Cause–Effect Relation Extraction from Documents in Metallurgy and Materials Science

by

Sachin Pawar, Raksha Sharma, Girish Palshikar, Pushpak Bhattacharyya, Vasudeva Varma

in

Transactions of the Indian Institute of Metals (Springer)

Report No: IIIT/TR/2019/-1

Centre for Language Technologies Research Centre
International Institute of Information Technology
Hyderabad - 500 032, INDIA
April 2019

Trans Indian Inst Met (2019) 72(8):2209–2217
https://doi.org/10.1007/s12666-019-01679-z

TECHNICAL PAPER

Cause–Effect Relation Extraction from Documents in Metallurgy and Materials Science

Sachin Pawar[1,3] · Raksha Sharma[2] · Girish Keshav Palshikar[1] · Pushpak Bhattacharyya[3] · Vasudeva Varma[4]

Received: 2 January 2019 / Accepted: 1 April 2019 / Published online: 17 April 2019

Abstract Given the explosion in availability of scientific documents (books and research papers), automatically extracting cause–effect (CE) relation mentions, along with other arguments such as polarity, uncertainty and evidence, is becoming crucial for creating scientific knowledge bases from scientific text documents. Such knowledge bases can be used for multiple tasks, such as question answering, exploring research hypotheses and identifying opportunities for new research. Linguistically complex constructs are used to express CE relations in text, which requires complex natural language processing techniques for CE relation extraction. In this paper, we propose two machine learning techniques for automatically extracting CE relation mentions from documents in metallurgy and materials science domains. We show experimentally that our algorithms outperform several baselines for extracting intra-sentence CE relation mentions. To the best of our knowledge, this is the first work for extraction of CE relations from documents in metallurgy and materials science domains.

Keywords Cause-Effect Relations · Relation Extraction · Materials Science · Unsupervised Learning · Distant Supervision

✉ Girish Keshav Palshikar
gk.palshikar@tcs.com

Sachin Pawar
sachin7.p@tcs.com

Raksha Sharma
rakshasharma.fcs@iitr.ac.in

Pushpak Bhattacharyya
pb@cse.iitb.ac.in

Vasudeva Varma
vv@iiit.ac.in

1 TCS Research, Tata Consultancy Services Limited, Pune, India

2 Indian Institute of Technology Roorkee, Roorkee, India

3 Indian Institute of Technology Bombay, Mumbai, India

4 International Institute of Information Technology Hyderabad, Hyderabad, India

1 Introduction

Much scientific and engineering knowledge is epistemically abstracted in the form of *causality* (or *causation*). Causality is also an integral and frequently used *primitive* in day-to-day linguistic communications among people (e.g., `Rains delayed the trains today`). For these reasons, causality is a core topic in philosophy and statistics [1], where different theories of causality are discussed, as also in the sciences and humanities. In particular, a *causal relation* (or *cause–effect (CE) relation*) connects a *cause* to an *effect*, along with other optional arguments such as the condition, polarity, uncertainty, evidence and exceptions for the CE relation. A CE relation is non-reflexive (nothing can be a cause of itself) and asymmetric. It has an implicit temporal flow: a cause must occur before its effect. Often, a mechanism or process is required to explain how an effect *flows* from (or is created due to) the cause. A cause can be necessary, sufficient or neither, for the effect to occur. One cause may lead to several effects, and an effect may happen due to any of the several possible causes. In a particular domain, such as metallurgy or biomedicine, there may be restrictions on what constitutes valid causes and valid effects.

With the advent of artificial intelligence (AI)-based intelligent computer systems, there is a need to automatically create knowledge bases consisting of stable and well-known CE relations which these systems can use and reason with. Since textbooks and research papers are authoritative sources of knowledge, and of causal knowledge in particular, one possibility is to explore natural language processing (NLP) techniques to automatically extract CE relations from such corpora and populate a knowledge base [2–11]. For example, any tools in integrated computational materials engineering, whose goals include assisting scientists and engineers in designing new materials, can use a knowledge base of extracted CE relations for tasks such as question answering, diagnosis, exploring and validating hypotheses, understanding the state of the art and identifying opportunities for new research.

Due to the flexibility, complexity and expressive power of natural language, there is a great variety in the ways in which CE relations are expressed in text. The simplest way of expressing CE relations in text is through the use of *causative* (or *causal*) verbs, such as `cause`, `lead`, `result`, e.g., `As the two-phase austenitization temperature is raised, the %C in the austenite will increase, and this can lead to reductions of as-quenched hardness.`. Apart from such generic causative verbs, different domains have their own causative verbs, which are either new verbs specific to that domain (e.g., `embrittle`, `austenitize`, `microsegregate`, `passivate`, `boronize`, `sinter` in metallurgy) or generic verbs that have a special causative sense specific to that domain (e.g., `deposit` in metallurgy, as in `Porous metal is created via depositing atomic metal on open-cell polymer foam, followed by eliminating polymers and sintering`). Moreover, every mention of a causative verb does not necessarily indicate a true CE relation mention. For example, the verb `achieve` indicates a CE relation in the sentence `Tumour cell killing was achieved by concerted action of necrosis apoptosis induction.`, but not in `90 patients achieved complete remission on the day of induction therapy.`.

There are other well-known problems with the linguistic expression of CE relations in text. First is the use of negation, which negates the apparent CE relation mention, e.g., `The laser surface texturing did not influence the tribological performance of the K390 Microclean tool steel.`. Next is the use of *coreference*, which refers to the use of a linguistic "variable" such as a pronoun, which is used to succinctly refer to something mentioned earlier. Coreferences are very frequent in natural language text, and it is important to obtain a resolution of every coreference in order to obtain the correct argument of a CE relation mention. For example, in `These factors are strongly influenced by the manganese content`, we need to analyze the text preceding this sentence to identify what `these factors` refers to. Additional complications may arise when cause and effects arguments of the same CE relation are mentioned in two different sentences; here, we focus on CE relation mentions which occur within a single sentence.

In this paper, we propose novel machine learning techniques to extract CE relation instances from English text documents (textbooks and research papers) in metallurgy and materials science. We only extract the CAUSE and EFFECT arguments for each CE relation mention, along with the trigger word that indicates the presence of causality. The first technique is an unsupervised technique that uses linguistic knowledge to identify CE relations in a sentence. The next technique trains bi-directional LSTM model using distant supervision and uses this model to identify CE relations in a sentence. Finally, we also train an LSTM model using labeled training data from another domain to identify whether or not a given sentence contains a CE relation mention. We show the comparative results of these techniques, along with some baselines, on a dataset of sentences in metallurgy and materials science domain. To the best of our knowledge, this is the first work for extraction of CE relations from documents in metallurgy and materials science domains.

The paper is organized as follows. Section 2 discusses related work. Section 3 discusses our unsupervised algorithms for discovering CE relation mentions. Section 4 proposes a bi-directional LSTM model for identifying CE relation mentions, trained using distant supervision. Section 5 proposes a supervised LSTM model for classifying a sentence as containing a CE relation mention or not. Section 6 elaborates the experimental setup and presents our results. Section 7 provides conclusions and outlines for further work.

2 Related Work

Extracting scientific knowledge from books and papers is a challenging problem, which is being explored by the NLP community. A workshop series (International Workshop on Mining Scientific Publications) is devoted to this research (https://wosp.core.ac.uk/). We have focused here on the literature related to extracting CE relations from text, but there are many other types of knowledge elements that can be extracted from scientific text, such as topics, e.g., Griffiths and Steyvers [12], Hall et al. [13], Song et al. [14].

Several researchers have proposed unsupervised knowledge-based methods for extracting CE relations. Joskowicz et al. [15] designed a specialized knowledge base and used it to build a causal analyzer for a Navy ship. Khoo et al. [4] used fixed pre-defined linguistic patterns to identify CE relation, where the patterns uses causal clues such as `hence, therefore, if-then, cause, break`. Girju [5] utilized grammatical patterns in order to analyze cause–effect questions in question answering system. Some other researchers have used grammatical patterns to identify CE relation targeting different applications [9, 16]. Some researchers have combined grammatical patterns with machine learning in order to extract semantic relation such as cause–effect. Chang and Choi [6] used cue phrases (cause triggering construct) with their probability to extract other lexical arguments of cause–effect relation. These probabilities are learned from raw corpus in an unsupervised manner. However, they did not use any linguistic information or domain knowledge. Do et al. [10] developed a minimally supervised approach, based on focused distributional similarity and discourse connectives. They have shown that combining lexical information, such as discourse connectives, with statistical measures provides additional improvement in CE relation extraction. Our unsupervised approach uses both linguistic knowledge and statistical measures derived from an unlabeled corpus.

3 Unsupervised Approaches for CE Relation Extraction

In this section, we have discussed three unsupervised approaches for extracting CE relation instances from given text. No labeled data are needed to train these algorithms, which instead rely on linguistic knowledge to extract CE relation mentions.

3.1 First Baseline Approach

In the first baseline approach (denoted as *cue_phrases*), we have identified several generic (domain-independent) *cue phrases* which are often used to trigger a CE relation mention. Examples of cue phrases include: `due to, because of, cause of, causes of, cause for, causes for, reason for, reasons for, reasons of, reason of, as a consequence, as a result`. The algorithm checks if the given sentence contains any of the given fixed set of cue phrases. If yes, then it extracts the subject of the verb as cause and the noun modifier of the cue phrase as the effect. The algorithm outputs only headwords of the noun phrases corresponding to a cause or an effect. It is easy to modify it to output the entire phrase instead.

Consider the sentence `The continued drop in hardness in the pearlite is due to the increased spacing of the plates of the pearlite at the lower cooling rates.`. The text fragments `the increased spacing of the plates of the pearlite at the lower cooling rates` and `The continued drop in hardness in the pearlite` are, respectively, the CAUSE and EFFECT arguments of this CE relation, which is triggered by the cue phrase `due to`. Figure 1 shows the dependency tree for this sentence generated by Stanford CoreNLP toolset [17], where `due` is connected to a verb (`is`) and also the CAUSE and EFFECT arguments of the CE relation are connected to `due` through the dependency relations *nmod* and *nsubj*, respectively. Note that every occurrence of a cue phrase may not indicate the presence of a CE relation, e.g., `Whatever money is due to John ....`

We can perform additional domain-specific checks on the text fragments corresponding to the CAUSE and EFFECT arguments, in order to retain only valid CE relations detected by a cue phrase. For example, in metallurgy and materials science domains, one may accept the CE relation mention as valid, if at least one argument contains a word of category material, property/state/condition, equipment, etc. For example, the WordNet hypernym hierarchy for the noun `hardness` contains the word **property**, and that for the noun `cooling` contains the word **process**.

3.2 Second Baseline Approach

The second baseline approach (denoted as *Girju*2003) is similar, except that it additionally uses the set of 61 verbs proposed by Girju [5] as cue phrases. Examples include: `give rise to, lead to, result from, bring on, induce, produce, generate, effect, trigger, entail, activate, actuate`. Again, not every occurrence of such a verb indicates a CE relation and additional checks are needed on the CAUSE and EFFECT arguments.

3.3 Unsupervised Co-discovery of Causal Triggers

We had earlier proposed an unsupervised co-discovery CE relation extraction algorithm (denoted *unsup_CE_codiscovery*) for the biomedical domain in [11], which we now adopt for metallurgy. The algorithm works in 3 phases. In phase-I, starting with a small list of known causal verbs ($S = \{$ `cause, result, lead`$\}$) and an unlabeled corpus, the algorithm automatically discovers many causes, effects and causal verbs. Since `causes` is a known causal verb, from the sentence `Each blow of the forging hammer causes some permanent deformation of the metal.`, we add its subject `blow` and direct object `deformation` to a list L of known causes and effects, respectively. In Phase-II, some of these nouns are eliminated from L, because they are not followed frequently enough by `of` or `in`, as computed using

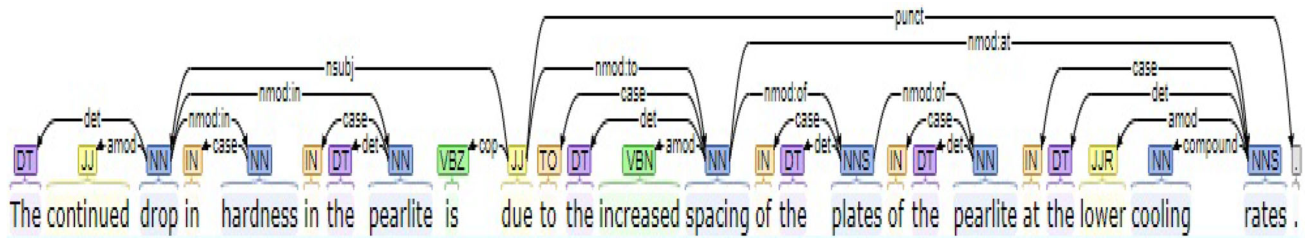

Fig. 1 Dependency graph generated by Stanford CoreNLP

pointwise mutual information (PMI). Finally, verb forms of these nouns are added to list *S* of new causal verbs, provided the WordNet hypernyms of the noun includes categories such as *growth, act, action, event, change, control, happening* which indicate that it is some kind of action.

Intuition behind Phase-III is: *if there is a verb whose subject and object in a sentence are both known (i.e., they have previously occurred with known causative verbs and hence are in list L), then that verb is also likely to be a causative verb*. Phase-III identifies such verbs, removing those that occur less than the given threshold times or those which are not actions (as per the above WordNet check). Once such new causal verbs are identified, we extract CE relation mentions from the given sentence *S* as follows. If a known causal verb *v* occurs in *S*, then we extract the subject *C* and object *E* arguments of *v* in *S* and report the tuple (*v*, *C*, *E*) as the CE relation mention. If no known causal verb occurs in *S*, then we say there is no CE relation mention in *S*. For details of the algorithm, please refer to [11].

Suppose we know that `produce` is a causal verb. Then, from the sentence `The wire-drawing operation produces residual surface compressive stresses that result in the improved toughness.`, we extract `stress` and `toughness` as (headwords of) CAUSE and EFFECT argument of this CE relation mention and add these two words to *L*. Then, when we examine the sentence `Steel toughness is enhanced by surface residual compressive stresses.`, we already know that `stress` and `toughness` are valid CAUSE or EFFECT arguments. Hence, we recognize that the verb `enhance` must be a causal verb. Some examples of causal verbs identified by this algorithm from an unlabeled corpus of documents from metallurgy domain: `quench, roll, microsegregate, decrease, austenitized`.

4 CE Relation Extraction using Distant Supervision

In order to train supervised machine learning models for CE relation extraction, a large amount of annotated (labeled) training data are needed. To manually create such an annotated dataset for training is quite effort-intensive and expensive. For metallurgy and materials science domain, there are no readily available training data where CE relation mentions are marked. Hence, we propose a distantly supervised approach for CE relation extraction and apply it to metallurgy and materials science domain. In *Distant Supervision* framework [18], training data are created automatically using some knowledge bases or domain-specific heuristics and an unlabeled corpus. Then, any supervised machine learning model can be trained using this automatically annotated data.

4.1 Distant Supervision

For distant supervision, we have used simple heuristics based on linguistic resources like FrameNet and WordNet and linguistic rules based on dependency parsing.

FrameNet [19] is a linguistic resource containing sentences annotated with certain predicate argument structures in various scenarios. Each predicate argument structure corresponds to a particular conceptual structure called a *Frame*. Each *Frame* contains *Frame Elements (FE)* corresponding to various possible semantic roles. Each *Frame* also identifies a set of predicates (can be verbal, nominal, adjectival etc.) which may invoke (trigger) the frame.

For automatically annotating causal triggers, we have focused on only "Causation"-related frames, some of which are listed in Table 1. We have manually inspected the predicate verbs for all the CE-related Frames and created a set of high-confidence causal verbs. We have then annotated the causal verbs automatically using a large unlabeled corpus in two phases. In the first phase, we found sentences containing at least one of these high-confidence causal verbs and annotated them as causal triggers if:

- In the dependency tree of the sentence, the causal verb should have at least one subject-like argument (connected to the verb with dependencies *nsubj* or *nmod* : *agent*) and at least one object-like argument (connected to the verb with dependencies *dobj*, *iobj*, *xcomp*, *nmod* or *nsubjpass*)
- The arguments should have one of the following as ancestors in WordNet hypernym hierarchy:

Table 1 Examples of cause–effect-related frames in FrameNet

Frame	Frame elements	Predicates	Example sentence
Causation	Cause, effect, actor	cause, because, due to, induce	If such [a small earthquake] $_{Cause}$ causes [problems] $_{Effect}$, just imagine a big one!
Cause_motion	Cause, agent, theme, path, goal	launch, move, push, throw	[She] $_{Agent}$ threw [her shoes] $_{Theme}$ [into the dryer] $_{Goal}$.

condition, state, physical_condition, attribute, quality, status, material, substance, artifact or event. Here, top 2 senses in WordNet of argument head words are considered for checking the hypernym hierarchy.

During the first phase, we have also recorded all the argument pairs (considering only head words of the arguments) where causal triggers are annotated. In the second phase, we annotate causal triggers in the sentences where:

- Sentence contains a pair of nouns matching with any of the recorded argument pairs such that the dependency path between the nouns is not longer than two edges.
- Additionally, the dependency path contains at least one edge labeled with any of the following dependency types: *nsubj*, *nmod* : *agent*, *dobj*, *xcomp*, *nsubjpass* or *advcl*
- The lowest common ancestor of the nouns is then annotated as the causal trigger.

Negative Instances In order to train a supervised machine learning model which identifies causal triggers from sentences, we also need some sentences in the annotated dataset without any causal trigger. For this purpose, we have identified a set of sentences from the corpus which contain no verb from any of the CE-related frames from the FrameNet. As this set is quite large as compared to the sentences with annotated causal triggers, we randomly select 20% sentences from this set as negative sentences.

Finally, we have an annotated dataset of 1258 sentences where 713 causal triggers are labeled automatically. Following is an example sentence from this dataset. Here, even if homogenizes is not from the initial high-confidence causal verbs list, it has been annotated as a causal trigger in the second phase.

In/O the/O second/O step/O ,/O the/O carbon/O composition/O in/O the/O newly/O formed/O austenite/O homogenizes/CAUSAL_TRIGGER by/O the/O process/O of/O diffusion/O ./O

4.2 BiLSTM-Based Causal Trigger Extraction

We have proposed an LSTM (Long Short Term Memory)-based [20] model to extract causal triggers (if any) from a given sentence. LSTM is a special kind of Recurrent Neural Network (RNN) which models long-range dependencies within a sequence and is designed to alleviate the exploding and vanishing gradient problems of general RNNs.

We have considered each sentence to be a sequence of words. A bi-directional LSTM [21] is then trained on the dataset created using distant supervision as explained earlier. It assigns an appropriate label to each word in a given sentence, where the label CAUSAL_TRIGGER is assigned if the corresponding word is a causal trigger and the label O is assigned otherwise. Bi-directional LSTM for a sentence captures representations for left as well as right context for each word. Each word is represented by a vector which is a concatenation of two vectors: (i) pre-trained word vector for the word using GloVe embeddings [22] and (ii) one hot encoded representation for POS tag of the word. We denote this model as *BiLSTM_distant_sup*; see Fig. 2 for the architecture of this model.

5 Causal Sentence Identification

Our unsupervised algorithm *unsup_CE_codiscovery* outputs a list of possible causal verbs or triggers. Any sentence containing any of these known causal triggers is considered to be a "causal" sentence. But we have observed that some verbs can denote causality in one sentence, but in another sentence they do not behave as causal verbs. For example, consider following sentences containing the verb achieve:

- S1: Further improvement in ductility is achieved by strain induced transformation of austenite to martensite, the so called TRIP effect.
- S2: We show that reasonable agreement can be achieved using a simple uncoupled model.

Here, even though both the sentences contain the verb achieve, the sentence S1 is a "causal" sentence, whereas S2 is not. Thus, depending on other context information in a sentence, verb may or may not trigger the CE relationship.

Hence, we have proposed a simple sentence classification approach which predicts whether any sentence

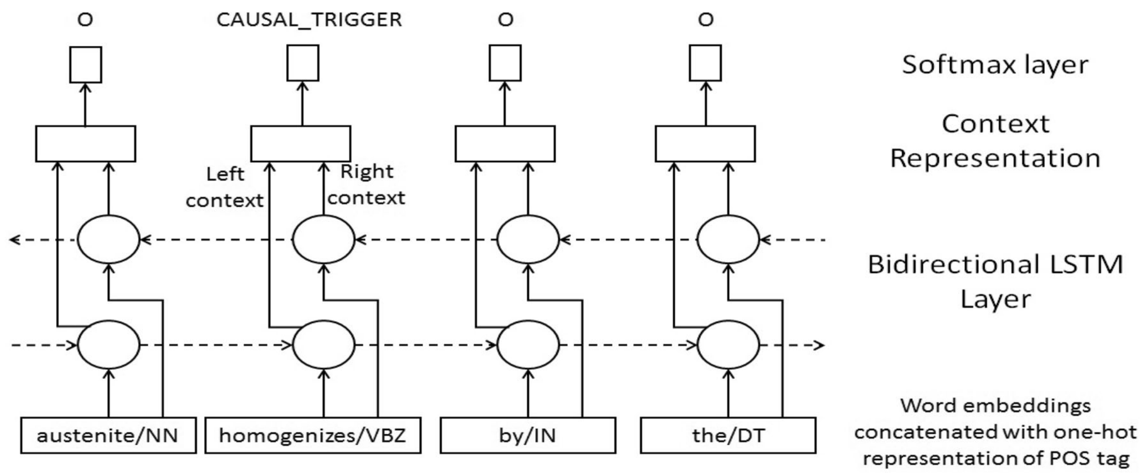

Fig. 2 Bi-directional LSTM architecture for causal trigger detection

expresses causality or not. We use LSTM [20]-based binary classifier for modeling this prediction. Each sentence is represented as a sequence of words where each word is represented using pre-trained word vectors. Sequence of words is passed through an LSTM layer, and the output of the final time step corresponds to the representation of whole sequence, i.e., the sentence. This output is then fed to a single node dense layer with sigmoid activation. We consider a sentence as "causal" if the output of the final node is 0.5 or more; otherwise, it is "non-causal." This model is trained on a manually tagged dataset of 2077 sentences from biomedical domain, containing 567 causal sentences. We apply this model as it is (without any adaptation) on our dataset of sentences from metallurgy and materials science domain. We denote this model as *LSTM_sup*; see Fig. 3 for the architecture of this model. It is to be noted that this *LSTM_sup* does not extract the CE relation trigger or arguments; it only marks a sentence as containing a CE relation mention or not.

6 Experimental Analysis

6.1 Dataset

There are no readily available annotated datasets in the metallurgy and materials science domain for CE relations. We have selected abstracts of 48 research papers in metallurgy and materials science from various journals and a subset of sentences from 4 chapters (1, 6, 12 and 14) of a book on steel metallurgy [23]. Three of the authors have manually annotated these sentences for cause–effect relations. Each annotated cause–effect relation instance is a triplet consisting of: (i) Causal trigger, (ii) Cause argument phrase, and (iii) Effect argument phrase. A few example sentences and their corresponding annotations are shown in Table 2. Total size of this dataset is 988 sentences, containing 266 causal sentences having 337 CE relation triplets. We refer this dataset as MST dataset, i.e., metallurgy and materials science Test dataset, because this dataset is used only for evaluation.

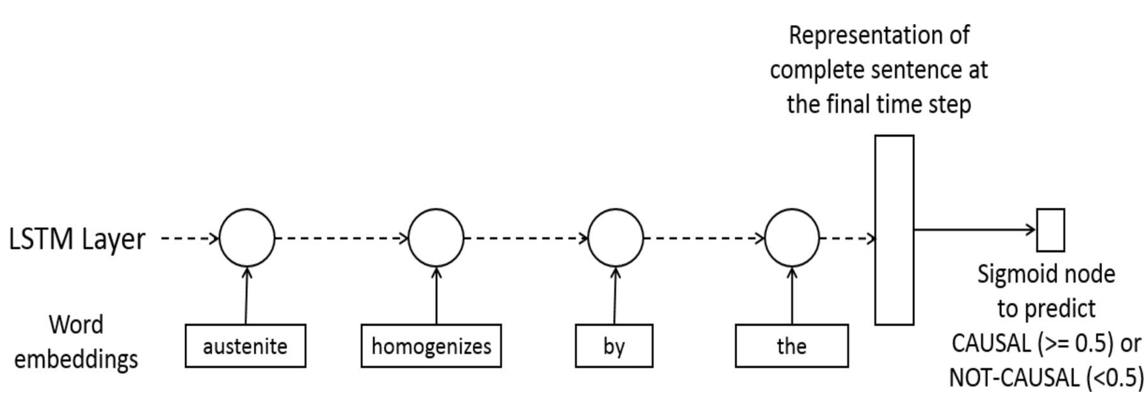

Fig. 3 LSTM architecture for causal sentence identification

Table 2 Example sentences from the manually annotated dataset in Materials Science domain with annotations

Sentence 1:	`Rapid post-deformation cooling at rates of approximately 100 C/s yields the desired UFF-martensite microstructure.`
Causal Trigger:	**yields**
Cause phrase:	`Rapid post-deformation` **cooling** `at rates of approximately 100 C/s`
Effect phrase:	`the desired UFF-martensite` **microstructure**
Sentence 2:	`The differences in transformation behavior can be attributed to the change in the intrinsic stacking-fault energy (ISFE) : in the compositional range studied , Cr and N additions cause an increase of the ISFE.`
Causal Trigger:	`can be` **attributed** `to`
Cause phrase:	`the` **change** `in the intrinsic stacking-fault energy (ISFE)`
Effect phrase:	`The` **differences** `in transformation behavior`
Causal Trigger:	**cause**
Cause phrase:	`Cr and N` **additions**
Effect phrase:	`an` **increase** `of the ISFE`

Headwords of the causal trigger, cause phrase and effect phrase are shown in bold

6.2 Identifying Cause–Effect Arguments

Both the proposed approaches do not directly extract complete triplet of (Causal trigger, Cause phrase, Effect phrase). Rather only causal triggers are identified by these approaches. Our semi-supervised approach creates a list of causal verbs. Given any sentence (from test dataset), causal trigger is extracted by simply looking up in this list of causal verbs. The distantly supervised approach directly extracts causal triggers from any new sentence by assigning the label CAUSAL_TRIGGER to appropriate words.

Once, any causal trigger is identified in a sentence, cause and effect arguments are determined by using dependency parse of that sentence, as follows:

- Phrases dependent on the causal trigger with dependency relations *nsubj* or *nsubjpass* are extracted as Cause candidates
- Phrases dependent on the causal trigger with dependency relations *dobj* or *nmod* are extracted as Effect candidates
- If at least one candidate cause and one candidate effect are extracted, then all possible triplets of causal trigger, cause candidates and effect candidates are output as valid cause–effect relation triplets

6.3 Evaluation of CE Relation Extraction

We have evaluated all the approaches for CE relation extraction on the MST dataset. To evaluate extracted triplets by any approach M, we compute following statistics by comparing with the annotated triplets in the MST dataset:

- True Positives (TP): Any predicted triplet in a sentence by M is counted as a true positive if a *similar* triplet exists in set of actual triplets annotated for that sentence in the MST dataset. Two triplets are deemed to be *similar* if headwords of their causal triggers are same and both the headwords of their cause–effect argument phrases are also same.
- False Positives (FP): Any predicted triplet in a sentence by M is counted as a false positive if no *similar* triplet exists in set of actual triplets annotated for that sentence in the MST dataset.
- False Negatives (FN): Any actual triplet annotated for a sentence in the MST dataset is counted as a false negative if no *similar* triplet is predicted by M for that sentence.

We then compute precision (P), recall (R) and F-measure (F) as follows:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \quad R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \quad F = \frac{2 \cdot P \cdot R}{P + R} \tag{1}$$

Table 3 shows the results of various algorithms for extracting CE relation mentions (trigger and both CAUSE and EFFECT arguments) on the MST dataset.

6.4 Evaluation of Causal Sentence Identification

We have used the same MST dataset for evaluating our LSTM-based approach for causal sentence identification. Any sentence in the MST dataset with at least one

Table 3 Comparative performance of various approaches for extracting causal trigger and CE arguments

Approach	Precision	Recall	*F*-measure
cue_phrases	41.2	2.1	4.0
*Girju*2003	**50.8**	19.9	28.6
unsup_CE_codiscovery	29.2	39.2	33.5
*Girju*2003 + *unsup_CE_codiscovery*	29.0	41.2	34.0
BiLSTM_distant_sup	42.8	**45.1**	**43.9**

The bold values indicate the best performing approach for each column, i.e. metric

annotated cause–effect relation triplet is considered as a "causal" sentence. All the remaining sentences are considered as "non-causal." To evaluate causal sentence predictions by any approach *M*, we compute following statistics by comparing with the actual causal sentences in the MST dataset:

- True Positives (TP): A predicted causal sentence by *M* is counted as a true positive if it is also a causal sentence in the MST dataset.
- False Positives (FP): A predicted causal sentence by *M* is counted as a false positive if it is not a causal sentence in the MST dataset.
- False Negatives (FN): An actual causal sentence in the MST dataset is counted as a false negative if it is not predicted as a causal sentence by *M*.
- True Negatives (TN): An actual non-causal sentence in the MST dataset is counted as a true negative if it is not predicted as a causal sentence by *M*.

We then compute precision, recall and *F*-measure using Eq. 1. We also compute accuracy as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{2}$$

Table 4 shows the results of two algorithms on MST dataset for labeling sentences as containing a CE relation mention or not. The algorithm *BiLSTM_distant_sup* is used to mark a sentence as containing a CE relation mention, if it identifies at least one word in it as a causal trigger.

6.5 Analysis of Results

Overall, the *BiLSTM_distant_sup* approach outperforms other approaches for both the tasks: (i) sentence-level causality identification and (ii) extracting causal triggers and CE arguments. One important aspect of *BiLSTM_distant_sup* is that, it utilizes the sentence context information. Other approaches such as *unsup_CE_codiscovery* do not use any sentence context information, rather it only relies on the presence of any known causal verb in the sentence. Consider following two sentences from the MST dataset:

S1 `Based on the information` **obtained** `from transmission electron microscopy and scanning electron microscopy, a computational simulation procedure is developed within the software package matcalc, which is capable of describing the experimental results in terms of the number density, composition, and type of precipitate phases.`

S2 `Further improvements in mechanical properties can be` **obtained** `by microalloying, especially with vanadium and nitrogen additions.`

Here, *unsup_CE_codiscovery* will extract `obtained` as a causal trigger in both the sentences S1 and S2, because `obtain` is a part of the list of causal triggers discovered in an unsupervised way using the *unsup_CE_codiscovery* algorithm. It can be observed that `obtained` is a correct causal trigger only in the sentence S2 but not in S1. Hence, such extractions result in lower precision for *unsup_CE_codiscovery*. On the other hand, the *BiLSTM_distant_sup* correctly identifies `obtained` as a causal trigger in S2 and also does not identify `obtained` as a causal trigger in S1. Hence, higher precision is observed for *BiLSTM_distant_sup* as compared with *unsup_CE_codiscovery* (42.8% vs 29.2%).

Table 4 Comparative performance of various approaches for identifying whether any sentence contains causality

Approach	Precision	Recall	*F*-measure	Accuracy
BiLSTM_distant_sup	41.5	**75.6**	**53.6**	64.7
LSTM_sup	**63.4**	19.5	29.9	**75.3**

The bold values indicate the best performing approach for each column, i.e. metric

7 Conclusions and Further Work

Cause–effect relations are crucial knowledge elements in any scientific discipline. With the explosion of books and research papers, it has become important to devise automated natural language processing techniques to extract CE relation mentions from a given corpus. In this paper, we have proposed two machine learning based algorithms for extracting CE relation mentions from documents: (i) an unsupervised co-discovery algorithm that uses linguistic knowledge and (ii) a bi-directional LSTM model trained using distant supervision. We have also proposed a supervised LSTM model for classifying a sentence as containing a CE relation mention or not. We have presented comparative results of these algorithms on a dataset containing sentences from metallurgy and materials science domain. We have observed that the bi-directional LSTM model generally performs well on this dataset. The problem of CE relation mention extraction from text is challenging due to the complex, expressive nature of natural language. To the best of our knowledge, this is the first work for extraction of CE relations from documents in metallurgy and materials science domains.

For future work, we are planning to use transfer learning techniques to build CE relation extraction algorithms that work on multiple scientific domains. We are also working on a joint machine learning model to simultaneously extract causal trigger as well as its arguments. Since cause and effect arguments for the same CE relation mention may occur in different sentences, it is important to extend the techniques to cross-sentence extraction. We are also working on extracting other arguments of the CE relation, such as polarity, uncertainty and evidence.

Acknowledgements The authors would like to express their gratitude to Dr. E.C. Subbarao for his guidance and support.

References

1. Pearl J, *Causality: Models, Reasoning, and Inference* (2000).
2. Kaplan R M, and Berry-Rogghe G, *Knowl Acquisition* **3** (1991) 317.
3. Garcia D, in *International Conference on Knowledge Engineering and Knowledge Management*, Springer (1997), p 347.
4. Khoo C S, Kornfilt J, Oddy R N, and Myaeng S H, *Lit Linguist Comput***13** (1998) 177.
5. Girju R, in *Proceedings of the ACL 2003 Workshop on Multilingual Summarization and Question Answering-Volume 12*, Association for Computational Linguistics (2003), p 76.
6. Chang D S, and Choi K S, in *International Conference on Natural Language Processing*, Springer (2004), p 61.
7. Beamer B, and Girju R, in *International Conference on Intelligent Text Processing and Computational Linguistics*, Springer (2009), p 430.
8. Cole S V, Royal M D, Valtorta M G, Huhns M N, and Bowles J B, in *SoutheastCon, 2006. Proceedings of the IEEE*, IEEE (2005), p 125.
9. Radinsky K, Davidovich S, and Markovitch S, in *Proceedings of Learning by Reading for Intelligent Question Answering Conference* (2011).
10. Do Q X, Chan Y S, and Roth D, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics (2011), p 294.
11. Sharma R, Palshikar G, and Pawar S, in *23rd International Conference on Natural Language and Information Systems (NLDB 2018)* (2018).
12. Griffiths T, and Steyvers M, in *Proceedings of 101st Meeting of the National Academy of Science* (2004), p 5228.
13. Hall D, Jurafasky D, and Manning C, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2008)* (2008), p 363.
14. Song M, Heo G, and Kim Y, *Scientometrics* (2014).
15. Joskowicz L, Ksiezyck T, and Grishman R, in *AI Systems in Government Conference, 1989., Proceedings of the Annual*, IEEE (1989), p 195.
16. Kim H D, Zhai C, Rietz T A, Diermeier D, Hsu M, Castellanos M, and Ceja Limon C A, in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, ACM (2012), p 2689.
17. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, and McClosky D, in *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (2014), p 55.
18. Mintz M, Bills S, Snow R, and Jurafsky D, 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-Volume 2*, Association for Computational Linguistics (2009), p 1003.
19. Baker C F, Fillmore C J, and Lowe J B, in *Proceedings of the 17th International Conference on Computational Linguistics-Volume 1*, Association for Computational Linguistics (1998), p 86.
20. Gers F A, Schmidhuber J, and Cummins F, *Learning to Forget: Continual Prediction with LSTM* (1999).
21. Lample G, Ballesteros M, Subramanian S, Kawakami K, and Dyer C, in *Proceedings of NAACL-HLT* (2016), p 260.
22. Pennington J, Socher R, and Manning C, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2014), p 1532.
23. Verhoeven J D, *Steel Metallurgy for the Non-Metallurgist*, ASM International (2007).