On the role of thematic roles in a historical event ontology

(Article begins on next page)

25 April 2024

On the role of thematic roles in a historical event ontology *

Anna Goy**, Diego Magro and Marco Rovera

Università di Torino, Dipartimento di Informatica, Corso Svizzera 185, 10149, Torino, ITALY

E-mail: {annamaria.goy, diego.magro, marco.rovera}@unito.it

Abstract. In this paper we discuss the issues related to the formal representation of *thematic roles* in an ontology modeling historical events. We start by analyzing the ontological distinctions between *thematic roles* and *social roles*, which suggest different formal representations. Coupling the study of existing approaches with an analysis of historical texts – available within the Harlock'900 project – we propose a formal representation of *thematic roles* in HERO (Historical Event Representation Ontology), based on *binary properties*, directly connecting the event to its participants. Moreover, we show that a fine-grained formal ontological model of participation in (historical) events should include *general* thematic roles (e.g., *agent*, *patient*) – able to capture the common aspects of the ways entities are involved in events – and *event-specific* roles (e.g., *sniper*), introduced in the ontology according to a specific criterion, that guarantees the needed expressivity without proliferating roles. We conclude the paper by discussing the benefits of our approach.

Keywords: Ontological analysis, Historical event ontology, Thematic roles, Ontology design patterns, Cultural heritage

**Corresponding Author: Anna Goy, Dipartimento di Informatica, Corso Svizzera 185, 10149, Torino, ITALY. E-mail: annamaria.goy@unito.it

---

1   Introduction

The notion of *role* has been discussed in different research areas, where it has received many different, sometimes overlapping, interpretations. What is clear is that it is an important challenge also in the applied ontology and in the Semantic Web communities.

In this paper, we try to clarify the notion of *role* in the representation of *historical events*, by providing an ontological analysis and a formal account of *thematic roles*, as well as a comparison between *thematic* and *social roles*. In the historical domain, in fact, roles are widely used to describe events: Historical texts, as well as multimedia documents (e.g., video, images), are full of expressions that refer to roles, e.g., *king*, *prime minister*, *leader*, *conqueror*, *loser, victim*.

Starting from the work presented by Goy, Magro, and Rovera (2017), we can identify three different notions that have been called "roles" in the literature, but show a quite different nature:

a) The role somebody or something plays when she/he/it **participates in an event** ("Duccio Galimberti was killed by a group of Fascists in December 1944", where Duccio Galimberti plays the role of *killed* – or *patient*, and the group of Fascists that of *killer* – or *agent*).

b) The role somebody plays within a given **social context** ("Alcide De Gasperi was the first Prime Minister of the Italian Republic", where no event is explicitly mentioned, but there is clearly somebody who plays a role, i.e., that of Prime Minister of the Italian Republic).

c) The role that can be attributed to someone/something from a specific **point of view** ("The liberators landed in Sicily", where the Anglo-American allies are seen as liberators, from a specific historical perspective).

The three different meanings of the notion of role mentioned here are not totally independent of each other, since many relations among them can be considered:

- As a consequence of killing someone (sense (a)), somebody can be socially considered a *killer* (sense (b)).
- Somebody can participate in an international meeting (sense (a)) "in the role of" Italian Prime Minister (sense (b)).
- Somebody can be considered a *liberator* (sense (c)) as far as its participation in a specific event is concerned (sense (a)).

However, we claim that these three senses have some important ontological differences that suggest they should be modeled in different ways. In particular, in this paper, we focus on case (a) and we analyze the historical domain to support our thesis, which can be summarized as follows:

1. Although sense (a) and sense (b) share several features, they also show important **ontological differences**, as already partially discussed by Goy and colleagues (2017).

2. As far as sense (a) is concerned, the notion of *participant* in an (historical) event is not enough and a more fine-grained formal ontological model of participation modality[1] is needed. In particular:

   a. Participation modality should be modeled by relying on a set of **general thematic roles** (such as *agent*, *patient*, etc.), which allow suitable generalizations.

   b. A criterion is needed to regulate the introduction of **event-specific roles**[2] (like *buyer* and *seller*, or *sniper*[3]).

3. As a consequence of the ontological characterization of thematic roles (sense (a)), the participation in an (historical) event should be modeled by means of **binary properties**, connecting the event to its participants.

In the following, we will briefly analyze two research areas in which the notion of role has been studied: formal computational ontologies modeling events (Section 2.1), and (computational) linguistic approaches accounting for thematic roles (Section 2.2). Then, by relying on the analysis by Goy and colleagues (2017), we will summarize the reasons supporting the ontological differences between the ways of participating in an (historical) event and social roles (Section 3.1), leaving the discussion of roles attribution according to points of view for a future investigation. We then present the results of our domain analysis of documents describing events related to the Italian 20th Century history (Section 3.2). These results provided us with the empirical basis of the model for representing the participation in historical events, presented in Sections 4.1 and 4.2. The benefits of the proposed model are discussed in Section 4.3. Finally, we summarize conclusions and future directions of our work in Section 5.

2   Related Work and Background

2.1   Event Ontologies and Participation in Events

In the following, we briefly sketch the modeling choices concerning roles that can be found in some existing and well-known event ontologies (or ontologies that *also* account for the concept of event).

Simple Event Model (SEM; van Hage, Malaisé, Segers, Hollink, & Schreiber, 2011) provides a pattern that allows one to specify that a participant in an event (i.e., an instance of the `Actor` class) plays a specific role (represented as an instance of the class `RoleType`) in the event; in particular, instances of the `Role` class can be used to represent the ternary relationship among an event, an actor and the actual role the actor plays in the event. The instances of the SEM `RoleType` class correspond to what we call (thematic) roles in this paper, while the instances of the SEM `Role` class are reifications of relationships specifying that a specific actor played a specific role ("role type", in SEM terminology) in a specific event. SEM also provides

---

[1]   Participation modality here means the way an entity participates in an event, i.e., the role it plays in the event.

[2]   Actually, these roles are event **type**-specific: e.g., *buyer* and *seller* are defined as participation modalities of commercial_transaction-type events. For the sake of simplicity, in the following we will use the expression *event-specific* roles to refer to this notion.

[3]   Here *sniper* (*cecchino* in Italian) means somebody who shoot from a hidden position (against an enemy); we do not refer to a marksman (specialized soldier).

a pattern to specify different points of view, possibly involving roles played by actors in events. Thus, we can say that SEM offers a representation schema for roles of type (a) and (c), in which roles (types) are reified as instances. However, it does not provide either a `RoleType` hierarchy nor specific role (type) instances, and thus its patterns require the use of external resources, like vocabularies of role types, to be instantiated.

LODE (Shaw, Troncy, & Hardman, 2009) provides two basic properties (`involved` and `involvedAgent`) to express a generic involvement of entities in events, but no specific support for any kind of roles.

CIDOC-Conceptual Reference Model (CIDOC-CRM; Doerr, 2003; Le Boeuf, Doerr, Ore, & Stead, 2015), offers two binary properties (`P12_occurred_in_the_presence_of`, `P11_had_participant`) to model general participation in events and a set of other binary properties to express a few participation modalities (e.g., `P14_carried_out_by`, `P16_used_specific_object`). Moreover, the model puts at disposal a set of "properties of properties" (Le Boeuf et al., 2015) that can be used to refine property assertions, for instance to characterize event participation modalities, i.e. type (a) roles (e.g., `P14.1_in_the_role_of` and `P16.1_mode_of_use` can be used to refine assertions made through `P14_carried_out_by` and `P16_used_specific_object`, respectively) or specific kinds of social roles (type (b) roles), such as the specific role of a member in a group (`P107.1_kind_of_member`).[4] In both cases, the actual roles should be taken from external resources, such as thesauri or other ontologies. Thus, as regards kind (a) roles in particular, CIDOC-CRM offers some means to express a few of them, although it does not account for all the general thematic roles usually listed in the literature (see Section 2.2).

The Event Model F (Scherp, Franz, Saathoff, & Staab, 2009) extends DnS (Gangemi & Mika, 2003), therefore it inherits the DnS capabilities of representing roles (in particular, type (b) roles, represented as instances of the DnS `Role` class); moreover, it provides a *participation* pattern that "enables to formally express the participation of objects in events" (Scherp et al., 2009, p. 140) and an *interpretation* pattern that could provide some support to type (c) roles. The `Participant` class in F is a subclass of `Role` and its subclasses and instances (possibly provided by an external resource, such as a domain ontology) can be exploited within the instantiations of the participation pattern to represent roles of event participants (type (a) roles). Further specification of roles played by participants is not provided in F.

The Event Ontology (EO; Raymond & Abdallah, 2007) does not explicitly claim to provide mechanisms for role representation, but it includes some properties and classes that can be used to express few participation modalities in events, namely: the `agent`, `factor`, and `product` properties, together with the `Agent`, `Factor` and `Product` classes.

Also in the Europeana Data Model (EDM; "EDM", 2016) roles are not modeled, but generic participation in events can be expressed through the `wasPresentAt` property.

---

4   It is worth noting that CIDOC "properties of properties" cannot be directly expressed in the current most common Web ontology languages, such RDFS and OWL.

The event ontology defined by Hyvönen, Lindquist, Törnroos, and Mäkelä (2012), and related to the CultureSampo project, supports the representation of people and objects involvement in events, but it does not provide a model to further specify the different types of involvement, i.e. the roles played by participants.

A little more fine-grained account of participation in events can be found in the ABC ontology (Lagoze & Hunter, 2001), which provides some support for type (a) roles (but no support for either types (b) or type (c) ones): In ABC, various properties can be used to express simple involvement in (`involves`) or presence at (`hasPresence`), general participation modalities (e.g. `hasParticipant`, `hasPatient`) and more specific ones (e.g., `destroys`, `creates`) in events; the participation modality can be further refined by means of the `participationType` property (although its actual formal status is left unspecified). However, ABC does not account for all the most relevant basic thematic roles, and the specific roles that can be used with the `participationType` property have to rely on some external resources.

An interesting perspective is provided by the Generalized Upper Model (GUM; Bateman, 1990; Bateman, Magnini, & Fabris, 1995; Bateman, Hois, Ross, & Tenbrink, 2010)[5], a linguistically-motivated ontology containing a backbone taxonomy of classes and a hierarchy of relations. The most relevant class, from our point of view, is `DoingAndHappening`, a subclass of `Configuration` ("the semantic correlate of some activity or state of affairs"; Bateman et al., 2010, p. 1042), representing "intentional actions and non-intentional happenings" (Bateman et al., 2010, p. 1043). The notion of *participation* in events, activities, or states is introduced as a set of relations (subrelations of `participatingInConfiguration`) taking two arguments: a `Configuration` and an `Element` (object). Such relations represent the different ways in which objects can participate in events, activities, or states. In particular, `Actor` is the relation linking a `SimpleThing` (subclass of `Element`) and a `DoingAndHappening`, while `Actee` is the relation linking a `SimpleThing` and an `AffectingAction` (a subclass of `DoingAndHappening`, representing actions in which an entity is affected by an actor). GUM also offers other relations expressing participation (e.g., *addressee*, *beneficiary*, etc.). The most interesting aspect of Bateman and colleagues' approach is its grounding in linguistic motivations (Bateman et al., 2010). As we will discuss in Section 4.2, our approach shares with it the account of (general) thematic roles as binary relations linking objects to events and expressing participation modalities.

Finally, Mizoguchi, Galton, Kitamura, and Kozaki (2015) divide `Occurrents` (i.e., entities that happen in time) into `Processes` and `Events`. The former are "open-ended", i.e. they do not have any completion criterion or finality and they could, in principle, run indefinitely (e.g., *singing*), while the latter do carry an intrinsic notion of completion (e.g., *singing a song*). `Entities` may participate in `Occurrents` (`Participates-in` property). Occurrents may provide contexts for a specific kind of roles, namely, the *occurrent-dependent roles* (e.g., *singer*). Occurrent-dependent roles may be either *original* (e.g., the role of *murderer*, intended as that role held by somebody at the exact moment when a killing event culminates) or *derived* (e.g., the role of *murderer*, intended as that role played by somebody that murdered someone in the past). The context of an original role is always present when the role is held (e.g., the killing event is present

when the murderer original role is held), while this is not true for a derived role (e.g., the context of the murderer derived role is no longer present when that role is held). A derived role may be either *retrospective* (e.g., the murderer derived role), or *prospective* (e.g., the role of *departing-passenger*, held by someone who is performing activities preparatory to the actual departure, but who has not departed yet). A context of a retrospective role is always present before (and only before) the role is held, while a context for a prospective role is always present only after the role is held. Moreover, the context for a prospective role may be missing (e.g., the actual departure might never occur). Mizoguchi and colleagues also introduce the notions of *definitional content* and *performable content* of roles. The former represents the set of necessary and sufficient conditions that a potential player must jointly satisfy in order to play the role (e.g., for somebody to play the murderer retrospective role, it is necessary and sufficient to have played the murderer original role in the past). The latter is the set of processes and events a role player may (typically) participate in, as a player of the role (e.g., somebody that plays the murderer retrospective role typically – although not always – is fugitive from justice, arrested, tried, etc.). Based on the two different kinds of content of a role, the authors refine the relation typically called "playing a role" by distinguishing the notion of *holding a role* from that of *performing a role*: Holding a role means satisfying the definitional content of the role, while performing a role means participating in some part of the performable content of the role. The analysis by Mizoguchi and colleagues is definitely relevant with respect to our approach, and in particular the notion of *retrospective derived role* provides useful insights to our distinction between thematic and social roles. We will discuss it with respect to our proposal in Section 3.

The analysis of existing event models, as far as the participation modality in events is concerned, shows that:

- All models provide some properties to represent the **participation** in an event.
- Only a subset of the models explicitly deal with the notion of **role**.
- None of the models analyzed account for **all** the different ways in which entities can participate in events. Some of them provide **some** means to represent general and/or specific participation modalities; others offer a pattern for representing roles that can be used to express fine-grained modalities of event participation, but only in conjunction with external resources (e.g., domain ontologies of roles).
- None of the analyzed models make any (formal) **distinction** between roles played by event participants (sense (a)), social roles (sense (b)), and roles attribution according to points of view (sense (c)). With respect to the distinction between roles played by event participants and social roles, the closest approach is the one presented by Mizoguchi and colleagues, that will be discussed in detail in Section 3.

2.2   Thematic Roles

In the wide literature about thematic roles belonging to the linguistics and computational linguistics fields, the terminology is sometimes confusing, and different expressions can be found, referring sometimes to the

same concepts, sometimes to different ones: *thematic roles*, *theta-roles*, *semantic roles*, *thematic relations*, and sometimes also simply *arguments*.

Probably, the most classical choice would be using *arguments* and/or *theta roles* at the syntactic layer, [6] and *semantic roles* and/or *thematic relations* at the semantic layer. The term *thematic roles* seems to be the most ambiguous one: It is sometimes used for syntactic relations between verbs and nouns (e.g., by Parsons, 1990), but it is also typically used for referring to the notions of *agent*/*patient*/etc., which – from our ontological perspective – represent modalities of participation in events, and, as such, they are concepts belonging to the semantic layer. Moreover, our approach is based on the claim that both participating in events (e.g., being the agent in a capture event) and playing social roles (e.g., being the Italian Prime Minister) are – ontologically speaking – *roles*, and thus we would prefer to use this term also for referring to the different modalities of participation in events, at the semantic level.

For these reasons, in this paper, we will use *thematic roles* for referring to the semantic relations in the formal representation of the meaning of a natural language expression. In this sense, thematic roles represent the distinguishing relations that link entities to events, thus providing a way to (partially) characterize events themselves on the basis of their participants, i.e. on the basis of *which* entities are involved and *how* they participate in a given event.

The majority of works in the (computational) linguistics area have faced the issue of the linking between syntactic structures and thematic role assignment in the semantic representation (also referred to as *semantic role labeling* or *theta-marking*). Although this is a very important issue, especially as far as automatic knowledge extraction is concerned (see Section 4), this paper does not discuss it and, instead, it focuses on thematic roles as relations defined in a semantic model – *conceptualization*, or *ontology* (Guarino, Oberle, & Staab, 2009) – of events and their participants. In this perspective, an important early contribution can be found in the work by Dowty (1989), who considered thematic roles (such as *agent*, *patient*, etc) as *prototypes* that can be used to classify events, i.e., as prototypical ways to participate in events (and, in fact, he defines two "proto-roles", called *proto-agent* and *proto-patient*). Another important work that deserves to be mentioned is the one by Parsons (1990), who lists the following roles: *agent*, *experiencer*, *theme*, *source*, *goal*, *instrument*, *benefactive*. Jackendoff (1990) groups thematic roles into three distinct tiers: the *Thematic Tier*, defining the role of a participant with respect to its agentivity/affectedness (thus providing agent/patient distinctions); the *Action Tier*, defining the role of a participant with respect to its movement and position (thus providing theme/goal distinctions); and the *Temporal Tier*, representing the temporal dimension of the event. Interestingly, a participant can play multiple roles, provided that they belong to different tiers (e.g., given the sentence "The car hit the tree", the tree can play both the *goal* and the *patient* roles).

One of the most influential studies of verbs, their arguments and corresponding thematic roles is the one by Levin and Rappaport (1991), where the authors argue for the intrinsic relational nature of thematic roles,

---

[6] Obviously, the term *arguments* can also be used to refer to the arguments of a relation at the (formal) semantic level: e.g., *hero:hasParticipant(x,y)* is a relation with two *arguments*, *x* and *y* (see Section 4.2).

which are not semantic primitives, or properties characterizing an entity, but rather relations between individuals and events (see also the paper by Jackendoff (1990), among the others).

An important debate concerns specificity of thematic roles: Are thematic roles specific for every type of event (*buyer*, *seller*, ...) – as claimed by McRae and Matsuki (2009) – or are they general ways of participating in events of different kinds (*agent*, *patient*, ...) – as maintained by Dowty (1989) and Levin and Rappaport (1991)? These two perspectives are not necessarily in competition, but, instead, they can be seen as complementary, and used together, defining verb-specific roles (e.g., *victim of a killing*) as specialization of general semantic roles (e.g., *patient*). However, as pointed out by Lebani, Bondielli, and Lenci (2015), the relationships between event-specific roles and general semantic roles is a complex issue, which deserves further study.

Another significant thread, in the thematic roles debate, is *Frame Semantics*, introduced by Fillmore (1982) – see also (Petruck, 1996). Fillmore claims that the meaning of a linguistic expression can be understood only within its context, and such a context is represented by a *frame*, i.e., a cognitive-grounded structure depicting a typical real-world scenario and supporting natural language understanding (Fillmore, 1982). A frame basically includes the entities involved and the relationships between those entities. In 1997 Fillmore and colleagues started the FrameNet project (framenet.icsi.berkeley.edu), a huge English lexical database, both human- and machine-readable, based on Frame Semantics. In FrameNet, word meanings are represented as frames: Each frame represents an event type (e.g., *cooking*), and participants playing different roles (e.g., *cook*, *food*, *heating_instrument*, *container*) are *frame elements* (FEs). Frames are linked by different types of relations (e.g., IS-A, *using* relation, *sub-frame* relation). FrameNet contains more than 1000 frames, and also provides an annotated corpus. Moreover, FrameNet entries have mappings onto other lexical resources, including VerbNet (verbs.colorado.edu/~mpalmer/projects/verbnet.html). VerbNet, the largest online verb lexicon for English, is organized into verb classes extending the work by Levin (1993) and its lexical entries include thematic roles and selectional restrictions on verb arguments. The complete list of thematic roles used by VerbNet includes 30 different roles (see verbs.colorado.edu/verb-index/vn/reference.php).

An interesting project, partially based on FrameNet, at least as far as thematic roles are concerned, is the Event and Situation Ontology (ESO; Segers et al., 2015). The main goal of ESO is to model implications of events, in terms of pre and post conditions, in order to enable a reasoner to infer *situations* holding before and after the occurrence of certain types of (static or dynamic) events. In ESO thematic roles are modeled as object properties connecting events to entities, and are mapped onto frame elements in the corresponding FrameNet frames.

As far as the semantic representation of linguistic expressions denoting events is concerned, again the literature, within the (computational) linguistics field, is extremely wide. A good, although not recent, survey is provided by Tenny and Pustejovsky (2000). Interesting suggestions, in particular concerning thematic roles, can be found in this literature, even though Tenny and Pustejovsky explicitly say that they refer to *events* as "grammatically or linguistically represented objects" and not as events in the world (Tenny &

Pustejovsky, 2000, p. 4). Our perspective is slightly different: Following the DOLCE cognitive approach (Masolo, Borgo, Gangemi, Guarino, & Oltramari, 2003), our goal is the definition of a historical event ontology modeling the knowledge that supports cognitive representations of events occurring in the world. Obviously, such a model can be used by people to understand linguistic expressions referring to events, but the same knowledge can also be exploited to interpret a picture, a movie or a real-world scene. This less "linguistically grounded" perspective over events is also justified by the fact that we will use the ontology to represent historical events that are accounted for in archival resources, which can be texts (books, documents, letters, ...), but also images, video, and even objects (e.g., dresses, or flags).

3 Analysis of Thematic Roles

Our ontological analysis of thematic roles has been supported by a mixed approach, including: (1) An analysis of the literature, coupled with a deep investigation of existing (event) ontologies (see Section 2) – which provided us a "top-down" hypothesis; (2) An analysis of historical texts describing events and their participants − which offered us a grounded "bottom-up" point of view. We describe these two processes, and their results, in the following sections.

3.1 Ontological Analysis of Event Participation Modality and Social Roles

In this section we analyze, from an ontological point of view, the similarities and the differences between playing a role in an event as participant (case (a) above) and playing a social role (case (b)), leaving the study of role attribution according to points of view (case (c) above), which is linked to the notion of (historical) *interpretation* (van den Akker et al., 2011), for a future work.

As described by Goy and colleagues (2017), we use the notions of *rigidity* and *foundedness*, defined by Welty and Guarino (2001) and used by Masolo and colleagues (2004), to characterize roles as *anti-rigid* and *founded* concepts.

Both social and thematic roles are *anti-rigid* since they are "concepts that can be 'played' (in a contingent and temporary way) by certain entities" (Masolo et al., 2004, p. 267); this means that individuals playing a role do not play it *necessarily*. Entities can start and stop playing a role and they can change role during their life. For example, no Prime Minister *necessarily* is a Prime Minister; any Prime Minister starts being a Prime Minister at a certain point in her/his life and, usually, (s)he stops being a Prime Minister before her/his death. Similarly, no entity playing the patient role in an event plays it *necessarily*; moreover, it/(s)he stops being a patient as soon as the event ends and there can be periods in which it/(s)he is not patient in any event; moreover, an entity can play the same role more than one time. Furthermore, an entity can play different roles simultaneously and a role can be played by different entities at the same time. [7]

---

[7] There are roles that cannot be played by different entities at the same time (e.g., Italian Prime Minister), but there are examples of roles that call for this possibility (e.g. Italian citizen), because of their very nature and independently from their formal representation as types or as individuals.

According to Masolo and colleagues (2004), roles (in particular social roles) are also *founded* concepts. The notion of foundedness captures the definitional dependency relation between concepts: Intuitively, a concept *x* is *founded* if and only if its definition mentions another concept *y*, "such that for each entity classified by *x*, there is an entity classified by *y* which is external to it" (Masolo et al., 2004, p. 273).[8] According to this definition, thematic roles, in particular, are *founded* concepts since their definitions necessarily mention the concept of *event* (e.g., the *beneficiary* role is usually defined as the role played by entities that gain some advantage *from an event*) and events are external entities with respect to entities playing thematic roles in them. As a consequence, thematic roles are *roles* in the specific sense formalized by Masolo and colleagues (2004), in the same way as social roles are. However, as maintained also by Loebe (2003), social roles are not the same as the roles played when participating in an event: If somebody is a musician (i.e., (s)he plays the *musician* social role), (s)he is a musician also when (s)he is not participating in any event in which (s)he acts as a musician (e.g., when (s)he is sleeping). On the basis of these considerations, Loebe identifies three different types of roles, depending on the ontological nature of the context defining them: *relational roles* (defined by the relation an entity is involved in), *processual roles* (defined by the event an entity participates in), and *social roles*. While Loebe's relational roles remain outside the scope of the current discussion, his distinction between processual roles and social roles mirrors our distinction between thematic roles and social roles. Such a distinction can be further supported by considering that the relations (represented by thematic roles) holding between an event and its participants are temporary delimited by the temporal boundaries of the event itself: If I participate in a kissing event as an *agent/kisser*, I stop being an *agent/kisser* when the event itself stops; if I play the *patient/killed* role in a killing event, after it I am probably dead, and therefore, I am not playing any role any more. This means that playing a thematic role does not imply becoming a participant *tout-court*, independently of the event itself. In other words, thematic roles express modalities of participation in events and thus, essentially, they represent *relations* between *specific* entities and *specific* events, lasting at most for the event time span.

The relation between participating in an event acting as a musician (thematic role) and being a musician (social role) can also be taken into account by referring to the distinction between *performable* and *definitional content* that can be found in the work by Mizoguchi and colleagues (2015); see Section 2.1. However, this issue seems to be more complex: The relation between playing a thematic role in an event − sense (a) in Section 1 − and playing, as a consequence, the "corresponding" social role even when the event itself is over − sense (b) in Section 1 − deserves a further discussion. In Section 1, we mentioned the fact that, as a consequence of killing someone, somebody can be socially considered a killer. This relation can be (at least partially) accounted for by the notion of *derived role*, and in particular by the concept of *retrospective derived role*, proposed by Mizoguchi and colleagues (2015): In their approach, the

---

8   The concept of external entity is a complex notion involving those of parts, qualities and constituents; for the purpose of the present discussion, it can be approximately defined as in the paper by Masolo and colleagues (2004): *y* is *external* to *x* iff *x* is not part of *y* and *y* is not part of *x*.

retrospective derived killer role[9] is played by any individual who played the *original event-dependent* killer *role* in a killing event, for the rest of his/her life, since such a role is a member of the *family of roles* including the original and the derived ones. Masolo, Vieu, Kitamura, Kozaki, and Mizoguchi (2011) call these roles *historical participation roles*: An *Everest-climber* is not just someone reaching the Everest top, but someone having such an achievement to his/her credit (Mizoguchi et al., 2015, p. 374).

We agree with Masolo and co-authors (2011) and with Mizoguchi and colleagues (2015) in recognizing a relation between the *original occurrent-dependent role* (i.e., our thematic role) and the *retrospective derived occurrent-dependent role*. However, we claim that such derived roles are full-fledged *social roles*, since, besides the context represented by the original event (e.g., the event of reaching the Everest top), they need a *social context*, i.e., a community that recognizes such events as credits for the agent who performed them. This means that, for a derived role to be played by the individual who played the original one, the original context is not enough and a current *social context*, defining such a role and attributing it to the individual, is needed. Consider, for example, a Partisan in a liberation war, or a war hero: S/he killed people, but s/he does not play the (derived) killer role, because s/he is not socially considered a killer. In order to play the killer social role, s/he needs to have killed someone *and* to be socially recognized as a killer. Therefore, differently from Mizoguchi and colleagues (2015), we do not think that, in order to be a killer (social role), having killed someone is a sufficient condition (although it is probably necessary): In fact, someone can become a killer (socially recognized as such), and after some time, maybe due to changes in the socio-political context, s/he can stop being (considered) a killer.

The ontological differences between thematic and social roles, just discussed, support our claim that in an ontology they should be conveniently formalized in two distinct ways.

Although in a few cases some social roles, such as jobs, are modeled as states or events, with properties expressing temporal boundaries (e.g., BiographyNet; Ockeloen et al., 2013), in computational ontologies, roles (as well as contexts) are often *reified* in order to place them in the domain of discourse, thus being able to "talk about" them, i.e., to explicitly represent their properties and the relations they are involved in (e.g., the link with the definitions that introduce or use them, the relationships with the contexts in which they hold, the reciprocal relationships among them, etc.). In this perspective, social roles are usually represented as instances of some sort of *Role* class. Sometimes, also *role attributions* are reified, and one of the most common reasons is that reification enables to express temporal boundaries for the relation between an entity and the role it plays, especially in those ontology languages in which only unary or binary predicates can be specified, such as OWL (see, for instance, the *RoleInTime* class in the Publishing Role Ontology, exploited also in PRoles; Daquino, Peroni, Tomasi, & Vitali, 2014).

As a consequence of the previous analysis, we claim that, in order to represent (historical) events:

- We need a formal representation that enables us to "use" both social and thematic roles – for instance, to formally represent the fact that De Gasperi played the role of Prime Minister of the

---

9   Mizoguchi and colleagues (2015), for the very same case, use the term *murderer* instead of *killer*: in the present discussion, we consider them as synonyms (both translations of the Italian word *assassino*).

Italian Republic (social role), and the fact that a group of Fascists played the role of *agent* (*killer*) in the murder of Galimberti (thematic role);

- We need a formal representation that enables us to "talk about" (i.e., predicate on) social roles − for instance, to formally represent the fact that the role of President of the Italian Republic (social role) is defined in the Italian Constitution.

- We do not need a formal representation enabling us to "talk about" (i.e., predicate on) thematic roles *per se*.

Therefore, reification is a suitable pattern to represent social roles in a first order logical theory[10], since it enables us to predicate on them. However, there is no reason to reify thematic roles (at least in a computational ontology of historical events). In fact, given their intrinsic relational nature − that inseparably binds them to the event in which they represent a participation modality − it does not make sense to "talk about" them as individuals (particulars) independently from the event they are bound to.

This claim enables us to state that thematic roles should be formally represented as binary properties, linking events and individuals participating in them, in line with some of the approaches discussed in Section 2.1, like, for instance, the one by Bateman and colleagues (2010). Such a representation, in fact, provides a more immediate account of the close relationships holding between events and participants, as we will show in Section 4.

3.2   Event Participation in the Historical Domain

With the purpose of verifying the ontological analysis described in Section 3.1, we examined written texts related to the Italian history of the 20th Century. This choice was driven by the context of Harlock'900, a project (running 2016-2018) involving the Department of Computer Science of the University of Torino and the Fondazione Istituto Piemontese Antonio Gramsci (www.gramscitorino.it), a non-profit institute promoting research on contemporary history, within the framework of the Polo del '900 initiative (www.polodel900.it). The project aims at implementing a semantic layer, based on computational ontologies of historical events, to enrich archive metadata with information about the content of resources. An overview of the overall approach adopted in Harlock'900 can be found in the paper by Goy, Magro, and Rovera (2015).

We performed a (manual) qualitative, in-depth analysis of 200 text fragments, extracted from books (biographies, war reports, testimonies, etc.) containing the narration of events which occurred in Piemonte, a North-West region of Italy, in the period 1943-1945, and belonging to the "Resistenza" (the partisans struggle against the Fascist regime and the Nazi occupation).

---

[10] The large majority of computational ontologies are expressed in first order (possibly modal) logic languages. This is true, in particular for DOLCE (Masolo et al., 2003), which we take as a reference framework and, in general, for application-oriented ontologies, where a trade-off between expressivity and computability/computational complexity must be reached.

Each fragment contains references to 1.7 events on average (with a lot of fragments referring to a single event and some fragments containing references to up to 4 or 5 events). For each event, we identified the *time period* and the *place* it occurred (when available in the text), as well as the *participants* (a total of 380), distinguishing persons, organizations and groups; moreover, we assigned each event a *typology* (see below).

To perform the overall task (identification of events – with typology, participants, time and place) we set up a small team of "annotators", with an average expertise about the historical period in focus. For the identification of events typology, each annotator was provided with the following rules: *(a) look at the available typologies (initially an empty set) and choose one (or more) of them, if suitable; (b) if no available typology is suitable, then define the new suitable ones; (c) try to assign the most specific typology it make sense to you.*

After a first set of fragments was "annotated" (i.e., events with typology, participants, time and places where identified), the annotators set up a panel to discuss their choices and to formulate a hypothesis about more general classes that could represent generalizations of the identified typologies, thus defining the backbone of the taxonomy of our historical event ontology HERO (Historical Event Representation Ontology). We continued iteratively, alternating annotation rounds and panel discussions until all the fragments were annotated, the event typologies identified and the more general classes defined. For each participant in an event, we also asked the annotators to assign a thematic role, starting from a standard list of general roles found in the literature (see Section 2.2), with the possibility of proposing new roles, if the available ones were not suitable.

A small excerpt of text fragments from our corpus can be found in Table 2, together with their (partial) formal representation (that will be described in the next section); a small part of the event class taxonomy is depicted in Fig. 1; the final list of (general) thematic roles used is shown in Table 1(a) and some examples of event-specific roles can be found in Table 1(b).

4 A Formal Model for Representing Participation Modality in Historical Events

4.1 Events in HERO

In this section we describe the ontological model for representing participation in historical events (thematic roles) in HERO, by relying on the results of the previously described analysis. Moreover, we evaluate the proposal by showing the benefits of the approach.

Our approach is inspired by the Davidsonian view of events (Davidson, 1967), in which events are ontologically treated as *individuals*, thus enabling quantification and predicate attribution over them; more precisely, our approach is inspired by the neo-davidsonian perspective (Parsons, 1990), which maintains that event types are represented as unary predicates (e.g., *stabbing(e)*), while participants are specified by means of binary predicates, expressing their participation modality and corresponding to "thematic roles" (e.g., *agent(e, a), patient(e, b)*). Within this perspective, we endorse the assumption that "neo-davidsonian events"

are the main conceptual entities through which we describe the world (Gangemi, Presutti, & Recupero, 2014).

In order to rely on a well-founded account for the notion of event, we refer to the `dolce:Event` class, which is a subclass of `dolce:Perdurant`, as characterized in the DOLCE ontology (Masolo et al, 2003).[11] Moreover, we assume that the characterization of events is based on (at least) three main dimensions, namely time, space, and participants: An event always starts and ends, lasting for a certain time; it always takes place in a given space; [12] and it has at least one participant. [13]

As discussed above, the general notion of *participant*, such as the one provided by DOLCE and by other event models (see Section 2.1), does not seem to be enough to describe how entities are involved in events, at least as far as the historical domain is concerned: In the corpus we analyzed (see Section 3.2), clearly distinguishing, for example, victims from perpetrators emerged to be a major requirement.

For this reason, we claim that a more fine-grained formal ontological characterization of participation modalities is needed, describing *what* are the entities involved and *how* they participate in a given (historical) event (i.e., what role they play in the event).

But what is the correct level? Does a good ontological model need general roles (such as *agent*, *patient*, *beneficiary*, etc.) or specific ones (*killers*, *makers*, *cooks*, etc.)?

We claim that participation modality in historical events should rely on a set of **general thematic roles** (such as *agent*, *patient*, etc.), which allows suitable generalizations. Moreover, we maintain that a criterion is needed to define when **event-specific roles** (like *buyer* and *seller*, or *sniper*) should be introduced. In the following sections, we describe and provide support for this position.

4.2   Thematic Roles in HERO

Event-specific thematic roles seem to be required from a cognitive perspective. For example, in Frame Semantics (Fillmore, 1982; Petruck, 1996) (see Section 2.2), the roles played by the involved entities correspond to prototypical participation modalities and define frame-specific roles (Narayanan, Baker, Fillmore, Petruck, 2003). In FrameNet, a very popular frame representing commercial transactions – *Commerce_sell* ("describing basic commercial transactions involving a buyer and a seller exchanging money and goods, taking the perspective of the seller"; see framenet.icsi.berkeley.edu) – contains *buyer*, *goods*, and *seller* as core participant roles, and a number of non-core roles (like, for example, *money*). Event-specific roles are also supported by psycholinguistic studies, again claiming that they are involved in language understanding (see, for example, the paper by McRae and Matsuki, 2009).

---

[11] Similar considerations could hold for stative perdurants (see the `dolce:Stative` class and its subclasses), but they deserve a deeper analysis, and thus we leave them outside of the scope of the current discussion.

[12] Space (as other event properties, such as time) can be often described in a rough and fuzzy way (e.g., "in Italy", "yesterday"), but this does not mean that events do not occur in specific spaces and at specific times: for example, a trans-atlantic telephone conversation takes place in a space region that is maybe scattered and difficult to be formally represented, but it does take place in some space.

[13] Sometimes participants are not so easy to be identified and described (e.g., what are the participants in a rainstorm?), but "something" seems to always participate (masses of air, rain drops, etc.).

Moreover, there seem to be formal reasons for supporting the representation of event-specific thematic roles in an ontology of events. There are, in fact, some cases in which two participants play the same (general) role, but with peculiar characteristics that are lost if more specific roles are not defined. For example, in a commercial transaction, it is impossible to distinguish between the *buyer* and the *seller* if only the *agent* role is used; in a trial (at least in the Italian legal system), different participants typically play an agentive role, with different characteristics: the State Attorney (*Pubblico Ministero*), the private prosecution attorney (*avvocato dell'accusa*), people bringing a civil action (*Parte Civile*); in a combat, among "shooters" there can be individuals playing the specific role of snipers.

However, a very specific characterization of participation modality alone is not enough, since it does not capture the common semantic aspects shared by specific roles (such as *killers*, *makers*, *cooks*, etc.), i.e., it fails in representing common general aspects of the ways entities are involved in events. Capturing such common aspects could be useful also from an application-oriented perspective, for example in cases in which a user may want to retrieve all participants playing a given general role (e.g., all beneficiaries) of a set of events.

On the basis of these considerations, we introduced in our model **general thematic roles**, representing generalizations of participation modalities in (historical) events. Moreover, we defined the following criterion for the introduction of **event-specific thematic roles**, which guarantees the needed expressivity while avoiding the useless proliferation of specific roles:

> **Criterion** for introducing event-specific thematic roles: (1)
> *An event-specific thematic role should be introduced iff there is an event typology (class)*
> *in whose instances two (or more) participants can play two (or more) different roles, but*
> *general roles are not suitable to capture such a difference.*

Thematic roles (both general and event-specific ones) are formally defined in our historical event ontology HERO; HERO top layer relies on the already mentioned foundational ontology DOLCE (Masolo et al, 2003) and its extension DnS (Gangemi & Mika, 2003). In particular, HERO includes the following axioms:[14]

```
hero:Object(x) → dolce:Endurant(x)
hero:Physical_object(x) → dolce:Physical_endurant(x)
hero:Agent(x) → dolce:Agent(x)
hero:Perdurant(x) → dolce:Perdurant(x)
hero:Event(x) → dolce:Event(x)
hero:Action(x) → dolce:Action(x)
hero:Phenomenon(x) → dolce:Phenomenon(x)
hero:Action(x) → hero:Event(x)
hero:Phenomenon(x) → hero:Event(x)
```

---

[14] We express them in First Order Logic, for the sake of clarity. Free variables should be intended as universally quantified.

`hero:Event` has two subclasses, i.e. `hero:Phenomenon`, representing an event that occurs without intentionality (e.g., natural phenomena like rain, landslides, etc.) and `hero:Action`, representing an event that occurs due to the intentionality of some agent (e.g., combats, captures, etc.).

Within the domain considered (historical events), although in the analyzed corpus also phenomena are described, the major role is played by actions. Therefore, we introduced a taxonomy of subclasses of `hero:Action`, defined on the basis of the domain analysis described in Section 3.2. Fig. 1 shows a small fragment of this taxonomy, including the classes that represent action typologies involved in the text fragments presented in Table 2.
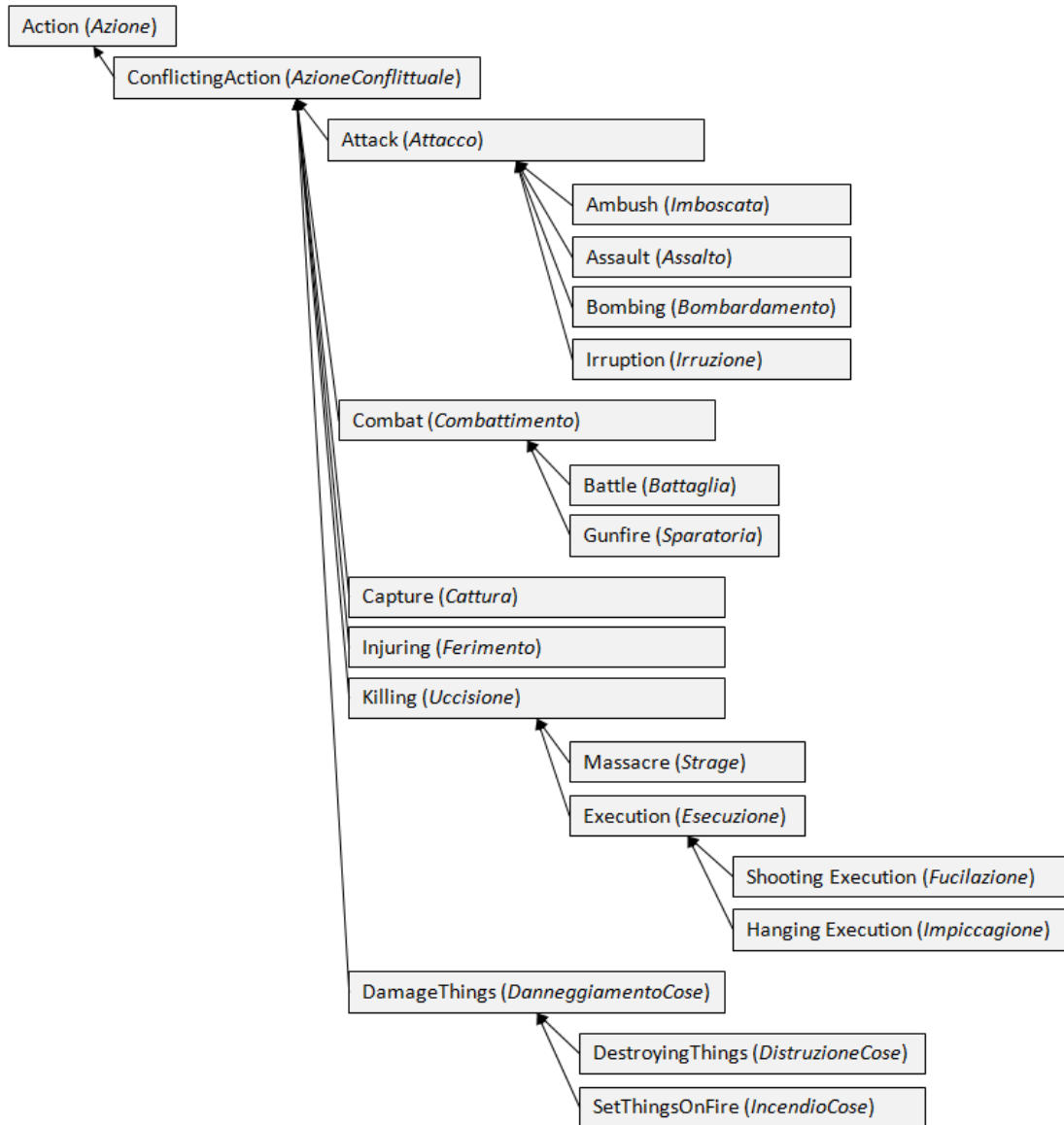

Fig. 1. A small fragment of the taxonomy of action types in HERO.

Based on the analysis discussed above, in HERO, we represented thematic roles as binary properties, connecting an event with a participant. More precisely, we have a hierarchy of thematic roles, where:

- The **top property** represents general event participation – in the DOLCE sense:

  `hero:hasParticipant(x,y)` → (∃t)(dolce:Timeinterval(t) ∧

$$\text{dolce:participatesIn}(y,x,t))$$

```
hero:hasParticipant(x,y) → hero:Event(x) ∧ hero:Object(y)
hero:Event(x) → (∃y)(hero:Object(y) ∧ hero:hasParticipant(x,y))
```

- The **intermediate layer** contains properties representing general thematic roles, identified on the basis of the mentioned analysis, and shown in Table 1(a).

- The **lower layer** includes event-specific thematic roles, introduced according to Criterion (1). Some examples of event-specific roles are shown in Table 1(b).

Table 1. (a) General thematic roles in HERO; (b) Some examples of event-specific thematic roles in HERO

(a)

agent	volitional causer of the event (an action necessarily has at least one *agent*)	`hero:hasAgent(x,y) → hero:Action(x) ∧ hero:Agent(y)`
		`hero:hasAgent(x,y) → hero:hasParticipant(x,y)`
		`hero:Action(x) → (∃y)(hero:Agent(y) ∧ hero:hasAgent(x,y))`

force	non-volitional causer of the event	`hero:hasForce(x,y) → hero:Phenomenon(x) ∧ hero:Object(y)` `hero:hasForce(x,y) → hero:hasParticipant(x,y)`
patient	participant affected by the (effects of the) event	`hero:hasPatient(x,y) → hero:Event(x) ∧ hero:Object(y)` `hero:hasPatient(x,y) → hero:hasParticipant(x,y)`
experiencer	participant experiencing (or being aware of) the event	`hero:hasExperiencer(x,y) → hero:Event(x)∧hero:Object(y)` `hero:hasExperiencer(x,y) → hero:hasParticipant(x,y)`
beneficiary	participant gaining advantage from the event	`hero:hasBeneficiary(x,y) → hero:Event(x) ∧ hero:Agent(y)` `hero:hasBeneficiary(x,y) → hero:hasParticipant(x,y)`
damaged	participant having some disadvantage from the event	`hero:hasDamaged(x,y) → hero:Event(x) ∧ hero:Agent(y)` `hero:hasDamaged(x,y) → hero:hasParticipant(x,y)`
instrument	tool used in the event	`hero:hasInstrument(x,y) → hero:Event(x) ∧ hero:Object(y)` `hero:hasInstrument(x,y) → hero:hasParticipant(x,y)`
result	end product of the event[15]	`hero:hasResult(x,y) → hero:Event(x) ∧ hero:Object(y)` `hero:hasResult(x,y) → hero:hasParticipant(x,y)`
theme	participant changing place, or changing owner, or being exchanged	`hero:hasTheme(x,y) → hero:Event(x) ∧ hero:Object(y)` `hero:hasTheme(x,y) → hero:hasParticipant(x,y)`

(b)

sniper	someone shooting at individuals from a hidden position	`hero:hasSniper(x,y) → hero:Combat(x) ∧ hero:hasAgent(x,y)`
in_favor_voter	someone voting in favor	`hero:hasInFavorVoter(x,y) → hero:Vote(x) ∧ hero:hasAgent(x,y)`
dissenting_voter	someone voting	`hero:hasDissentingVoter(x,y) → hero:Vote(x) ∧`

---

[15] This property can be seen as a shortcut for: Event *x* causes object *y* to come into existence.

	against	`hero:hasAgent(x,y)`
vote_abstainer	someone not expressing any vote	`hero:hasVoteAbstainer(x,y) → hero:Vote(x) ∧` `hero:hasAgent(x,y)`
meeting_speaker	someone speaking (taking the floor)	`hero:hasMeetingSpeaker(x,y) → hero:Meeting(x)` `∧ hero:hasAgent(x,y)`

General thematic roles are defined as properties with `hero:Event` as domain; `hero:hasAgent` and `hero:hasForce` are slightly more specific, having – respectively – `hero:Action` and `hero:Phenomenon` as domain. All subclasses of `hero:Event` "inherit" the general roles, i.e. their instances can be the first argument of the relations; for example, an execution (*e*) where somebody (*a*) shot someone else (*b*) can be represented as follows:

```
hero:Execution(e) ∧ hero:hasAgent(e,a) ∧ hero:hasPatient(e,b)
```

Event-specific thematic roles are formally introduced only when needed, according to the Criterion (1); for example, a gunfire *e* (instance of `hero:Gunfire`, subclass of `hero:Combat`) – where, among the agents *a*, *b*, and *c*, *a* plays the more specific role of sniper (i.e., s/he shoots from a hidden position) – can be represented as follows:

```
hero:Gunfire(e) ∧ hero:hasSniper(e,a) ∧ hero:hasAgent(e,b) ∧
hero:hasAgent(e,c)
```

The intermediate layer, hosting general thematic roles, is worth a final remark. As already mentioned, the proposed list is the result of the convergence of two analytical processes, taking into account the literature and the existing ontologies (top-down), and historical texts (bottom-up). Not surprisingly, such a result is not far from "standard" thematic role lists that can be found in the literature (see Section 2.2). Only a few peculiarities may deserve a further comment:

- We included both the *patient* and the *theme* roles, because we think that these two roles, with their intended meaning in HERO, can capture different relevant aspects of participation modality. Basically, we consider *patients* all entities affected by the event ("A group of Fascists killed *Galimberti*"), while we use *theme* to describe the role of participants that change place or owner in the event ("The shepherd gave the two Partisans *some bread*").[16]
  The choice of having both *patient* and *theme* roles is supported by authors like Jackendoff (1983, 1990), with his tier-based approach (see Section 2.2), as well as by projects like VerbNet (verbs.colorado.edu/~mpalmer/projects/verbnet.html), where both roles are present.
- We added the *damaged* role, coupling the traditional *beneficiary* role, in order to capture cases in which the side-effects of an event cause some disadvantage to a participant ("The collapse of the bridge represented a great handicap for *the Allied troops*"). A similar role – although not very common – can be found in the literature, for instance in the approach by Smith and Grenon (2004),

---

[16] Typically patients are "affected" because they change their (intrinsic) state, while themes only change their position in space (see, for example, the approach by Bateman and colleagues, 2010). However, the issue is a complex one, and it would deserve a deeper investigation, which is out of the scope of this paper.

where the authors distinguish between *facilitation* and *hindrance* as specifications of a more general *influence* role.

We did not include properties referring to time and place in the list of HERO general thematic roles since we prefer considering thematic roles as properties representing only *participation modalities* in a strict sense. Obviously, HERO provides properties expressing the time and the place an event occurs (although they are out of the scope of the current discussion).

4.3   Benefits of the Proposed Approach

In order to evaluate the benefits of our representation model, we formulated, with the help of domain experts involved in the project, a set of questions that an application, offering a smart access to the content of archival resources about the "Resistenza" in Piemonte, should be able to answer. From the dialog with historians, it emerged that – when performing different activities related to historical research based on access to library and archive resources – both generalization and specification mechanisms can be useful in order to face typical questions researchers try to answer during their investigation activity. For instance, as far as conflicting actions are concerned, according to the domain experts we interviewed, the application should be able to answer questions representing generalization needs (G), as well as questions representing specialization needs (S); for example:

(G) Examples of questions expressing generalization needs:

Who were the victims of conflicting actions?
Who perpetrate conflicting actions?
How many people have been executed (by shooting, hanging, etc.)?
Which buildings have been destroyed (by fires, or bombing, or ...)?

(S) Examples of questions expressing specialization needs:

In which combats did snipers take part?
Who voted in favor of the Grandi's Order of the Day, in the Grand Council of Fascism, on July 25th 1943?
Who took the floor in the Grand Council of Fascism, that took place on July 24-25th 1943?

Considered these requirements, the **research hypothesis** we aimed at verifying is the following. Given an application based on an ontology that contains a taxonomy of event typologies (like the fragment in Fig. 1):

- If the ontology includes
    - only a **generic** (top) **property** representing general event participation, or
    - only **event-specific thematic roles**, or
    - only **general thematic roles**

    then the application cannot answer both kinds of questions, (G) and (S), mentioned above.

- If the ontology includes both **general thematic roles** (as defined in Table 1(a)) and **event-specific thematic roles introduced according to the Criterion (1)** to avoid useless proliferation, then the application can answer both kinds of questions, (G) and (S).

In order to explain how the proposed approach supports an application in answering questions like those listed in (G) and (S), we discuss in detail the first two, namely:

a) *Who were the victims of conflicting actions?*

b) *In which combats did snipers take part?*

Table 2 shows some examples of text fragments extracted from our corpus (see Section 3.2); expressions referring to events that are classified as conflicting actions (i.e., instances of the `hero:ConflictingAction` class) are highlighted; expressions denoting participants playing a *patient* role in those events (i.e., individuals filling the second argument of the `hero:hasPatient(x,y)` relation) are underlined; expressions denoting participants playing a *sniper* role ("subrole" of *agent*) are double underlined. Each fragment is coupled with a partial formal representation, based on the HERO ontology; the complete representations are far more complex: for the sake of readability, only the part representing the highlighted events and roles is shown.

Table 2. Text fragments from the Harlock'900 corpus, with their semantic representation

Id	Text Fragment	Semantic Representation (partial)

15	Ernesto Valabrega, fratello di mia madre, dirigeva, come aveva fatto il nonno, la fabbrica di mobili; [...] il 24 marzo 1944, fu catturato dai fascisti proprio nello stabilimento. [...]	`hero:Capture(e15) ∧`
		`hero:hasPatient(e15,E_Valabrega) ∧`
		`hero:Person(E_Valabrega)`

18	Franco [...] ad Agliasco con altri due partigiani era di corvée per la legna: camminavano nel bosco e lui aveva sentito dei movimenti più in basso, aveva lanciato un fischio, ma intanto si era accorto che quelle che intravedeva erano divise tedesche; non aveva fatto in tempo a nascondersi che un proiettile l'aveva colpito.	`hero:Injuring(e18) ∧`
		`hero:hasPatient(e18,F_Diena) ∧`
		`hero:Person(F_Diena)`

20	Il 26 settembre '44, giunta all'"appuntamento", appresi che era appena avvenuto un breve violento combattimento presso Pancalieri, due dei nostri erano morti. [...] Solo il giorno dopo, allo stesso "appuntamento", Gustavo Comollo mi disse: "Uno dei due caduti è tuo fratello" [...] Com'era possibile che Franco fosse morto?	`hero:Combat(e20a) ∧`
		`hero:Killing(e20b) ∧`
		`hero:hasPatient(e20b,F_Diena) ∧`
		`hero:hasPatient(e20b,p20b) ∧`
		`hero:Person(F_Diena) ∧`
		`hero:Person(p20b) ∧`
		`(F_Diena ≠ p20b)`

27	il 27 aprile ero già a Torino mentre ancora agivano i cecchini fascisti. Ricordo un'intensa sparatoria [...] in via Cernaia angolo corso Siccardi	`hero:Gunfire(e27) ∧`
		`hero:hasSniper(e27,c27) ∧`
		`hero:Person(c27)`

111	Fine dicembre 1944 - Prima metà gennaio 1945. [...] All'alba del 12 dicembre tre partigiani del Servizio Intendenza sono stati sorpresi in un casolare al Bric delle Scialle, sopra Bagnolo; una donna [...] è stata obbligata dai fascisti ad assistere alla fucilazione: sono caduti Angelo Savoldi, Francesco Graziadei, Giuseppe Bianco e il novantenne abitante della casa, Antonio Fenoglio.	`hero:ShootingExecution(e111) ∧`
		`hero:hasPatient(e111,A_Savoldi) ∧`
		`hero:hasPatient(e111,F_Graziadei) ∧`
		`hero:hasPatient(e111,G_Bianco) ∧`
		`hero:hasPatient(e111,A_Fenoglio) ∧`
		`hero:Person(A_Savoldi) ∧`
		`hero:Person(F_Graziadei) ∧`
		`hero:Person(G_Bianco) ∧`
		`hero:Person(A_Fenoglio)`

112	Fine dicembre 1944 - Prima metà gennaio 1945. [...] [al Bric delle Scialle, sopra Bagnolo] poco più in là venivano	`hero:Capture(e112a) ∧` `hero:Killing(e112b) ∧`

	catturati e uccisi <u>Ignazio Salvemini</u> e <u>Luciano Druetto</u>.	`hero:hasPatient(e112a,I_Salvemini) ∧`
		`hero:hasPatient(e112a,L_Druetto) ∧`
		`hero:hasPatient(e112b,I_Salvemini) ∧`
		`hero:hasPatient(e112b,L_Druetto) ∧`
		`hero:Person(I_Salvemini) ∧`
		`hero:Person(L_Druetto)`

115	Fine dicembre 1944 - Prima metà gennaio 1945. [...] A Piossasco sono caduti, durante un'imboscata, il commissario di battaglione della 104a Brigata <u>Claudio Lugero (Rino)</u> e l'austriaco <u>Franz Debentz</u>; <u>Giuseppe Bevione</u>, fatto prigioniero, sarà deportato in Germania [...]	`hero:Ambush(e115a) ∧`
		`hero:Killing(e115b) ∧`
		`hero:Capture(e115c) ∧`
		`hero:hasPatient(e115a,C_Lugero) ∧`
		`hero:hasPatient(e115a,F_Debentz) ∧`
		`hero:hasPatient(e115a,G_Bevione) ∧`
		`hero:hasPatient(e115b,F_Debentz) ∧`
		`hero:hasPatient(e115b,C_Lugero) ∧`
		`hero:hasPatient(e115c,G_Bevione) ∧`
		`hero:Person(C_Lugero) ∧`
		`hero:Person(F_Debentz) ∧`
		`hero:Person(G_Bevione)`

135	Seconda metà di gennaio - Prima metà di febbraio 1945. [...] Alle 4 del mattino del 5 febbraio a Villafranca Piemonte [...] vengono catturati nei loro letti l'ispettore divisionale della polizia <u>Ennio Carando (Silvio)</u>, il fratello di lui <u>Ettore (Arturo)</u>, capo di Stato maggiore della I Divisione, il vice commissario della Divisione <u>Leo Franco (Carlo)</u> [...] verso le 3 del pomeriggio, li portano a fucilare in piazza sotto la tettoia del mercato [...]	`hero:Capture(e135a) ∧`
		`hero:ShootingExecution(e135b) ∧`
		`hero:hasPatient(e135a,En_Carando) ∧`
		`hero:hasPatient(e135a,Et_Carando) ∧`
		`hero:hasPatient(e135a,L_Franco) ∧`
		`hero:hasPatient(e135b,En_Carando) ∧`
		`hero:hasPatient(e135b,Et_Carando) ∧`
		`hero:hasPatient(e135b,L_Franco) ∧`
		`hero:Person(En_Carando) ∧`
		`hero:Person(Et_Carando) ∧`
		`hero:Person(L_Franco)`

In order to answer questions (a) and (b), an application (*App*) providing access to resource content is faced with the following situations:

- If the participants in conflicting actions are simply represented by some **generic *participant* relation** (hasParticipant(x,y)), it is clearly impossible to distinguish between victims and perpetrators, thus *App* cannot properly answer question (a); moreover, it is unable to distinguish generic participants from snipers, and thus it cannot answer question (b).

- If **only event-specific thematic roles** are available, in order to answer question (a), *App* should separately query for all the individuals playing the role of *captive* (hasCaptive(x,y)), *killed* (hasKilled(x,y)), *executed* (hasExecuted(x,y)), *injured* (hasInjured(x,y)), and so on; to do that, it has to know that all such roles represent specific ways of participating in events as *victims*, but – without general thematic roles – it does not have such information. As a consequence, *App* cannot answer question (a). In this case, it is obviously easy to answer question (b).

- If **only general thematic roles** are available – for instance, the model includes the *agent* (hasAgent(x,y)) and *patient* (hasPatient(x,y)) roles (see Section 4.1) – *App* can query for all those conflicting actions (e.g., instances of the ConflictingAction class, subclass of Action) holding a *patient* (hasPatient(x,y)) relation with an individual that is a human being (or a more generic volitional agent, instance of an Agent class); this enables *App* to answer question (a).

However, *App* is clearly unable to answer question (b), since it has no means to distinguish – in combat actions – snipers among agents.

As a consequence, in order to be able to answer both questions, *App* should be provided with **general thematic roles** and **specific roles introduced according to Criterion (1)**: This result represents a validation of the research hypothesis stated above.

5 Conclusions and Future Work

The main goal of this paper was to clarify the different meanings of the notion of *role* and to provide an ontological analysis and a formal representation of *thematic roles* in (historical) events. In particular, this paper has discussed an ontologically-grounded distinction between social and thematic roles, leading to the claim that, in an (historical) event ontology, thematic roles should be represented as binary properties, connecting events to their participants. The model resulting from our analysis includes a set of general thematic roles coupled with a criterion for the introduction of event-specific roles. The paper also presents a validation of the model, by showing the benefits it would provide to an application that supports a content-based access to resources from historical archives.

There are also interesting open issues we plan to address. In particular, we are building an OWL version of the HERO ontology to be used in the already mentioned Harlock'900 project and in PRiSMHA, a national project started in May 2017 and funded by Compagnia di San Paolo and Università di Torino; it involves the Computer Science and the Historical Studies Departments of the same university, and it is based on a close collaboration with the Polo del '900 and, in particular, with the archives and library of the Fondaz. Ist. Piemontese A. Gramsci. Both projects aim at enriching metadata with semantic knowledge about historical events: This will enable us to test our approach in a running application, with real users.

Moreover, we are investigating the opportunities offered by automatic information extraction (Rovera, 2016; Rovera, Nanni, Ponzetto, & Goy, 2017), and in particular by event mining and thematic role labeling approaches. In this perspective, the possibility of mapping HERO classes and properties onto frames and frame elements in FrameNet will be taken into consideration. This task is both interesting and challenging, due to the peculiarity of the domain (the Italian history of the 20th Century, and the "Resistenza" in particular) and the language used in our corpus, represented by biographies, war reports, testimonies, etc.

Finally, the case of role attribution on the basis of specific (historical) perspectives (case (c) in Section 1) represents an interesting open research issue to investigate.

References

Bateman, J. A. (1990). Upper Modeling: Organizing Knowledge for Natural Language Processing. In Proc. 5th Int. Natural Language Generation Workshop, Pittsburgh, PA, 54−60.

Bateman, J. A., Magnini, B., & Fabris, G. (1995). The Generalized Upper Model Knowledge Base: Organization and Use. In N. J. I. Mars (Ed.). Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing. IOS Press, 60−72.

Bateman, J. A., Hois, J, Ross, R., & Tenbrink, T. (2010). A linguistic ontology of space for natural language processing. Artificial Intelligence, *174*, 1027−1071.

Daquino, M., Peroni, S., Tomasi, F., & Vitali, F. (2014). Political Roles ontology (PRoles): enhancing archival authority records through Semantic Web technologies. Procedia Computer Science, *38*, 60–67.

Davidson, D. (1967). The logical form of action sentences. In N. Rescher (Ed.). The Logic of Decision and Action. University of Pittsburgh Press, Pittsburgh, PA, 81–120.

Doerr, M. (2003). The CIDOC Conceptual Reference Model: An Ontological Approach to Semantic Interoperability of Metadata. AI Magazine, *24(3)*, 75−92.

Dowty, D. R. (1989). On the semantic content of the notion of 'thematic role. In G.Chierchia, B. Partee, R. Turner (Eds.). Properties, types and meaning. vol. II, Kluwer Academic Press, Dordrecht, 69−129.

EDM. (2016). EDM − Definition of the Europeana Data Model v.5.2.7. Retrieved October 30, 2017 from pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Docu mentation//EDM_Definition_v5.2.7_042016.pdf.

Fillmore, C. J. (1982). Frame semantics. In The Linguistic Society of Korea (Ed.). Linguistics in the Morning Calm. Hanshin, Soeul, 111−137.

Gangemi, A., & Mika, P.(2003). Understanding the Semantic Web through Descriptions and Situations. InG. Goos, J. Hartmanis, J. van Leeuwen (Eds.). On The Move to Meaningful Internet Systems 2003: CoopOS, DOA, and ODBASE, LNCS 2888. Springer, Heidelberg, 689−706.

Gangemi, A., Presutti, V., & Recupero, D. R. (2014). Frame-Based Detection of Opinion Holders and Topics: A Model and a Tool. IEEE Computational Intelligence Magazine, *9(4)*, 20−30.

Goy, A., Magro, D., & Rovera, M. (2015). Ontologies and historical archives: A way to tell new stories. Applied Ontology, *10(3-4)*, 331−338.

Goy, A., Magro, D., & Rovera, M. (2017). An ontological perspective on thematic roles. In P.Ciancarini, F. Poggi, M. Horridge, J. Zhao, T. Groza, M. C. Suárez-FigueroaM. d'Aquin, V. Presutti (Eds.). Knowledge Engineering and Knowledge Management, LNAI 10180. Springer, Heidelberg, 123−126.

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology?. In S. Staab, R.Studer (Eds.). Handbook on Ontologies − 2nd Edition. Springer, Heidelberg, 1−17.

Hyvönen, E., Lindquist, T., Törnroos, J., & Mäkelä, E. (2012). History on the Semantic Web as Linked Data − An Event Gazetteer and Timeline for the World War I. In Proc. of CIDOC 2012 – Enriching Cultural Heritage, Helsinki.

Jackendoff, R. (1983). Semantics and Cognition. MIT Press, Cambridge, MA.

Jackendoff, R. (1990). Semantic structures. MIT Press, Cambridge, MA.

Lagoze, C., & Hunter, J. (2001). The ABC Ontology and Model. InK. Oyama, H. Gotoda (Eds.). Proc. DCMI'01. National Institute of Informatics, Tokyo, 160−176.

Le Boeuf, P., Doerr, M., Ore, C.E., & Stead, S. (Eds.). (2015). Definition of the CIDOC Conceptual Reference Model, v. 6.2.1. ICOM/CIDOC CRM SIG.

Lebani, G. E., Bondielli, A., & Lenci, A. (2015).You Are What you Do. An Empirical Characterization of the Semantic Content of the Thematic Roles for a Group of Italian Verbs. Journal of Cognitive Science, *16(4)*, 401–430.

Levin, B. (1993). English Verb Classes and Alternations. University of Chicago Press, Chicago, IL.

Levin, B., & Rappaport, M. (1991). Wiping the slate clean: a lexical semantic exploration. InB. Levin, S. Pinker, S. (Eds.). Lexical and conceptual semantics. Blackwell, Cambridge, MA, 123–151.

Loebe, F. (2003). An analysis of roles. Toward ontology-based modelling. Master's Thesis, University of Leipzig.

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., & Oltramari, A. (2003). WonderWeb Deliverable D18. Technical Report, CNR.

Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., & Guarino, N. (2004). Social Roles and Their Descriptions. In D. Dubois, C. Welty, M. A. Williams (Eds.). Proc. KR2004. AAAI Press, CA, 267–277.

Masolo, C., Vieu, L., Kitamura Y., Kozaki, K., & Mizoguchi, R. (2011). The Counting Problem in the Light of Role Kinds, In E. Davis, P. Doherty, E. Erdem (Eds.). Proc. CommonSense 2011 – 10th Int. Symposium on Logical Formalization on Commonsense Reasoning. AAAI Press, CA, article n. 24.

McRae, K., & Matsuki, K. (2009). People Use their Knowledge of Common Events to Understand Language, and Do So as Quickly as Possible. Language and Linguistics Compass, *3(6)*, 1417–1429.

Mizoguchi, R., Galton, A., Kitamura, Y., & Kozaki, K. (2015). Families of Roles: A new theory of occurrent-dependent roles. Applied Ontology, *10(3-4)*, 367–399.

Narayanan, S., Baker, C. Fillmore, C., Petruck, M.(2003). FrameNet Meets the Semantic Web: Lexical Semantics for the Web. In D. Fensel, K. Sycara, J. Mylopoulos (Eds.).The Semantic Web – Proc. ISWC 2003. LNCS 2870. Springer, Heidelberg, 771–787.

Ockeloen, N., Fokkens, A., terBraake, S., Vossen, P., de Boer, V., Schreiber, G., & Legêne, S. (2013).BiographyNet: managing provenance at multiple levels and from different perspectives. In E. Blomqvist, T. Groza (Eds.). Proc. 3rd Int. Conf. on Linked Science, Vol. 1116. CEUR-WS, 59–71.

Parsons, T. (1990). Events in the semantics of English: A study in subatomic semantics. MIT Press, Cambridge, MA.

Petruck, M. R. L. (1996). Frame Semantics. In J. Verschueren, J.-O. Östman, J. Blommaert, C. Bulcaen (Eds.). Handbook of Pragmatics. Benjamins, Philadelphia, 1–13.

Raymond, Y., & Abdallah, S. (2007). The Event Ontology. Retrieved October 30, 2017 from http://purl.org/NET/c4dm/event.owl.

Rovera, M. (2016). A Knowledge-Based Framework for Events Representation and Reuse from Historical Archives. In H. Sack, E.Blomqvist, M. d'Aquin, C. Ghidini, S. P.Ponzetto, C. Lange (Eds.).The Semantic Web –Proc. ESWC 2016. LNCS 9678, , Springer, Heidelberg, 845–852.

Rovera, M., Nanni, F., Ponzetto, S. P., & Goy, A. (2017). Domain-specific named entity disambiguation in historical memoirs. In Proc. Fourth Italian Conference on Computational Linguistics (CLIC-IT 2017), Rome, Italy.

Scherp, A., Franz, T., Saathoff, C., & Staab, F. (2009). F – A Model of Events based on the Foundational Ontology DOLCE+DnSUltralite. InY. Gil Y., N. Noy (Eds.). Proc. K-CAP 2009. ACM, New York, 137–144.

Segers, R., Vossen, P., Rospocher, M., Serafini, L., Laparra, E., & Rigau, G. (2015). ESO: a Frame based Ontology for Events and Implied Situations. In Proc. MAPLEX 2015, Yamagata, Japan.

Shaw, R., Troncy, R., & Hardman, L.(2009). LODE: Linking Open Description of Events. In: Gómez-Pérez, A., Ying Ding, Y.Y. (Eds.). The Semantic Web – ASWC 2009. LNCS 5926, Springer, Heidelberg, 153–167.

Smith, B., & Grenon, P. (2004). The Cornucopia of Formal-Ontological Relations. Dialectica, *58(3)*, 279–296.

Tenny, C., & Pustejovsky, J. (2000). History of Events in linguistic Theory. InC. Tenny, C.,J. Pustejovsky (Eds.). Events as Grammatical Objects. Cambridge University Press, Cambridge, MA, 3–37.

van den Akker, C., Legêne, S., van Erp, M., Aroyo, L., Segers, R., van der Meij, L., van Ossenbruggen, J., Schreiber, G., Wielinga, B., Oomen, J., & Jacobs, G. (2011). Digital hermeneutics: Agora and the online understanding of cultural heritage. InD. De RoureS. Poole (Eds.). Proc. WebSci'11. ACM Press, New York, NY, 10:1–10:7.

van Hage, W.R., Malaisé, V., Segers, R., Hollink, L., & Schreiber, G. (2011). Design and Use of the Simple Event Model (SEM). Journal of Web Semantics, *9*, Elsevier, Amsterdam, 128–136.

Welty, C., & Guarino, N. (2001). Supporting Ontological Analysis of Taxonomic Relationships. Data and Knowledge Engineering, *39*, 51–74.