The Circumstantial Event Ontology (CEO) and ECB+/CEO: an Ontology and Corpus for Implicit Causal Relations between Events

Roxane Segers*, **Tommaso Caselli** ∓, **Piek Vossen** *
*Vrije Universiteit Amsterdam, The Netherlands
{r.h.segers, p.t.j.m.vossen}@vu.nl
∓Rijksuniversiteit Groningen, The Netherlands
t.caselli@rug.nl

Abstract

In this paper, we describe the Circumstantial Event Ontology (CEO), a newly developed ontology for calamity events that models semantic circumstantial relations between event classes, where we define circumstantial as inferred implicit causal relations. The circumstantial relations are inferred from the assertions of the event classes that involve a change to the same property of a participant. Our model captures that the change yielded by one event, explains to people the happening of the next event when observed. We describe the meta model and the contents of the ontology, the creation of a manually annotated corpus for circumstantial relations based on ECB+ and the first results on the evaluation of the ontology.

Keywords: Ontology, Event Modeling, Event Chaining, Causality, Annotated Corpora, Text Mining, Semantic Role Labeling

1. Introduction

Suppose we read a sequence such as: *"Today was the burial of Mary Johnson, that was broadcasted live on TV. The pop star died last week when her yacht capsized and sunk after hitting a tanker. Johnson was not wearing a life jacket and drowned."* As it is clear to most readers, but implicit in this sentence, there must be some relation between "hit", "capsize", "sinking", "drown", "die" and "burial". The interpretation of this sentence as a text, i.e. a unitary message (De Beaugrande and Dressler, 1981), requires some coherence relations between the events, that are not explicitly expressed. In the context of this occurrence, it is normal for a human reader to interpret the events as a chain of *consequences*. This coherence is the result of the fact that the events imply changes on a set of shared properties.

We consider this type of relations between event pairs as a case of *circumstantial relations*, i.e. relations between events which allows interpreting their occurrence in the world, and in a text, as coherent. A circumstantial relation makes clear "why" something happened, without necessarily predicting it. Circumstantial relations are a set of relations which include temporal, causal, entailment, prevention and contingency relations, among others.

We distinguish two types of circumstantial relations: episodic and semantic. An *episodic* circumstantial relation is a relation that holds between a pair of specific actual event instances in a specific context, where their connection is necessary to understand what is described in a meaningful and coherent way. For instance, the relation between the events "[a yacht] sunk" and "hitting [a tanker]" is a case of an episodic circumstantial relation: both events may happen independently without implying the other necessarily, but when described in the same context, or circumstance, a connection is created that explains their occurrence as a dependent relation.

On the other hand, we define *semantic* circumstantial relations as a relation that holds between event classes (abstracting from actual event instances), where an event of class A gives rise to another event of class B or vice versa, based on shared properties in the formalization of the classes.

For instance: the class "ceo:Shooting" has a semantic circumstantial relation with the class "ceo:Impacting", because they both share the property of translocation of an object from location X to Y. The latter as the outcome of the event, and the former as a condition to take place. Likewise, an "Impacting" event may, but not necessarily, lead to "ceo:Injuring" or "ceo:Damaging", which is based on the shared property of some object being damaged.

Modeling these relations provides a means to track chains of logically related events and their shared participants within and across documents. Semantic circumstantial relations define possible explanatory sequences of events, but not the actual explanatory sequences. Episodic relations, on the other hand, define circumstantial relations that are dependent on the actual occurrences of events in the world.

The Circumstantial Event Ontology (CEO) (Segers et al., 2017) [1], described in this paper, models such semantic relations, based on *shared properties* of the event classes with the aim to support the detection of episodic circumstantial relations in texts.

Modeling these semantic relations in an ontology will allow us to 1.) abstract over the different lexical realizations of the same concept (i.e. at an event mention level); and 2.) facilitate reasoning between event classes and enrich the extraction of information for event knowledge and event sequences.

The remainder of this paper is organized as follows: in section 2., we describe related work; in section 3. we explain the meta model and the development of CEO. Section 4. describes an annotated corpus of episodic circumstantial relations, that has been used to run preliminary experiments for the evaluation of the CEO. Experiments and results are

[1] CEO is publicly available with a CC-BY-SA license at `https://github.com/newsreader/eso-and-ceo`.

described in 5.. Finally, conclusions and future work are reported in section 6.

2. Related Work

Existing ontologies and models such as SUMO (Niles and Pease, 2001) and FrameNet (Ruppenhofer et al., 2006) provide explicit causal relations between event classes (SUMO), or preceding and causal relations (FrameNet). These causal relations are strict, meaning that if A happens, then B must happen as well. However, our relations are circumstantial, meaning that some instance of event class C and D can happen independently, but given the circumstance that they coincide, C likely implies D or D is likely implied by C because they share a property or a set of properties. The implication is however not necessary.

Previous work on the encoding of semantic relations between event pairs has focused on specific subsets of circumstantial relations. For instance, one example is the encoding of the entailment relations in WordNet (Fellbaum, 1998). With respect to the WordNet approach, we abstract from various event types (i.e. lexical items) and do not depend on relations defined at a synset level, by formalizing event knowledge and relations in an ontology. We also provide more details on the property involved.

Another related approach are narrative chains (Chambers and Jurafsky, 2010), that provide chains of various event mentions. However, the relation between these mentions is not specified explicitly but based on co-occurrence of participants and a basic precedence relation. Manual inspection of these chains revealed that dissimilar relations are implied within these chains, varying from temporal ordering, to episodic, up to causal.

The Penn Discourse TreeBank (PDTB) (Prasad et al., 2007) annotates contingency relations, of which causal relations are a subclass. In PDTB, the focus of the annotation is between two Abstract Objects (called Arg1 and Arg2), corresponding to discourse units, rather than event mentions. The contingency relation is annotated either in presence of an explicit connective, i.e. a lexical item, between the two abstract objects, or implicitly, by adjacency in discourse. In our approach, contingency relations are one of the possible values which express circumstantial relations, and, most importantly, they are independent of the presence of connectives or adjacency in discourse, but grounded on (shared) properties of events.

A related resource is the Rich Event Ontology (REO) (Brown et al., 2017), that provides an independent semantic backbone to different lexical resources such as FrameNet and VerbNet. REO will have explicit causal relations between event classes as well as predefined pre- and post conditions. However, these relations are more strictly defined and on class level. On the other hand, CEO maintains a looser definition in terms of causality, and takes into account the roles affected by the event and the circumstantial relation.

A resource such as CEO is envisioned to be of added value for several NLP tasks such as script mining, question answering, information extraction, and textual entailment, among others. Furthermore, the explicitly defined relations between events can be of help in reconstructing storylines (Van den Akker et al., 2010; Vossen et al., 2015) and improve the coherence of existing narrative chain models (Chambers and Jurafsky, 2010).

pre situation:	damaging-undergoer	inState	damaging-state-1
	damaging-state-1	hasRelativeValue	"+"
post situation:	damaging-undergoer	inState	damaging-state-2
	damaging-state-2	hasRelativeValue	"-"
	damaging-undergoer	isDamaged	true
	damaging-undergoer	hasDamage	damaging-damage
	damaging-damage	hasNegativeEffectOn	activity

Figure 1: The ESO assertions for the class eso:Damaging

3. The Circumstantial Event Ontology

CEO builds upon an existing event ontology called the Event and Implied Situation Ontology (ESO) (Segers et al., 2016). ESO is designed to run over the output of Semantic Role Labeling systems by making explicit both the ontological type of the predicative element and the situation that holds before, during and after the predicate. Each so called pre-, post- and during situation consists of a set of properties and roles that define what holds true. For instance, as can be seen in Figure 1, the pre- and post-situations of the event class "eso:Damaging" define:

- that something is in a "relatively plus (+)" state (pre-situation);
- that this something is in a "relatively less (-)" state, i.e. it underwent a loss or a negative change, relatively to the state before the damaging (post-situation);
- that some object is in a state "damaged" after the event (post-situation);
- that something has some damage which has some negative effect on some activity (post-situation).

ESO allows to track chains of states and changes over time, whether explicitly reported or inferred. However, ESO does not provide any explicit definition on what event class logically precedes or follows some other event class, i.e. the pre-, post- and during situations provide only descriptions of properties of the participants of the event in analysis. In CEO, we further developed the event hierarchy of ESO, and the expressiveness of the pre-, post-, and during situations in order to infer the circumstantial semantic relations between the classes.

3.1. The CEO Meta Model

CEO is an OWL2 ontology and its meta model fully adopts and extends the ESO model (Segers et al., 2016). The reasons to reuse and extend it are twofold: 1.) The ESO classes and roles are mapped to FrameNet, therefore we can rely on existing SRL techniques and models to instantiate CEO (Björkelund et al., 2009; de Lacalle et al., 2016); 2.) ESO provides a model that defines what situation, or state, is true before and after an event, thereby already providing the

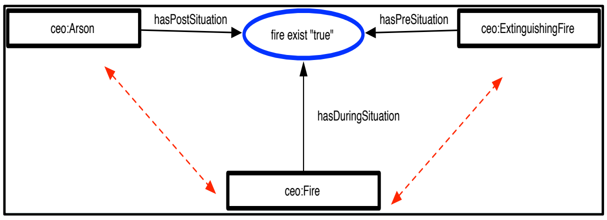

Figure 2: The meta model of CEO and the chaining of classes by shared properties in the pre-, during-, and post situations.

initial hooks to infer the circumstantial semantic relations. This principle is illustrated in Figure 2. The black boxes represent event classes in CEO; each class has at least one assertion (ceo:fire exist ”true”) that is shared with two other classes. In the case of “ceo:Arson” it is part of its post situation; it is the during situation of “ceo:Fire” and the pre situation of “ceo:ExtinguishingFire”. Based on these shared properties we can infer a semantic circumstantial relation that is in this model represented by the red arrows. Whether the shared property is in a pre-, during-, or post situation implicitly defines the logical order of the events.

The full expressiveness of a class in CEO.owl is illustrated in Figure 3, where we transcribed the class “ceo:Arson” and its assertions in a human readable format. Each class has a subclass relation (subclassOf) and a definition (Definition). Furthermore, the class “ceo:Arson” is mapped to FrameNet (fn:Action) and SUMO (sumo:Arson). All mappings were created manually. Next, we show the assertions in the pre- (pre situation), during- (during situation), and post situation (post situation). Each assertion consists of a property and one or more roles that are mapped to FrameNet (role mappings are not shown).

CEO properties consist of 1.) binary properties where two roles are connected, e.g. (hasPurpose, deteriorates), 2.) unary properties that connect a role with a boolean expression ”true” or ”false” (e.g. inDanger), or a relative value ”+” or ”-” (e.g. hasRelativeValue). For some roles, we defined an OWL existential restriction if no instance can be found in a text. In this case, the role will be instantiated with a blank node and some URI. In Figure 3, this occurs for the roles “damaging-state-1” and “damaging-state-2”.

Figure 4 illustrates the inference capabilities of CEO using FrameNet-based role labeling. Only those assertions can be fired and instantiated if an instance of the CEO role is found via the FrameNet mappings. In this case, there is no Frame element and instance found for the CEO role ”damage”, hence the assertion can not be instantiated. In line 2, we see how a blank node is created for the role ”damaging-state-1”, encoded here as ”abc123”.

In short, the assertions in Figure 4 define that 1.) the fire does not exist before the Arson (line 5), but it does during (line 10) and after (line 21); 2.) Mary is in offense during (line 14) and after (line 22) the arson of the stables, 3.) the stables and the village are in danger during (lines 12 and 13) and after (lines 19 and 20) the arson; and 4.) the stables are damaged after the arson (line 18).[2]

```
0   -Arson
1   subclassOf: IntentionalDamaging
2   Definition: “The subclass of IntentionalDamaging where someone
3   deliberately sets some object on fire.”

4   CLASS MAPPINGS:
5   closeMatch: fn:Arson
6   closeMatch: sumo:Arson

7   ASSERTIONS:

8   pre situation
9   agent              hasPurpose              purpose
10  undergoer          inState                 damaging-state-1
11  damaging-state-1   hasRelativeValue        "+"
12  undergoer          inDanger                false
13  fire               exist                   false

14  during situation
15  agent              deteriorates            undergoer
16  undergoer          inDeterioration         true
17  agent              hasIntention            “deterioration-of-undergoer”
18  fire               exist                   true
19  agent              endangers               undergoer
20  undergoer          inDanger                true
21  place              inDanger                true
22  agent              inOffense               true

23  post situation
24  undergoer          inState                 damaging-state-2
25  damaging-state-2   hasRelativeValue        "-"
26  undergoer          isDamaged               true
27  undergoer          inDanger                true
28  place              inDanger                true
29  fire               exist                   true
30  undergoer          hasDamage               damage
31  damage             hasNegativeEffectOn     activity
32  agent              inOffense               true
```

Figure 3: The expressiveness of an event class in CEO, including subclass relation, mappings and assertions and roles in the pre-, post, and during situation.

3.2. Semantic Circumstantial Relations between Event Classes

CEO is modeled in such way that it allows for inferencing, chaining classes, and reasoning over the assertions, roles, and role instances.

For chaining the event classes, the most basic way is to track paths trough the ontology, based on shared properties in the class assertions. This is illustrated in Figure 5. Here, in each box we show eight different sentences related to the same Arson incident. The property in red (inOffense ”true”) is in the post situation of “ceo:Arson” and in the pre- situation of the event class “ceo:Arresting”. Likewise, the property ”fire exist true”, which is marked here in orange, ties a circumstantial relation from “ceo:Arson” to the class “ceo:Fire”, and from this latter class to “ceo:ExtinguishingFire”. As such, we can chain the event mentions based on shared semantic properties. To exploit the model at its maximum, a reasoner will have to take into account the properties and their values, the roles, as well as the role instances.

[2]A full transcription of the CEO classes including all assertions, the inherited assertions and example sentences that show the instantiation can be found at `https://github.com/newsreader/eso-and-ceo`.

Arson
"[Mary] deliberately **set** [three stables] **on fire** in [John's village]."

1	*pre situation*		
2	three stables	inState	"abc123"
3	"abc123"	hasRelativeValue	"+"
4	three stables	inDanger	false
5	fire	exist	false
6	*during situation*		
7	Mary	deteriorates	three stables
8	three stables	inDeterioration	true
9	Mary	hasIntention	"deterioration-of-undergoer"
10	fire	exist	true
11	Mary	endangers	three stables
12	three stables	inDanger	true
13	John's village	inDanger	true
14	Mary	inOffense	true
15	*post situation*		
16	three stables	inState	"abc456"
17	"abc456"	hasRelativeValue	"-"
18	three stables	isDamaged	true
19	three stables	inDanger	true
20	John's village	inDanger	true
21	fire	exist	true
22	Mary	inOffense	true

Figure 4: Example of what the CEO assertions infer from a SRL labeled sentence for the pre-, during and post situation of the event.

3.3. Building the CEO

CEO is designed to capture chains of events in newswire, more specifically calamity events. We define a calamity event as any event where some situation turns from relatively positive to some relatively negative state due to changes in the world, either intentional or not. Event classes that define processes where some agent tries to improve some situation in reaction to some calamity are also modeled in CEO, e.g. going from a relatively negative situation back to a relatively positive situation. Examples of calamity event classes are "CyberAttack" and "Earthquake". Examples of event classes where an attempt to some improvement of a situation is made are "Repairing" and "Evacuation".
ESO already provides event classes for calamities, though the coverage is rather limited, because it was designed for the economic-financial domain. As such, we massively extended the hierarchy from the initial 63 event classes in ESO to the 223 event classes in CEO 1.0. To the best of our knowledge, no formal ontology specific for calamities and the inter-event relations exist. Some thesauri such as the IPTC [3] contain terms for calamities but these are not formalized and provide few relations. Therefore, we decided to define a new model, reusing existing resources as much as possible.
As a starting point for the identification of instances of the calamity classes in CEO, we used Chamber's narrative chains (Chambers and Jurafsky, 2010). This selection was made manually, based on at least three calamity events per event chain. We also manually selected FrameNet frames that capture calamity events and we used the SUMO ontology as a backbone for modeling our initial list of verbs and frames. Finally, we defined SKOS mappings from each CEO event class to FrameNet and SUMO[4], thus providing the opportunity to use CEO on SRL labeled text as well as to find the vocabulary expressing calamities by means of the lexical units mapped to frames in FrameNet and the mappings to Princeton WordNet that are defined in SUMO. An overview and specification of all modeling decisions regarding class selection, class hierarchy and defining the assertions, properties, roles and role mappings to FrameNet can be found in the CEO documentation. [5]

3.4. Contents of CEO

In January 2018, we released CEO 1.0. The ontology consists of 223 event classes of which 189 are fully modeled with pre-, during and post situations. For 34 classes, we have a minimal set of assertions. These classes pertain to natural disasters and will be modeled for CEO version 1.1. Further, we defined 92 binary properties and 29 unary properties. In total, 189 unique situation rules were defined that consist of 192 binary situation rule assertions and 264 unary rule assertions.
Further, all classes are mapped to FrameNet frames (265 mappings) and SUMO classes (195 mappings), and the CEO roles to FrameNet elements (265 mappings).

4. The ECB+/CEO Corpus

In addition to the CEO, we developed a corpus of annotated circumstantial event relations. For this, we build upon an existing corpus, specifically annotated for event coreference: the ECB+ Corpus (Cybulska and Vossen, 2014). ECB+ consists of 984 news articles divided over 42 topics. From these topics, we manually selected 22 topics (508 articles) that cover calamities such as earthquakes, murders, hijacks and arson. In ECB+, only the most relevant event mentions are manually annotated. For ECB+/CEO, we automatically extended the set of annotated event mentions by applying a state-of-the art machine learning based system [6]. Two linguistically trained annotators were hired for the selection of relevant calamity events and the annotation of circumstantial relations.
More specifically, the annotation procedure consisted of the following steps:

1. Select event mentions denoting calamity events and generate corresponding event instances;
2. Extending existing ECB+ coreference sets with new mentions;
3. Creating new coreference sets for new calamity mentions;
4. Creating circumstantial relations (CEO links) between the event instances where each instance refers to a set of coreferential mentions.

[3] https://iptc.org/

[4] https://www.w3.org/2004/02/skos/

[5] https://github.com/newsreader/eso-and-ceo

[6] (Caselli and Morante, to appear) https://github.com/cltl/TimeMLEventTrigger

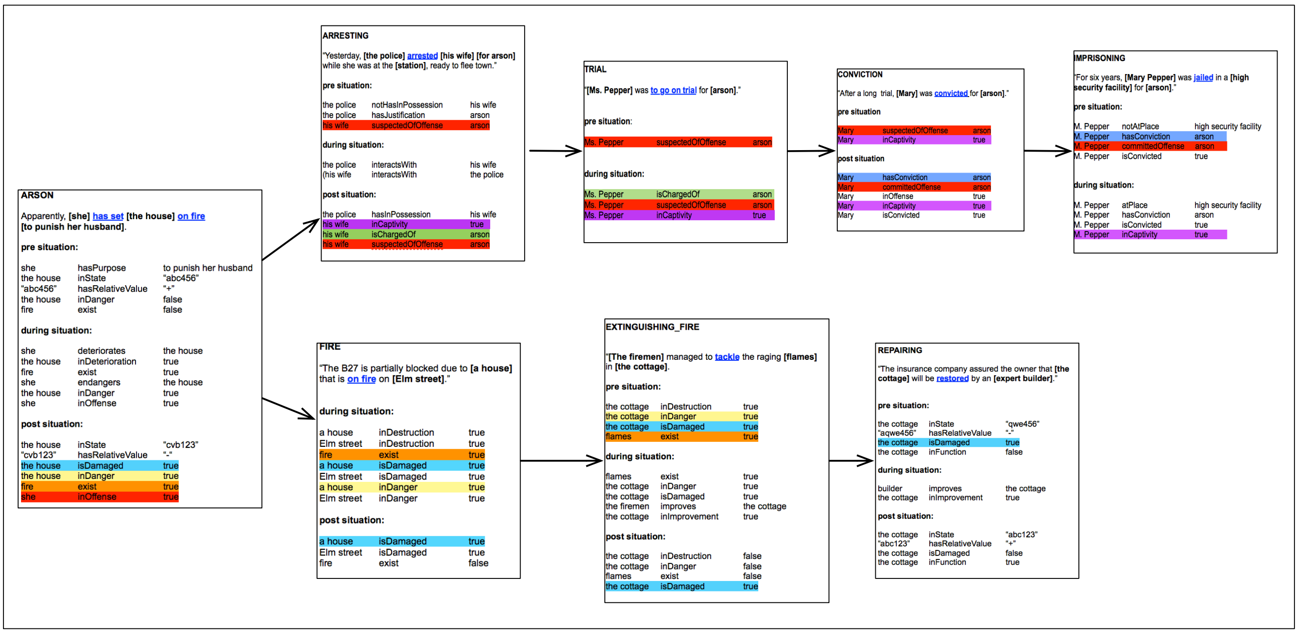

Figure 5: Inferring circumstantial relations from shared properties in the pre-, post-, and during assertions between event expressions in eight sentences.

Annotators were asked to connect pairs of calamity event instances with a CEO link if one event instance could be used to explain the occurrence of the other.

For the value of a CEO relation, the annotators could opt for the default value (has circumstantial post event - `HCPE`) or the subset relation (`hasSubevent`).[7] The `HCPE` relation is directional and is defined from a source, or trigger, event to a target, or consequence, event.

We followed the original ECB+ annotation guidelines where applicable and we deviated on certain points. For instance, we only annotated calamity event mentions; the participants, locations and time expressions were not annotated. Furthermore, speech acts and events expressing cognition, perception and emotions were excluded for the annotation.

Negated events are annotated and added to the CEO links, as a statement that something did *not* happen points at the fact that it usually does happen (e.g. he was *shot* but not *injured* severely).

For the definition of coreference, we specified that two event mentions are coreferential if they (more or less) denote the same concept, and they share the same participants, time, and location. Event coreference was only annotated within document, and not across documents, like in ECB+.

In table 1 we show the results of the annotation. In total, 508 articles were annotated for ECB+/CEO which resulted in 3038 new event instances expressing calamities. Further, 3448 new event coreference sets were created. Not every instance and coreference set ends up in a CEO link as for many events no circumstantial event or subevent is present in the text. As such, 2437 CEO links were created of which 2244 circumstantial ones and 193 subevents. On average, every ECB+/CEO article contains about 7 new coreference sets and about 5 different circumstantial relations.

	ECB+	ECB+/CEO
Instances	3323	3038
Coreference sets	3323	3448
CEO relations	-	2437
- of which Circumstantial	-	2244
- of which subEvent	-	193

Table 1: Overview of the annotations made for ECB+/CEO in contrast with ECB+ for the topics annotated

For the annotation, we used the CAT annotation tool (Bartalesi Lenzi et al., 2012) which outputs the annotations in XML. In terms of annotation effort, a single article took about 30 minutes to annotate on average. The corpus and the annotation guidelines are publicly available at `https://github.com/newsreader/eso-and-ceo`.

Inter Annotator Agreement For the calculation of the Inter Annotator Agreement (IAA), we selected 25 articles from five different topics in ECB+/CEO covering variation in article length and complexity. The evaluation was carried out on the CEO links. Agreement was calculated on the existence, or identification, of CEO links.

CEO links are created between event instances, where each instance points to a set of event mentions in the document. These sets are defined as coreference relations. To evaluate the quality and reliability of the CEO links, we calculated the inter-annotator agreement (IAA) by means of Cohen's Kappa score (Cohen, 1960). We obtained a value of 0.54. To better understand the reasons behind such a score, we randomly inspected some annotated articles. As an outcome of this inspection, it appeared that the major differences beween the annotators were due to mismatches

[7] Subevents are currently not modeled in CEO, but they were annotated for future experiments and evaluations.

in the coreference sets, rather than in actual disagreements on the presence/absence of a CEO link. As such, we manually added a post processing step to align those coreference sets where either one or both annotators missed one or more mentions. To avoid introducing bias, we harmonized the coreference sets only if there were no conceptual differences between them. To clarify, if annotator A created one coreference set with three different mentions, and annotator B created two sets with the same mentions, we did not merge the sets of annotator B. With this post processing step, we solved 107 cases of partial disagreements on event coreference.

After this, we calculated again the IAA Cohen's kappa and reached a score of 0.76. Following Landis and Koch (1977), a score between 0.61 and 0.80 is considered substantial.

Both reported kappa scores are based on 21 out of the 25 initial articles. For four articles, the annotators agreed that there were no CEO links at all, and thus we excluded them.

Analysis of the disagreements We inspected some cases of clear disagreements in the annotated CEO links. These disagreements relate to differences in interpretation and to some unavoidable errors. For differences in interpretation, we see that the annotators disagreed whether some mention denoted the same concept or not. For instance, A1 created a CEO link between "suicide, hang" and "dead", while A2 interpreted all three mentions as denoting the same concept and did not create a CEO link. Further, there are disagreements on whether or not some mention still expresses a calamity and aftermath. As such, most agreements where e.g. A1 added an additional CEO relation and A2 did not, the relation leans towards a episodical one and not a semantic one. For those CEO relations for which there is agreement, these episodical relations are sparse. Further, we did not see any cases where the annotators disagreed on the type of the relation (HCPE or subEvent), or disagree on the directionality of the relation.

Creation of an initial CEO vocabulary For the annotation of ECB+/CEO, the annotators have focused on the creation of circumstantial relations between event instances. The instances themselves were not typed with a CEO class as it was thought to be too difficult for the annotators to do this. In order to know what class an event mention refers to, we extracted all mentions from the event coreference sets in the corpus. All mentions have been mapped manually to a CEO class. In total, 650 unique mentions were annotated with a total frequency of 3982. 14 unique mentions could not be mapped as they were too polysemous, 25 unique mentions were not mapped as they were out of domain. In terms of coverage, the vocabulary extracted from the corpus covers about 50% of the classes in CEO, meaning that 111 classes modeled in the ontology are not represented in the corpus. Likewise, the vocabulary points at 78 mentions that potentially can be added to the ontology, e.g. 'peace' and 'bankruptcy'. Most of these mentions however, point at very fine grained sub events related to trials and are basically out of domain.

5. Experiment and Evaluation on the ECB+/CEO corpus

We ran a first experiment to analyse to what extent CEO is able to connect events by means of semantic circumstantial relations, based on shared situation properties only. That implies that for this experiment, we deliberately did not take into account the CEO roles, the property values or the role instances to further fine tune the event chaining. The reason for this was twofold: 1.) we wanted to be able to analyse what CEO can achieve without any advanced reasoner and with just simple heuristics and 2.) we did not want to be affected by error propagation coming in from a NLP pipeline.

For this experiment, we developed the CEO-Pathfinder[8] (version 0.1) that checks for possible relations between events based on shared event properties in the pre-, post-, and during situations. CEO-Pathfinder compares all the mentions of events within a specified context window and checks the pre-, post- and during properties for matches. It uses a lexicon of 650 mentions that have been mapped to one or more CEO classes. The properties of associated classes (C1) and (C2) are compared as follows:

1. from a post situation in C1 to a pre situation in C2;
2. from a during situation in C1 to a pre situation in C2;
3. from a post situation in C1 to a during situation in C2;

We count the number of matching properties across classes of two mentions in both directions, assuming that the order of mention is not necessarily the order of the events in time. The software uses a threshold for the minimal matching properties. If below the threshold, no circumstantial relation is extracted. For both directions: C1 is circumstantial to C2 or C2 is circumstantial to C1, we then take the highest number of shared properties. If the shared properties are equal, the order of the mentions determines the direction of the circumstantial relation. The software can use the directly expressed properties or the inherited properties as well. We experimented with both options but got the best results with the directly expressed properties.

Finally, we implemented different context strategies for comparing mentions of events: 1) mentions within the same sentence (most strict), 2) one preceding and following sentence, 3) two preceding and following sentences, 4) all mentions in the full document.

Baseline system As a baseline, we compared all the mentions within the previous context windows 1, 2, 3 and 4 sentences, by assuming a CEO relation between all of them following the mention order. Table 2 shows the precision, recall and F1 results considering the order of the relation and ignoring the order (loose). B-1s is the baseline where we compare only mentions within the same sentence. B-3s is the baseline considering also one preceding and one following sentence, B-5 two preceding and following, and B-all the full document.

Not surprisingly, the precision results are all very low, both for order sensitive and loose matching. Highest recall is

[8] https://github.com/cltl/ceopathfinder

Baseline	B-1s	B-3s	B-5s	B-all
Precision order	0.236	0.202	0.188	0.144
Recall order	0.072	0.140	0.200	0.511
F1 order	0.110	0.166	0.194	0.225
Precision loose	0.556	0.432	0.386	0.282
Recall loose	0.169	0.300	0.409	0.999
F1 loose	0.259	0.354	0.397	0.439

Table 2: Result of the baseline system with different context windows

obtained for comparing all mentions ignoring the order: 0.99. When we take the order into account, we see that recall drops to 0.502. This means that about 50% of the event pairs with a CEO relation also are mentioned in their causal order. This pattern also holds for the other baselines where we compare mentions within limited contexts: recall drops by more or less 50% in all cases. Obviously, recall drops when we restrict the context, while precision increases. This means that there is a substantial amount of circumstantial relations expressed beyond the sentence boundary and event a context of five sentences that appears to be relevant.

Evaluation results In Table 3, we show the results for the CEO-Pathfinder exploiting the shared assertions from the ontology. The upper part represents the results when setting the threshold to one matching assertion and the lower part setting the threshold to two matching assertions. The different columns show the different context windows for comparing mentions similar to the previous baseline results. Overall, the precision and F1 results of the CEO-based approach outperform the baseline. We can see that the recall is much lower as can be expected.

1 assertion	CEO-1s	CEO-3s	CEO-5s	CEO-all
Precision order	0.455	0.400	0.379	0.311
Recall order	0.011	0.023	0.043	0.086
F1 order	0.021	0.044	0.077	0.135
Precision loose	0.650	0.563	0.512	0.420
Recall loose	0.015	0.033	0.058	0.117
F1 loose	0.029	0.062	0.104	0.183
2 assertions	CEO-1s	CEO-3s	CEO-5s	CEO-all
Precision order	0.645	0.498	0.464	0.405
Recall order	0.006	0.011	0.020	0.040
F1 order	0.011	0.021	0.038	0.073
Precision loose	0.710	0.556	0.509	0.439
Recall loose	0.006	0.012	0.021	0.044
F1 loose	0.012	0.023	0.041	0.079

Table 3: Results of Pathfinder using different settings (1 or 2 shared assertions) and varying context windows

The highest precision (0.710P) is achieved using the same sentence as a context window and, remarkably, ignoring the order. We also see that 2 shared assertions instead of 1, increases precision. Increasing the context window lowers precision and increases recall, where we have the highest recall (0.117R) and F1 (0.183F1) using the complete document and 1 shared assertion but ignoring the order.

To analyze the low recall, we collected all mentions for which the lexicon did not provide a CEO class to see if this could explain the difference in recall between the baseline and the CEO-version. The baseline does not use any external resource and is not dependent on the lexicon to map mentions to CEO classes. We found 3246 out-of-vocabulary cases that represent 12,999 mentions. Note that the event mentions are generated using ECB+ gold data and silver-data generated from the full text documents. We analyzed the most frequent of these mentions and did not observe any major gaps in the lexicon (an exception being *drunken driving* and *drunk driving* occur 8 and 9 times) that could explain the drop in recall.

We also abstracted from the assertions by only considering the property predicate. When ignoring the order (loose), we get 0.114P, 0.165R and 0.135F1. We thus see a slight drop in precision but higher recall and F1. Nevertheless, the difference is small and does not outweigh the value of using full assertions to connect events using circumstantial causal relations with specific implications for the involved participants.

To conclude: there is still substantial ground to cover in the CEO to increase the recall but the results for precision of the relations without using any further information on time and participants are promising. Especially, as the CEO appears to capture relations far beyond the context of sentences and even paragraphs.

6. Conclusion and Future Work

We have described our work on an event ontology that captures calamity events in newswire and the semantic circumstantial relations that hold between event classes, based on shared properties in the pre-, post- or during situations defined for each class.

First experiments and evaluations show that applying very basic heuristics to retrieve circumstantial relations based on assertions properties gives promising results with respect to precision. For increasing both recall and precision, adjustment and extension of the defined situation assertions will be needed as well as developing reasoner that can take into account the roles, property values and role instances to further scope the chaining of event instances.

Future work includes developing a reasoner and additional experiments on finding more sophisticated heuristics for salient circumstantial paths in the ontology. Further, we will evaluate the added value of our model extrinsically, by means of a QA task. For this, we are designing a Question-Answering task, where systems will have to provide answers to questions ”why” a certain event has taken place rather than factoid questions by providing the most relevant and direct preceding event that can be seen as an explanation.

7. Acknowledgements

We thank Nynke Visscher and Joy van Wooning for their work and input on the annotation task.

The work presented in this paper was funded by the Netherlands Organization for Scientific Research (NWO) via the Spinoza grant, awarded to Piek Vossen in the project ”Understanding Language by Machines”.

8. Bibliographical References

Bartalesi Lenzi, V., Moretti, G., and Sprugnoli, R. (2012). Cat: the celct annotation tool. In *LREC*, pages 333–338.

Björkelund, A., Hafdell, L., and Nugues, P. (2009). Multilingual semantic role labeling. In *Proceedings of CoNLL-2009*, Boulder, CO, USA.

Brown, S., Bonial, C., Obrst, L., and Palmer, M. (2017). The rich event ontology. In *Proceedings of the Events and Stories in the News Workshop*, pages 87–97, Vancouver, Canada, August. Association for Computational Linguistics.

Caselli, T. and Morante, R. (to appear). Agreements and Disagreements in Temporal Processing: An Extensive Error Analysis of the TempEval-3 Systems. In *Proceedings of Language Resources and Evaluation Conference (LREC 2018)*.

Chambers, N. and Jurafsky, D. (2010). A database of narrative schemas. In *Proceedings of the 9th Language Resources and Evaluation Conference (LREC2010)*.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

Cybulska, A. and Vossen, P. (2014). Using a sledgehammer to crack a nut? lexical diversity and event coreference resolution. In *Proceedings of the 9th Language Resources and Evaluation Conference (LREC2014)*, Reykjavik, Iceland, May 26-31.

De Beaugrande, R. and Dressler, W. (1981). *Textlinguistics*. New York: Longman.

de Lacalle, M. L., Laparra, E., Aldabe, I., and Rigau, G. (2016). A multilingual predicate matrix. In *Proceedings of Language Resources and Evaluation Conference (LREC 2016)*.

Fellbaum, C. (1998). *WordNet: an electronic lexical database*. MIT Press.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, pages 159–174.

Niles, I. and Pease, A. (2001). Towards a standard upper ontology. In *Proceedings of FOIS-Volume 2001*. ACM.

Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A., Robaldo, L., and Webber, B. L. (2007). The penn discourse treebank 2.0 annotation manual.

Ruppenhofer, J., Ellsworth, M., Petruck, M., Johnson, C., and Scheffczyk, J. (2006). *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California.

Segers, R., Rospocher, M., Vossen, P., Laparra, E., Rigau, G., and Minard, A. (2016). The event and implied situation ontology: Application and evaluation. In *Proceedings of Language Resources and Evaluation Conference (LREC 2016)*.

Segers, R., Caselli, T., and Vossen, P. (2017). The circumstantial event ontology (ceo). In *Proceedings of the Events and Stories in the News Workshop*, pages 37–41, Vancouver, Canada, August. Association for Computational Linguistics.

Van den Akker, C., Aroyo, L., Cybulska, A., van Erp, M., Gorgels, P., Hollink, L., Jager, C., Legene, S., van der Meij, L., Oomen, J., van Ossenbruggen, J., Schreiber, G., Segers, R., Vossen, P., and Wielinga, B., (2010). *Historical Event-based Access to Museum Collections*, pages 1–9. CEUR-WS (online).

Vossen, P., Caselli, T., and Kontzopoulou, Y. (2015). Storylines for structuring massive streams of news. In *Proceedings of the 1st Workshop on Computing News StoryLines (CNewS 2015) at the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2015)*, Bejing, China.

9. Language Resource References

VV.AA. (2017). *WordNet*. unspecified, ISLRN 379-473-059-273-1.