Annotating Entity and Causal Relationships on Japanese Vehicle Recall Information

Hsuan-Yu Kuo, Youmi Ma, and Naoaki Okazaki

Tokyo Institute of Technology,
2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550 Japan
{kuo.hsuanyu,youmi.ma}@nlp.c.titech.ac.jp
okazaki@c.titech.ac.jp

Abstract

A vehicle recall system is a process of recalling and repairing vehicles with defective designs or potential for accidents and failures. The recall document concisely explains the circumstances and causes of product defects. This paper presents two types of annotations on public vehicle recall reports, part entities and their relations, and causality. We annotated 6,394 car-recall text documents. Named entity and relation annotation suggests a relationship between the elements of an automobile, and causality annotation indicates the cause of a malfunction. The entity and relation annotation and causality annotation allow the system to automatically extract knowledge in the automotive design domain. Subsequently, we present the experimental results for named entity recognition and relation extraction and causality extraction of our annotated corpus to verify the feasibility of building a system for extracting part information and causality. Finally, the experimental results show that employing named entity and relation information as the external knowledge improves causality extraction.

1 Introduction

A defect/bug in a product causes significant losses and damages to both users and manufacturers. Although manufacturers conduct design/code reviews to ensure the quality of a product, manual reviews involve various challenges, such as correctness, comprehensiveness, cost, and development of human experts. Therefore, we expect computers to automate or assist in the review process.

Information extraction from unstructured text is a straightforward approach for computers to learn expert knowledge, and researchers have applied information extraction in various fields such as news (Chinchor, 1998), biomedical (, 2002), clinical (Demner-Fushman et al., 2009; Rumshisky et al., 2016), and business (Bahja, 2020). However, no previous work has explored its applications in the manufacturing industry.

In this study, we explore scenarios for information extraction in the car industry. As the design and review records of each company in the car industry are kept strictly confidential, we cannot share a corpus and dataset created for this domain. Instead, we focus on vehicle recall reports published by the Japanese government.

A recall system is a system in which an automobile manufacturer, at its discretion, notifies the Minister of Land, Infrastructure, Transport and Tourism (MLIT)[1] in advance of a recall or repair of a product due to a problem in the design or manufacturing process to prevent further accidents and problems. The text describing the situation of each recall is available on the MLIT website[2]. A recall text briefly describes the circumstances and causes of product defects, possibly useful for extracting information from design reviews in the manufacturing process.

Useful information in a car recall text includes entity mentions, entity relations, and causal relationships. For example, consider the following sentence: "Due to inappropriate electrical circuitry in the aux-

[1] https://www.mlit.go.jp/en/

[2] https://www.mlit.go.jp/jidosha/recall.html

iliary braking device (electromagnetic retarder), the braking light does not turn on when the electromagnetic retarder is activated." There are entity mentions (e.g., "auxiliary braking device," "electromagnetic retarder," and "braking light") and entity relations (e.g., "electromagnetic retarder" *is an* "auxiliary braking device"; "braking light" *is connected with* "electromagnetic retarder"). The text also contains a causal relationship between "inappropriate electrical circuitry in the auxiliary braking device" and "the braking light does not come on when the electromagnetic retarder is activated." These causal relationships are extremely useful for specifying the reason for a malfunction, thus helping to avoid related problems in product design or reviews.

Recognizing causality requires knowledge of individual components and their relations in vehicles. In addition, relation instances extracted from recall information can be used to build a knowledge base (KB) for manufacturing cars.

In this study, we build a corpus of Japanese vehicle recall information, where 6,394 documents are annotated with named entities (NEs), their relations, and causal relations to build a system that assists the review process when designing and manufacturing vehicles. To verify the feasibility of building a system for extracting part information and causality, we employ a joint NER/RE model to extract information from the annotated corpus. The main contributions of this paper can be summarized as follows:

- According to our research, this is the first work annotating NEs, relations, and causalities on vehicle recall information[3].

- We report the experimental results on named entity recognition, relation extraction, and causality extraction. We also show that incorporating knowledge about entities and their relations improves the performance of causality extraction.

- We summarize issues in building the corpus, hoping that these findings will be useful for building corpora in other manufacturing fields.

[3] We will release the corpus to the public after this paper is accepted.

2 Related work

Considering that the ultimate goal of this study was to assist the reviewing process of product design, our goal was to build a model and KB to infer possible defects in a given design. Therefore, research on causality extraction was the most relevant to our study. A common approach for causality extraction is to build an annotated corpus and train a model on the corpus. In this section, we describe existing corpora for causality extraction in general and specific domains.

SemEval 2007 Task 4 (Girju et al. , 2007) considered the task of recognizing a semantic relation (including a cause-effect relation) between simple nominals as a binary classification problem. SemEval 2010 Task 8 (Hendrickx et al., 2010), a direct successor of SemEval 2007 Task 4, also addressed the same task but formalized the task as a multi-class classification problem. The datasets of these two tasks use Wikipedia as the source documents.

BECauSE 1.0 (Dunietz et al., 2015) annotated causality instances in the New York Times (NYT) corpus (Sandhaus, 2008). BECauSE 2.0 (Dunietz et al., 2007), a successor of BECauSE 1.0, includes relations overlapping with causality. In addition, CaTeRS (Mostafazadeh et al., 2016) is an annotation scheme that captures a set of temporal and causal relations between events, and the authors annotated a total of 1600 sentences sampled from ROCStories (Mostafazadeh et al., 2016). Inspired by TimeML (Pustejovsky et al., 2003), Mirza et al. (Mirza et al., 2014) proposed guidelines for annotating the causality relation in the TempEval-3 corpus and a rule-based algorithm for automatic annotation.

Biomedical literature is the most explored domain for causality extraction. BioInfer (Pyysalo et al., 2007) presents an annotation scheme and corpus capturing NEs and their relationships, along with a dependency analysis of a sentence. BioCause (Mihuailua et al., 2013) is an annotated corpus with open-access full-text biomedical journal articles belonging to the subdomain of infectious diseases. The corpus annotates linguistic causality instances consisting of a causal trigger (usually a connective), cause, and effect. Using the defined scheme, the researchers added 851 casual relations annotations

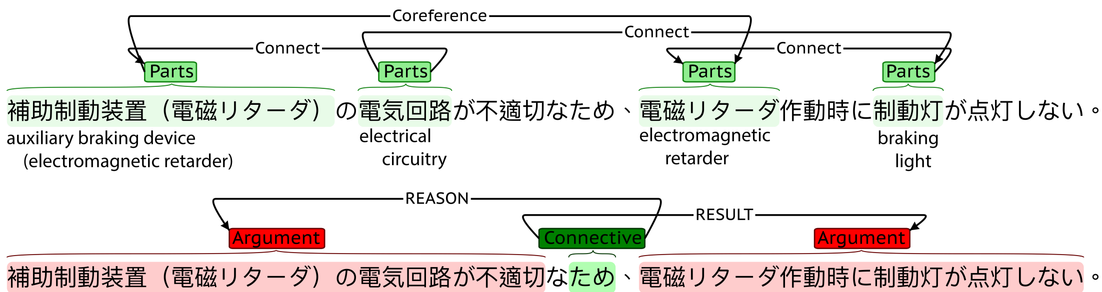

Figure 1: Upper figure shows an example document annotated with NEs and relations. Lower figure shows the same document annotated with causality. BRAT (Stenetorp et al., 2012) is used for annotation and visualization. Translation: Due to inappropriate electrical circuitry in the auxiliary braking device (electromagnetic retarder), the braking light does not come on when the electromagnetic retarder is activated.

to the collection of articles. The corpus is pre-annotated with NEs and events of genes and their interactions (e.g., positive and negative regulations).

Inspired by efforts in the biomedical domain, this study assumes that interactions between car components refer to the causality chain of a malfunction; therefore, we annotate causal relations on top of NEs and their relationships in the car recall text.

3 Corpus

We explain the target text in Section 3.1, followed by the annotations of NEs and their relations in Section 3.2 and causal annotations in Section 3.3, respectively.

3.1 Data

We crawled 6,394 Japanese documents reporting car recall information from the MLIT website[4]. The average length of documents is approximately 135 letters, 1.7 sentences, and 74 tokens after the tokenization using the Japanese tokenizer MeCab (Kudo, 2006).

3.2 NEs and relations

We define a single entity type PART and five relation types, CONTACT, CONNECT, PART-WHOLE, COREFERENCE, ONEWAY-COREFERENCE, between the two parts. The upper figure of Figure 1 illustrates an example of a document annotated with entities and relations.

[4]https://www.mlit.go.jp/jidosha/recall.html

3.2.1 Entity type

This study uses a single entity type PART to annotate car parts (components). We do not distinguish between the granularity of car parts (e.g., "cylinder head" and "engine") and semantic differences of mentions (e.g., "oil filler" as a car component or as a location in a car). We also include the names of car models and other necessary components as PART, but exclude the following text spans:

- A part that cannot be interpreted as a component but only stands for a specific location, for example, "joint section".
- Design of structures and methods, for example, "water immersion prevention structure" and "four-wheel-drive".
- Air, for example, "put *air* into a tire". Similarly, we exclude "electricity" from the annotations.
- A modifying clause of a part entity. For example, we only annotate "program" as PART entity in "program to calculate the amount of particulate matter deposition."

3.2.2 Relation types

We define five relation types that frequently appear in recall texts and commit to causal relations.

- CONTACT: Part_1 is located next or attached to Part_2. This relation is useful for causality extraction because it expresses direct contact between the two parts. Figure 2 illustrates an example of the CONTACT relation: "the steering

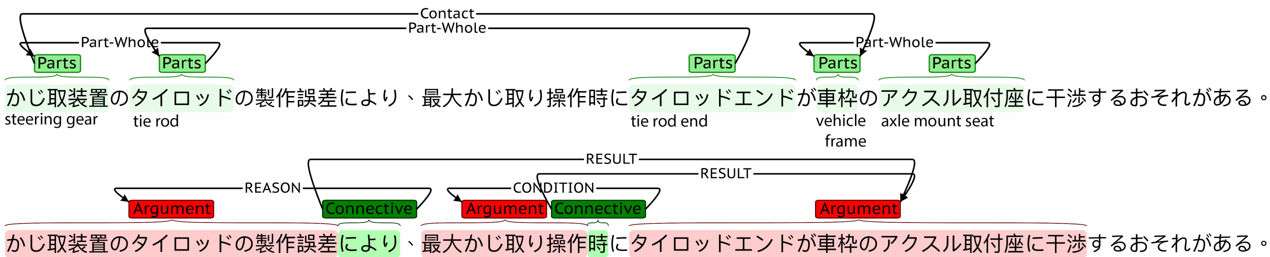

Figure 2: Translation: Due to a manufacturing error in the tie rod of the steering gear, the tie rod end may interfere with the axle mount seat of the vehicle frame during maximum steering.

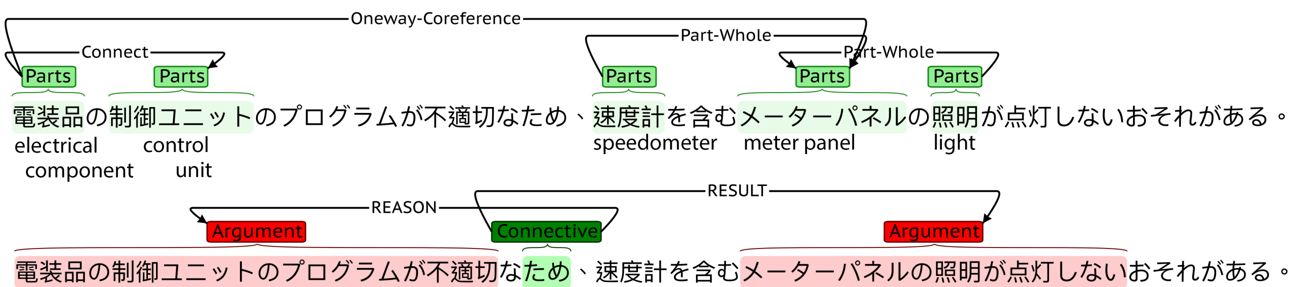

Figure 3: Example for showing Oneway-Coreference and Part-Whole relations are related to causality.
Translation: Due to an inappropriate program in the control unit of the electrical components, the lights on the meter panel, including the speedometer, may not turn on.

gear" and "the vehicle frame" are in contact with each other. In this example, "a manufacturing error in the tie rod of the steering gear" leads to "the tie rod end may interfere with the axle mount seat of the vehicle frame," indicating that the defect of "the steering gear" can influence "the vehicle frame." In addition, we also annotate *implicit* contact relations between two parts; in other words, contact relations that are not explicitly stated in the text but can be inferred by the context or external knowledge. Consider the example in Figure 2. The CONTACT relation between "the steering gear" and "the vehicle frame" is not explicitly mentioned in the text. If a model for causality extraction is aware of this contact relation, the model can extract causality even without a connective of a causal relation.

- CONNECT: Something connects Part_1 and Part_2 (electrically, by transfer of matter, or some other forms of transmissions), for example, an "accelerator pedal" is connected to an "engine." Connect relations are also useful clues for recognizing causality relations, expressing an association between the two parts. Consider an example, "the battery is connected with the headlights." If the battery charge is insufficient, we can infer that the headlights are not working. Similar to contact relations, we also annotate the *implicit* connect relations between two parts; for instance, in Figure 1, "the electrical circuity" and "braking light" are connected, but the relation is not explicitly described in the text.

- PART-WHOLE: We consider the following relationships as a part-whole relation:

  1. Part_1 is composed by Part_2.
  2. Part_2 is a part of Part_1.
  3. Part_2 is spatially contained in Part_1.

  The direction of PART-WHOLE is from Part_2 to Part_1. Figure 2 shows that there is a PART-WHOLE relation from "the tie rod end" to "the tie rod," and the text explains that a manufacturing error in the tie rod of the steering gear leads to the tie rod end possibly interfering with

Entities	Parts	43,158
Relations	Coreference	14,116
	Oneway-Coreference	337
	Part-Whole	13,561
	Contact	8,176
	Connect	4,510
	Total	40,700

Table 1: Statistics of entity and relation annotation.

Entities	Argument	42,312
	Connective	34,369
	Total	76,681
Relations	REASON	30,098
	RESULT	35,957
	CONDITION	4,510
	Total	70,565

Table 2: Statistics of causality annotation.

the axle mount seat of the vehicle frame. In this example, a manufacturing error in the tie rod causes a problem in the tie rod end, showing that PART-WHOLE relations are related to causality.

- COREFERENCE: Part_1 and Part_2 refer to the same *part* entity. This relation is also important for recognizing causal relations where the same entity appears in multiple states and events. The example in Figure 1 illustrates how the COREFERENCE relation can be useful for causality extraction. "The auxiliary braking device (electromagnetic)" and "electromagnetic" refer to the same vehicle *part*, confirming a COREFERENCE relation between them. These two mentions are in clauses indicating the reason and result of a malfunction of the same entity. In the recall report data, clauses containing mentions that refer to the same entity may have causality relations.

- ONEWAY-COREFERENCE: Part_2 refers to Part_1, but Part_1 does not necessarily refer to Part_2 (Part_1 is the subset of Part_2). The direction of ONEWAY-COREFERENCE is from Part_2 to Part_1. In Figure 3, the entity "electrical component" can refer to the "meter panel," but not vice versa. Similar to the COREFERENCE relation in Figure 1, these two annotated entities are in the clauses explaining the reason and the corresponding result. The reason for extracting ONEWAY-COREFERENCE relations is much the same as that for extracting COREFERENCE.

Table 1 summarizes the statistics of entity and relation annotations for car parts. In total, we annotated 43,158 part entities and 40,700 relations.

3.3 Causality annotation

Following the PDTB 3.0 annotation manual (Webber et al., 2019), we annotated the causality relationships between arguments. We define two argument types: ARGUMENT and CONNECTIVE, and three relation types: REASON, RESULT, and CONDITION. We annotated a causality relation as a combination of REASON and RESULT relations:

$$\begin{aligned} \text{ARGUMENT}_1 &\xleftarrow{\text{REASON}} \text{CONNECTIVE} \\ \text{CONNECTIVE} &\xrightarrow{\text{RESULT}} \text{ARGUMENT}_2 \end{aligned} \quad (1)$$

ARGUMENT_1 presents a cause (reason) of the causal relation, and ARGUMENT_2 presents its effect (result). They are connected to each other by CONNECTIVErelation. Figure 1 shows a real example of a causal relation annotated in a recall document. Here, ARGUMENT_1 is 「補助制動装置（電磁リターダ）の電気回路が不適切な」"the electrical circuit of the auxiliary braking device (electromagnetic retarder) is inappropriate"; ARGUMENT_2 is 「電磁リターダ作動時に制動灯が点灯しない。」"braking light does not turn on when the electromagnetic retarder is activated"; CONNECTIVE is 「ため」 "due to" . When a causal relation holds for a specific condition, we also annotate the condition relation:

$$\text{ARGUMENT}_3 \xleftarrow{\text{CONDITION}} \text{CONNECTIVE}$$

Considering the text in Figure 2 as an example, the condition argument is "maximum steering."

Table 2 summarizes the statistics of causality annotations in the corpus regarding the number of arguments and causal relations.

3.4 Issues during the annotation work

This section describes several issues and ambiguous cases during the annotation work.

3.4.1 No causal connective

In English, a causal relationship is usually expressed with a connective. In contrast, a connective is often dropped in Japanese by simply placing two predicates in the same sentence. In this study, a causal relation is composed of two ARGUMENTS and a CONNECTIVE . A common problem with the above rule is the annotation of connectives. For instance, 「ため」 "due to" or 「場合」"in the case of", we can simply annotate the connective. However, for examples like 「エンジンが焼き付き走行不能になる」, there is no connective. We annotate the last character of the predicate indicating the reason, 「（焼き付）き」, as a connective to resolve the problem.

3.4.2 Handling of liquids in inter-component relationships

In many cases, it was difficult to determine whether the relation was PART-WHOLE or CONTACT for substances such as fuel and engine oil that go through multiple vehicle parts. In this study, we designed predefined rules to define these relations. Specifically, for the relation between liquid or gas and the tank where it is stored, we annotated the relation as PART-WHOLE. For the relation between the pathways (e.g., fuel pump) and parts where it is used, we annotated the relation as CONTACT .

3.4.3 Annotation for materials

In the recall report texts, mentions of material components such as "zinc" sometimes appeared. In this study, we ignore them during entity annotations; however, it may be necessary to annotate material components because they are often involved in the cause of recalls.

4 Experiments

As stated before, corpora are collected to build an automatic information extraction system and KB to promote better and safer vehicle design. In this section, we present experiments conducted to evaluate the extent to which our corpus can help build systems that extract NEs, relations, and causalities. The goals of the experiments are summarized as follows:

1. Evaluate the extent of accuracy of an information extraction model to automatically extract parts information from the text.

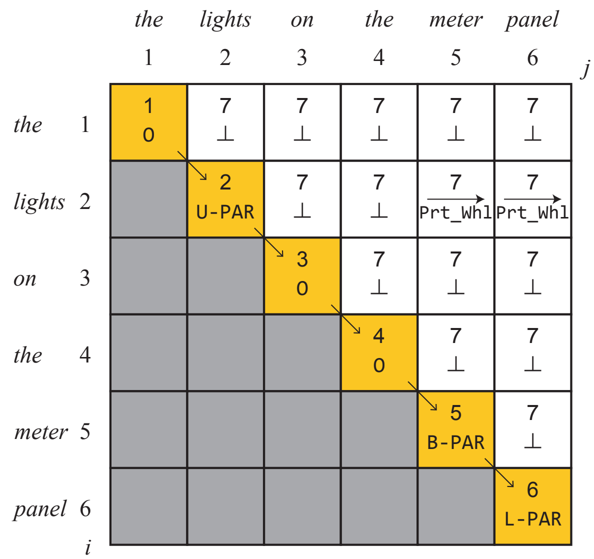

Figure 4: Overview of the table filling strategy.

2. Evaluate model accuracy on causality extraction.

3. Check if part information helps causality relation extraction.

We categorize the causality extraction task into named entity recognition (NER) and relation extraction (RE) tasks. Thus, we use the NER & RE models for part and causal extraction tasks.

4.1 Methodology

To extract NEs and relations jointly, we apply TablERT (Ma et al., 2022), a joint NER/RE model that achieves state-of-the-art performance on the CoNLL04 (Roth & Yih, 2004) and ACE05 English datasets. The model follows the table-filling framework proposed by Miwa & Sasaki (2014) to cast the join extraction of entities and relations as a table-filling problem. As shown in Figure 4, a table is used to represent the label space of both entities and relations. Diagonal cells are entity labels using BILOU notation (Ratinov & Roth, 2009); for example, the entity label for "lights" is *U-Part* meaning "lights" is a unit length PART entity. Upper triangular cells are relation labels, for example, there are PART-WHOLE relations pointing from "lights" to “meter” and “panel”. The model fills the table according to the order indicated by the number in each cell. The model first predicts the entity labels (the diagonal cells) sequentially using span features computed from contextualized representations and then predicts the relation labels (the off-diagonal cells) simultaneously using the scores of each word pair computed by the tensor dot-product (Ma et al., 2022).

We trained two models to extract the part entity and relations and to extract causality. Subsequently, we used part relations as external knowledge when training the table-filling model to verify how the entity and relation knowledge help improve the causality extraction. Employing the PURE (Zhong & Chen, 2021) method, we can add the information of part entities and relations to the end of the input text without modifying the architecture of the table-filling model. Specifically, we added a sequence containing all entity pairs with relations. For each entity pair, we added markers to the end of the text:

<E:S></E:S>relation_type<E:O></E:O>.

Here, <E:S> and </E:S> share the same position embedding with the start and end tokens of the subject (head) entity. Similarly, <E:O> and </E:O> specify the start and end positions of the object (tail) entity, respectively.

4.2 Experimental Settings

The annotated data were randomly divided into training and evaluation data in an 8:2 ratio. The open-source implementation of the table-filling model[5] was used in the experiments. We employed Japanese Bert as a pre-trained model[6]. The experiments were conducted on a single GPU of NVIDIA GTX 1080 Ti (11 GiB.) For hyperparameter settings, we set the learning rate to 5×10^{-5}, the dropout rate to 0.3, and the maximum length of the input tokens to 250.

4.3 Evaluation results

First, we present the experimental results for NER and RE tasks on the corpus annotated with NEs and relations in Table 3. Precision (P), recall (R), and F1-score (F1) were used for the evaluation. For part entity extraction, the model reached a high F1 score of 0.9710, and the precision and recall were also as high as the F1-score. The high F1-score indicates that the model can recognize the vehicle entity spans. Next, we observe the best overall performance for COREFERENCE and the lowest accuracy for ONEWAY-COREFERENCE. Their performance reflects the number of their occurrences in the corpus as the type COREFERENCE is the most frequent and type ONEWAY-COREFERENCE is the least frequent in the corpus (Table 1). Notably, although the number of relation instances annotated as COREFERENCE and PART-WHOLE are similar, the F1-score for PART-WHOLE is approximately 0.19 less than that for COREFERENCE. Moreover, the model does not perform well in predicting CONNECT and CONTACT, with f1-score 0.5045 and 0.6371, respectively.

Type	P	R	F1
NER			
PARTS	0.9674	0.9746	0.9710
RE			
CONNECT	0.5636	0.4566	0.5045
COREFERENCE	0.8972	0.9230	0.9099
ONEWAY-COREFERENCE	0.6042	0.3187	0.4173
PART-WHOLE	0.7221	0.7018	0.7118
CONTACT	0.6630	0.6131	0.6371
All	0.7585	0.7270	0.7424

Table 3: Experimental results on NER and RE.

	P	R	F1
TablERT	0.7120	0.7341	0.7229
TablERT + parts_info	**0.7193**	**0.7389**	**0.7290**

Table 4: Experimental results on causality extraction

Next, we present the experimental results for the causality extraction task in Table 4. The first row summarizes the results of causality extraction without entity and relation knowledge, and the second row summarizes the results of causality extraction with the external entity and relation information. Compared to the model without any external knowledge, including the part entity and relation knowledge improves all evaluation metrics. Although the

[5] https://github.com/YoumiMa/Enhanced_TF

[6] https://github.com/cl-tohoku/bert-japanese

improvement is modest, the results show that entities and relations can enhance the performance of causality extraction on the vehicle recall corpus.

4.4 Discussion

It is challenging for the model to extract the implicit relations. Although the model performs well on the NER task on the part-relation corpus, there is still room for improvement regarding RE, especially in predicting certain relation types.

For type COREFERENCE, the model reached an impressive f1-score of 0.9099. One possible reason for the high performance is the property that two PARTS entities paired with a COREFERENCE relation usually share a common word span. As presented in Table 3, the model can recognize word spans with high accuracy, resulting in a high accuracy in extracting COREFERENCE tuples as these tasks are similar. For instance, in Figure 1, PARTS "auxiliary braking device (electromagnetic retarder)" and "electromagnetic retarder" have a COREFERENCE relation and they share the common word span "electromagnetic retarder."

In contrast, the F1 scores for the other relation types were relatively low, possibly because knowledge about these relation types is usually absent from the document. Again, take the annotated document illustrated in Figure 1 as an example, "electromagnetic retarder" and "braking light" has a CONNECT relation; however, recognizing the relation is difficult even for a non-expert human. Thus, it is understandable that the problem is also difficult for a machine learning system without access to expertise in vehicles.

5 Conclusion

In this work, we have presented two annotations on the vehicle recall dataset: NE and relation annotation and causality annotation. We trained NER and RE models for the NER and RE and causality extraction tasks on both annotated corpora and presented the results. The results demonstrate the feasibility of building a causality relation system using an annotated corpus. Subsequently, we used the part entity and relation annotated corpus to improve causality extraction on the car recall corpus. The experimental results show that incorporating part entity knowledge improves the performance of causality extraction.

In future work, we will investigate a more effective approach to utilize NE and relation information to improve causality extraction.

Acknowledgments

This paper is based on the results obtained from a project, JPNP18002, commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

Jesse Dunietz, Lori Levin, and Jaime Carbonell. 2015. Annotating Causal Language Using Corpus Lexicography of Constructions. In *Proceedings of The 9th Linguistic Annotation Workshop*, pages 188–196.

Nancy A. Chinchor. 1998. Overview of MUC-7/MET-2, In *Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, April 29 - May 1, 1998.*

2002. In *Proceedings of the ACL-02 Workshop on Natural Language Processing in the Biomedical Domain.*, Association for Computational Linguistics, Phildadelphia, Pennsylvania, USA, edition.

Dina Demner-Fushman, Wendy W Chapman, and Clement J McDonald. 2009. What can natural language processing do for clinical decision support?, *Journal of the Association for Computing Machinery.*

Anna Rumshisky, Kirk Roberts, Steven Bethard, and Tristan Naumann. 2016. *Proceedings of the Clinical Natural Language Processing Workshop (ClinicalNLP)*, The COLING 2016 Organizing Committee, Osaka, Japan, edition.

2020. Natural language processing applications in business. Mohammed Bahja, ed. by *E-Business-Higher Education and Intelligence Applications*, IntechOpen.

Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. brat: a Web-based Tool for NLP-Assisted Text Annotation. In *EACL*, pages 102–107.

Roxana Girju, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney, and Deniz Yuret. 2007. SemEval-2007 Task 04: Classification of Semantic Relations between Nominals. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 13–18.

Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Sebastian Padó, Marco

Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 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,* pages 33–38.

Evan Sandhaus. 2008. The New York Times Annotated Corpus LDC2008T19. Web Download. Philadelphia: Linguistic Data Consortium, 2008.

Jesse Dunietz, Lori Levin, and Jaime Carbonell. 2017. The BECauSE Corpus 2.0: Annotating Causality and Overlapping Relations. In *Proceedings of the 11th Linguistic Annotation Workshop,* pages 95–104.

Nasrin Mostafazadeh, Alyson Grealish, Nathanael Chambers, James Allen, and Lucy Vanderwende. 2016. CaTeRS: Causal and Temporal Relation Scheme for Semantic Annotation of Event Structures. In *Proceedings of the Fourth Workshop on Events,* pages 51–61.

Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* pages 839–849.

James Pustejovsky, José Castano, Robert Ingria, Roser Sauri, Robert Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir Radev. 2003. TimeML: Robust specification of event and temporal expressions in text. *New directions in question answering,* page 28-34.

Paramita Mirza, Rachele Sprugnoli, Sara Tonelli, and Manuela Speranza. 2014. Annotating Causality in the TempEval-3 Corpus. In *Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language (CAtoCL),* pages 10–19.

Sampo Pyysalo, Filip Ginter, Juho Heimonen, Jari Björne, Jorma Boberg, Jouni Järvinen, and Tapio Salakoski. 2007. BioInfer: a corpus for information extraction in the biomedical domain. *BMC Bioinformatics,* 8(1), pages 1-24.

Claudiu Mihăilă, Tomoko Ohta, Sampo Pyysalo, and Sophia Ananiadou. 2013. BioCause: Annotating and analysing causality in the biomedical domain. *BMC Bioinformatics,* 14(1), pages 1-18.

Taku Kudo. 2006. Mecab: Yet another part-of-speech and morphological analyzer.

Bonnie Webber, Rashmi Prasad, Alan Lee, and Aravind Joshi. 2019. The Penn Discourse Treebank 3.0 Annotation Manual. *Philadelphia, University of Pennsylvania.*

Youmi Ma, Tatsuya Hiraoka,and Naoaki Okazaki. 2020. Named Entity Recognition and Relation Extraction Using Enhanced Table Filling by Contextualized Representations. *Journal of Natural Language Processing,* 29(1), pages 187–223.

Dan Roth and Wen-tau Yih. 2004. Annotating Causality in the TempEval-3 Corpus. In *Proceedings of the Eighth Conference on Computational Natural Language Learning (CoNLL-2004) at HLT-NAACL 2004,* pages 1–8.

Makoto Miwa and Yutaka Sasaki. 2014. Modeling Joint Entity and Relation Extraction with Table Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP),* pages 1858–1869.

Lev Ratinov and Dan Roth. 2009. Design Challenges and Misconceptions in Named Entity Recognition. In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL-2009),* pages 147–155.

Zexuan Zhong,and Danqi Chen. 2021. A Frustratingly Easy Approach for Entity and Relation Extraction. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* pages 50–61.