Beyond Causality: Representing Event Relations in Knowledge Graphs

Youssra Rebboud[0000−0003−3507−5646], Pasquale Lisena[0000−0003−3094−5585], and Raphael Troncy[0000−0003−0457−1436]

EURECOM, Sophia Antipolis, France
{youssra.rebboud, pasquale.lisena, raphael.troncy}@eurecom.fr

Abstract. Dynamic environments can be modeled as a series of events and facts that interact with each other, these interactions being characterised by different relations including temporal and causal ones. These have largely been studied in knowledge management, information retrieval or natural language processing, leading to several strategies aiming at extracting these relationships in textual documents. However, more relation types exist between events, which are insufficiently covered by existing data models and datasets if one needs to train a model to recognise them. In this paper, we use semantic web technologies to design FARO, an ontology for representing event and fact relations. FARO allows representing up to 25 distinct relationships (including logical constraints), making it a possible bridge between (otherwise incompatible) datasets. We describe the modeling decision of this ontology resource. In addition, we have re-annotated two already existing datasets with some of the FARO properties.

Keywords: Semantic Web · Ontology · Event Relations

1 Introduction

In our experience of the world, we observe continuous occurrences of events. We may connect new events to one or more previous ones, giving birth to relationships of several types, such as cause-effect, relatedness, co-occurrence in time or space, etc. Even restricting our research to causality, we need to take into account several scenarios like *preemption* (causing an event which was going anyway to happen) and *disconnection* (making an event happening by removing the cause of not-happening) [23]. Events can influence each other (reciprocally or not), even without being recognised as cause-consequences. An event can be made of sub-events, each of them potentially relating to others. Being able to represent and exploit those relationships can be beneficial for different applications, involving the general public and domain experts.

The semantic web provides methods and tools to represent facts in Knowledge Graphs (KG) generally expressed in RDF. Some KGs are even specialised for representing event-centric information [7]. In Temporal Knowledge Graphs (TKG), each edge of the graph includes time information for identifying the temporal validity of a triple [20, 6]. It is possible to use the event and time information for inferring edges [27, 10]. However, while TKGs are capable to represent an event occurrence, there are not suitable to represent inter-events relationships, making it hard to retrieve flow of events.

In this paper, we introduce the *Facts and Events Relationship Ontology (FARO)*, a data model for representing events relationships in Knowledge Graphs. In particular, we aim to design a structure which make possible to navigate through semantic links between events, exploring the flow of events backwards (searching for the causes or conditions of an event), forward (looking at consequences) or passing through other kind of connections. In other words, we want to make possible the creation of interconnected timelines of events, in which the connections between two consecutive points have explicit semantics. A such created graph would serve to improve the performance of downstream task (namely link prediction) and the explainability in decision making systems. We present several contribution:

- We compare a multitude of partially overlapping models, in order to understand which relationships should be represented because of interest of the community – Section 2;
- As an outcome of the literature review, we introduce the FARO model – Section 3;
- In other to foster future research, we realise a first Event Relation dataset that includes numerous event relations. This dataset has been obtained by re-annotating two existing datasets and by extending the TimeML format [22] with a new `RLINK` tag – Section 4.

We conclude in Section 5, summarising the contribution and the resources.

2 Related Work

In the literature, several works have studied event relationships, the most common type of relationships being **temporality**. Fan et al. [4] identified 13 temporal relations – to be used in the context of 3D simulation –, including simultaneity (*equal*) and 6 other asymmetric (directed) properties, with their respective inverse – e.g. *before / after*. Equivalent relations are included in [9], with the addition of *Vagueness*. **Mereology** in the context of events – i.e. the interaction between sub-events and super-events – is also often represented [6, 9, 13, 24, 29]. Finally, the literature mentions more kinds of relation that we can group under the name of **contingency**. Wolf distinguishes the causality relations in four different concepts [31]:

- CAUSE: event A that leads to an event B;
- ENABLE: condition C to make an event B possible;
- PREVENT: event A that avoids an event B;
- DESPITE: event A did not succeed in avoiding an event B.

Hong et al. [9] designed one of the most complete event-event relationship classification, including 5 types (Inheritance, Expansion, Contingency, Comparison, Temporality) and 21 sub-types, with possible overlaps between classes. To the best of our knowledge, this is the only work including **comparative relations** which cover three kinds of relation types, such as *Opposition*, when two events are improbable to be both true (*parole* → *sentenced*), *Negation*, when two events can be both true in different time slots, but not simultaneously (A is behind bars →A left). However, several relations between events are not accompanied by proper definitions, while still some relation types are missing.

Several ontologies have been published using semantic web technologies. While some of them do not include relations between events (e.g. LODE [25]), most of them include at least the concept of sub-events, such as in the *Event Pattern* [13], the *Event Ontology*[1], and the *Simple Event Model (SEM)* [29].

Event Model F is an ontology created to support the response in emergency events [24]. It includes three kind of event relationships: mereological, causal and correlation. Its `Justification` class enables to support the relationship with provenance – e.g. opinion, scientific law, etc. However, this is modeled by including classes – e.g. `EventCompositionSituation` and `EventCompositionDescription` – with the only purposes of connecting events and defining their roles. As a consequence, there are no direct links between the composite super-event and its components sub-events (same for cause-effect). Furthermore, only 1-1 relations are foreseen, so additional instances must be created for aggregating causes/effects. All this led to a complex model, hard to understand and to adopt.

One of the most popular models among libraries and cultural institutions is *CIDOC CRM* [3]. It is an event-centric model, in which everything is represented though the interlinking of events of creation, production, movement, destruction, etc. Among its properties, there are some which intend or allow to interlink events, instantiating temporal relations (e.g. `P176 starts before the start of`), mereological relations (`P9 consists of`), causal relations (`P17 was motivated by`), and even include intentionality (`P20 had specific purpose`).

It is evident from the literature the necessity to represent, next to proper events, also some *state* or *condition*, lasting in time. This concept has been modeled as a sub-class of event [11] or as a completely separate class [5].

Several datasets for the detection of events and event relations are available, focusing mostly on temporal relations or on pure causality. Temporal relations have been largely investigated since 2009 in the TempEval shared task [28], which used the standard TimeML format and the TimeBank corpus [22]. The latter has been extended in CausalTimeBank [18] that follows the {CAUSE, ENABLE, PREVENT} model. In addition, events are marked as *factual* (happened), *counterfactual* (not happened) or *non-factual* (possibilities), while their relation can be *certain* or *uncertain*. On top of TimeML, the EventStoryLine dataset is proposed in [1], and includes the representation of causes and consequences in the context of `PLOT_LINKs`, for tagging events that are relevant in a plot.

[1] http://motools.sourceforge.net/event

EventKG [6] is a knowledge graph of harmonised and interlinked events extracted from several resources, such as Wikidata and YAGO [8]. It includes over 1,3 million events, linked to their spatial and temporal coordinates. Only the connection between sub-events and super-events is represented in this dataset. For instance, it includes events such as *"Covid-19 lockdowns"* and *"Covid-19 pandemic in UK"*, with no direct relation between[2]. In the medical field, the datasets CSci [32] and EurekAlert [33] have been annotated according to four levels of causal relation: no relationship (c0), causal (c1), conditional causal (c2), and correlational (c3).

Table 1 summarises these models and datasets, showing which kind of relations are included in each of them. In addition to those, it is important to mention CausalNet, a common sense graph of actions, with weights between them indicating the likelihood that they are in a cause-effect relation [17]. Finally, it is worth to cite CausalBank – including 314 million sentence-level cause-effect pairs – from which it has been generated the Cause Effect Graph, in which links between events are weighted based on their co-occurrence in the text [15].

The table shows clearly that none of the existing resources is able to represent the entirety of the possible relations, calling for a more complete data model.

3 FARO: an Event Relation Ontology

Not all event relationships involve just events. For instance, one may want to describe that being tall is helping a player to score in a basketball game. The player's height is of course not an event, but rather a *condition* which supported the happening of an *event*. For this reason, FARO includes two different classes, *Condition* – transcendent, possibly can result in a RDF statement – and *Event* – immanent, following the categorisation in [23] – that are direct children of the more general class *Relata*, as in Figure 1. The latter is not intended to be directly

[2] We can logically imagine here that the spread the pandemic *caused* the lockdown, which is in its turn a measure for *preventing* the worsening of pandemic.

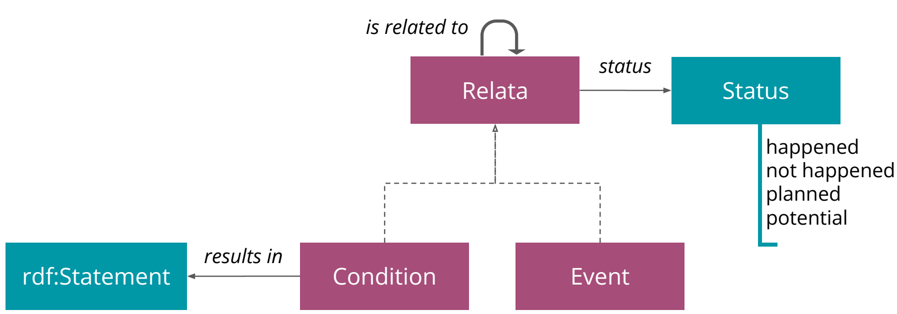

Fig. 1. Core elements of the FARO ontology

	Fan, 2008 [4]	Hong, 2016 [9]	CIDOC CRM [3]	EventModelF [24]	Event Pattern [13]	Simple Event Model [29]	Wolff, 2007 [31]	Event Ontology	TimeBank [28]	Causal TimeBank [18]	EventStoryLine [1]	EventKG [6]	CSci [32] / EurekAlert [33]	**FARO**
Temporal relations														
Before (after)	✔	✔	✔						✔	✔	✔			✔
Immediately Before (Immediately After)									✔	✔	✔			✔
Equal / Simultaneous	✔	✔	✔						✔	✔	✔			✔
Meets (is met by)	✔	✔	✔						✔	✔	✔			✔
Overlaps (is overl. by) / During	✔	✔	✔						✔	✔	✔			✔
Contains (is cont. by)	✔	✔	✔						✔	✔	✔			✔
Starts (is started by) / Begins	✔	✔	✔						✔	✔	✔			✔
Finishes (is finished by) / Ends	✔	✔	✔						✔	✔	✔			✔
Vague		✔												
Mereological relations														
Sub-event (super-Event)		✔	✔	✔	✔	✔		✔				✔		✔
Re-emergence		✔												✔
Coreference		✔												✔
Variation		✔												✔
Confirmation / Ev. type		✔				✔								
Contingent relations														
Cause		✔	✔	✔			✔			✔	✔*		✔+	✔
Enable / Condition		✔					✔			✔-				✔
Prevent							✔			✔-				✔
Despite / Concession		✔					✔							✔
Correlation		✔		✔									✔	✔
Intention / Purpose			✔											✔
Not cause													✔	✔
Comparative relations														
Comparison		✔												✔
Conjunction / Similarity		✔												✔
Disjunction / Dissimilarity		✔												✔
Opposite		✔												✔
Negation / Alternative		✔												✔
Competition / Contrasting		✔												✔

Table 1. Type of event relations which is possible to instantiate using certain schemas/ontologies or to find in certain datasets. In EventStoryLines (✔*), it is possible to find plot actions which may be interpreted as cause of other events. In CSci and EurekAlert (✔+), it is also possible to express the conditional causation. In Causal TimeBank (✔-), both Enable and Prevent relations are separately considered in the process, but they are not distinguished in the dataset.

use for instantiate entities, but is rather an abstraction layer for the other two main classes, allowing to define relations which connects indiscriminately any combination of them.

We found interesting to allow to define the *Status* of a *Relata* entity, to be chosen between four different options:
1. *happened* for sure at some moment in the past;
2. *not happened* for sure, we can exclude any happening of it in the future;
3. *potential*, meaning it is still uncertain if it will happen or not;
4. *planned*, sort of stronger potentiality, due to a will to this to happen.

This *Status* is intended to see an evolution in time, until it reaches either the *happened* or *not happened* status. We decided to leave possible to even define unforeseen statuses, apart to the four ones defined by the ontology.

Two *Relata* instances can be connected with a *is related to* property, which suggests general relatedness without further specification. The *is related to* property is further extended by 25 more specific properties, organised around four direct sub-classes of *is related to*, namely:

- comparatively related to
  - alternative to
  - compared to
    - dissimilar to
      - opposite to
    - similar to
  - contrasting version of
- contingently related to
  - causes
  - correlates with
  - does not cause
    - prevents
  - does not prevent (despite)
  - enables
  - intends to cause
- mereologically related to
  - coreference of
  - part of
  - re-emerges in
  - variation of
- temporally related to
  - before
    - immediately before
      - meets
  - contains
    - ends
    - starts
  - overlaps
  - simultaneous to

Differently from other works, we decided to structure these properties hierarchically, in order to enable reasoning. This hierarchy has been realised following the definition of the individual relations. For the same purpose, we included logic constraints – such as `owl:cardinality` and `owl:propertyDisjointWith` – and further define property characteristics – using `owl:SymmetricProperty` and `owl:Transitive Property`. Please note that FARO is only intended to be used for representing the relationships between events, leaving the event description to be represented using other vocabularies or ontologies.

Figure 2 shows two contingent relations, which represent using the FARO ontology the following text snippet: "A tight monetary program **caused** a temporary downturn but **prevented** a monetary meltdown"[3].

[3] The text sample has been taken from https://economynext.com/sri-lanka-will-repay-bonds-holders-should-appreciate-efforts-made-cabraal-83785/. Last visited: 10/06/2022

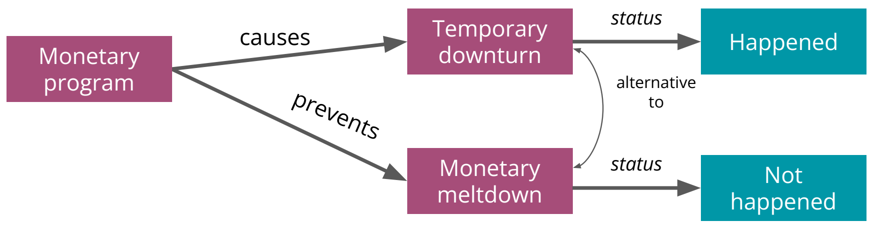

Fig. 2. A relata causing an event and preventing another one, represented using FARO.

Looking a second time at Table 1, it is possible to appreciate that FARO is covering most of the listed relations, proposing itself as central ontology for the harmonisation of different data models. We decided to not include in our ontology the *Vague* temporal relation: even if valuable from the point of view of information extraction, these kind of properties are not common in semantic web environments, where a more generic super-property can be used – in this case, *temporally related to*. Similarly, FARO is not including any *Confirmation / Event Type* property, because it can be expressed directly with an `rdf:type` statement. Alternatively, it is possible to use FARO in combination with other data models for event description – such as SEM [29], which allows typing events.

4 An Event Relation Dataset

In this section, we describe a dataset that includes some of the relations described in FARO, focusing on the contingent relations and in particular *Cause*, *Intend*, *Prevent*, *Enable*, *Not Cause*. The choice of targeting only a subset of the relations is due to time and resource constraints. However, we believe that a first version of a multi-relation event dataset is crucial to start designing new automatic methods for extracting them. Note that this is the first dataset incorporating *Intend*, and differentiating between *Cause*, *Prevent*, and *Enable*.

We developed this dataset by extending and re-annotating two existing datasets with new event relations types, namely intention, enabling, prevention, and explicit negation of causality. The choice of the datasets were based on their format (TimeML), which was convenient for extending it with other relation link.

- TimeBank [30], published by Brandeis University, providing 183 English news articles with over 27,000 event and temporal annotations about events, times and temporal links between events and times. The annotation respects the TimeML 1.2.1 specification.
- EventCausality [21], the dataset comes with causal and temporal annotations on 25 news articles collected from CNN7, giving at the end 1.3k events, 3.4k temporal links and 172 causal relations between events.

Both selected datasets are represented using the TimeML format [22], which we kept it as a base. This format enables to annotate events in the text and to declare possible connections between them using one among:

- `TLINK`, a temporal relation between events (or between an event and a time expression). Ex: "John **left (ei1)** 2 days *before (s1)* the **attack (ei2)**" →

```
<TLINK eventInstanceID="ei1" signalID="s1"
relatedToEvent="ei2" relType="BEFORE" magnitude="t1" />
```

- `ALINK`, a relationship between an "aspectual" event (events that add a notion about an action whether it begins, finishes, continues, etc.) – normally represented by phrasal verbs, e.g. *start to*– and its argument event: initiation, continuation, etc. Ex: John **started (ei5)** to **read (ei6)** →

```
<ALINK eventInstanceID="ei5"
relatedToEventInstance="ei6" relType="INITIATES" />
```

- `SLINK`, refers as a Subordination Link, which is used for contexts introducing relations between two events, or an event and a signal. Ex. "John **said (ei2)** that he **taught (ei3)** on Monday." →

```
<SLINK   eventInstanceID="ei2"
subordinatedEventInstance="ei3" relType="EVIDENTIAL" />
```

While TimeBank uses all 3 types of links, EventCausality instantiates explicit TLINK relation tags, with causal links are represented separately in another file – not following TimeML, so hard to re-use in other dataset. We kept the temporal links and we enriched it by new event relation tags.

4.1 A generic relation link: `RLINK`

Following the experience described in [18] with the addition of the causal link `CLINK`, we extended TimeML with a new relation type `RLINK`, which we designed as a generalisation of the existing ones (`TLINK`, `ALINK`, `CLINK`), and enriched the previously described datasets accordingly. `RLINK` – or relation link – is a description of a generic relationship between two events, that can be further specified. A `RLINK` instance has 4 attributes as following:

- Link Identifier (`lid`) represents an ID for the relation, unique at the document level;
- Relation type (`relType`) refers as the type of relation between two events or the predicate of the triple, which can be one of the property of FARO, e.g. *Cause*, *Prevent*, etc.;
- Event instance Identifier (`eventInstanceID`) is the relata with the role of subject of the triple;
- Related event instance Identifier (`relatedEventInstance`) is the relata with the role of object of the triple.

Example. "Subcontractors will be offered a **settlement (ei264)** and a swift **transition (ei265)** to new management is expected to avert an **exodus(ei268)** of skilled workers from Waertsilae Marine's two big shipyards."

```
<RLINK eventInstanceID="ei264"
lid="l42" relType="prevention" relatedEventInstance="ei268" />

<RLINK eventInstanceID="ei265"
lid="l43" relType="prevention" relatedEventInstance="ei268" />
```

4.2 Candidate generation

We re-annotated each of the mentioned datasets applying a semi-automatic procedure, based on expression matching as first step, followed by a manual check to validate the extracted annotations.

First, we collected a set of potential signal words for each of the 5 studied relations. We searched in the text these signals and extracted the sentences containing them, which we consider potential candidates. Each candidate sentence is dispatched according to the number of possible event pair combinations of relata that can construct the relation, among all the already annotated events for that specific sentence in the original datasets. In other words, we created a table in which each line contains a unique combination of two events, the signal word, the document id and the full sentence, as in Table 2.

Event1	eid1	Event2	eid2	signal	Annotation	DocumentID	Sentence
settlement	e44	expected	e14	avert	0	$wsj_0 187.tml$	Subcontractors will ...
...	...	...	...	...	0	$wsj_0 187.tml$	Subcontractors will ...
settlement	e44	exodus	e46	avert	1	$wsj_0 187.tml$	Subcontractors will ...
...	...	...	...	...	0	$wsj_0 187.tml$	Subcontractors will ...
transition	e45	exodus	e46	avert	1	$wsj_0 187.tml$	Subcontractors will ...

Table 2. Table of the candidate pairs for a specific relation type (prevention), with manual annotation (1 = correct, 0 = wrong).

In the following, we detail the strategy applied for the signal collection and the extraction for each relation type, together with some examples.

Causality. We adopted the manually defined causal signals and causal verbs in [19], in which causal signals are nominal phrases that express causality (e.g: because of, in order to, as a result of). However, causal verbs are a set of verbs representing the act of causing, such as: cause, bribe, push, etc. The first automatic selection results in 1790 candidate causal relation for TimeBank dataset, and 697 for EventCausality dataset. After dispatching, we ended up with 9658 and 1205 possible event pair causal relation for TimeBank and EventCausality datasets respectively.

Example. "Ocean Drilling amp Exploration Co. will sell its contract-drilling business, and took a $50.9 million loss from discontinued operations in the third quarter **because** of the planned sale."

$$planned \xrightarrow{\textbf{causes}} sell\ ,\ planned \xrightarrow{\textbf{causes}} took\ ,\ planned \xrightarrow{\textbf{causes}} loss$$

Intention. To capture intention, we manually created a list of possible intention signals (e.g: want, plan, aim). Additionally, we adopted another set of events as signals taken from the TimeBank dataset belonging to the class I-action. I-action (Intentional action), is an argument for those events that express an action of intention to do something.

Example. "Companies such as Microsoft or a combined worldcom MCI are **trying** to monopolize Internet access."

The I-action is in bold face and, and the related event is underlined. The selection was manually performed after observing that some of these I-actions can alert the existence of this type of relation in a sentence.

Example. "Courtaulds PLC announced **plans** to spin off its textiles operations to existing shareholders in a restructuring to boost shareholder value."

$$spin \xrightarrow{\textbf{Intends to}} boost$$

As a result of automatic intention signals matching, we got 412 candidate expression for holding intention for TimeBank and 154 for EventCausality dataset. However, after extracting all possible event pair combinations, we ended up with 4028 and 230 intention candidate expression for TimeBank and EventCausality datasets respectively.

Prevention. We integrate prevention signals as defined in [19], in which are initially included into the causal verbs list and claimed to express prevention, e.g: block, bar, deter, etc. After the exploitation of these signals, we could extract 120 and 25 candidate expression, which lead to 988 and 53 event pair combination from TimeBank and EventCausality respectively.

Example. "In addition to the estimated 45,000 Marines to ultimately be part of Operation Desert Shield, Stealth fighter planes and the aircraft carrier John F. Kennedy are also headed to Saudi Arabia to **protect** it from Iraqi expansionism."

$$headed \xrightarrow{\textbf{Prevents}} expansionism$$

Enabling. For this event relation, we defined a list of verbs that alert the existence of enabling, such as authorize, warrant, entitle, etc.. We extended this list with enable signals as defined in [19], e.g: help, permit, empower, etc. to guarantee a high coverage. As a result, we obtained 41 and 17 candidate expression and 328 and 16 candidate event pairs combination for TimeBank and EventCausality datasets respectively.

Example. "In addition, Courtaulds said the moves are logical because they will **allow** the textile businesses to focus more closely on core activities."

$$moves \xrightarrow{\textbf{Enables}} focus$$

Not Causality. To extract the explicit not cause relation, we rely on the previously extracted causal relations, in which, we first naively pick those expression having both negation and causality at the same time, than manually validate the right ones. Consequently, we obtained 230 and 124 candidate expression and 1640 and 255 candidate event pairs from TimeBank and EventCausality datasets respectively.

Example. "He also rejected reports that his departure stemmed **from** disappointment the general manager's post had not also led to a board directorship at the London-based news organization."

$$disappointment \xrightarrow{\textbf{NOT cause}} departure,\ post \xrightarrow{\textbf{NOT cause}} directorship$$

4.3 Manual annotation

The described process extracted a long list of candidate relations, most of them being incorrect and to be filtered out. The structure in Table 2 has been then used by two fluent English speakers annotators, which manually checked the candidate sentences. The process is summarised in the following steps:

1. Each annotator reads and annotates 300 lines for each type of relation.
2. On this preliminary annotation, we compute Cohen's kappa inter annotator agreement (IAA) [12] between the two annotations.
   - If the IIA does not show a substantial agreement (> 0.6), the annotators meet, check the contrasting annotations and agree on a strategy. Then, 300 different lines are chosen and the process goes back to point 1.
   - Otherwise, we progress to next point.
3. The annotation is completed for the rest of the datasets, each annotator taking a unique portion.

During annotation, only relations with precised relata have been considered as correct, while others have been marked as not correct. The annotation process relied on an IAA = 0.7112, which is considered a substantial agreement.

In the following example, the signal word is marked in **bold**, events have been marked using *italic*, but only the underlined ones have been considered part of relationships of type *Prevent* by the annotators in Table 2.

Example. "Subcontractors will be offered a <u>*settlement*</u> and a swift <u>*transition*</u> to new management is *expected* to **avert** an <u>*exodus*</u> of skilled workers from Waertsilae Marine's two big shipyards, government officials *said*."

4.4 Results and discussion

Table 3 reports the number of candidate sentences, event pairs and final correct relations for each dataset and relation type, while in Table 4 we show the number of occurrences for each relation to have an insight about the balance of the final dataset, which can be useful in multi label classification tasks.

		Extraction		**Annotation**
Dataset	**Relation types**	**n. of candidate sentences**	**n. of candidate event pairs**	**n. of correct relations**
TimeBank	Cause	1790	9658	217
	Intend	412	4028	42
	Enable	41	328	11
	Prevent	120	988	17
	Not Cause	230	1640	3
EventCausality	Cause	697	1205	66
	Intend	154	230	2
	Enable	17	16	2
	Prevent	25	53	1
	Not Cause	124	255	0

Table 3. Number of candidate sentences, event pairs and final correct relations for each dataset and relation type.

Relation type	Cause	Intend	Prevent	Enable	Not-Cause
Number of relations	283	44	13	18	3

Table 4. Total number of relations validated by annotators for each relation type. These relations are present in the released Event Relation dataset.

Due to the applied strategy, we were able to only extract relations between events which have been explicitly tagged in the original datasets. This consequently affected the number of extracted links within each relation type, which is particularly low for *Not Cause*, *Prevent* and *Enable* – the subject of the latter not always being an event. The explicit negation of causality is not very expressed in the datasets that we covered, besides the native way of extracting them was not very efficient: indeed, collecting all sentences with causal signal and a negation has lead of lots of (false) candidate sentences.

5 Conclusion and Future Work

In this paper, we introduced FARO, an ontology for representing event relations. FARO includes a structured set of properties, which cover most of the relation types which can be found in the literature. In addition, we re-annotated two existing datasets in order to include some of the relation defined in FARO, releasing a new Event Relation dataset which can be used as ground truth for new multi-event extraction systems. FARO has been implemented in OWL and publicly documented[4]. The Event Relation dataset[5] is released in TimeML format. Both resources are published under an open source license.

We believe that empowering Knowledge Graphs with event relationship information will improve knowledge discovery and link prediction. At the same time, this kind of semantic representation can sensibly improve the explainability in decision making systems and the quality of text generation from graphs [26]. For this reason, we aim to realise a KG of events interconnected using semantically precise relations according to the FARO ontology, in which would be possible to follow relation chains and compute new ones. This KG should be populated by both extracting information from text and by interlinking with existing event-based KGs, such as EventKG [6] and YAGO [8].

In order to do it, an improved version of the Event Relation dataset should be realised. A first enhancement would come by offering a better coverage of different relation types. The used annotation methods can be improved, for example applying event detection techniques such as [2] in the candidate generation. We aim to use the annotated dataset within multi classification supervised tasks for event relation detection, in which we are considering the exploit and the adaption of previously implemented binary event relation extraction approaches – e.g. [16, 14] – also enriching event representation with the involvement of common sense knowledge.

Acknowledgements

This work has been partially supported by the French National Research Agency (ANR) within the kFLOW project (Grant n°ANR-21-CE23-0028).

References

1. Caselli, T., Vossen, P.: The Event StoryLine Corpus: A New Benchmark for Causal and Temporal Relation Extraction. In: Events and Stories in the News Workshop. pp. 77–86. Association for Computational Linguistics, Vancouver, Canada (08 2017). https://doi.org/10.18653/v1/W17-2711
2. Deng, S., Zhang, N., Li, L., Hui, C., Huaixiao, T., Chen, M., Huang, F., Chen, H.: OntoED: Low-resource Event Detection with Ontology Embedding. In: 59^{th} Annual Meeting of the Association for Computational Linguistics and the 11^{th}

[4] https://purl.org/faro/
[5] https://github.com/ANR-kFLOW/EventRelationDataset

International Joint Conference on Natural Language Processing. vol. 1, pp. 2828–2839. Association for Computational Linguistics, Online (Aug 2021)
3. Doerr, M.: The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata. AI magazine **24**(3), 75–75 (2003)
4. Fan, H., Meng, L.: Analysis of events in 3D building models. In: Liu, L., Li, X., Liu, K., Zhang, X., Chen, A. (eds.) Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Geo-Simulation and Virtual GIS Environments. vol. 7143, pp. 1047 – 1058. International Society for Optics and Photonics, SPIE (2008)
5. Galton, A.: States, processes and events, and the ontology of causal relations. Frontiers in Artificial Intelligence and Applications **239**, 279–292 (Jan 2012)
6. Gottschalk, S., Demidova, E.: EventKG – the hub of event knowledge on the web – and biographical timeline generation. Semantic Web **10**, 1039–1070 (2019), 6
7. Guan, S., Cheng, X., Bai, L., Zhang, F., Li, Z., Zeng, Y., Jin, X., Guo, J.: What is Event Knowledge Graph: A Survey. CoRR **abs/2112.15280** (2021), https://arxiv.org/abs/2112.15280
8. Hoffart, J., Suchanek, F.M., Berberich, K., Weikum, G.: YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. Artificial Intelligence **194**, 28–61 (2013). https://doi.org/10.1016/j.artint.2012.06.001
9. Hong, Y., Zhang, T., O'Gorman, T., Horowit-Hendler, S., Ji, H., Palmer, M.: Building a Cross-document Event-Event Relation Corpus. In: 10th Linguistic Annotation Workshop 2016 (LAW-X 2016). pp. 1–6. Association for Computational Linguistics, Berlin, Germany (Aug 2016). https://doi.org/10.18653/v1/W16-1701
10. Jin, W., Qu, M., Jin, X., Ren, X.: Recurrent Event Network: Autoregressive Structure Inference over Temporal Knowledge Graphs. In: Conference on Empirical Methods in Natural Language Processing (EMNLP) (2020)
11. Kaneiwa, K., Iwazume, M., Fukuda, K.: An Upper Ontology for Event Classifications and Relations. In: AI 2007: Advances in Artificial Intelligence. pp. 394–403. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
12. Kılıç, S.: Kappa testi. Journal of Mood Disorders **5**(3) (2015)
13. Krisnadhi, A., Hitzler, P.: A Core Pattern for Events. In: 7th Workshop on Ontology and Semantic Web Patterns (WOP@ISWC). IOS Press, Kobe, Japan (2016)
14. Li, P., Mao, K.: Knowledge-oriented convolutional neural network for causal relation extraction from natural language texts. Expert Systems with Applications **115**, 512–523 (2019). https://doi.org/10.1016/j.eswa.2018.08.009
15. Li, Z., Ding, X., Liu, T., Hu, J.E., Van Durme, B.: Guided Generation of Cause and Effect. In: 29^{th} International Joint Conference on Artificial Intelligence (IJCAI) (2020)
16. Liu, J., Chen, Y., Zhao, J.: Knowledge enhanced event causality identification with mention masking generalizations. In: 29^{th} International Conference on International Joint Conferences on Artificial Intelligence. pp. 3608–3614 (2021)
17. Luo, Z., Sha, Y., Zhu, K., Hwang, S.w., Wang, Z.: Commonsense Causal Reasoning between Short Texts. In: Fifteenth International Conference on Principles of Knowledge Representation and Reasoning (KR). p. 421–430. AAAI Press (2016)
18. Mirza, P., Sprugnoli, R., Tonelli, S., Speranza, M.: Annotating Causality in the TempEval-3 Corpus. In: EACL 2014 Workshop on Computational Approaches to Causality in Language (CAtoCL). pp. 10–19. Association for Computational Linguistics, Gothenburg, Sweden (4 2014). https://doi.org/10.3115/v1/W14-0702
19. Mirza, P., Tonelli, S.: CATENA: Causal and temporal relation extraction from natural language texts. In: 26^{th} International Conference on Computational Linguistics. pp. 64–75. ACL (2016)

20. Motik, B.: Representing and Querying Validity Time in RDF and OWL: A Logic-Based Approach. In: Patel-Schneider, P.F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J.Z., Horrocks, I., Glimm, B. (eds.) The Semantic Web – ISWC 2010. pp. 550–565. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
21. Ning, Q., Feng, Z., Wu, H., Roth, D.: Joint Reasoning for Temporal and Causal Relations. In: 56^{th} Annual Meeting of the Association for Computational Linguistics. vol. 1, pp. 2278–2288. Association for Computational Linguistics, Melbourne, Australia (Jul 2018). https://doi.org/10.18653/v1/P18-1212
22. Pustejovsky, J., Castaño, J.M., Ingria, R., Sauri, R., Gaizauskas, R.J., Setzer, A., Katz, G., Radev, D.R.: TimeML: Robust Specification of Event and Temporal Expressions in Text. In: Maybury, M.T. (ed.) New Directions in Question Answering. pp. 28–34. AAAI Press (2003)
23. Schaffer, J.: The Metaphysics of Causation (2016), https://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/
24. Scherp, A., Franz, T., Saathoff, C., Staab, S.: F–a Model of Events Based on the Foundational Ontology Dolce+DnS Ultralight. In: 5th International Conference on Knowledge Capture (K-CAP). p. 137–144. Association for Computing Machinery, New York, NY, USA (2009). https://doi.org/10.1145/1597735.1597760
25. Shaw, R., Troncy, R., Hardman, L.: LODE: Linking Open Descriptions of Events. In: Gómez-Pérez, A., Yu, Y., Ding, Y. (eds.) The Semantic Web. pp. 153–167. Springer, Berlin, Heidelberg (2009)
26. Suchanek, F.: The Need to Move beyond Triples. In: Text2Story — Third Workshop on Narrative Extraction From Texts (ECIR) (2020)
27. Trivedi, R., Dai, H., Wang, Y., Song, L.: Know-Evolve: Deep Temporal Reasoning for Dynamic Knowledge Graphs. In: 34th International Conference on Machine Learning (ICML). vol. 70, p. 3462–3471. JMLR.org (2017)
28. UzZaman, N., Llorens, H., Derczynski, L., Allen, J., Verhagen, M., Pustejovsky, J.: SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations. In: 7^{th} International Workshop on Semantic Evaluation (SemEval). pp. 1–9. Association for Computational Linguistics, Atlanta, USA (2013)
29. van Hage, W., Ceolin, D.: The Simple Event Model, pp. 149–169. Springer, New York, USA (2013). https://doi.org/10.1007/978-1-4614-6230-9_10
30. Verhagen, M., Mani, I., Sauri, R., Littman, J., Knippen, R., Jang, S.B., Rumshisky, A., Phillips, J., Pustejovsky, J.: Automating temporal annotation with TARSQI. In: ACL interactive poster and demonstration sessions. pp. 81–84 (2005)
31. Wolff, P.: Representing Causation. Journal of experimental psychology. General **136**, 82–111 (03 2007). https://doi.org/10.1037/0096-3445.136.1.82
32. Yu, B., Li, Y., Wang, J.: Detecting Causal Language Use in Science Findings. In: 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). pp. 4656–4666 (2019), https://www.aclweb.org/anthology/D19-1473.pdf
33. Yu, B., Wang, J., Guo, L., Li, Y.: Measuring Correlation-to-Causation Exaggeration in Press Releases. In: 28th International Conference on Computational Linguistics (COLING). pp. 4860–4872 (2020)