otherwise known as cause and effect is crucial for building efficient Natural Language Understanding (NLU) supported AI systems such as Event tracking and it is considered as a complex semantic relation under discourse theory. A cause-effect relation consists of a linguistic marker and its two arguments. The arguments are semantic arguments where the cause is the first argument (Arg1) and the effect is the second argument(Arg2). In this work we have considered the causal relations in Tamil Newspaper articles. The analysis of causal constructions, the causal markers and their syntactic relation lead to the identification of different features for developing the language model using RBMs (Restricted Boltzmann Machine). The experiments we performed have given encouraging results. The Cause-Effect system developed is used in a mobile App for Event profiling called "Nigalazhvi" where the cause and effect of an event is identified and given to the user.

Key words: Causality extraction• Explicit intra-sentential causality • Implicit causality • Inter-sentential causality• Cause- effect, Event extractionThis work presents an automatic identification of explicit connectives and its arguments using supervised method, Conditional Random Fields (CRFs).

1. Introduction

In the last three decades researchers have successfully proved how to extract facts from unstructured text and also have developed large repositories which are focusing on is-a (Hearst, 1992) and part-of (Girju et al., 2003) relations. Information Extraction has many task which extracts facts such as Named Entity Recognition (NE), Relation Extraction (RE) and Event Extraction (EE). Cause–effect extraction is a relational extraction, a challenging task which requires semantic understanding and contextual knowledge of the unstructured text. Cause-effect relations appear frequently in any text. Consider the example which contains a cause-effect relation "heavy rain inundated the city." In this sentence "heavy rain" is the cause and the effect is "inundation of the city". A traditional definition of Cause-Effect relation can be as follows: An Event or Events that come first and results in the existence of another Event, ie, whenever the first event (the Cause) happens, the second event (the Effect) essentially or certainly follows.

The published work in this area can be classified into three approaches : knowledge-based, statistical/ML based , and deep-learning-based. And each method has its advantages and weaknesses. The knowledge-based approach uses linguistic patterns by using pre-defining hand-crafted rules or keywords (Garcia et al., 1997; Khoo et al., 2000; Radinsky et al., 2012; Girju et al., 2009; Kang et al., 2014; Bui et al., 2010). Statistical approach uses probabilistic models over features extracted (Girju, 2003; Do et al., 2011). Using CRFs cause-effect arguments were identified in (Menaka. S, et al., 2011). (Sindhuja G and Lalitha Devi, S 2017) where they consider the identification of causal relations across clauses and sentences using discourse connectives. This approach was applied on BIONLP/NLPBA corpus and identified the causal relations and causal entities. The most frequently used deep learning approaches are feed-forward network (Ponti and Korhonen, 2017), convolutional neural networks (Jin et al., 2020; Kruengkrai et al., 2017) and recurrent neural networks (Yao et al., 2019). Later unsupervised training model such as BERT (Devlin et al., 2018; Sun et al., 2019) and RoBERTa (Becquin, 2020) are also used.

In this work, we base our model on RBMs (Restricted Boltzmann Machine), a deep learner for identifying the cause-effect (arguments) and a CRFs (Conditional Random Fields) Model for identifying the event. The rest of this paper is organized as follows. In Section 2 we present an analysis of causal constructions in Tamil and the data. Section 3 describes the method used for extracting causal relations from News wire text in Tamil. Results and discussions are presented in Section 4. At the end we give the conclusion.

2. Analysis of Cause –Effect in Tamil

The cause-effect relation in Tamil is characterized by the cause, the effect and an optional marker. The marker indicates the presence of a cause-effect relation. The cause is the event that is the reason for the other event called the effect to happen. There is a dependency of one event on the other. One event causes the other event. In other words, an event is a consequence of a preceding event. The cause might be just one of the reasons for the effect to happen in real-world, but what matters in the context of cause-effect relations is the way it is expressed in text. The text may express more than one event as the reasons for the effect to happen, which is a case of multiple causes.

WILDRE-7 Workshop, pages 47–53

2.1 Types of Cause-Effect Relations.

The cause-effect relation in Tamil is classified broadly into explicit cause-effect relations and implicit cause-effect relations. An explicit cause-effect relation is an expression which contains a cause-effect marker explicitly. Certain morphological or syntactic elements bring out the causal meaning. The cause-effect marker denotes the presence of a cause-effect relation. An implicit cause-effect relation is inferred from the context and the world knowledge i.e., there is no explicit cause-effect marker to denote the presence of a cause-effect relation. Ex1 shows an explicit cause-effect relation and Ex.2 shows the same cause-effect relation as in Ex.1, but not connected by an explicit cause-effect marker. Based on the semantics of the context, the reader infers a cause-effect relation. Cause is marked as "C" and effect as "E".

Ex1. [kaaRRu aTi-tt-ataal]C [tuNikaL paRa-nt-ana]E
Wind blow-Pst-Cause clothes fly-Pst-3pn
'Because the wind blew,the clothes flew.'

Ex2. [kaaRRu aTi-tt-atu]C
. [tuNikaL paRantana]E.
Wind blow-Pst-3sn
clothes fly-Pst-3pn
' The wind blew. The clothes flew.'

In the corpus, it was observed that the cause and effect are not always as simple as shown in the examples.

2.2 Text Span of Cause/Effect

The span of text denoting cause or effect does not always coincide with clause boundaries and sentence boundaries. The identification of the text span of the cause and the effect is not very straightforward. The following examples illustrate the point.

Ex3. [atai naan kaNTataal]C ["atellaam nii een

paarkkiRaay(finite verb)" enRaaL]E.

'[As I saw that]C, ["Why are you seeing

those?" said she']E.

It can be noted that the first finite verb following the causal marker is not necessarily the end of the effect because of the verb occurring within the quotes in direct speech. The text span of cause or effect can stop at the boundary of the first verb in reported speech as well (Ex4).

Ex4.[appaTip paTippataal]C [ivvaLavu aRivu vaLarntiruntat]Eai uNarnteen.

'I realized that [by studying so]C, [my knowledge improved so much]E.

In Ex.4, the effect does not extend up to the end of the sentence. In addition, it can be noted that the end of the text span does not coincide with the end of a token.

2.3 Interdependency of Cause-Effect Relations.

Sometimes cause-effect relations form a chain with the effect of the first relation being the cause of the second. Two cause-effect relations occurring in close proximity can be interdependent.

Ex5. [vaNTikkaararkku ippootu varuvaay kuRaintupoo^nataal]C [[avarkaL kutiraikaLai na^nRaaka vaittiruppatillai]E]C. aakaiyaal [ippootuLLa kutiraikaLum mu^npool paarppataRku azakaaka illai]E.

'[Because the cart-owners' incomes have reduced these days]C, [[they do not care for the horses well] E]C. So, [the horses these days don't look as beautiful as those before] E.'

2.4 Anaphors

Most often, though the cause and effect are found in close proximity to the marker, complete sense cannot be made with this information alone due to the presence of anaphors. Thus anaphors have to be resolved for complete comprehension of the cause-effect relation. In Ex.6, the pronominal anaphors, nii and avan should be resolved to completely understand the two events.

Ex6. [nii anpaaka pazakiyataal]C [avan appaTi eNNiviTTaan]E.

'[Because you interacted lovingly]C, [he thought so]E.'

The above issues are some of the major ones which have to be resolved for identification of the cause and effect of a cause-effect relation. From the linguistic analysis we have arrived at the following

1. A cause-effect relation consists of the cause, the effect and an optional marker and can have multiple causes and/or multiple effects.

2. The cause-effect relation can be classified as explicit and implicit cause-effect relations based on the presence or absence of a marker.

3. In an implicit cause-effect relation, the subordinate clause has a non-finite verb in the infinitive form and Explicit cause-effect relations is marked by a grammatical marker or a lexical marker.

4. Explicit cause-effect markers can be intra-sentential or inter-sentential. Intra-sentential markers can be inter-clausal or intra-clausal. Also Intra-sentential markers are grammatical markers “Grammatical markers” get inflected with a noun or a verb.

5. The grammatical marker for cause-effect that inflect with a noun is -aal. This is a polysemous marker denoting instrumentality and cause-effect among others. This ambiguity in sense is resolved by the verb phrase of the clause in which the marker occurs.

6. The grammatical markers for cause-effect that inflect with a non-finite verb are -ataal, -ata^naal, -ati^naal, -amaiyaal, -aamaiyaal. They denote the cause in the subordinate clause and the effect in the main clause.

7. There are Inter-sentential discourse connectives like ata^naal, ita^naal, aa^napaTiyaal, aakaiyaal, aakaiyi^naal, aatalaal, aakavee, e^navee are lexical markers denoting cause-effect.

8. There are other lexical markers such as kaaraNam, kaaraNamaaka and kaaraNattaal and they occur in complex patterns.

9. Certain verbs inherently denote cause.

2.5 Benchmark Datasets

As we all know that data is the foundation of experiment. There is a number of datasets available for English which are used for evaluating cause-effect models. The

SemEval-2007 task 4, it is part of SemEval (Semantic Evaluation), the 4th edition of the semantic evaluation event (Girjuet.al 2007). This task provides a dataset for classifying semantic relations between two nominals. Within the set of seven relations, the organizers split the Cause–Effect examples into 140 training with 52.0% positive data, and 80 test with 51.0% positive data. SemEval-2010 task 8, unlike its predecessor, SemEval-2007 Task 4, which has an independent binary-labelled dataset for each kind of relation, this is a multi-classification task in which relation label for each sample is one of nine kinds of relations (Hendrickx I 2010). PDTB 2.0, the second release of the penn discourse treebank (PDTB) dataset from Prasad et al. (Prasad R 2007) is the largest annotated corpus of discourse relations. It includes 72,135 non-causal and 9190 causal examples from 2312 Wall Street Journal (WSJ) articles. TACRED, similar to SemEval, the Text Analysis Conference (TAC) is a series of evaluation workshops about NLP research. The TAC Relation Extraction Dataset (TACRED) contains 106,264 newswire and online text that have been collected from the TAC KBP challenge.1 during the year from 2009 to 2014 (Zhang Y 2017). BioInfer (Pyysalo et al.2007) introduce an annotated corpus, BioInfer (Bio Information Extraction Resource), which contains 1100 sentences with the relations of genes, proteins, and RNA from biomedical publications. There are 2662 relations in the 1100 sentences, of these 1461 (54.9%) are causal-effect. ADE, the corresponding ADE task aims to extract two entities (drugs and diseases) and relations about drugs with their adverse effects (ADEs) (Hidey C, and McKeown K 2016).

2.6 Tamil Data

There are no standard annotated dataset for cause-effect for Tamil and for any Indian languages. The data we have used is annotated in house from different genres such as novels and new wires. The details of causal markers and their distribution in the corpus is given below (Table -2). In this work the data is collected through crawling the content from 5 major online Tamil News portals. The data is collected over a period of time, by performing daily crawling. The online News portals used for crawling are listed below:

1. Dinamani – https://www.dinamani.com/
2. Dinathanthi- ttps://www.dailythanthi.com/
3. Dinamalar – https://www.dinamalar.com/
4. The Hindu (Tamil)– https://www.hindutamil.in/
5. Maalaimalar - https://www.maalaimalar.com/

There are a total of 2000 documents. Each document is a News article. The average size of a News article is 25 Sentences. Along with this we have also taken data from a few Tamil story and travelogue blogs and Novels. In Table 1 the data statistics is given. The column #sentences shows the total number of sentences. The second column #Relations, shows the number of causal relations. And Table 2 describes causal markers distribution in the corpus (data statistics based on different causal markers). In this table 2 the second column gives number of times the causal marker has occurred in the corpus and third column gives the number of instances where a cause-effect relation has occurred.

Table 1. **Overall Corpus Statistics**

SNo	Corpus Type	#Sentences	#Relations
1	News Corpus	50300	3590
2	Web blogs	488	45
3	Tamil Novels (Akalvilakku, Civakamiyin Sapatham, Kurinchi Malar)	31741	1345
	Total	**83529**	**4980**

We have annotated the data manually using trained linguists. The cause is marked by “C” and effect by “E”. We calculated the inter-annotators agreement using Kappa score and the score was 96%.

Table 2. Causal Markers Distribution in the corpus

SNo	Causal Marker	Total no. of occurrences	No. of Cause-Effect relations
1	*-ataal -atanaal, -itanaal, -paTiyaal, -amaiyaal, -aamaiyaal*	720	660
2	*atanaal, itanaal, aakaiyaal, aanapaTiyaal, aata laal, aakavee, enavee*	1470	1450
3	*kaaraNattaal*	230	210
4	*kaaraNamaaka*	230	210
5	*kaaraNam*	720	490
6	*-aal*	6360	1010
7	*eenenil, eenenRaal*	980	950
	Total	**10710**	**4980**

3. Our Method

In this work we have followed the two step approach,

Step 1: Event Identification using Conditional Random Fields (CRFs), a machine learning algorithm.

Step 2: The Cause-Effects (Causal relations) related to the event are extracted using Restricted Boltzmann Machine (RBM), an unsupervised deep learning algorithm.

Before doing the Event Identification and Cause-Effect identification the documents are pre-processed for syntactic and semantic information enrichment.

Syntactic Pre-Processing: The data obtained from crawling online news portals is cleaned and the text content alone is extracted. After the content is extracted and cleaned, the syntactic pre-processing of the data is performed. We use in house developed POS Tagger (Arulmozhi & Sobha., 2006), and Chunker (Pattabhi et al., 2007) for pre-processing. The text that is split into sentences and then tokenized is send to POS tagger for tagging the POS and the POS tagged data is given to the chynker for chunking. The performance of the POS tagger is 93.26% accuracy and for chunking, the accuracy is 92.73%.

Semantic pre-processing: The semantic pre-processing of the data includes named entity (NE) tagging and anaphora resolution(AR). It is observed that any event there is an involvement of a human/non-human entity, location and time. Thus the entity identification is important.

3.1 Event Identification

The event extraction is performed using Conditional Random Fields (CRFs). The challenge in developing an event extraction system using ML techniques lies in designating the striking features and designing of feature template. A window size of 5 is used in this work. We describe in detail the features used in developing the event identification system.

Lexical features and Syntactic features: Word, Parts of Speech (PoS) and chunk are used. PoS help in disambiguating the sense of the word in a sentence. PoS is an important feature for extracting the events as most of the arguments of an event are proper noun and event trigger belongs to noun and verb category. Hence PoS is a key feature for event extraction task. Most of the event trigger and arguments are descriptive i.e., they occur as a phrase. Hence chunk tag will help in argument and event trigger extraction.

Named Entities (NEs): Named Entity Recognition (NER) is the task of extraction of NEs such as place names, organization names, person names, facilities names etc. from a given text document.

The combination of all the above described features is used to develop the feature template for training the CRFs for Event identification and the language model is obtained.

3.2 Cause-Effect Identification

For each event identified, there will effects of the event and causes of the event. For example for an “earthquake event”, the causes could be tectonic plate shifts and effects are the severe damages to the people, animals and other properties, facilities etc. For automatic identification of causal relations Restricted Boltzmann Machine (RBM), an unsupervised deep learning algorithm is used. RBM is a type of Boltzmann Machines (BMs), to make them powerful enough to represent complicated distributions which go from the limited parametric setting to a non-parametric one. We consider that some of the variables are never observed (they are called hidden). By having more hidden variables (also called hidden units), it can increase the modelling capacity of the Boltzmann Machine (BM). Restricted Boltzmann Machines (RBMs) further restrict BMs to those without visible-visible and hidden-hidden connections. Unlike other unsupervised learning algorithms such as clustering, RBMs discover a rich representation of the input. RBMs are two-layer neural nets. The first layer of the RBM is called the visible, or input, layer, and the second is the hidden layer.

These data in the visible layer (or input layer) is converted to vectors of n-dimension and passed to the hidden layer of the RBM. The word vectors are the vector representations. These are obtained from the word2vec. These are also called as word embedding. Word embedding, in computational linguistics, referred as distributional semantic model, since the underlying semantic theory is called distributional semantics. The real valued n-dimensional vector for each level is formed using the word2vec algorithm. Word2vec creates or extracts features without human intervention and it includes the context of individual words/units provided in the projection layer. Word2vec is a computationally-efficient predictive model for learning word embedding’s from text. The context comes in the form of multiword windows. Given enough data, usage and context, Word2vec can make highly accurate word associations. Word2vec expects a string of sentences as its input. Each sentence – that is, each array of words – is vectored and compared to other vectored lists of words in an n-dimensional vector space. Related words and/or groups of words appear next to each other in that space. The output of the Word2vec neural net is a vocabulary with a vector attached to it, which can be fed into the next layer of the deep-learning net for classification. We make use of the DL4J Word2vec API for this purpose [Mikolov 2013].

We have obtained optimal hyper parameters for good performance by performing several trials. The main hyper parameters which we need to tune include choice of activation function, number of hidden units, learning rate, dropout value, and dimensionality of input units. We used 20% of training data for tuning these parameters. The optimal parameters include: 200 hidden units, rectilinear activation function, 200 batch size, 0.025 learning rate, 0.5 dropout and 25 training iterations. We obtained best development set accuracy at 80 dimensional words. The output layer uses Softmax function for probabilistic multi-class classification. The Softmax function classifies into two classes: A Causal relationship (Cause/Effect) or not a causal relationship. We train the RBM and using the language model obtained during the training of RBMs on a given new text document.

4. Results and Discussion

This section describes the performance of our system in terms of Precision, Recall and F-score. Precision is the number of Events/Causal relations correctly identified by the system from the total number of Events/Causal relations identified by the system. available in the gold standard. Recall is the number of Events/Causal relations correctly detected by the system to the total number of Events/Causal relations available in the corpus (gold standard). F-score is the harmonic mean of precision and recall.

Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
F score = (2 × Recall × Precision / (Recall + Precision))

where, TP means true positives, FN means false negatives and FP means false positives.
In this work for the evaluation, only the single causal relations are considered and not the embedded ones. A causal relation that is inside or embedded inside another causal relation is not considered as separate a distinct causal relation.
A 10-fold cross validation experiment is performed on the data, by splitting the data into 10 equal parts. A set of 9 equal parts is concatenated to form training partition and 1 part is used for testing. Table 3 describes the results obtained for 10-fold cross validation experiment. We have obtained an average precision of 84.35% and average recall of 81.04%.

Table 3. Causal Relation – 10 Fold Experiment results

n-Fold Number	Total C-E Relations in the test set (Gold tagged)	Total C-E Relations Identified by the system	Total C-ERelations Correctly identified by the system	Precision %	Recall (%)
1	460	451	367	81.37	79.75
2	410	401	325	81.05	79.18

3	424	407	335	82.31	79.03
4	454	437	369	84.42	81.23
5	446	425	377	88.70	84.55
6	498	467	399	85.44	80.18
7	464	459	389	84.75	83.78
8	476	474	387	81.64	81.32
9	497	463	403	87.04	81.07
10	418	387	336	86.82	80.33
			Average	**84.35**	**81.04**

5. Conclusion

We have described in detail the cause and effect in Tamil with linguistic analysis, how automatically the cause and effect can be identified using a 2 step process where we have used CRFs to identify the events and RBM for identify the cause and effect of the event. The system works on News paper articles and it is a real time application. The system works with 84.35% precision and 81.04% recall.

6. Acknowledgments

We acknowledge Bhashini, NLTM project under MeitY on Discource Integrated Dravidian to Dravidian Machine Translation (DL-DiscoMT) for the funding to do this research.

7. Bibliographical References

Marti Hearst. 1992. Automatic acquisition of hyponyms from large text corpora. In Proc. of the 14th conference on Computational linguistics, pages 539–545.

Roxana Girju, Adriana Badulescu, and Dan Moldovan. 2003. Learning semantic constraints for the automatic discovery of part-whole relations. In Proc. of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, pages 1–8.

Daniela Garcia et al. 1997. Coatis, an nlp system to locate expressions of actions connected by causality links. In International Conference on Knowledge Engineering and Knowledge Management, pages 347– 352. Springer.

Christopher SG Khoo, Syin Chan, and Yun Niu. 2000. Extracting causal knowledge from a medical database using graphical patterns. In Proceedings of the 38th annual meeting of the association for computational linguistics, pages 336–343.

Kira Radinsky, Sagie Davidovich, and Shaul Markovitch. 2012. Learning causality for news events prediction. In Proceedings of the 21st international conference on World Wide Web, pages 909–918

Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, and Deniz Yuret. 2009. Classification of semantic relations between nominals. Language Resources and Evaluation, 43(2):105–121.

Ning Kang, Bharat Singh, Chinh Bui, Zubair Afzal, Erik M van Mulligen, and Jan A Kors. 2014. Knowledge-based extraction of adverse drug events from biomedical text. BMC bioinformatics, 15(1):1– 8.

Quoc-Chinh Bui, Breanndán Ó Nualláin, Charles A Boucher, and Peter Sloot. 2010. Extracting causal relations on hiv drug resistance from literature. BMC bioinformatics, 11(1):1–11.

Quang Do, Yee Seng Chan, and Dan Roth. 2011. Minimally supervised event causality identification. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, pages 294– 303.

Edoardo Ponti and Anna-Leena Korhonen. 2017. Eventrelated features in feedforward neural networks contribute to identifying implicit causal relations in discourse.

Xianxian Jin, Xinzhi Wang, Xiangfeng Luo, Subin Huang, and Shengwei Gu. 2020. Inter-sentence and implicit causality extraction from chinese corpus. In Pacific-Asia Conference on Knowledge Discovery and Data Mining, pages 739–751. Springer.

Canasai Kruengkrai, Kentaro Torisawa, Chikara Hashimoto, Julien Kloetzer, Jong-Hoon Oh, and Masahiro Tanaka. 2017. Improving event causality recognition with multiple background knowledge sources using multi-column convolutional neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 31.

Liang Yao, Chengsheng Mao, and Yuan Luo. 2019. Graph convolutional networks for text classification. In Proceedings of the AAAI conference on artificial intelligence, volume 33, pages 7370–7377.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language

Cong Sun, Zhihao Yang, Ling Luo, Lei Wang, Yin Zhang, Hongfei Lin, and Jian Wang. 2019. A deep learning approach with deep contextualized word representations for chemical–protein interaction extraction from biomedical literature. IEEE Access, 7:151034–151046.

Guillaume Becquin. 2020. Gbe at fincausal 2020, task 2: span-based causality extraction for financial documents. In Proceedings of the 1st Joint Workshop on Financial Narrative Processing and MultiLing Financial Summarisation, pages 40–44.

Sindhuja Gopalan, Sobha Lalitha Devi 2017 Cause and Effect Extraction from Biomedical Corpus Comp. y Sist. vol.21 no.4 Ciudad de México oct./dic. https://doi.org/10.13053/cys-21-4-2854

Menaka. S, Rao, P.R.K., Lalitha Devi, S. (2011). Automatic Identification of Cause-Effect Relations in Tamil Using CRFs. In: Gelbukh, A.F. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2011. Lecture Notes in Computer Science, vol 6608. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-19400-9_25

Roxana Girju, Nakov P, Nastase V, Szpakowicz S, Turney P, Yuret D (2007) Semeval-2007 task 04: classification of semantic relations between nominals. In: Proceedings of the 4th international workshop on semantic evaluations. Association for Computational Linguistics, USA, SemEval '07, pp 13–18

Hendrickx I, Kim SN, Kozareva Z, Nakov P, Ó Séaghdha D, Padó S, Pennacchiotti M, Romano L,Szpakowicz S (2010) SemEval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In: Proceedings of the 5th International Workshop on Semantic Evaluation, Association for Computational Linguistics, Uppsala, Sweden, pp 33–38

Prasad R, Miltsakaki E, Dinesh N, Lee A, Joshi A (2007) The penn discourse treebank 2.0 annotation manual. IRCS technical reports series 203 Philadelphia: University of Pennsylvania ScholarlyCommons, p 105

Zhang Y, Zhong V, Chen D, Angeli G, Manning CD (2017) Position-aware attention and supervised data improve slot filling. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Copenhagen, Denmark, pp 35–45. https://doi.
org/10.18653/v1/D17-1004 [97]

Pyysalo S, Ginter F, Heimonen J, Björne J, Boberg J, Järvinen J, Salakoski T (2007) Bioinfer: a corpus for information extraction in the biomedical domain. BMC Bioinform 8:50. https://doi.org/10.1186/1471-2105-8-50

HideyC, McKeown K (2016) Identifying causal relations using parallel Wikipedia articles. In: Proceedings of the 54th annual meeting of the association for computational linguistics (volume 1:longpapers).Association for computational linguistics, Berlin, Germany, pp 1424–1433. https://doi.org/10.18653/v1/P16-1135