HAL
open science

Functional and Semantic Roles in a High-Level Knowledge Representation Language

Gian Piero Zarri

▶ To cite this version:

Gian Piero Zarri. Functional and Semantic Roles in a High-Level Knowledge Representation Language. Artificial Intelligence Review, 2017, 51 (4), pp.537-575. 10.1007/s10462-017-9571-5 . hal-02286170

HAL Id: hal-02286170
https://hal.sorbonne-universite.fr/hal-02286170

Submitted on 13 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gian Piero ZARRI

Functional and Semantic Roles in a High-Level Knowledge Representation Language

Sorbonne University
STIH Laboratory
Maison de la Recherche
28, rue Serpente
75006 Paris, France
zarri@noos.fr

Abstract

We describe in this paper a formalization of the notion of "role" that involves a clear separation between two very different sorts of roles. Semantic roles, like student or customer, are seen as (pre-defined) transitory properties that can be associated with (usually animate) entities. From a formal point of view, they can be represented as standard concepts to be placed into a specific branch of a particular ontology; they formalize the static and classificatory aspects of the notion of role. Functional roles must be used, instead, to model those pervasive and dynamic situations corresponding to events, activities, circumstances etc. that are characterized by spatio-temporal references; see, e.g., "John is now acting as a student". They denote the specific function with respect to the global meaning of an event/situation/activity… that is performed by the entities involved in this event/situation... and formalize the dynamic and relational aspects of the notion of role. A functional role of the subject/agent/actor/protagonist… type is used to associate "John" with the notion of student or customer (semantic roles) during a specific time interval. Formally, functional roles are expressed as primitive symbols like subject, object, source, beneficiary. Semantic and functional roles interact smoothly when they are used to deal with challenging knowledge representation problems like the so-called "counting problem", or when we need to set-up powerful inference rules whose atoms can directly denote complex situations. In this paper, the differentiation between semantic and functional roles will be illustrated from an NKRL point of view. NKRL (Narrative Knowledge Representation Language) is a high-level conceptual tool used for the computer-usable representation and management of the inner meaning of syntactically complex and semantically rich multimedia information. But, as we will see, the importance of this distinction goes well beyond its usefulness in a specific NKRL context. In particular, the use of functional roles is of paramount importance for the set-up of those evolved *n*-ary forms of knowledge representation that allow us to get rid from the limitations in expressiveness proper to the standard (binary) solutions.

Keywords

Knowledge representation, functional roles, semantic roles, *n*-ary structures, qua entities, inference techniques.

1 Introduction

In the everyday ontological practice, *roles* are normally *reified* by representing them as *standard conceptual units* that can be inserted into a specific branch of a particular ontology. They are assimilated, in practice, to usual ontological entities (concepts and individuals) like "human entity", "John" or "physical object". This is true in general for high-level foundational ontologies like OpenCyc (Web ref.1), DOLCE (Masolo et al. 2003), Topic Maps (Garshol and Moore 2008) or BFO (Spear 2006). In OpenCyc for example, Role is a specialization of ObjectPredicate that derives, through intermediary concepts like Predicate, ThruthFunction and MathematicalObject, from IntangibleIndividual and, eventually, from the top level entity, Thing. Moreover, this is also true for many domain-related ontological applications as, to give only some examples, the medical engineering-related CASCOM project (Cáceres et al. 2006), the Task-based Service Navigation Ontology (Fukazawa et al. 2006), the Domain Ontology for Land Administration (Soon 2013), the Publishing Roles Ontology (Shotton and Peroni 2013) or the Organization Ontology (Reynolds 2014). This way of dealing with roles as *static objects/entities* is unsatisfactory[1] given that, chiefly, the essential *relational properties* of this notion – "…a role is meaningful only in the context of a relationship" (Steimann 2000: 86) – are largely hidden. Moreover, following this *static entities* view, it becomes difficult to represent and to manage correctly all those *spatio-temporal constrained and dynamic structured entities* that are centered on the notion of role and that correspond to complex events (episodes, situations, activities), temporal-grounded information, narratives (in the widest meaning of this word, from textual narratives like news to multimedia-based ones like chronicles of sporting events), corporate memory management, cultural heritage, see sub-section 2.1.2 below.

In the context of the work to set up the Narrative Knowledge Representation Language (NKRL), a more structured approach to the notion of role has been adopted. NKRL (Zarri 2009) is a high-level conceptual tool used for the computer-usable representation and management of the *inner meaning* of syntactically complex and semantically rich multimedia information. Two sorts of roles are differentiated in NKRL, *semantic* and *functional* roles – note, however, that these notions are quite general and not strictly limited to a use within an NKRL context. Semantic roles, like student (or customer, passenger, employee), are seen as *pre-defined properties* that can be associated with (usually animate) entities; they formalize the *static and classificatory aspects* of the notion of role. From a formal point of view, they can be represented, according to the traditional ontological practices evoked above, as standard concepts to be associated with a branch of a particular ontology. Let us take into account now a dynamic and real-life context corresponding to some event, situation, activity, characterized by precise spatio-temporal references like, e.g., "John is *now acting* as a student". Functional roles are used here to denote the specific *function* with respect to the *global meaning of the event/situation/activity* that is performed by the static and pre-defined entities (semantic roles) involved in the event. A *functional role* of the subject/agent/actor/protagonist type will be used to associate John, during a specific time interval, with the *semantic role* corresponding to the notion of student. Therefore, functional roles are expressed as primitive symbols like subject, object, source, beneficiary, allowing us to formalize, e.g., that John is the subject/actor of the situation of being a student; they correspond to the *dynamic and relational aspects* of the notion of role. A preliminary, short paper on the semantic/functional distinction is (Zarri 2011a); this was deprived, however, of all the necessary ontological/cognitive background needed to fully understand the interest and the novelty of this differentiation.

In the following, Section 2 defines more precisely the notions of semantic and functional roles and introduces the theoretical principles that explain and justify this distinction. Section 3 develops further the notion of functional roles. Sub-section 3.1 discusses the relationships between functional roles and the associated notion of thematic roles in Computational Linguistics. Sub-section 3.2 illustrates the high-level modelling possibilities introduced by the use of

[1] For example, Loebe (2007) and Mizoguchi and colleagues (2007, 2012) contest vigorously the possibility of any monolithic interpretation of the notion of role. Loebe differentiates between relational, processual and social roles – where relational and processual roles are then unified into abstract roles – and supplies formal models for the three. The role model of Mizoguchi and colleagues, called Hozo, is based on the use of four basic role-related categories: *role concept*, *potential player*, *role-playing thing* and *role holder*. These categories take into account both the static (role concept, role holder) and dynamic (potential player, role-playing thing) aspects of the role notion. Steimann introduces a separation between an IsA structured type hierarchy and a proper role hierarchy where, e.g., the roles Customer and Supplier are subsumed by Agent (2000: 89, 94-98). In the same vein, Fan and colleagues (2001) assert that "…entities and roles are not related taxonomically" (2001: 30). Therefore, ROLES and ENTITIES are collected into two separate sub-hierarchies included within a global ontology having THING as top level concept. Two relations between roles and entities are introduced: *played-by* and *purpose*. The former is used to connect every instance of a role with an instance of an entity, while the latter represents a role that the entity is intended to play. With respect now to Guarino (Guarino 1992, Guarino and Welty 2000), his formalization of the notion of role is based on two principles: *foundation* (there cannot be a role of student without the presence of a school or university and of at least a subject matter) and *anti-rigidity* (the state of being a student is a transient one). Operationally, however, roles for Guarino are still traditional ontological entities (concepts) that can be inserted into an IsA hierarchical structure see, e.g., Fig. 1 in (Guarino 1992) and Fig. 1 and 2 in (Guarino and Welty 2000).

the functional roles through a critical analysis of the "qua entity" mechanism normally suggested as standard solution of the so-called "counting problem". Section 4 will point out concisely that the notion of functional role is not only of theoretical import but that it also allows us to implement actually complex and powerful query/inference procedures. Section 5 is a short review of the state of the art mentioning some knowledge representation approaches that show possible relationships with the methodologies illustrated in this paper. Section 6 is a brief Conclusion.

2 General properties of semantic and functional roles

According to what stated in the previous Section, semantic roles take into account the static, classificatory aspects of the notion of role and can be seen as *properties* that specify the behavior of external entities. To clarify this assertion, let us see how the semantic role notion is concretely implemented in an HClass context: HClass, the Hierarchy of Classes, is the *standard ontology of concepts of NKRL*, bearing (at least some) similarities to upper level ontologies like DOLCE, BFO, OpenCYC and SUMO (Niles and Pease 2001). Semantic roles like student_, passenger_ or customer_ represent (direct or indirect) specializations of the *concept* semantic_role of HClass[2]. semantic_role is, in turn, a specialization of property_ via the intermediate concept qualifier_; property_ is a specific term of the high-level non_sortal_concept of the HClass hierarchy, see Fig. 1. This means that all the semantic role concepts are part of the non_sortal_concept sub-hierarchy of HClass. As we will see in more detail in sub-section 2.1, this means also that semantic roles like student_, customer_, employee_ or boyfriend_ *cannot be endowed with direct instances*: in NKRL, the creation of an individual like STUDENT_1 (or JOHN_{qua}harvard-student, see sub-section 3.2) is *syntactically and semantically forbidden*. This is quite new in the context of the theoretical studies about the notion of role, which quite unanimously assume, following Guarino's approach, that a role can be instantiated.

Therefore, the representation of the semantic roles as *concepts inserted in an ontology* conforms (at least partly) to the common ontological practice; in contrast, the definition of functional roles is more unusual in a traditional ontological context. As already stated, functional roles denote the logic-semantic functions that the involved entities have in the formal representation of the sense/meaning of events/episodes/situations/actions/circumstances and, formally, they are dealt with as *linking operators* represented by *primitives* instead of *concepts*. To represent a simple event or situation like "John gives a book to Mary", we make use in NKRL of the three *functional roles* SUBJ(ect), OBJ(ect) and BEN(e)F(iciary) to introduce, respectively, the *individuals* JOHN_, BOOK_1 and MARY_ (the entities involved in the event) as *arguments* of a semantic predicate like MOVE. They denote the *specific function* of each of the three individuals in the context of the capture of the meaning of an event/episode/situation… characterized by some sort of "shift" (MOVE). In NKRL, *individuals* like the three above are *instances* of HClass *sortal* concepts such as human_being and book_. Modelling a situation like "John is a student" will imply in turn the use of the SUBJ(ect) and MODAL(ity) functional roles to introduce, respectively, the *individual* JOHN_ and the *semantic role* student_ (a non-sortal concept, see above) as the two arguments of the predicate BEHAVE. We can note immediately that:

- According to the dynamic characterization of the notion of functional role, temporal (and spatial) information must be added to the formal structures built around this notion. This will allow us to state that the transfer of the book is occurring now or that John has kept his student status during a given time interval see, e.g., Zarri (1998).
- As it will be explained more in depth in the next sub-sections, the presence of additional relational elements like the functional roles increase the syntactic and conceptual complexity of the ontological structures devoted to the formal modelling of the time-dependent, dynamic information. This precludes the possibility of making use, for these structures, of the traditional binary-like arrangements. Accordingly, they structures will be *of the n-ary type*.
- The two categories of roles are disjointed from a *strict epistemological point of view* – (quite traditional) concepts on the one hand and relational operators represented as primitives on the other hand. However, as we will see, the syntactic/semantic interactions between functional and semantic roles (and between the corresponding binary and *n*-ary ontological structures) are essential to characterize the originality and the practical utility of our approach.

Before giving additional formal details about the two categories of roles and introducing some concrete and detailed examples, we will illustrate briefly the general semantic/cognitive background of these notions.

2.1 "Permanent/static" and "transitory/dynamic" knowledge

The differentiation between semantic and functional roles can be fully understood only by discussing another, more general, fundamental ontological distinction. This last concerns what – in the absence of unambiguous designations within the ontological/computer science domains – we will denote here as *permanent/static* (basic, primitive, a-temporal…) information/knowledge on one hand versus *transitory/dynamic* (structured, composite, spatio/temporal constrained…) information/knowledge on the other.

[2] According to the NKRL's conventions, in this paper *concepts* are denoted in lower case and *individuals* (instances of concepts) in upper case. Moreover, the symbolic names of NKRL concepts and individuals always include *at least* one underscore symbol.

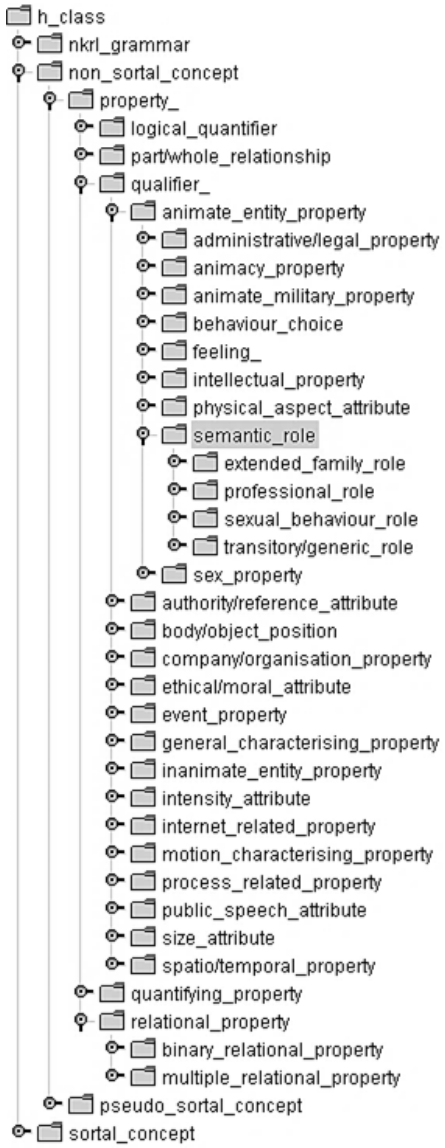

Fig. 1. semantic_role in HClass as a specialization of non_sortal_concept.

Note that the partition between permanent and transitory knowledge has been evoked several times, implicitly or explicitly, in the ontological and cognitive literature – see also the (more restricted) differentiation between endurants and perdurants in Masolo et al. (2003). However, from a concrete modelling point of view, this difference has been usually neglected and a single representation schema has been often used to describe the two sorts of knowledge. Along with the uniform way of representing the notion of role already evoked, this is an example of the *one and only syndrome* that often affects the knowledge representation community.

2.1.1 Permanent/static knowledge

Permanent/static knowledge corresponds to those *self-contained, mandatory and (at least relatively) stable notions (concepts)* that must be taken necessarily into account to represent correctly the general cognitive background of any given, practical application. These notions can be very general (like human being, amount, color or artefact) – and thus relevant to a number of application domains – or specially linked to a given application/set of applications (like control room operator, level of temperature, heat exchanger, business meeting or press conference). They are all characterized by the following general properties:

- These notions concern an (at least provisionally) *stable* vision of the world and/or some specific application domains. They are, in a sense, *a-temporal*, *permanent* and *universal*. Therefore, they can be fully defined, described and classified *a priori* in an autonomous way, independently from their specific modalities of use. In particular, when formal definitions of any kind are added, they can be simply built up by using, under the form of

a set of properties, *simple terminological-like descriptions* in the style of, for example, "an artefact is a physical object created by human beings for cultural/technical etc. use" or "a business meeting is a management-directed gathering companies use to review company's operating principles". And this without any need for specifying actual situations/events/circumstances, spatio-temporally characterized, where a specific artefact is produced/sold/bought/used or a particular meeting is convened.

- Therefore, a first level, basic type of definition for these notions can be expressed simply in terms of *genus/species terminological relationships (IsA links)* – an artefact is a physical object and a business meeting is a sort of gathering. Even in case more structured definitions/descriptions should be needed, the *self-contained, universal and stable* character of these notions implies that any possible definition must be considered *as fixed and not subject to change*. They do not need to be updated in real-time making use of contingent information like, e.g., spatio-temporal data. Of course, this stable character is only valid in the short term – e.g., in the context of a specific on-going application – and these descriptions/definitions can evolve in the long term according, for example, to the progress of our knowledge about a specific entity/domain.
- Taking into account the previous remarks, it can be reasonably affirmed that, in case formal definitions/descriptions for the permanent/static notions should be added to those expressed *implicitly* in terms of genius/species relationships, they can be kept as relatively simple and represented merely by some bundle of properties/attributes with their associated values. In short, these formal definitions/descriptions could be implemented according to some sort of simple *binary* model. In this model, the defining properties or attributes associated with a given notion/concept can only *link an individual to another individual or a value*. And this independently from the fact that these binary relationships are structured as sets of property statements specifying a class in a W3C language like OWL/OWL 2 (Bechhofer et al. 2004, Hitzler et al. 2009) or in the form of template slots describing a class frame according to the OKBC specifications (Chaudhri et al. 1998, Noy et al. 2000).

NKRL, in particular, follows the *usual binary approach* with respect to the formal definitions that – in addition to the implicit genus/species ones – are/could be used in the context of this first category of knowledge. It makes use, in this context, of its proper ontology of (permanent/static) concepts – called HClass (hierarchy of classes) as already stated. Basically structured according to an OKBC approach, see Zarri (2009: 43-55, 123-137) for further details, HClass includes presently (July 2017) more than 7,500 "*concepts*".

From an ontological content point of view, however, HClass presents some interesting properties. Its upper level is partitioned, in fact, between sortal_concept and non_sortal_concept, i.e., between "(sortal) notions that can be instantiated directly into enumerable specimens (individuals)", like chair_ (a physical object) and "(non-sortal) notions that cannot be instantiated directly into specimens", like gold_ (a substance), white_ (a color) or student_ (a property, more precisely, a semantic role). The specializations of sortal_concept can have *direct instances* (CHAIR_27, PARIS_). On the contrary, the non_sortal_concept like gold_, white_ or student_ can admit *further specializations*, see red_gold, whitish_ or university_student, but do not have direct instances. With respect to the NKRL's analysis and implementation of controversial notions like substances and colors, see Zarri (2009: 123-137).

A (severely reduced) representation of the non_sortal_concept sub-hierarchy of HClass where, for simplicity and legibility's sake, only the genus/species relationships are shown, is reproduced in Fig. 1 introduced above. As already stated, it appears clearly that the permanent/static notions corresponding to *semantic roles* are indeed specializations – through animate_entity_property, qualifier_ and eventually property_ – of non_sortal_concept. From Fig. 1, we can also see that the sub-categorizations of semantic_role that are *currently* included in HClass are:

- extended_family_role, with specializations like boyfriend_, spouse_, parent_, child_ (daughter_ and son_), sibling_. The inclusion of the kinship-like terms in the role (semantic role in NKRL) category is accepted by at least some of the researchers of the domain see, e.g., Guarino (1992), Guizzardi (2006), Almeida and Guizzardi (2007), Loebe (2007). See also, on the contrary, the comments in Masolo et al. (2011). Some of these terms pose, in fact, problems from an anti-rigidity (making use of Guarino's terminology) point of view, see for example Mizoguchi et al. (2012, 2015). If boyfriend_, or spouse_ with its specializations (husband_, wife_ …), can be accepted as examples of *provisional* (anti-rigid) properties, it seems difficult to affirm the same for parent_, child_ or sibling_. Practical considerations concerning, mainly, the possibility of creating *interesting* inference rules (concerning, for example, the importance of the family relationships on the behaviors of many people) suggest preserving the unity of this sub-tree of semantic_role. Nothing will prevent us, however, from introducing in the future, if really needed, further differentiations into, e.g., provisional and continuous extended_family_role.
- professional_role. The HClass terms presently included in this sub-category of semantic_role are relatively numerous and distributed into several sub-trees (civil_service_role, military_service_role, industry/manufacturing_role, medical/hospital_role …) that reflect mainly the different application domains where the NKRL technology has been used.
- sexual_behaviour_role (drag_queen, heterosexual_, sadomasochist_ …). The presence in HClass of this sub-hierarchy goes back to the NKRL EUFORBIA project, financed by the European Commission in the framework of their Safer Internet Programme that was intended to build up new tools for filtering out questionable web sites, see in this context Zarri (2009: 234-239).

- transitory/generic_role (candidate_, customer_, passenger_, user_, witness_ …). They seem to correspond well to Mizoguchi et al. (2012: 495) event-dependent roles (including speaker and invited speaker). Guarino's pedestrian non-relational role (Guarino 1999) could be also included in this category.

Note that the low-level structure of HClass must be updated, whenever a new NKRL application is started, to take the specific notions of the new domain into account. Other sub-categorizations of semantic_role – or even a restructuring of the present sub-classes – could be needed, possibly, in the future.

From what expounded up to now, the main conceptual and operational characteristics associated with the semantic_role HClass concepts can be summarized as follows:

- As specializations of the upper-level non_sortal_concept entity introduced above, see Fig. 1, they *cannot be instantiated*. Formally: $\forall x\ (C_{ns}(x) \rightarrow \neg\exists z(InstanceOf\ (z, x)))$. This is in contrast with the usual ontological practice where role classes can be instantiated into individuals. See in this context, e.g., the traditional way (Masolo et al. 2005: 103, Guarino 2009: 64) of relating *student* (a non-sortal property that cannot be instantiated) to *person* (a sortal animate entity that, of course, can be instantiated) using a (contradictory) subsumption link.
- HClass' semantic roles can be *associated/linked* (see next point) *exclusively* with/to specific terms and instances of the (sortal) HClass concept living_entity and of its specializations (animal_, human_being_or_social_body, living_plant…). *Exceptions* due to the personification of physical entities or natural phenomena are admitted when these entities or phenomena can be seen as characters acting as protagonists within a specific application. See, e.g., a story relating the many troubles suffered by a nuclear submarine that, seen like an animated entity, becomes the proper "hero" of this story.
- Associated/linked means here that the semantic role concepts will be used to specify that a particular living_entity is acting in a giving way or is perceived in a certain state (as student_, passenger_ of a giving flight, or as drag_queen during a lesbian and gay pride parade, or as republic_president). Formally, this association is expressed making use of one of those *structured/dynamic n-ary expressions* built up thanks to the use of functional roles, see the next sub-section and the concrete examples in 2.3 and 3.2 below.

2.1.2 Structured/dynamic knowledge

In the permanent/static knowledge described above, every *basic formal element (concept) of the digital/virtual world* corresponds to the formal representation of a *single conceptual entity (notion) of the real world*. In the structured/dynamic knowledge case, the correspondence real world/digital world is neatly more complex.

In conformity with the non-deterministic, deeply-structured and temporally-influenced characteristics of this knowledge, its basic real-world units are represented, in fact, by *elementary events* (Zarri 2015) – corresponding, at least partly, to Davidsonian (Davidson 1967) or monadic (Kim 1996) events. Elementary events are characterized, in the first place, by the presence of a *generalized predicate*[3] that indicates the basic kind of event/action/situation/behavior expressed by the events. The entities involved in this event – *denoted by concepts of the permanent/static type or by their instances* – are associated with the predicate by conceptual links (*functional roles*) like subject, object, beneficiary, modality. They define the *function* these entities have in the context of the global meaning of the elementary event. Moreover, the *general spatio-temporal framework* of this event should be also explicitly or implicitly defined. Therefore, to generate a complete formal representation of a real world structured/dynamic situation (an elementary event) like "The Control Room operator presses a button" in the context of the initialization procedures of a turbine – this example concerns a recent NKRL's application dealing with industrial accidents in a gas/oil domain (Zarri 2011b) – several sorts of digital entities must be used. These include:

- A conceptual predicate that allows us to specify the basic type of event/action/situation/behavior… represented by the elementary event. This predicate correspond to the translation, at deep conceptual level and in digital world terms, of a surface/linguistic level (generalized) predicate like "take part" in the present example.
- The digital world items corresponding to the permanent/static entities involved in the elementary event like, in our example, CONTROL_ROOM_OPERATOR_1, instance of the concept control_room_operator pertaining to the professional_role branch of the semantic_role sub-tree of HClass, and BUTTON_1, instance of the start/stop_device concept of HClass.

[3] The *generalized predicates* can correspond to the usual surface tensed/untensed verbs like press, move, exist, produce, but also to adjectives ("… worth several dollars…", "…a dormant volcano…"), nouns ("…Jane's amble along the park…", "…a possible attack…) when they have a *predicative function*. In a structured/dynamic knowledge context, to each generalized predicate of the real world corresponds an elementary event in the digital world. Therefore, a situation like "The Control Room operator presses a button to start the auxiliary lubrication pump" implies the presence of two elementary events corresponding to the two verbs "press" and "start". Situations like "Lucy was looking for a taxi" or "Peter lives presently in Paris" denote, in contrast, the presence of a unique elementary event – see in this context, for example, Matsuyoshi et al. (2010).

- The functional role SUBJ(ect), used to specify that CONTROL_ROOM_OPERATOR_1 is at the origin of the pressing elementary event and the functional role OBJ(ect), used to indicate that BUTTON_1 is the target of this pressing activity.

This sort of structured/dynamic knowledge is particularly pervasive and, therefore, also economically important. It has been rarely studied as such, even if it deals with the symbolic modelling of those dynamic and complex information structures denoted in Computer Science, according to the context as, for example, structured temporal entities (Pustejovsky et al. 2005), eChronicles (Westermann and Jain 2006), complex events (Stojanovic et al. 2009, Zarri 2015), computational narratives (Finlayson et al. 2010), and so on.

To express now in a more formal and complete way the NKRL representation of the structured/dynamic knowledge entities, we can note first that the digital world description of real world elementary events corresponds to a *general* n-*ary* schema whose core is denoted by Eq. 1:

$$(L_i\ (P_j\ (R_1\ a_1)\ (R_2\ a_2)\ \ldots\ (R_n\ a_n))) \qquad (1)$$

where:

- L_i is a *symbolic label* identifying (*reifying*) the global *n*-ary formal representation of a particular elementary event (e.g., the event corresponding to: "The Control Room operator presses a button"). As already noticed by William Woods (1975: 58) in his seminal paper about the conceptual/ontological 'links', understanding the precise meaning of the formal rendering of an elementary event becomes possible only when *all the constituents of this formalism are dealt with as a unique and coherent block through the use of a global handle in the* L_i *style.*
- P_j is a *conceptual predicate*, i.e., a deep level abstraction, independent from a particular natural language, of the surface generalized predicates mentioned above.
- R_k is a generic *functional role*, like SUBJ(ect), OBJ(ect), SOURCE, MODAL(ity), i.e., the formalization of one of the conceptual relationships that exist between the predicate and one of its arguments a_i. This relationship denotes the specific *function* of a_i within the formal representation of the global meaning denoted by the elementary event.
- a_i represents, eventually, one of the *arguments* of the predicate (e.g., the individuals CONTROL_ROOM_OPERATOR_1 or BUTTON_1 in the previous example) necessarily introduced by a specific *functional role*. Argument a_i correspond to permanent/static entities (HClass concepts or individuals) or to structured associations of these entities.

Spatio-temporal information like "yesterday", "December 15", "Paris" that contributes strongly to characterize the dynamic properties of this second type of knowledge must also, normally, be added to the specific instantiations (formalized elementary events) of Eq. 1 core. For example, in the case of the previous example, an instance of the extended_location concept can be used to denote the building where the turbine is located – geographical_location and extended_location are the two specializations of the location_ concept of HClass. The (implicit or explicit) temporal information associated with the event requires the use of the specific temporal data representation system of NKRL, see later in the paper.

2.2 The two NKRL ontologies

We can now introduce what represents, from an ontological point of view, the main characteristic of NKRL: this language makes use, in fact, of *two structurally quite dissimilar but strictly integrated ontologies*. The first one concerns the (binary) HClass ontology that, as stated in sub-section 2.1.1, takes into account the permanent/static knowledge, and the second is an *n*-ary ontology of events that deals with the structured/dynamic knowledge. This last is an unconventional sort of hierarchical organization where the nodes take the form of *n*-ary structures, called *templates* and *congruent with the general conceptual schema represented by Eq. 1*. This ontology of events is labelled as HTemp (hierarchy of templates) in NKRL. Templates (see Table 1 below for a concrete example) supply the *formal, n-ary digital world representation of general classes of elementary events* like "move a physical object", "be present in a place", "produce a service", "send/receive a message", "make a change of state happen". Therefore, HTemp corresponds to a sort of catalogue of basic structured/dynamic entities (more than 150 presently, July 2017), which are very easy to extend and customize, see in this context, e.g., Zarri (2009: 66). The availability of this catalogue is very important for practical applications. It implies, in particular, that: i) a system-builder does not have to create himself the formal structures needed to describe and exploit a (sufficiently) large set of structured/dynamic knowledge; ii) it becomes easier to secure the reproduction or the sharing of previous results.

More precisely, in an NKRL template, the predicate (P_j in Eq. 1) must correspond to a term of the set {BEHAVE, EXIST, EXPERIENCE, MOVE, OWN, PRODUCE, RECEIVE}, and the functional roles (R_k) to a term of the set {SUBJ(ect), OBJ(ect), SOURCE, BEN(e)F(iciary), MODAL(ity), TOPIC, CONTEXT}. The arguments of the predicate (a_i in Eq. 1) consist of references (using *variables*, see the next sub-section) to an HClass permanent/static concept or to a structured association (*expansion*) of several HClass elements. The logical integration of the two ontologies is assured by the use of HClass elements to fulfil the function of arguments of the predicate within HTemp

templates. Predicates and functional roles are then limited in number and represented as *primitives* – see Zarri (2009: 57-61) for an in-depth discussion about this choice. An immediate, pragmatic justification of this option concerns the desire of avoiding any danger of possible *combinatorial explosion* and *logical inconsistency*. Phenomena of these types could appear by leaving completely free the choice of the elements pertaining to the three main classes of conceptual entities – predicates, roles and arguments – that appear in Eq. 1, see Section 5 below and Zarri (2014a).

2.3 Examples of structured/dynamic knowledge representation

When a particular elementary event pertaining to one of the general classes included in the HTemp hierarchy must be represented, the corresponding template is instantiated to produce a "*predicative occurrence*". Predicative occurrences supply the formal representation of all the elementary events concerning a given NKRL application.

As an example, let us consider a more complete version of the elementary event mentioned in sub-section 2.1.2 and concerning the application in the gas/oil industry domain described in Zarri (2011b). This event reads now, "On October 16th, 2008, the Control Room operator pushes the button labelled as SEQ1_BUTTON in the framework of the first sequence of operations, the 'greasing pump sequence', which is part of the start-up procedures of the GP1Z turbine". To formalize this situation in NKRL's terms we must select first, in the HTemp hierarchy, the template corresponding to "perform a task or an activity" that is represented in the upper part of Table 1. This template is located on the Produce: sub-hierarchy of HTemp that includes all the templates characterized by the presence of the PRODUCE conceptual predicate (P_j in Eq. 1).

Table 1. Deriving a predicative occurrence from a template.

name: Produce:PerformTask/Activity
father: Produce:
position: 6.3
natural language description: 'Execution of Intellectual or Industrial Procedures, of Economic Interest Activities, etc.'

```
PRODUCE     SUBJ        var1: [var2]
            OBJ         var3
            [SOURCE     var4: [var5]]
            [BENF       var6: [var7]]
            [MODAL      var8]
            [TOPIC      var9]
            [CONTEXT    var10]
            { [modulators], ≠abs }

var1        =   human_being_or_social_body
var3        =   activity_, process_, temporal_development
var4        =   human_being_or_social_body
var6        =   human_being_or_social_body
var8        =   activity_, artefact_, process_, temporal_sequence
var9        =   pseudo_sortal_concept, sortal_concept
var10       =   situation_, symbolic_label
var2, var5, var7 = location_

virt2.c32)  PRODUCE  SUBJ     INDIVIDUAL_PERSON_102: (GP1Z_MAIN_CONTROL_ROOM)
                     OBJ      button_pushing
                     TOPIC    SEQ1_BUTTON
                     CONTEXT  (SPECIF SEQ1_GREASING_PUMP (SPECIF member_of GP1Z_STARTUP_PROCEDURE))
                     date-1:  2008-10-16-08:26
                     date-2:
```

Fig. 2 shows a fragment of HTemp displaying, in particular, some offsprings of the Produce: and Move: sub-hierarchies. As it appears from this figure, *HTemp is organized into seven branches*, where each of them contains only the templates structured around one of the seven conceptual predicates admitted by NKRL. All the templates of the HTemp hierarchy follow the general syntax of Eq. 1. A detailed description of many of these templates is supplied in Zarri (2009: 149-177); see also, in this context, Zarri (2014b), where the templates of the Behave: branch of HTemp are discussed in some depth.

From Table 1 we can see that, in the templates, the predicate arguments (the a_i terms in Eq. 1) are represented actually by variables with associated constraints. The elements included in square brackets are optional, i.e., they can be present or not in the instances of this template. The format of the constraints is particularly simple in this example; see Zarri (2009: 64) for the complete syntax of the constraint expressions. As we can note from Table 1, the constraints

are specified as HClass concepts or combinations of concepts, i.e., by using the terms of the standard ontology, HClass. This confirms that the two ontologies, HClass and HTemp, are strictly associated. When creating a predicative occurrence like virt2.c32 to represent a particular elementary event, see the lower part of Table 1, the role fillers in this occurrence must conform to the constraints of its father-template.

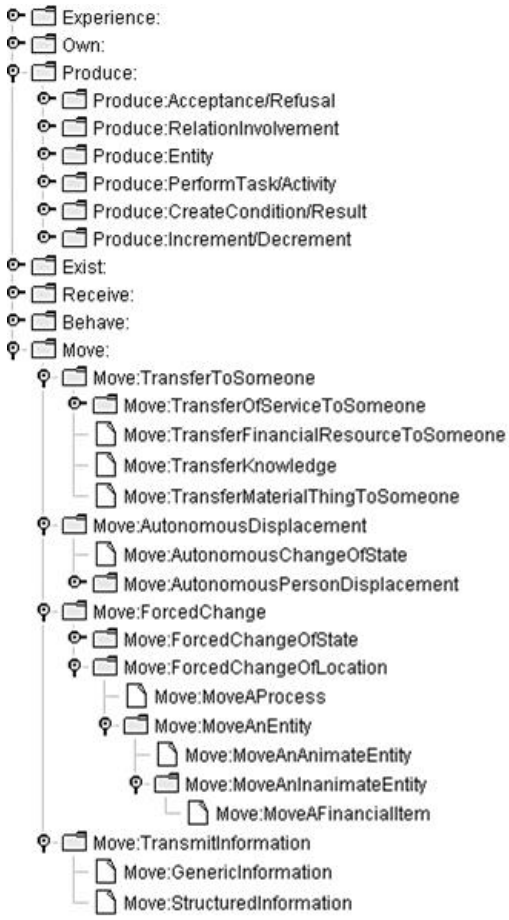

Fig. 2. The seven branches of the HTemp hierarchy; some offsprings of the Produce: and Move: branches are shown.

For example, in the predicative occurrence virt2.c32, INDIVIDUAL_PERSON_102 is an instance of the HClass concept individual_person. This last is a specialization of human_being, specialization in turn of human_being_or_social_body see, in Table 1, the specific constraint on the argument *var1* associated with the SUBJ(ect) role. GP1Z_MAIN_CONTROL_ROOM is an instance of the concept control_room. This is a specialization of the HClass concept location_[4], see the constraint on the variable *var2* in Table 1, through a specialization chain of concepts including, e.g., office/room_ and extended_location. button_pushing is a specialization of activity_ through device_use and other terms. The meaning of the "(SPECIF SEQ1_GREASING_PUMP (SPECIF member_of GP1Z_STARTUP_PROCEDURE))" expression can be explained in these terms. The general framework (functional role CONTEXT) of the action of pushing the button is a particular process_phase (i.e., SEQ1_GREASING_PUMP) that is a constituent (member_of, an HClass non-sortal concept) of the specific technical/industrial_procedure represented by the individual GP1Z_STARTUP_PROCEDURE. The attributive operator, SPECIF(ication), is one of the four operators that make up the so-called "AECS sub-language" of NKRL, expressly created and formalized for

[4] Determiners (attributes) can be added to templates and predicative occurrences to introduce further details about the basic core, see Eq. 1, of their formal representation. In Table 1, several types of determiners are mentioned. The variables *var2*, *var5* and *var7* denote, e.g., determiners/attributes of the *location* type represented, in the predicative occurrences, by specific terms of the HClass concept location_ or by individuals derived from these terms. *Modulators* represent an important category of determiners/attributes that *apply to a full, well-formed template or predicative occurrence* to particularize its meaning. They are classed into three categories, temporal (begin, end, obs(erve)), deontic (oblig(ation), fac(ulty), interd(iction), perm(ission)) and modal (abs(olute), against, for, main, ment(al), wish, etc.), see Zarri (2009: 71-75). An example of use of the temporal modulator obs(erve) is reproduced in Table 4. A last category of attribute/determiners concerns the two operators date-1, date-2. They can only be associated with full predicative occurrences (e.g., the occurrence virt2.c32 in Table 1) and are used to materialize the temporal interval normally associated with the elementary event corresponding to the occurrence. A detailed description of the formal system used in NKRL for the representation and management of temporal information can be found, e.g., in Zarri (1998).

the set-up of *structured arguments* (expansions) of the predicate (Zarri, 2009: 68-70). In addition to SPECIF(ication) = S, AECS includes the disjunctive operator ALTERN(ative) = A, the distributive operator ENUM(eration) = E and the collective operator COORD(ination) = C. The interweaving of these operators within a structured argument is controlled by a *priority rule* that forbids, e.g., to use a list of the ALTERN type within the scope of a list COORD.

What we have expounded illustrates the NKRL solutions to the problem of providing a complete representation of single elementary events. New modelling problems arise when, in the context of larger structured/dynamic arrangements like narratives (Bal 1997) and complex scenarios (Schank and Abelson 1977), several elementary events are linked through *connectivity phenomena* operators as causality, goal and indirect speech. In this case, we make use of *second order structures* created through *reification* of the predicative occurrences corresponding to the events to be associated. The reification is implemented utilizing the *symbolic labels* of the occurrences (the L_i terms of Eq. 1); reification corresponds here, as usual, to the creation of new objects ("first class citizens") out of already existing entities and of *saying something* about them without making explicit reference to the original entities.

Therefore, in the digital world, several predicative occurrences denoted by their symbolic labels L_i can be associated within the scope of second order structures called *binding occurrences*, i.e., labelled lists formed of a binding operator Bn_i and its L_i arguments. The general, formal expression of a binding occurrence bc_i is:

$$(Lb_k\ (Bn_i\ L_1\ L_2\ \ldots\ L_n))\ , \qquad (2)$$

where Lb_k is the symbolic label identifying the whole binding structure. The Bn_i operators are: ALTERN, COORD, ENUM (see the AECS operators introduced above), CAUSE (the strict causality operator), REFER (the weak causality operator), GOAL (the strict intentionality operator), MOTIV(ation, the weak intentionality operator), COND(ition). Some restrictions must be respected to set up *well-formed* binding occurrences. For example while, for the binding occurrences of the ALTERN, COORD and ENUM type, no restrictions are imposed on the cardinality of the list (i.e., on the possible number of arguments L_i), only two arguments L_m and L_n are admitted for the binding occurrences labelled with CAUSE, REFER, GOAL, MOTIV and COND. The binding occurrences that make use of these binding operators are simply of the type: Lb_k (Bn_i L_m L_n). Note also that, in Eq. 2, the arguments of Bn_i can i) correspond directly to L_i labels, denoting simply the presence of particular elementary events, or ii) correspond recursively to sets of labelled lists in Eq. 2 format, i.e., to complex combinations of COORD, CAUSE, GOAL etc. clauses.

To give a concrete example, and returning to Table 1's situation, let us suppose we would now state that: "… the production activities leader pushes the SEQ1_BUTTON … *in order to* start the auxiliary lubrication pump". Note that the specific elementary event corresponding to the action of pushing is still represented by the predicative occurrence virt2.c32 in Table 1. To encode correctly the new information, see Table 2, we must introduce first an *additional predicative occurrence* labelled, e.g., as virt2.c33 and meaning that: "[the aim of the previous action is to …] move the auxiliary lubrication pump from an 'idle' to a 'running' state". We will eventually add *a binding occurrence virt2.c30*, labelled using the GOAL binding operator and involving only two arguments L_m and L_n, see above, to link together the conceptual labels virt2.c32 (L_m, the planning activity) and virt2.c33 (L_n, the intended result). The global meaning of virt2.c30 is: "the activity described in virt2.c32 is focalized towards (GOAL) the realization of virt2.c33". Note also that, in agreement with the semantics of the GOAL operator (Zarri 2009: 71), virt2.c33, the result, is characterized by the presence of an uncertainty attribute code, "*", to specify that, at the moment of 'pushing', the real instantiation of a situation corresponding to "pump running" cannot be definitely confirmed.

Table 2. Binding and predicative occurrences**.**

```
virt2.c32)   PRODUCE   SUBJ       INDIVIDUAL_PERSON_102: (GP1Z_MAIN_CONTROL_ROOM)
                       OBJ        button_pushing
                       TOPIC      SEQ1_BUTTON
                       CONTEXT    (SPECIF SEQ1_GREASING_PUMP (SPECIF member_of
                                                              F17_STARTUP_SEQUENCE))
                       date-1:    2008-10-16-08:26
                       date-2:

Produce:PerformTask/Activity (6.3)

*virt2.c33)  MOVE      SUBJ       AUXILIARY_LUBRICATION_PUMP_M202: (idle_)
                       OBJ        AUXILIARY_LUBRICATION_PUMP_M202: (running_)
                       CONTEXT    (SPECIF SEQ1_GREASING_PUMP (SPECIF member_of F17_STARTUP_SEQUENCE))
                       date-1:    2008-10-16-08:26
                       date-2:

Move:AutonomousChangeofState (4.32)

virt2.c30)     (GOAL  virt2.c32  virt2.c33)
```

From this example and from Table 2 we can also deduce that the top-level representation in NKRL terms of a complex, structured/dynamic knowledge entity (narrative, complex event, scenario) *has the general form of a binding occurrence*, see virt2.c30 in Table 2. This means that narratives, complex events etc. can always be represented as tree structures whose nodes correspond to the NKRL binding operators. This remark is valid in general, independently from the type of formalization used see, to give only an example, the story trees of Mani and Pustejovsky (2004). A recent paper dealing in detail with the NKRL approach to the formalization of the connectivity phenomena, the syntax and semantics of the associations of elementary events, the binding operators and related topics is Zarri (2014a). An example of formalization of a scenario including several elementary events can be found in Zarri (2014b)[5].

3 Deepening the notion of functional roles

In the semantic/functional roles partnership, the notion of functional role represents certainly the main element of interest. In this Section, we will try to clarify further the significance and the scope of this notion from two different points of view. The first concerns the relations of functional roles with thematic roles in Linguistics/Computational Linguistics. This comparison will allow us, among other things, to better define the *operational meaning* of the seven primitives introduced as functional roles in 2.2 above[6]. The second discusses the use of the semantic/functional roles association to find *concrete solutions* to (apparently) unsolvable problems like the so-called "counting problem".

3.1 Functional and thematic roles

The origins of the use in a Computer Science environment of an (at least initially) *pure* linguistic notion like that of thematic roles can be traced back to the publication of Charles Fillmore's ground-breaking paper (1968) about the Case Grammars theory. Fillmore's intention was to highlight the existence of *universal, deep grammatical relationships*. These are related both to the *explicit* inflected cases associated with nominals in Latin and other highly-inflected languages like German or the Slavic ones and to the personal pronouns, prepositions and positional clues denoting *implicit* "case distinctions" in English and other non-inflected languages. Rather than introducing for all these deep relationships – denoted, e.g., as Agentive, Instrumental, Beneficiary, Factitive, Locative – a new distinctive label, he used the term *case*, familiar from the case-inflecting languages. Fillmore's *cases* are now usually denoted as *thematic roles* (Palmer et al. 2010) in Linguistics/Computational Linguistics. Fillmore associates a set of cases with each verb sense. Some of them are obligatory and others optional; a case is obligatory if the sentence would be ungrammatical when omitted. For example, according to Fillmore's theory, "John gave the book" is ungrammatical, given the absence of the "Beneficiary" component.

Fillmore was a linguist and, as such, mainly interested in the description of the syntactic-semantic relationships proper to the *surface* aspects of any sort of linguistic utterance, i.e., of the relationships between natural language (NL) *words* in a specific language. Thanks to one of those shifts of sense that are not uncommon in the history of sciences, Fillmore's system was adopted by the Artificial Intelligence researchers of the late sixties to represent formally the *deep meaning* underlying the surface expression of events, situations, states, actions; see Bruce (1975), Woods (1975), Goldstein and Papert (1977), Rosner and Somers (1980) in this context. The *cases*, even if still labelled with terms in the Fillmore's style, are now *deep cases* that link together *conceptual entities* (concepts, concept instances, semantic predicates, spatio-temporal abstractions etc.) instead of *words*. Given that they refer to a common *deep meaning*, the formal expressions obtained in this way are independent from a particular NL, even if they are often drafted in a sort of Basic English for human understanding. Pure surface phenomena like the idiosyncrasies in the lexical choices, the active/passive alternation, the morphology, are ignored at deep level.

When trying, however, to find some concrete and computationally-exploitable lists of thematic roles to be used in an Artificial Intelligence mode, it appears clearly there *no consensus exists*. This situation is paralleled, by the way, by that existing in the proper Linguistics/Computational Linguistics domain, even when we take into account more

[5] The binding occurrence mechanism allows us to associate together, through reification operations and the use of connectivity phenomena operators, events (or even full narratives and scenarios) originally created as *independent entities*. Another *second order* NKRL mechanism permits to make reference to an elementary event (or full narrative), *reified* using its conceptual label, as an *argument* of another event. An example can correspond to an event *X* that denotes someone speaking about *Y*, where *Y* is an elementary event or a logically coherent set of events. This relational device is called *completive construction* in NKRL, see Zarri (2009: 87-91, 2014a). The completive construction mechanism indicates clearly that all the second order features of NKRL, *reification-based*, correspond to powerful *Higher Order Logic (HOL) structures*. According to HOL, in fact, a predicate can take one or more other predicates as arguments, allowing quantification over super-predicate symbols (Enderton 2012).

[6] Note, in spite of this excursion in the Linguistics/Computational Linguistics field, that the work described in this paper is deeply rooted in the Artificial Intelligence and Knowledge Representation research. Like all the similar/related approaches described in sub-section 5.2 below it is based, in fact, on the hypothesis that is possible and reasonable to develop a formal notation for expressing "deep meaning" contents in a way totally independent from the search for an optimal form of correspondence with the surface, linguistic form these contents can assume in a particular natural language; see Zarri (1997) in this context and, for example, Jackendoff's *θ-Criterion* (1990) for a completely different approach.

theoretical-oriented work that deals with related notions as the so-called "thematic relations" in Jackendoff (1990) or Van Valin (1999). And, according to Dowty (1989), it is doubtful that such a consensus could ever be found.

However, interesting proposals that have *a clear computational and operational flavor* can be found in Bruce (1975), Spark Jones & Boguraev (1987) and Sowa (1999)[7]. Note that, in all of these proposals, deep cases/thematic roles are considered as *primitives*, in agreement with the approach chosen for the functional roles in NKRL. In this context, we must also mention the work done by the Expert Advisory Group on Language Engineering Standards (EAGLES), an initiative promoted by the European Commission and aimed at accelerating the provision of standards for Computational Linguistics formalisms. In a report on Lexical Semantic Encoding (Web ref.2), the EAGLES researchers supply "…a list of the most popular roles and the properties usually associated with them", widely reproduced in the literature as a sort of consensus list about thematic roles. This list includes 7 items: Agent ("a participant which the meaning of the verb specifies as doing or causing something, possibly intentionally"), Patient ("a participant which the verb characterizes as having something happen to it, and as being affected by what happens to it"), Experiencer ("a participant who is characterized as aware of something"), Theme ("a participant which is characterized as changing its position or condition, or as being in a state or position, ex: syntactical objects of verbs like "give, hand", or subjects of verbs like "walk, die"), Location, Source ("objects from which motion proceeds"), Goal ("objects to which motion proceeds", ex: subjects of verbs like "receive, buy", dative objects of verbs like "tell, give"). A Beneficiary role ("for whose benefit action is performed") has been added in Palmer et al. (2010: 4).

When comparing the seven functional roles introduced in sub-section 2.2 with the above solutions a first, fundamental principle to be kept in mind is that our functional roles are strictly relative to frameworks adhering to Eq. 1, i.e., their duty consists solely in denoting *the functional relationships of the* a_i *arguments with respect to the predicate* P_j *in a modelling of elementary events context*. This principle allows us to discard, e.g., all the relationships that, in the previous proposals, can be associated with notions in the CAUSE (e.g., Force in Grimes and Spark Jones & Boguraev, Reason in Spark Jones & Boguraev) or GOAL style (e.g., Goal in Spark Jones & Boguraev and EAGLES and, at least partially, Completion, Destination and Result in Sowa), see also Note 7. These last roles do not concern, in fact, the representation of the *internal structure of an elementary event*: they are related, instead, to the representation of the *mutual relationships between two (or more) of these events*, see sub-section 2.3 above. Let us consider, e.g., examples like "The girl died from an accident" and "John went to town in order to buy a shirt", introduced by Spark Jones & Boguraev to illustrate the use of their Force and Goal roles, see again Note 7. For each of them we are confronted, actually, with *logical/temporal relationships* of the CAUSE/GOAL type between two (or more) elementary events. The above roles refer, in reality, to that set of surface syntactic constructions like causality, goal, indirect speech, co-ordination, subordination that denote, at *deep level*, the logical/temporal relationships collectively gathered in our approach under the term connectivity phenomena and endowed with their *proper representation tools*.

A second principle prevents the inclusion, in our list of functional roles, of all those "roles" dealing with temporal and spatial notions, like Locus and Time in Simmons, After, Before, Location… in Spark Jones & Boguraev, Duration, PointInTime in Sowa, Location in EAGLES. According to what expounded in the previous Section, in fact, the (single) semantic predicate, the seven functional roles and the (simple or complex) arguments of the predicate of Eq. 1 are the three *basic building blocks* strictly necessary to give rise to a meaningful representation, in the digital world, of structured/dynamic entities under the form of elementary events or of classes of elementary events. These three blocks, however, cannot receive *separately* an interpretation in terms of entities of the real world: a valid interpretation will only arise after their (mandatory) assembling through a L_i label has been carried out. This implies also that all the residual information-carrying elements – formally denoted as locations, modulators and temporal information – that could be necessary to associate with specific elementary events/templates, are dealt with simply as *determiners/attributes*, see Note 4. As already stated, they (can) introduce further details about the basic core represented by Eq. 1, but they are *never strictly necessary* for a meaningful interpretation of this core.

---

[7] Bruce describes several deep cases systems typical of the state of research in knowledge representation and computational linguistics of the seventies. Among them, the original version of Fillmore's case system includes eight relationships: Agent, Counter-Agent, Object, Result, Instrument, Source, Goal (defined here as "the place to which something moves") and Experience. Simmon's case system consists of seven relationships: Actant, Theme, Source, Goal ("place or state of the termination of the act"), Instrument, Locus, Time. A more articulate system is proposed by Grimes, who introduces a differentiation among Orientation Roles like Object, Source, Goal, Process Roles like Patient, Material, Result, Agentive Roles like Agent, Instrument, Force and the unique Benefactive Role, Benefactive. Spark Jones & Boguraev supplies a long list (28) of deep cases. These include, e.g., Activity, After, Agent, Before, Destination, Force ("The girl died from an accident"), Goal ("John went to town in order to buy a shirt"), Instrument, Location, Manner, Mental-Object, Object, Quantity, Reason ("John is afraid of being apprehended by the police"), Recipient, etc. In the Appendix B of his 1999 book, Sowa supplies his own list of thematic roles. They are for example: Agent, Beneficiary, Completion, Destination, Duration, Instrument, Location, Origin, Path (ex: "The pizza was shipped via Albany and Buffalo"), Patient, PointInTime, Recipient, Result, Start, Theme ("an essential participant that may be moved, said, or experienced, but is not structurally changed"), and so on.

Having supplied, through a comparison with the Computational Linguistics thematic roles, additional information about the origin and the intrinsic characteristics of our seven functional roles, we give their definition in Table 3. Note that the operational use of these roles depends also on the syntactic/semantic features of the templates they are inserted in, and on the practical modalities of use of these templates and of the related predicative occurrences.

Table 3. NKRL's functional roles.

Role	Acronym	Mnemonic Description
Subject	SUBJ	The *main actor* (the "agent") of the elementary event, independently from the grammatical/ syntactic form of the corresponding expression in natural language, see "Caesar has been stabbed by Brutus (the SUBJ)". The filler (argument of the predicate) of this role is normally (but not necessarily) an animate entity or a group of animate entities (e.g., a social body).
Object	OBJ	The entity, *animate* (e.g., Caesar, the "patient", in the previous example) *or inanimate* (e.g., the book moved from John to Mary), which is *acted upon* in the context of the elementary event.
Source	SOURCE	The animate entity (group of entities), *if any*, who is *responsible for* the behavior, situation, state etc. of the SUBJ of the elementary event.
Beneficiary	BENF	The animate entity, ("Mary" in the book example), or the group of entities, which denotes the *addressee* (the *recipient*) of the OBJ mentioned in the event (or, more generally, the addressee of the global behavior of the SUBJ of the event).
Modality	MODAL	The (often inanimate) *entity* (e.g., the knife) or the *process* (e.g., stabbing, if the event to be represented was "Brutus killed Caesar by stabbing him") that is *instrumental* in producing the situation described in the elementary event.
Topic	TOPIC	Additional *local* details or explanations (about, à propos of…, with respect to…) that concern the *main fact(s) or situation(s)* represented in the elementary event (e.g., the particular button that is affected by the button_pushing action in Table 1, see also Table 4 and 5 below).
Context	CONTEXT	The *general context* (in the context of…) of the fact(s) or situation(s) that is/are represented in the elementary event, e.g., "Roman Senate's fears about Caesar's ambitions" or "Mary's birthday". The filler of the CONTEXT role normally denotes other events or streams of events. It can also be represented by a single *reified event* denoted by a concept, like us_presidential_address, or an individual, like for example US_PRESIDENTIAL_ADDRESS_1 (Zarri 2009: 137).

3.2 Functional roles and qua entities

The *qua-entities paradigm* – loosely derived from the qua-concepts approach (concepts defined as functions of other concepts) described in a KL-ONE framework (Freeman 1982) – is quite popular today in relation to the theoretical debates about the notion of role (Masolo et al. 2004, 2005; Guizzardi 2006; Almeida and Guizzardi 2007; Loebe 2007; Vieu et al. 2008). This popularity is largely due to the widely shared belief that the qua-entities paradigm can supply the *only possible solution* to deal formally with a set of complex problems which question the standard view of roles as anti-rigid (temporary) and relationally dependent entities introduced, e.g., in Guarino (1992) and Guarino and Welty (2000). In this respect, the *counting problem* (Gupta 1980; Wieringa et al. 1995; Masolo et al. 2004, 2005, 2011; Guizzardi 2005, 2006; Almeida and Guizzardi 2007; Loebe 2007; Vieu et al. 2008) is considered "… one of the most difficult to solve" (Masolo et al. 2005: 103). And Loebe asserts (2007: 143) that he is not aware of a solution of the counting problem that could concretely work without introducing some kind of role instances, qua-individuals, or instances of the role universals *Q* in his own theory – note, however, that other solutions to this problem exist, see in this context, e.g., Barlatier and Dapoigny (2012: 348-349). To complete our discussion about the characteristics and advantages of the *dichotomy/association* between semantic and functional roles introduced in this paper, we will discuss in this sub-section how this particular approach can supply a *complete, plain and coherent solution* to this complex modeling problem. And this without any need to introduce additional conceptual tools with respect to those already proposed above to formalize the permanent/static and transitory/dynamic real world entities. This discussion will also allow us to show and comment other interesting examples of conceptual representations in NKRL terms.

Expressed according to its classical *passenger formulation* – see, e.g., Masolo et al. (2004, 2011) for different formulations – the counting problem consists of an argument formed of three sentences:

- Air France-KLM served about 93.4 million passengers in 2016
- Every passenger is a person.
- Ergo, Air France-KLM served about 93.4 million persons in 2016.

As there certainly have been persons who flew Air France-KLM more than once in 2016, simply counting the passengers cannot be considered as *universally equivalent* to counting persons that have played the role of Air France-KLM's clients in 2016; the argument is *surely flawed*. The natural way of getting rid of the counting problem and of determining the correct number of clients consists, accordingly, in counting not the passengers but the *carrying*

situations of the type "Air France-KLM carried a specific person *x* on the flight *y* on day *z*". This is, in practice, universally accepted and suggested by, among others, Masolo et al. (2004, 2011) and Vieu et al. (2008). However, severe difficulties rise when trying to translate this informal intuition into formalized models and procedures.

According to the qua-entities/qua-individuals approach, we must associate with ordinary individuals like LUC_, JOHN_, PETER_ and JANE_ that play a specific role (respectively customer_, husband_, passenger_, student_) with respect to external entities (FIAT_, MARY_, AIR_FRANCE/KLM, HARVARD_UNIVERSITY) *additional qua-individuals* under the form of, e.g., LUC_{qua}fiat-customer, JOHN_{qua}mary-husband, $\text{PETER}_{\text{qua}}$airfrance/klm-passenger and JANE_{qua}harvard-student. These qua-individuals are intended to solve difficulties in the counting problem style given that they can take into account the particular situations where the original individuals are playing a specific role by specifying exactly, for example, what the above different carrying situations are. In the *standard* interpretation of qua-entities, these new individuals are *instances (reification) of the particular roles involved* – they are, in contrast, instances of Quality in Guizzardi (2006) and Almeida and Guizzardi (2007). In the standard version, the qua-instances *inherit* the properties of the corresponding role and, according to the formulation used, e.g., in Vieu et al., (2008), Masolo et al., (2011), they also *inhere in* the original players. This means that they are *existentially dependent on, but different from*, the original individuals LUC_, JOHN_, PETER_ and JANE_. Moreover, they are also *existentially dependent* on (possibly a multitude of) other individuals (FIAT_, MARY_, AIR_FRANCE/KLM) characterizing the circumstances where they play a role. *Role types* (*role templates*) as $\text{Person}_{\text{qua}}$Customer, $\text{Person}_{\text{qua}}$Husband, $\text{Person}_{\text{qua}}$Passenger can be found in Genilloud and Wegmann (2000) and Almeida and Guizzardi (2007).

Doubts have been raised, from both a theoretical and a practical point of view, about the *general effectiveness* of this solution. In particular:

- The ontological status of the qua-entities is not well defined. According to their standard interpretation, see above, qua-individuals are considered as part of *ordinary ontologies* where they appear as *instances of concepts* like customer_, husband_, passenger_, student_ corresponding to *roles* and inherit the properties of these roles. Instantiating roles is, however, a *particularly controversial kind of ontological operation*. According to a large consensus in the ontological and Semantic Web milieus, in fact, "a role is … an anti-rigid externally dependent unary *property* [emphasis added]" – see, e.g., Guarino (1992, 2009). Therefore, roles should correspond to *provisional properties*. However, *instantiating a property does not coincide* with the usual ontological and Artificial Intelligence notion of instantiation, which means creating a *new* concrete example (an individual) *I* of an abstract notion represented by concept *C*. In the case of properties, the usual way of defining an *instantiation* is to state that the *specific property P* is *instantiated* when it is possible to identify an (*already existing*) individual *I* that is characterized by this property (Swoyer and Orilia 2014). This means that JANE_ – created independently from any property considerations – *instantiates* the property student_ or harvard_student by introducing a sort of dyadic relationship with this property that *does not correspond*, however, to any sort of traditional *subsumption* notion. In the qua-individuals case, the situation is even more complex given that these individuals, as stated above, are also *existentially dependent* on bunches of disjoint entities (FIAT_, MARY_, HARVARD_UNIVERSITY, AIR_FRANCE/KLM in the previous examples) that characterize the circumstances according to which they play a specific role. For other problems (e.g., Bradley's regress) linked with the notion of instantiation of a property see again Swoyer and Orilia (2014). As we have already seen in sub-section 2.1.1, NKRL gets rid of any theoretical problem about instantiation of properties by regarding all the *semantic roles* concepts in the style of customer_, husband_, passenger_, student_ as *non-sortal concepts* (i.e., deprived of direct instances).
- From a practical point of view, the criticism often raised against the qua-entity solution concerns the *not only inelegant but, mainly, actually unworkable proliferation of individuals* that will occur in case of systematic adoption of this approach; see also Masolo et al. (2004, 2005), Loebe (2007), Vieu et al. (2008) in this context. It is difficult, in fact, to imagine the practical exploitation of a non-toy system of reasonable dimensions where the users are regularly obliged to create new individuals for specifying all the possible everyday behaviors of Luc, John, Peter, Jane and the others. Moreover, in large and operational knowledge-based systems, the coherent management of large amounts of *quasi-identical individuals* proves to be particularly difficult. Taking up this approach could then be in contradiction with any possible interpretation of the Occam's razor or of the *parsimony principle* where this last, as stated in particular in Vieu et al. (2008: 122), "… is a principle that no-one, philosopher or computer scientist, renounces to".

To deal with situations in the qua-entity and counting problem style, NKRL does not require the addition of ad-hoc entities and makes use instead of its knowledge representation tools largely based on the semantic/functional roles approach. Let us consider the Behave:Role template (an offspring of Behave:HumanProperty that admits specializations like, e.g., Behave:User), see the upper part of Table 4. Behave:Role requires the use of semantic_role and of its specializations as filler of the (mandatory) MODAL functional role (Zarri, 2014b). If, as in the case of the qua-entity examples above, the precise identification of the external elements characterizing the circumstances according to which the SUBJ(ect) plays a semantic role is required, these elements are inserted as fillers of the (non-

compulsory in general) TOPIC functional role. The presence of the OBJ(ect) functional role is obviously forbidden, *(OBJ), in the predicative occurrences derived from this template; the date-1, date-2 temporal determiners can be used to specify how long the SUBJ(ect) stuck to this role, or the beginning and the end of this particular situation. Using this template and its specializations, one or more predicative occurrences of the Behave:Role type can be generated to represent qua-entity-like situations without requiring the introduction of additional conceptual tools and without inflating the ontologies with an extra amount of ad-hoc entities (individuals).

Table 4. The Behave:Role template and an example of its use in a qua-entity-like context**.**

name: Behave:Role
father: Behave:HumanProperty
position: 1.11
natural language description: 'A Human Being or a Social Body Acts in a Particular Role'

BEHAVE	SUBJ	*var1*: [*var2*]
	*(OBJ)	
	[SOURCE	*var3*: [*var4*]]
	[BENF	*var5*: [*var6*]]
	MODAL	*var7*
	[TOPIC	*var8*]
	[CONTEXT	*var9*]
	[modulators]	

var1 = human_being_or_social_body
var3 = human_being_or_social_body
var5 = human_being_or_social_body
var7 = semantic_role
var8 = pseudo_sortal_concept, sortal_concept
var9 = situation_, symbolic_label
var2, *var4*, *var6* = location_

virt2.c40)	BEHAVE	SUBJ	INDIVIDUAL_PERSON_102: (GP1Z_MAIN_CONTROL_ROOM)
		MODAL	production_activity_leader
		TOPIC	GP1Z_COMPLEX
		{ obs }	
		date-1:	2008-10-16-08:26
		date-2:	

As an example, the predicative occurrence represented in the lower part of the Table 4 relates the condition of INDIVIDUAL_PERSON_102, already mentioned in Table 1, as (i.e., "qua") production_activity_leader with respect to the GP1Z_COMPLEX, an installation for the separation of LPG (Liquefied Petroleum Gas) into butane and propane. obs(erve), see also Note 4, is a temporal modulator (Zarri, 1998, 2009: 71-72). It is used to state that, at the specific date associated with the date-1 determiner, the situation described in the predicative occurrence surely *holds*, even when it is impossible (or simply deprived of interest) to try to specify the real extent of the temporal interval where this situation continues to be true. We can now change, in Table 4, INDIVIDUAL_PERSON_102 into PETER_ (the *sole* Peter individual that it would be necessary to introduce in the context of a given application), the NKRL production_activities_leader semantic role into the semantic role passenger_ and the individual GP1Z_COMPLEX into the AIR_FRANCE/KLM individual. After having introduced a correct specific date, this will be sufficient to supply all the main elements needed to model a given *carrying situation* and to get rid of any counting problem. Note, moreover, that the details of the flight can be introduced by filling the CONTEXT role (see Table 3 above and Zarri (2009: 137)) with an individual like, e.g., AF_9786 – an instance of reified_event, specific term in turn of situation_. The precise circumstances of AF_9786 could be detailed, if necessary, in a proper set of predicative and binding occurrences.

Eventually, we will evoke here a particular case of counting problem – see, e.g., papers like Masolo et al. (2004, 2005, 2011) – where, according to the Authors, even the standard qua-individual approach is deemed to fail. According to Masolo et al. (2004: 276) in fact, this technique has difficulties "…with any case in which the event type doesn't select the specific role that the participant is playing during the event". In a situation where a former Italian Prime Minister participates in the same meeting both as Prime Minister and as President of its own Mediaset conglomerate, three qua-individuals should be created, INDIVIDUAL_281, INDIVIDUAL_281$_{qua}$Italian-prime-minister, INDIVIDUAL_281$_{qua}$Mediaset-president (Masolo et al 2004: 277) to denote the same physical person. The conclusion by Masolo et al. is that, in such cases, no *reasonable* solution to the counting problem seems *even to be possible*.

Once again, the (apparent) impossibility of generating a satisfactory formal solution for situations of this type is due to the lack of adequate knowledge representation tools. Let us consider Table 5, where INDIVIDUAL_PERSON_281 denotes the (*unique*) NKRL representation of the former Italian Prime Minister that is needed in the context of the previous situation – separate occurrences in the style of virt2.c40 in Table 4 can specify

his different appointments *without requiring the creation of new specific individuals*. participant_ is an HClass concept, a semantic_role pertaining to the transitory/generic_role sub-hierarchy, see Fig. 1 above. According to the definition associated with the TOPIC *functional role* in Table 3, the TOPIC filler (in our case, an expansion or structured argument) introduces further local details about the main theme (here, the participation in a meeting) of the corresponding elementary event. In ex1.c1, this filler indicates that INDIVIDUAL_PERSON_281 took part in this meeting, at the same time (COORD(ination)), as italian_prime_minister (a semantic role concept of the civil_service_role type) and as mediaset_president (a semantic role concept of the professional_role type). acting_as is a (non-sortal) concept included in the binary_relational_property branch of the relational_property sub-hierarchy of HClass, see Fig. 1. Other concepts pertaining to the same branch are, e.g., compared_to, exemplified_by, related_to, replacement_for; an example of multiple_relational_property concept is between_. The attributive operator, SPECIF(ication), and the collective operator, COORD(ination), have been introduced in 2.3. According to the characteristics of the situation considered, further details about the context of the event, MEETING_213, could be introduced in the CONTEXT's filler (see the example of Table 1) or specified in separate predicative occurrences.

Table 5. Participation in the same event according to different functions.

ex1.c1)	BEHAVE	SUBJ	INDIVIDUAL_PERSON_281: (MILAN_)
		MODAL	participant_
		TOPIC	(COORD (SPECIF acting_as italian_prime_minister) (SPECIF acting_as mediaset_president))
		CONTEXT	MEETING_213
		date-1:	2002-10-21
		date-2:	

Behave:Role (1.11)

4 Inference in NKRL

In this Section, we would like to emphasize the practical utility of the semantic/functional roles differentiation by discussing briefly its importance in the context of the implementation of the rule system of NKRL. The interested reader is referred to Zarri (2005, 2009: 183-243, 2013) for more complete information in this context.

4.1 Search patterns and the Filtering Unification Module

The basic building block for all the NKRL querying/inference procedures is represented by *Fum*, the *Filtering Unification Module*. *Fum* takes as input specific data structures called *search patterns*, p_i, whose task consists in specifying the features of the information to be searched for within a knowledge base (*KB*) of NKRL expressions.

Search patterns are used according to two different modalities:

- A search pattern p_i can be *directly created* by the user making use of the appropriate NKRL editing tools. In this case, it must be interpreted as a direct formal query used, in an advanced information retrieval system mode, to retrieve the corresponding (if any) information stored in the *KB*.
- A search pattern p_i formally identical to a *manual* pattern can also be *created automatically* by the different versions of the *InferenceEngine*. This occurs when a particular phase of a reasoning procedure (a specific step of an inference rule) must be validated by examining the contents of the *KB*. The validation of this step corresponds to the positive retrieval of one/more predicative occurrences c_j compatible with p_i (i.e., where c_j is matched to p_i).

Actually, search patterns p_i correspond to specialized/partially instantiated templates where all the *explicit variables* – identified by conceptual labels in the var_i style, see Table 1 and 4 above – have been replaced by HClass concepts/individuals compatible with the original constraints imposed on these variables. In particular, the HClass *concepts* introduced in this way play now the role of *implicit variables* in the sense that, when the comparison between p_i and the predicative occurrences c_j of the *KB* takes place, the p_i concepts can successfully match not only the identical c_j concepts, if any, but also *the instances (individuals) of these last concepts and all their specific terms in HClass with the corresponding instances*. Therefore, search patterns make use of "semantic expansion" techniques.

Let us consider the search pattern p_i of Table 6, to be interpreted here as an external query posed by the user. This pattern clearly derives from the template Produce:PerformTask/Activity of Table 1. The two mandatory functional roles SUBJ(ect) and OBJ(ect) of the template, along with the optional CONTEXT, have been conserved. *var1* has been replaced by the concept human_being (a specific term of the *var1* restriction, human_being_or_social_body), *var3* has been replaced by device_use, a specific term of the constraint activity_, while industrial_temporal_sequence, a

specialization of situation_, takes the place of *var10*. The *temporal information* to be associated with the query is expressed through the two temporal attributes *date1* and *date2*, which represent now the *search interval* to be used to limit the search to the slice of time that is considered as appropriate to explore (Zarri 2009: 194-202).

Table 6. A simple example of search pattern (query) to be dealt with by *Fum*.

PRODUCE
SUBJ : human_being :
OBJ : device_use :
CONTEXT : industrial_temporal_sequence
{}
date1 : 2008-10-16
date2 : 2008-10-16

Try to see whether, on October 16, 2008, there is any information in the system concerning the use of any type of device in an (unspecified, but temporally characterized) industrial context.

The search pattern of Table 6 returns as an answer, among all the possible answers congruent with the syntax of the pattern and the state of both Class and the *KB*, the predicative occurrence virt2.c32 of Table 1. Note that, even if the binding occurrences of the *KB* are not directly involved in the p_i/c_j unification process, the query interface can show all the binding occurrences associated with a given c_j returned as an answer and, in case, the complete narrative/episode/scenario that includes c_j.

The main principle underpinning the matching operations performed by *Fum* can be expressed plainly as follows: "All the terms (predicate, roles, simple arguments, expansions with their proper operators, locations and modulators) used to build up a search pattern p_i must be necessarily retrieved in the same order in the matched occurrences c_j. The retrieved terms can be either *strictly identical* to the corresponding terms of the pattern (as in the case of the predicate and the roles), or be represented by *subsumed concepts or instances of the concepts used as role fillers and location fillers within the pattern*. Additional terms (e.g., roles and role fillers) with respect to those included in the pattern can be freely retrieved in the c_j without affecting the result of the match". Special directions must be followed for the comparison of temporal data see, e.g., Zarri (2009: 194-201). The above principle can be summarized by saying that, for a successful match, i) the search pattern must be *fully included* within the matched occurrences – therefore, the unification is *oriented*, from the pattern (i.e., the query) towards the occurrences (i.e., the targets), and ii) the implicit variables (concepts) of the pattern must all find *a correspondence with some of their subsumed HClass terms*, concepts or individuals, within the matched occurrences. Further details can be found in Zarri (2009: 183-201, 2013).

4.2 The high-level inference procedures

Two kinds of high-level rules, *transformations* and *hypotheses*, are systematically used in the NKRL applications. See Zarri (2009: 201-243) for the technical details and for the description of other kinds of high-level inference rules.

Given a *KB* of well-formed NKRL occurrences, the querying system mentioned in sub-section 5.1 is *complete* with respect to the particular version of HClass used. This does not mean, of course, that at least a c_j is systematically returned by *Fum* for every submitted p_i, either manually or automatically created. Transformation rules try, then, to adapt a search pattern p_i that failed (that was unable to find a unification within the KB) to the real contents of this base. They attempt to *transform* p_i into one or more *different* p_1, p_2 … p_n that *are not strictly equivalent but only semantically close* (analogical reasoning) to the original one.

A transformation rule is made of a left-hand side, the *antecedent* – the formulation, in search pattern format, of the query to be transformed – and of one or more right-hand sides, the *consequent(s)* – the representation(s) of one or more search patterns to be substituted for the given one. Denoting with A the antecedent and with Cs_i all the possible consequents, these rules can be expressed as:

$$A(var_i) \Rightarrow Cs_i(var_j), \quad var_i \subseteq var_j \qquad (3)$$

The restriction $var_i \subseteq var_j$ corresponds to the usual *safety condition* that guarantees the logical congruence of the rules by stating that all the variables declared in the antecedent A must also appear also in the consequent Cs_i accompanied, in case, by additional variables. The transformation arrow of Eq. 3 can be interpreted according to a double reading:

- Operationally speaking, the arrow indicates the direction of the transformation. The original search pattern (corresponding to a specialization of the left-hand side A of the transformation rule) is *removed* and *replaced* by one or several new search patterns obtained through the updating, using the parameters of the original pattern, of the right-hand side Cs_i.

- From a semantic point of view, we assume that between the information retrieved through Cs_i and the information we wanted to obtain through the original p_i there is a sort of *implication relationship*. This, normally, denotes solely a possible (*weak*) implication.

Let us look at an example concerning the oil/gas application already mentioned (Zarri 2011b). We want to ask whether a given oil extractor is running making use of a *KB* where all the events and situations related to the activation of a gas turbine are stored. In the absence of a direct answer we can reply by supplying, thanks to a rule like *t11* of Table 7, other related events stored in the *KB*, e.g., information stating that the site leader has heard the working noise of the oil extractor. This last result could be paraphrased in natural language as: "The system cannot assert that the oil extractor is running, but it can certify that the site leader has heard the working noise of this extractor".

Table 7. An example of transformation rule.

t11: *"working noise/condition" transformation*

antecedent:

OWN SUBJ *var1*
OBJ property_
TOPIC running_

var1 = consumer_electronics, hardware_, surgical_tool, diagnostic_tool/system, small_portable_equipment, technical/industrial_tool

first consequent schema (*conseq1*):

EXPERIENCE SUBJ *var2*
OBJ evidence_
TOPIC (SPECIF *var3 var1*)

var2 = individual_person
var3 = working_condition, working_noise

second consequent schema (*conseq2*):

BEHAVE SUBJ *var2*
MODAL industrial_site_operator

Being unable to demonstrate directly that an industrial apparatus is running, the fact that an operator hears its working noise or notes its working aspect can be a proof of its running status.

In this example, as soon as the transformation *InferenceEngine* has verified (by successfully unifying the antecedent *A* with the original p_i) that a transformation rule like *t11* can supply (at least in principle) a valid answer to the problem posed, it tries to validate the first consequent schema, *conseq1*, by converting it into a valid search pattern p_j. The *automatic construction* of a search pattern from the formal description of a specific inference step implies two sorts of operations. In the first, the values previously bound to some of the explicit variables that appear in this formal description are substituted to these variables. In the Table 7 example, we can suppose that, during the initial unification of *A* with the search pattern p_i to be transformed, *var1* has been bound to the individual oil_extractor_7; this value is substituted to *var1* in *conseq1*. For the residual variables, the associated constraints are associated with the variables according to a chronological backtracking approach. This means that, whereas *var2* in *conseq1* can be replaced only by individual_person, a first search pattern will be built up from this consequent schema by using the value working_condition for *var3* and, in case of failure (no unification with information in the *KB*), it will be replaced by a new one with *var3* = working_noise. If the validation operations for the reasoning step *conseq1* are successful, *InferenceEngine* will pass to the validation of *conseq2* according to the same principles; a high-level inference rule composed of several reasoning steps is validate *iff all its reasoning steps are validated.*

With respect now to the hypothesis rules, these allow us to build up automatically a sort of *causal explanation* for an elementary event (a predicative occurrence c_j) retrieved using a given p_j within a NKRL *KB*. These rules can be expressed as *biconditionals* of the type:

$$X \text{ iff } Y_1 \text{ and } Y_2 \ldots \text{and } Y_n \,, \qquad (4)$$

where the head *X* of the rule corresponds to a predicative occurrence c_j to be explained and the reasoning steps Y_i – called *condition schemata* in a hypothesis context – must all be satisfied. This means that, for each of them, at least one successful search patterns p_i must be (automatically) set-up by the hypothesis *InferenceEngine* in order to find, using *Fum*, a successful unification with some information of the base. In this case, the set of $c_1, c_2 \ldots c_n$ predicative

occurrences retrieved by the condition schemata Y_i thanks to their conversion into p_i can be interpreted as a *context/causal explanation* of the original occurrence c_j (X).

To mention a well-known NKRL example let us suppose we have directly retrieved, in a querying-answering mode, the information: "Pharmacopeia, a USA biotechnology company, has received 64,000,000 dollars from the German company Schering in connection with an R&D activity"; this information, see Table 8, corresponds to c_j (X).

Table 8. Predicative occurrence retrieved through a direct user query.

conc2.c34)	RECEIVE	SUBJ	(SPECIF PHARMACOPEIA_ (SPECIF biotechnology_company USA_))
		OBJ	(SPECIF money_ usa_dollar (SPECIF amount_ 64,000,000))
		SOURCE	(SPECIF SCHERING_ (SPECIF pharmaceutical_company GERMANY_))
		TOPIC	r_and_d_activity
		date1 :	
		date2 :	

A hypothesis rule (*h1*) able to supply a sort of causal explanation for this event is represented in Table 9; even if the final results produced by the hypothesis version of *InferenceEngine* must be interpreted in a different way from those obtained through the transformation version, the execution strategy is not very different.

Table 9. An example of hypothesis rule.

h1: *"payment order hypothesis"*

premise:

RECEIVE	SUBJ	*var1*
	OBJ	money_
	SOURCE	*var2*

var1 = company_
var2 = human_being, company_

A company has received some money from another company or a physical person.

first condition schema (*cond1*):

PRODUCE	SUBJ	(COORD *var1 var2*)
	OBJ	*var3*
	BENF	(COORD *var1 var2*)
	TOPIC	(SPECIF manufacture_ *var4*)

var3 = mutual_commitment, business_agreement
var4 = artefact_

The two parties mentioned in the premise have concluded an agreement about the creation of a product.

second condition schema (*cond2*):

PRODUCE	SUBJ	*var1*
	OBJ	*var4*
	MODAL	*var5*
	CONTEXT	*var3*

var5 = industrial_process, technological_process

The company that received the money has actually created the product mentioned in the first condition schema.

The processing of *h1* starts by trying to unify the premise of the hypothesis *h1* and the event conc2.c34 to be explained. Having validated the bindings *var1* = PHARMACOPEIA_ and *var2* = SCHERING_, these values are used to derive a valid search pattern p_i from the first condition schema (*cond1*), see Table 9. This can lead to the creation of deadlocks, linked mainly to the chronological use of the constraints. For example, when trying to utilize a pattern derived from *cond1* where the variable *var3* has been substituted by its first constraint mutual_commitment, see Table 9, a failure is generated. The occurrences we must retrieve in the *KB* about the links between Pharmacopeia and Schering concern, in fact, some sorts of business_agreement (the second constraint on *var3*) and not a *private* arrangement as mutual_commitment. We will, eventually, find in the base an instantiation of *cond1* that corresponds to an event as, "Pharmacopeia and Schering have signed two agreements about the production by Pharmacopeia of a new compound, COMPOUND_1". The values associated with the variables *var3* (r_and_d_agreement and sale_agreement, both specific HClass terms of business_agreement) and *var4* (COMPOUND_1) in *cond1* will be used to create the search patterns derived from *cond2*. An occurrence corresponding to the event, "In the framework

of an R&D agreement, Pharmacopeia has actually produced the new compound", will be retrieved. Therefore, the global information obtained can supply a sort of plausible explanation of the Schering's payment: Pharmacopiea and Schering have concluded some agreements for the production of a given compound, and this compound has been actually produced by Pharmacopeia.

Note that (as usual in a hypothesis context) the explication proposed by the above rule corresponds to *only one* of the possible reasons that can be interpreted as the cause of the original event. A particular hypothesis rule must be conceived always as a member of a *family* of possible explications. Moreover it is evident than, in an NKRL high-level inference rules context *in general*, an important collaborative work involving the active participation of *domain experts* should be anticipated to be able to conceive and formalize really useful and problem fitted rules.

An in-depth description of the different versions of *InferenceEngine* is beyond the possibilities of the present paper. In brief, these versions operate according to a backward chaining approach with chronological backtracking, see Clocksin and Mellish (1981). The differences with respect to other applications of this well-known approach (e.g., Prolog) are mainly linked to the *unusual complexity proper to the NKRL data structures.* This implies for example the need, after a deadlock, to carry out particularly complex operations of restoration of the previous program environment (based on the use of *environment variables*) to be able to return to the previous choice point in the execution tree. A detailed description of these operations can be found in Zarri (2009: 205-212).

4.3 Integrating transformation within hypotheses

An interesting recent development of NKRL concerns the possibility of using the two modalities of inference above in an integrated way (Zarri 2009: 216-234): this means that it is possible to make use of transformations when working in a hypothesis context. Therefore, whenever a search pattern p_i is derived from a condition schema Y_i of a hypothesis (see Eq. 4) to implement a step of the reasoning process, we can use it both *as is* – i.e., as originally built up by *InferenceEngine* from its father condition schema – but also in a *transformed form* if the appropriate transformation rules exist. In this way, a hypothesis that was deemed to fail because of the impossibility of deriving a successful p_i from one of its condition schemata can now continue if a new p_j, obtained using a transformation, will find an effective unification within the base, getting new values for the hypothesis variables.

We supply an informal example, see Zarri (2011b). Using the hypothesis of Table 10 we can try to show, see the premise, why an operator has activated a (particularly spectacular and costly) piping segment isolation procedure in the context of a gas leakage (or of other important failures). A possible explication is based on these main elements: i) a previous milder maintenance procedure has been executed (*cond1*), but this was unsuccessful (*cond2*); ii) the accident is a serious one (*cond5*). In the absence of information corresponding exactly to this last condition, the p_i derived from *cond5* can be transformed to obtain *indirect* confirmations of the gravity of the accident. We will then get information in the style of, e.g., "The gas leakage has a gas cloud shape" or "An alarm situation has been validated (*conseq1*) *and* the level of this alarm is 30% LEL, Low Explosion Level (*conseq2*)", see the last lines of Table 10.

Table 10. Gas/oil hypothesis in the presence of transformations.

(***premise***)	An individual has carried out an *isolation* procedure in the context of an industrial accident.
(***cond1***)	A different individual had carried out previously a (milder) *corrective maintenance* procedure.
(***cond2***)	This second individual has experienced a failure in this corrective maintenance context.
(***cond3***)	The first individual was a control room operator.
(***cond4***)	The second individual was a field operator.
(***cond5***)	The industrial accident is considered as a serious one.
	– (**Rule t6, *Consequent***) *The leakage has a gas cloud shape …*
	– (**Rule t8, *Consequent***) *A growth of the risk level has been discovered …*
	– (**Rule t9, *Conseq1***) *An alarm situation has been validated,* ***and***
	– (**Rule t8, *Conseq2***) *the level of this alarm is 30% LEL (Low Explosion Level).*

4.4 Functional roles and inference rules

We can conclude about the NKRL querying/inference procedures by noticing that, with reference to the main topics of this paper, the use of *functional roles* allows us to build up inference rules that are both concise, powerful and particularly expressive. From what stated above and looking at the previous examples it appears clearly, in fact, that the shaping elements of these rules (antecedents, consequents, condition schemata…) follow strictly a (partially instantiated) *templates format*. Therefore, they conform to the *n*-ary approach of Eq. 1, whose *n*-ary properties depend largely by the introduction of functional roles as structural links between the semantic predicate and its arguments. The NKRL rule system can consequently profit from the improved expressiveness properties generally associated

with the *n*-ary knowledge representation approaches, see also Section 5 below. In particular, the use of an *n*-ary approach has a very important consequence: we are no more constrained to express the *atoms* of these rules in terms of binary clauses using just monadic (unary) and dyadic (binary) predicates (as usually carried out in an Artificial Intelligence context) and these atoms can directly represent complex and meaningful situations see, e.g., the logical schemata used to build up the rules of Table 7 and 9.

Note that modeling these atoms according to a simple binary format implies, among other things, the impossibility of stating even relatively simple implications of the "if-then" type making use of a reduced number of clauses. This leads in practice to the necessity of the (systematic) decomposition of the original formulation of the rules into (possibly very large) amounts of simple *binary clauses*, implying serious problems both from a logical (and operational) point of view. It is evident, in fact, that techniques of this type are relatively simple to use when the rules to be built up are also quite simple, but they can be out of place in more realistic situations. Returning, e.g., to the relatively simple example of Table 9, the decomposition techniques would be unable to deal with this situation using a limited number of binary clauses. In NKRL on the contrary, expressing in a complete way the *common sense meaning* that supports this inference rule only requires to make use of two *AND*-linked *n*-ary clauses and of an indirect implication *IF-THEN* involving the premise. This and related issues are dealt with exhaustively in Zarri (2013).

5. Some comparisons

Given the present tendency in the Knowledge Representation domain, see Section 1, to formally model roles according to an approach not too different from the use of a simple attribute-value *binary* model, it is not very easy to find operational systems sharing the chiefly *n*-ary approach to deal with roles detailed in this paper.

5.1 Recent popular binary models

The W3C languages, or Semantic Web (SW) languages, OWL, OWL 2, RDF(S) etc. are characterized by a low level of *expressiveness* given that, being based principally on a simple binary model, they are unable to represent *natively* in a compact and natural way the different relationships between *multiple related entities* within a given *spatio-temporal framework*. Because of the difficulties of making use of W3C/SW languages to model important real world entities like situations/scripts/scenarios, the SW scholars have tried to introduce some sorts of *n*-ary features in these languages – with limited success, given the *deep and irreducible differences* between binary and *n*-ary paradigms.

The most well-known SW *n*-ary proposal is titled "Defining *N*-ary Relations on the Semantic Web" (Noy and Rector 2006); this proposal is analyzed in depth in Zarri (2009: 17-21). We can note in this proposal, for example, the invasive addition *of ad-hoc binary links and questionable RDF/OWL classes and individuals* the Authors use to take into account relatively simple situations like "Christine has breast tumor with high probability". A recent analysis of the different efforts – RDF reification, named graphs, RDF blank nodes plus rdf:value properties, quad-tuples, *n*-ary design patterns, F-events – developed these last years to try to transform binary tools into *n*-ary ones taking the temporal dimension into account is presented in Trame at al. (2013). Their conclusions (not really optimistic) go from "… RDF reification is one of the most confusing and controversial constructs in the RDF specification" (p. 208) to "… the F-Event model … add far more complexity than required … [it] does not suggest how to distinguish among different kinds of participants in an event" (p. 215) passing through, e.g., "… n-ary patterns … increase the number of statements and lead to a proliferation of objects … [they are] frequently used in an arbitrary fashion, lacking any design rationale" (p. 211). These conclusions are particularly interesting as *intrinsically unbiased*, given that the Authors are enthusiastic supporters and users of the W3C/SW languages. As explained in a Section of the paper, they have chosen to deal with the representation and management of a rather simple type of narrative/biographic information making use, in fact, of a typical W3C/SW tool, the DOLCE+DnS Ultralite (DUL) ontology (Web ref.3).

Other recent, popular approaches developed in a general Semantic Web context concern *linked data*, LD (Berners-Lee 2006, Heath and Bizer 2011) and *ontology design patterns*, ODPs (Blomqvist 2009, Falbo et al. 2013).

In difference of the traditional Web where the documents can be crawled by following *hypertext links*, in the Linked Data (LD) Web they are crawled by following *RDF links* to gather information stating that one piece of data has some kind of relationship to another piece of data. See in this context, e.g., the popular LOD (Linked Open Data) diagram, Web ref.4. By September 2011, this diagram had already grown to 31 billion RDF triples, interlinked by around 504 million RDF links; an updated version has been published in April 2014. From a general point of view, the LD approach has been often criticized because of the *necessity of associating all the exploitable data with HTTP URLs that point at RDF descriptions*. Taken to its extreme consequences, this means that all the real world entities should be characterized by HTTP URLs supplying RDF data when fetched. From our more limited viewpoint and without challenging the practical utility of LD in many domains, it is evident that LD and NKRL are used according to two different perspectives. While LD operate in a *big data* context by trying to achieve useful results through the mechanical application of shallow surface techniques, NKRL works according to a typical *symbolic* approach by exploring carefully in some depth the meaningful characteristics of relatively limited conceptual domains.

With respect now to the ODPs, they could have, at least in principle, something in common with NKRL, given that NKRL templates can be seen as kinds of *formal patterns*. The standards ODPs definition (Blomqvist, 2009: 140) says, in fact, that "… an ontology pattern is a set of ontological elements, structures or construction principles that intend to solve a specific engineering problem and that recurs, either exactly replicated or in an adapted form, within some set of ontologies …". In practice, however, current ODPs denote simply *small fragments of ontologies in the DOLCE's and SW style* that should provide a modelling solution to very specific sets of problems in the legal, financial, medical, commercial, technical, social domains and that could be composed, specialized and reutilized. Moreover, when looking at the patterns collected in the ODP portal (Web ref.5), it appears clearly that the existing ODPs, at the difference of templates, are characterized by a *high level of heterogeneity* and by the *lack of shared theoretical principles* for their construction and use. They evoke those idiosyncratic patterns that "lack compatibility with others" and lead "to decrease the semantic interoperability of ontologies" mentioned in a well-known Kozaki et al. paper (2007: 39) – see also the conclusions of Trame et al. (2013) about the "*n*-ary patterns" mentioned above.

We conclude this sub-section by mentioning the very recent Schema.org initiative (Web ref.6), launched jointly, in June 2011, by the search engines Google, Bing, Yahoo and the Russian Yandex. Schema.org aims at defining *an interconnected vocabulary of terms*. These can be added to the HTML markup of a web page to clarify the *meaning* of the concepts on that page and to improve the results obtained by the search engines. Concretely, the vocabulary is a sort of *ontology of schemas* that define the different items of interest by i) classifying an item as a specific type of resource and ii) by defining a set of properties/attributes to provide a description of the main characteristics of the item. The hierarchy of types of Schema.org has Thing as top level. The types deriving directly from Thing are Action, BroadcastService, CreativeWork, Event, Intangible, MedicalEntity, Organization, Person, Place and Product, see Web ref.7. There has been no effort in structuring the hierarchy according to some sort of semantic/logic criteria. Information associated with the different schemas is expressed according to a standard *binary* approach, i.e., as *sets of properties/values*, some peculiar to the entity described in the schema, and other inherited from higher level schemas. The meaning of a property is expressed in natural language. Its possible values are constrained by denoting the expected type, i.e., by indicating the terms of the hierarchy – their schemas with their proper associated properties – that can be used recursively as fillers. For example, the schema for Book, specific type of CreativeWork, includes peculiar properties like isbn (expected type = Text), properties derived from CreativeWork like Author (expected type = Person or Organization) and dateCreated (expected type = Date), properties derived from the top-level Thing like image (expected type = URL or ImageObject). This sort of information is expressed in machine readable format using tools such as microdata, RDFa Lite and JSON-LD (JSON-LD = JSON for Linked Data). When the different schemas are instantiated to represent specific entities, the final values of the different attribute consist normally of LN strings, digits, proper names, URLs etc., i.e., of *non-formal items*. The Schema.org metalanguage is surely interesting as a particularly rich markup language but, at least for the time being, it is difficult to understand it as a full-fledged inference framework allowing reasoning under logical consistency. See also, in this context, several incoherencies and misunderstandings from a semantic/conceptual point of view concerning some Schema.org definitions, those, for example, about the use of subtypes of Action like agent, object and participant (Web ref.8).

5.2 Old and more recent *n*-ary approaches

The literature of all possible approaches – often purely theoretical – proposed for the set-up of *n*-ary conceptual structures is particularly copious. Here we will limit to discussing some solutions that have directed influenced the development of NKRL or that have a strong affinity with this language. For a larger panorama see, e.g., Mueller (2013); interestingly, this last has been drafted from a "computational narratives" point of view.

Starting from the fifties-sixties, the solutions proposed by Silvio Ceccato (1961, 1964) in a Mechanical Translation context are largely based on the use of two associated tools, the *tabellone* (the big table, the main chart) and the *sfera nozionale* (the notional sphere, the network of notions). The former is a collection of (more than 150) syntactic/semantic triadic structures (*correlations*) arranged around a central *correlator*. Correlators can be *explicit* and represented by terms like "through", "while" or "with the exception of"; they can also be denoted *implicitly* by high-level conceptual relationships like coordination and subordination, apposition, subject-predicate development, subject-result of her/his/its activity. The correlators materialize the linking element of the triadic correlations, allowing us to associate a left (e.g., "Mary") and a right (e.g., "sings") *correlatum*; correlations are denoted globally by numerical labels having a function similar to that of L_i in Eq. 1. The *correlata* can also correspond to full correlations. This permits the creation of associations of embedded triadic structures able to represent situations as "Mary sings because she loves Peter". The *tabellone* concerns the formal representation of *structured/dynamic knowledge* (see sub-section 2.1.2) where the correlators act in practice as *functional roles*; the *sfera nozionale* concerns the *permanent/static knowledge* (2.1.1). Used to disambiguate the selection of the different *correlata*, this last identifies the properties of the entities (the nodes of the networks) listed in the dictionary proper to a specific application by associating them through a set of conceptual relationships (the edges). Examples can be "[to eat]

Activity/Its Usual Place [table]", "[table] Species/Genus [furniture]", "[to seat] Activity/Instrument [chair]", "[bed] Principal Thing/Accessory [pillow]; the network can be recursively traversed to deduce automatically new relationships[8].

The Conceptual Dependency theory of Roger Schank (Schank 1973, Schank and Abelson 1977) is an *n*-ary model that was very popular in the seventies. In this theory, at least partially inspired by Fillmore's work, the underlying meaning (*conceptualizations*) of narrative-like utterances is expressed as combinations of semantic predicates chosen from a set of twelve formal *primitive actions* (like INGEST, MOVE, ATRANS, the transfer of an abstract relationship like possession, ownership and control, PTRANS, physical transfer, etc.) plus states and changes of states, and seven *role relationships* (that can be likened to functional roles) in a Case Grammar style: Object (in a state), Object (of a change of state), Object (of an action), Actor, Recipient/Donor, From/To, Instrument. In one of Schank's well-known examples (1973: 201), the representation of the narrative "John eats the ice cream with a spoon" is obtained under the form of a conceptualization where (schematically) John, the Actor, INGEST the Object (of an action) ice cream via (Instrument deep case) a PTRANS of another Object, spoon, from the soup to the mouth of John. Unfortunately, Schank's theory was, on the one hand, insufficiently specified and, on the other, unnecessarily complicated because of the influence of psychological (introspective) considerations according to a characteristic trend of AI in those years. Nevertheless, Schank's work had a particularly strong influence on the development of formalized (and at least partly computerized) systems for the representation and management of high-level scripts, scenarios, narratives see, e.g., (Dyer 1983, Kolodner 1984). The SnePS (Semantic Network Processing System) proposal of Stuart Shapiro (1979) belongs roughly to the same period and allows us, e.g., to annotate situations like "Sue thinks that Bob believes that a dog is eating a bone" by associating labeled graphs in a way similar to NKRL's completive construction style, see sub-section 2.3 and Note 5 above.

Among the recent *n*-ary proposal, we must mention at least two well-known systems: CYC (Web ref.1, Lenat et al. 1990) and Topic Maps (Garshol and Moore 2008, Pepper 2010).

CYC concerns one of the most controversial endeavors in the history of Artificial Intelligence. Started in the early '80, the project ended about 15 years later with the setup of an enormous knowledge base containing about a million of hand-entered logical assertions including both simple statements of facts and rules about what conclusions could be inferred if certain statements of facts were satisfied. A detailed analysis of the origins, developments and motivations of CYC can be found in Bertino et al. (2001: 275-316). The upper level of the CYC ontology (OpenCyc) is now freely accessible on the Web, see Web ref.1. CYC represents a really amazing achievement; it has been, however, also severely criticized. For example, when examining the CYC's collection of entities, people are often struck by its *complexity* (i.e., the number of branches in the CYC's ontology and their inter-relationships) and wonder whether this complexity is really necessary. From a more technical point of view, a criticism often addressed to CycL – the *n*-ary knowledge representation language of CYC – concerns its *uniform use* of the same representation model (substantially, a frame system rewritten in logical form) to represent phenomena conceptually very different (see the *one and only syndrome* already mentioned in 2.1). A domain where the CYC researchers are actually (and successfully) very active is Question Answering see, e.g., Lenat *et al.* (2010).

Topic Maps (TMs) were developed in the '90s. Their key notions concern *topics*, *occurrences* and *associations*. A "topic" is used to represent any possible specific notion that could be interesting to speak about, like the play Hamlet, the playwright William Shakespeare, or the authorship relationship. Topics can have names, and each individual topic is an instance of one or more classes of topics (topic types) that may or may not be indicated explicitly. They can also have "occurrences", that is, information elements – e.g., links to external resources – that are considered to be relevant in some way to the subjects the topic reify. Links from topics to occurrences can be, e.g., discussed-in, mentioned-in or depicted-in. Topics can participate in relationships with other topics, called "associations"; an association can be thought of as an *n*-ary aggregate of topics. Associations consist of a number of "association roles" each of which has a topic attached as a role player; the association role defines how its role player topic takes part in the association according to a mechanism that can evoke functional roles' machinery. Topic Maps became ISO standard in 2000 as ISO/IEC 13250:2000.

From the point of view discussed in this paper, the most interesting modern *n*-ary proposal is certainly represented by *Conceptual Graphs (CGs)*. CGs are the powerful knowledge representation and knowledge management environment introduced by John Sowa in the late seventies and described in depth in two Sowa's books (1984, 1999) and in several papers. CGs are built according to an *n*-ary conceptual model structured around a *predicate* and its

[8] It can be interesting to note the similarity of Ceccato's *sfera nozionale* with a recent popular tool developed in a MIT context like ConceptNet (Liu and Singh, 2004, Speer and Havasi, 2012, Web ref.9). ConceptNet is a semantic network whose nodes represent *compound concepts*: these are made up of words or short phrases in semi-structured English and of labeled relationships between them. The relationships (21, including IsA) are in the form of, e.g., CreatedBy, PartOf, UsedFor, PrerequisiteOf, DefinedAs, LocatedNear. (Recursive) compound concepts can be "[wake up in the morning] PrerequisiteOf [eat breakfast]", "[kitchen table] UsedFor [eat breakfast]" "[chair] LocatedNear [kitchen table]" (Liu and Singh, 2004: 213), see the sfera nozionale's examples above. John Sowa (1984) had already noticed that Ceccato's sfera nozionale was one of the first concrete examples of semantic network and that it predates the publication (1968) in abridged form of the Ross Quillian's Ph.D. thesis (1966) that is conventionally assumed as the starting point of the semantic network developments.

arguments (the entities involved in the action, state, situation, event… denoted by the predicate), similar to the model utilized in the NKRL's templates. In CGs, the arguments are associated with the predicate through relational operators called *conceptual relations*, corresponding clearly to the functional roles of NKRL. For example, a CG describing the elementary event "John is going to Boston by bus" is represented by a conceptual structure where a "concept node", Go (having a function similar to that of an NKRL conceptual predicate, but denoted by a natural language term) is associated with three "relation nodes" (functional roles) like Ag(e)nt, Dest(ination) and Instr(ument). These last introduce the three arguments of the predicate i.e., three new concept nodes representing, respectively, the constant John (the agent) as an instance of the concept Person, the constant Boston (the destination) as an instance of the concept City, and the concept Bus (the instrument). Other similarities with NKRL tools concern the use of second order conceptual structures for dealing with contexts and scenarios, and the way of implementing some basic inference procedures. Important differences involve, however, the organization of the hierarchy of concepts (relatively simple in a CGs context with respect to the complexity of HClass in NKRL) and the choice and the *meaning* of the primitives used to denote, respectively, conceptual relations and functional roles.

The main divergence between NKRL and CGs concerns, however, John Sowa's choice of leaving completely free, for the sake of generality, the choice of the predicate in the CGs conceptual structures. Normally, the predicate is then represented as a *NL surface verb*, see "Go" in the example above – it would be MOVE in NKRL, one of the seven primitives used as conceptual predicates in this language. Given the several thousands of NL verbs that exist in English, it is well evident that an exhaustive and authoritative list of CGs *canonical graphs*, roughly equivalent to the (about 150) NKRL's templates, cannot be established for reasons of combinatorial explosion. From a practical point of view, this means that the CGs developed in the context of a given application are *possibly different and logically conflicting* with those utilized in a different application. This is in contrast with a fundamental characteristic of NKRL, where the catalogue of basic templates (HTemp, the hierarchy of templates) is, in practice, part and parcel of the definition of the language securing, therefore, the reproduction and/or the sharing of the results.

6 Conclusion

In this paper, we have described in some depth a knowledge representation proposal intended to deal, in a complete and coherent way, with the multifaceted notion of "role". In particular:

- The proposal concerns the adoption of the neat differentiation between two conceptually separate (but practically coordinated from an operational point of view) understandings of the notions of roles:
  - Semantic roles are interpreted as *non-sortal concepts* to be included within a *static* – i.e., that can be defined *a priori* – standard (binary) ontology. Non-sortal means that semantic roles cannot be endowed with direct instances. They are used to formalize the *transient properties*, like student, passenger or customer, of living (or assimilated to living) entities.
  - Functional roles are represented as *primitives* – like subject, object, beneficiary, modality – that specify the *functions* the entities involved in a *dynamic* spatio-temporal characterized situation like "Today, Mary has taken the flight Air France 3628 to New York" have in the context of this situation. For example, Mary is the "subject" of the *elementary event* represented by the flight; elementary events (characterized by a conceptual predicate) and functional roles necessarily co-occur. These events must be represented according to an *n*-ary formalism to take fully into account their context and the complexity of the mutual interrelationships (formalized by the functional roles) among their involved entities.
- The separation between semantic and functional roles is, of course, valid in general, and *not solely relevant to the use of a particular implementation formalism*. In this paper, we have used NKRL (Narrative Knowledge Representation Language) as support for illustrating the two types of roles through practical examples. This choice is justified by these reasons:
  - The knowledge representation language used in NKRL is particularly *expressive*, given that it includes a collection of formal tools able to take into account *a large set of information structures*, from temporal data to connectivity phenomena (coordination, subordination, indirect speech etc.).
  - In particular, NKRL is endowed with formal tools proper to the representation of both binary and *n*-ary data. These can be used, therefore, to model correctly both the *static* binary standard concepts and the *n*-ary structures that are needed to represent the *dynamic* elementary events.
- A key aspect of the paper concerns the emphasis given to the importance of the *relational aspects* of the notion of role and, as a consequence, to the importance of the notion of functional role. This importance has been concretely demonstrated by *the use of functional roles and the associated conceptual structures* to offer a complete, structured and logically coherent solution to the counting problem. The solution has been implemented

without any need for introducing extra *ad hoc* tools to deal with this problem, in addition to those already natively proposed by NKRL and based on the semantic/functional differentiation.

On a more practical level, we can note that our understanding of the notion of role is not only at the root of a conceptual representation language, NKRL, deeply grounded on the distinction between semantic and functional roles, but also of a *fully implemented computational environment* that puts into practice concretely this separation. This environment exists in two versions, an SQL-based version and a (lighter) file-oriented one. Technical details about the two (Java) implementations can be found in, e.g., Zarri (2009: Appendix A).

Eventually, with respect to the *one and only syndrome* several times mentioned in this paper, we can note that NKRL (as recalled also in the previous paragraphs) makes use of a large variety of modelling tools, where each of them intends to supply *the best fitting solution* to take into account a *specific aspect* of complex, real-world situations.

Acknowledgement

I gratefully acknowledge the two unknown referees who reviewed a first version of this paper for their useful suggestions that helped me to improve considerably the technical quality and the legibility of my text.

References

Almeida JPA, Guizzardi G (2007) On the Foundation for Roles in RM-ODP, Contributions from Conceptual Modelling. In: Workshops Proceedings of the 11th International IEEE Enterprise Distributed Object Computing Conference. Computer Society Press, Los Alamitos (CA), pp. 205-215

Bal M (1997) Narratology: Introduction to the Theory of Narrative, 2nd edition. University Press, Toronto.

Barlatier P, Dapoigny R (2012) A Type-Theoretical Approach for Ontologies: The Case of Roles. Applied Ontology 7: 311-356

Bechhofer S, van Harmelen F, Hendler J, Horrocks I, McGuinness DL, Patel-Schneider PF, Stein LA (eds) (2004) OWL Web Ontology Language Reference, W3C Recommendation 10 February 2004 (http://www.w3.org/TR/owl-ref/, accessed 2 July 2017)

Berners-Lee T (2006) Linked Data - Design Issues, W3C Note 27 July 2006 (http://www.w3.org/DesignIssues/LinkedData.html, accessed 2 July 2017)

Bertino E, Catania B, Zarri GP (2001) Intelligent Database Systems. Addison-Wesley and ACM Press, London

Blomqvist E (2009) Semi-automatic Ontology Construction Based on Patterns (PhD thesis). School of Engineering of the Linköping University, Linköping

Bruce B (1975) Case Systems for Natural Language. Artificial Intelligence 6: 327-360

Cáceres C, Fernández A, Ossowski S, Vasirani M (2006) Agent-Based Service Discovery for Healthcare: An Organizational Approach. IEEE Intelligent Systems 21(6): 11-20

Ceccato S (ed) (1961) Linguistic Analysis and Programming for Mechanical Translation (Technical report RADC-TR-60-18). Feltrinelli, Milano

Ceccato S (1964) Automatic Translation of Languages. Information Storage and Retrieval 2: 105-158.

Chaudhri VK, Farquhar A, Fikes R, Karp PD, Rice JP (1998). OKBC: A Programmatic Foundation for Knowledge Base Interoperability. In: Mostow J, Rich C (eds) Proceedings of the 1998 National Conference on Artificial Intelligence, AAAI/98. MIT Press/AAAI Press, Cambridge (MA), pp. 600-607

Clocksin WF, Mellish CS (1981) Programming in PROLOG. Springer-Verlag, Berlin

Davidson D (1967) The Logical Form of Action Sentences. In: Rescher N (ed) The Logic of Decision and Action. University Press, Pittsburg, pp. 81-95

Delugach HS (ed) (2007) Common Logic (CL): A Framework for a Family of Logic-Based Languages (ISO/IEC 24707:2007). International Organization for Standardization (ISO), Geneva

Dowty D (1989) On the Semantic Content of the Notion of 'Thematic Role'. In: Chierchia G, Partee BH, Turner R, (eds) Properties, Types, and Meaning, Volume 2: Semantic Issues. Kluwer, Dordrecht, pp. 69-130

Dyer MG (1983) In-Depth Understanding. The MIT Press, Cambridge (MA)

Enderton HB (2012) Second-order and Higher-order Logic. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy, Fall 2012 Edition (http://plato.stanford.edu/archives/fall2012/entries/logic-higher-order/, accessed 2 July 2017)

Falbo RA, Barcellos, MP, Nardi JC, Guizzardi, G (2013) Organizing Ontology Design Patterns as Ontology Pattern Languages. In: Cimiano P, Corcho O, Presutti V, Hollink L, Rudolph S (eds) The Semantic Web: Semantics and Big Data - Proceedings of the 10th ESWC International Conference (LNCS 7882). Springer, Berlin, pp. 61-75

Fan J, Barker K, Clark P, Porter B (2001) Representing Roles and Purpose. In: Handschuh S, Dieng R, Staab S (eds) Proceedings of the K-CAP 2001 Workshop on Knowledge Markup and Semantic Annotation. ACM, New York, pp. 38-43

Fillmore CJ (1968) The Case for Case. In: Bach E, Harms RT (eds) Universals in Linguistic Theory. Holt, Rinehart and Winston, New York, pp. 1-88

Finlayson MA, Gervás P, Mueller E, Narayanan S, Winston P (eds) (2010) Computational Models of Narratives – Papers from the AAAI 2010 Fall Symposium (Technical Report FS-10-04). AAAI Press, Menlo Park (CA)

Freeman MW (1982) The QUA Link. In: Schmolze JG, Brachman RJ (eds) Proceedings of the 1981 KL-ONE Workshop (BBN Report n. 4842). Bolt Beranek and Newman, Cambridge (MA), pp. 54-64

Fukazawa Y, Naganuma T, Fujii K, Kurakake S (2006) Construction and Use of Role-ontology for Task-based Service Navigation System. In: Cruz I, Decker S, Allemang D, Preist C, Schwabe D, Mika P, Uschold M, Aroyo L (eds) Proceedings of the 5th International Semantic Web Conference, ISWC 2006 (LNCS vol. 4273). Springer, Berlin, pp. 806-819

Garshol LM, Moore G (eds) (2008) Topic Maps – Data Model (ISO/IEC JTC1/SC34 Document) (http://www.isotopicmaps.org/sam/sam-model/2008-06-03/, accessed 2 July 2017)

Genilloud G, Wegmann A (2000) A New Definition for the Concept of Role, and Why it Makes Sense. In: Baclawski K, Kilov H (eds) Proceedings of the 2000 Workshop on Behavioral Semantics - co-located with the OOPSLA 2000 Conference on Object-oriented Programming, Systems, Languages, and Applications (http://infoscience.epfl.ch/record/269/files/GenilloudW00a.pdf, accessed 2 July 2017)

Goldstein I, Papert S (1977) Artificial Intelligence, Language, and the Study of Knowledge. Cognitive Science 1: 84-123

Guarino N (1992) Concepts, Attributes and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases. Data and Knowledge Engineering 8: 249-261

Guarino N (2009) The Ontological Level: Revisiting 30 Years of Knowledge Representation. In: Borgida A, Chaudhri V, Giorgini P, Yu E (eds) Conceptual Modelling: Foundations and Applications. Essays in Honor of John Mylopoulos (LNCS vol. 5600). Springer, Berlin, pp. 52-67

Guarino N, Welty C (2000) A Formal Ontology of Properties. In: Dieng R, Corby O (eds) Knowledge Acquisition, Modeling, and Management – Proceedings of EKAW'2000 (LNCS vol. 1937). Springer, Berlin, pp. 97-112

Guizzardi G (2005) Ontological Foundations for Structural Conceptual Models (PhD Thesis). Centre for Telematics and Information Technology of the University of Twente, Enschede

Guizzardi G (2006) Agent Roles, Qua Individuals and The Counting Problem. In: Garcia A, Choren R, Lucena C, Giorgini P, Holvoet T, Romanovsky A (eds) Software Engineering for Multi-Agent Systems IV, Research Issues and Practical Applications (LNCS vol. 3914). Springer, Berlin, pp. 143-160

Gupta A (1980) The Logic of Common Nouns: An Investigation in Quantified Modal Logic. Yale University Press, New Haven (CT)

Heath T, Bizer C (2011) Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool, San Rafael (CA)

Hitzler P, Krötzsch M, Parsia B, Patel-Schneider PF, Rudolph S (eds) (2009) OWL 2 Web Ontology Language Primer, W3C Recommendation 27 October 2009 (http://www.w3.org/TR/owl2-primer/, accessed 2 July 2017)

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

Kim J (1996) Events as Property Exemplifications. In: Casati R, Varzi AC (eds) Events (International Research Library of Philosophy, 15). Dartmouth Publishing, Aldershot, pp. 117-135

Kolodner JL (1984) Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model. Lawrence Erlbaum, Hillsdale (NJ)

Kozaki K, Sunagawa E, Kitamura Y, Mizoguchi, R (2007) Role Representation Model Using OWL and SWRL. In: Proceedings of the 2nd Workshop on Roles and Relationships in Object Oriented Programming, Multiagent Systems and Ontologies – co-located with ECOOP 2007 (http://www.ei.sanken.osaka-u.ac.jp/pub/kozaki/Role07kozaki.pdf, accessed 2 July 2017)

Lenat DB, Guha RV, Pittman K, Pratt D, Shepherd M (1990) CYC: Toward Programs With Common Sense. Communications of the ACM 33(8): 30-49

Lenat DB, Witbrock M, Baxter D, Blackstone E, Deaton C, Schneider D, Scott J, Shepard B (2010) Harnessing Cyc to Answer Clinical Researchers' Ad Hoc Queries. AI Magazine 31 (3): 13-32

Liu H, Singh P (2004) ConceptNet - A Practical Commonsense Reasoning Tool-Kit. BT Technology Journal 22(4): 211-226

Loebe F (2007) Abstract vs. Social Roles - Towards a General Theoretical Account of Roles. Journal of Applied Ontology 2: 127-158

Mani I, Pustejovsky J (2004) Temporal Discourse Models for Narrative Structure. In: Webber B, Byron D (eds) Proceedings of the ACL Workshop on Discourse Annotation. Association for Computational Linguistics (ACL), Stroudsburg (PA), pp. 57-64

Masolo C, Borgo S, Gangemi A, Guarino N, Oltramari A (2003) WonderWeb Deliverable D18 (IST Project 2001-33052). Laboratory for Applied Ontology, Trento (http://www.loa.istc.cnr.it/old/Papers/D18.pdf, accessed 2 July 2017)

Masolo C, Guizzardi G, Vieu L, Bottazzi E, Ferrario R (2005) Relational Roles and Qua-Individuals. In: Boella G, Odell J, van der Torre L, Verhagen H (eds) Roles, an Interdisciplinary Perspective - Papers from the AAAI 2005 Fall Symposium (Technical Report FS-05-08). AAAI Press, Menlo Park (CA), pp. 103-112

Masolo C, Vieu L, Bottazzi E, Catenacci C, Ferrario R, Gangemi A, Guarino N (2004) Social Roles and their Descriptions. In: Dubois D, Welty C, Williams MA (eds) Proceedings of the 9th International Conference on Knowledge Representation and Reasoning, KR2004. AAAI Press, Menlo Park (CA), pp. 267-277

Masolo C, Vieu L, Kitamura Y, Kozaki K, Mizoguchi R (2011) The Counting Problem in the Light of Role Kinds. In: Davis E, Doherty P, Erdem E (eds) Proceedings of Commonsense-2011, the Tenth International Symposium on Logical Formalization on Commonsense Reasoning (AAAI Spring Symposium, Technical Report SS-11-06). AAAI Press, Menlo Park (CA), pp. 76–82

Matsuyoshi S, Eguchi M, Sao C, Murakami K, Inui K, Matsumoto Y (2010) Annotating Event Mentions in Text with Modality, Focus, and Source Information. In: Calzolari N, Choukri K, Maegaard B, Mariani J, Odijk J, Piperidis S, Rosner M, Tapias D (eds) Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010. European Language Resources Association (ELRA), Paris, pp. 1456-1463

Mizoguchi R, Galton A, Kitamura Y, Kozaki K (2015) Families of roles: A New Theory of Occurrent-Dependent Roles. Journal of Applied Ontology 10: 367-399

Mizoguchi R, Kozaki K, Kitamura Y (2012) Ontological Analyses of Roles. In: Ganzha M, Maciaszek LA, Paprzycki M (eds) Proceedings of the Federated Conference on Computer Science and Information Systems, FedCSIS 2012. IEEE Xplore, New York, pp. 489-496

Mizoguchi R, Sunagawa E, Kozaki K, Kitamura Y (2007) A Model of Roles within an Ontology Development Tool: Hozo. Journal of Applied Ontology 2: 159-179

Mueller ET (2013) Computational Models of Narrative. Sprache und Datenverarbeitung - International Journal for Language Data Processing (SDV) 37(1-2): 11-39

Niles I, Pease A (2001) Towards a Standard Upper Ontology. In: Welty C, Smith B (eds) Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, FOIS-2001. IOS Press, Amsterdam, pp. 2-9

Noy FN, Fergerson RW, Musen MA (2000) The Knowledge Model of Protégé-2000: Combining Interoperability and Flexibility. In: Dieng R, Corby O (eds) Knowledge Acquisition, Modeling, and Management – Proceedings of EKAW'2000 (LNCS vol. 1937). Berlin, Springer, pp. 17-32

Noy FN, Rector A (eds) Hayes P, Welty C (contributors) (2006) Defining N-ary Relations on the Semantic Web, W3C Working Group Note 12 April 2006 (http://www.w3.org/TR/2006/NOTE-swbp-n-aryRelations-20060412/, accessed 2 July 2017)

Palmer M, Gildea G, Xue N (2010) Semantic Role Labeling. Morgan and Claypool, San Rafael (CA)

Pepper S (2010) Topic Maps. In: Bates MJ (editor-in-chief) Encyclopedia of Library and Information Sciences, 3rd edition. Taylor & Francis, Abingdon, pp. 5247-5260

Pustejovsky J, Ingria R, Saurí R, Castaño J, Littman J, Gaizauskas R, Setzer A, Katz G Mani I (2005) The Specification Language TimeML. In: Mani I, Pustejovsky J, Gaizauskas R (eds) The Language of Time: A Reader. University Press, Oxford, pp. 545-558

Quillian MR (1966) Semantic Memory (PhD dissertation). Carnegie Institute of Technology, Pittsburgh. An abridged version has been published in: Minsky M (ed) (1968) Semantic Information Processing. The MIT Press, Cambridge (MA), pp. 227-270

Reynolds D (ed) (2014) The Organization Ontology, W3C Recommendation 16 January 2014 (http://www.w3.org/TR/vocab-org/, accessed 2 July 2017)

Rosner M, Somers H (1980) Case in Linguistics and Cognitive Science (ISSCO Working Paper 40). ISSCO, Geneva

Schank RC (1973) Identification of Conceptualizations Underlying Natural Language. In: Schank RC, Colby KM, Computer Models of Thought and Language. W.H. Freeman and Co., San Francisco, pp. 187-247.

Schank RC, Abelson RP (1977) Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. Lawrence Erlbaum, Oxford

Shapiro SC (1979) The SNePS Semantic Network Processing System. In: Findler NV (ed) Associative Networks: Representation and Use of Knowledge by Computers. Academic Press, New York, pp. 179-203

Shotton D, Peroni S (2013) PRO, the Publishing Roles Ontology - Creative Commons, 1.5.3 Version (http://www.sparontologies.net/ontologies/pro/source.html, accessed 2 July 2017)

Soon KH (2013) Representing Roles in Formalizing Domain Ontology for Land Administration. In: Proceedings of LADM2013: 5th FIG International Land Administration Domain Model Workshop. International Federation of Surveyors (FIG), Copenhagen, pp. 203-222

Sowa JF (1984) Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley, Reading (MA)

Sowa JF (1999) Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks Cole Publishing Co, Pacific Grove (CA)

Spärck Jones K, Boguraev B (1987) A Note on a Study of Cases. Computational Linguistics 13: 65-68.

Spear AD (2006) Ontology for the Twenty First Century: An Introduction with Recommendations. Institute for Formal Ontology and Medical Information Science, Saarbrücken

Speer R, Havasi C (2012) Representing General Relational Knowledge in ConceptNet 5. In: Proceedings of the 8th Conference on Language Resources and Evaluation (LREC'12). European Language Resources Association (ELRA), Paris, pp. 3679-3686

Steimann F (2000) On the Representation of Roles in Object-Oriented and Conceptual Modeling. Data and Knowledge Engineering 35: 83-106

Stojanovic N, Abecker A, Etzion O, Paschke A (eds) (2009) Proceedings of the 2009 AAAI Spring Symposium on Intelligent Complex Event Processing (Technical Report SS-09-05). AAAI Press, Menlo Park (CA)

Swoyer C, Orilia F (2014) Properties. In: Zalta EN (ed) The Stanford Encyclopedia of Philosophy, Fall 2014 Edition (http://plato.stanford.edu/archives/fall2014/entries/properties/, accessed 2 July 2017)

Trame J, Kessler C, Kuhn W (2013) Linked Data and Time - Modeling Researcher Life Lines by Events. In: Tenbrink T, Stell J, Galton A, Wood Z (eds) Proceedings of the 11th International Conference on Spatial Information Theory, COSIT 2013 (LNCS vol. 8116). Springer, Berlin, pp. 205-223

Van Valin RD Jr (1999) Generalized Semantic Roles and the Syntax-Semantics Interface. In: Corblin F, Dobrovie-Sorin C, Marandin JM (eds) Empirical Issues in Formal Syntax and Semantics 2. Thesus, The Hague, pp. 373-389

Vieu L, Borgo S, Masolo C (2008) Artefacts and Roles: Modeling Strategies in a Multiplicative Ontology. In: Eschenbach C, Grüninger M (eds) Proceedings of the 5th International Conference on Formal Ontology in Information Systems - FOIS 2008. IOS Press, Amsterdam, pp. 121-134

Westermann U, Jain R (2006) Events in Multimedia Electronic Chronicles (E-Chronicles). International Journal on Semantic Web and Information Systems (IJSWIS) 2(2): 1-23

Wieringa R, de Jonge W, Spruit P (1995) Using Dynamic Classes and Role Classes to Model Object Migration. Theory and Practice of Object Systems 1: 61-83

Woods WA (1975) What's in a Link: Foundations for Semantic Networks. In: Bobrow DG, Collins AM (eds) Representation and Understanding: Studies in Cognitive Science. Academic Press, New York, pp. 35-82

Zarri GP (1997) NKRL, a Knowledge Representation Tool for Encoding the 'Meaning' of Complex Narrative Texts. Natural Language Engineering - Special Issue on Knowledge Representation for Natural Language Processing in Implemented Systems 3: 231-253.

Zarri GP (1998) Representation of Temporal Knowledge in Events: The Formalism, and Its Potential for Legal Narratives. In: Martino AA, Nissan E (eds) Special Issue on Formal Models of Legal Time: Law, Computers and Artificial Intelligence, Information & Communications Technology Law 7: 213-241

Zarri GP (2005) Integrating the Two Main Inference Modes of NKRL, Transformations and Hypotheses. Journal on Data Semantics (JoDS) 4: 304-340

Zarri GP (2009) Representation and Management of Narrative Information, Theoretical Principles and Implementation. Springer, London (https://link.springer.com/book/10.1007%2F978-1-84800-078-0, accessed 2 July 2017)

Zarri GP (2011a) Differentiating Between "Functional" and "Semantic" Roles in a High-Level Conceptual Data Modeling Language. In: Murray RC, McCarthy PM (eds) Proceedings of the 24th International Florida AI Research Society Conference, FLAIRS-24. AAAI Press, Menlo Park (CA), pp. 75-80

Zarri GP (2011b) Knowledge Representation and Inference Techniques to Improve the Management of Gas and Oil Facilities. Knowledge-Based Systems (KNOSYS) 24: 989-1003

Zarri GP (2013) Advanced Computational Reasoning Based on the NKRL Conceptual Model. Expert Systems With Applications (ESWA) 40: 2872-2888

Zarri GP (2014a) Conceptual and Content-based Annotation of (Multimedia) Documents. Multimedia Tools and Applications (MTAP) 72: 2359-2391

Zarri GP (2014b) Sentiments Analysis at Conceptual Level Making Use of the Narrative Knowledge Representation Language. In: Hussain A, Cambria E, Schuller B, Howard N (eds) Special Issue on Affective Neural Networks and Cognitive Learning Systems for Big Data Analysis, Neural Networks (NEUNET) 58: 82–97

Zarri GP (2015) A Structured and In-Depth Representation of the Semantic Content of Elementary and Complex Events. International Journal of Metadata, Semantics and Ontologies (IJMSO) 10: 12-27

Web References

Web ref.1: http://www.cyc.com/platform/opencyc (accessed 2 July 2017)
Web ref.2: http://www.ilc.cnr.it/EAGLES96/EAGLESLE.PDF (accessed 2 July 2017)
Web ref.3: http://ontologydesignpatterns.org/wiki/Ontology:DOLCE%2BDnS_Ultralite (accessed 2 July 2017)
Web ref.4: http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/ (accessed 2 July 2017)
Web ref.5: http://ontologydesignpatterns.org/wiki/Ontology_Design_Patterns_._org_%28ODP%29 (accessed 2 July 2017)
Web ref.6: http://schema.org/ (accessed 2 July 2017)
Web ref.7: https://schema.org/docs/full.html (accessed 2 July 2017)
Web ref.8: https://lists.w3.org/Archives/Public/public-vocabs/2014Oct/0170.html (accessed 2 July 2017)
Web ref.9: http://conceptnet5.media.mit.edu/ (accessed 2 July 2017)