Neural Networks 58 (2014) 82–97

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2014 Special Issue

Sentiments analysis at conceptual level making use of the Narrative Knowledge Representation Language

Gian Piero Zarri *

Sorbonne University – STIH Laboratory, 1, rue Victor Cousin – 75005 Paris, France

ARTICLE INFO

Article history:
Available online 2 June 2014

Keywords:
Behaviors
Representation of deep sentiments
Narratives
Standard ontologies
Ontologies of events
Querying and inference techniques

ABSTRACT

This paper illustrates some of the knowledge representation structures and inference procedures proper to a high-level, fully implemented conceptual language, NKRL (Narrative Knowledge Representation Language). The aim is to show how these tools can be used to deal, in a sentiment analysis/opinion mining context, with some common types of human (and non-human) "behaviors". These behaviors correspond, in particular, to the concrete, mutual relationships among human and non-human characters that can be expressed under the form of non-fictional and real-time "narratives" (i.e., as logically and temporally structured sequences of "elementary events").

1. Introduction

"Sentiment analysis" (or "opinion mining") concerns all the possible computer-based applications that try to identify and extract "subjective information" (opinions, beliefs, emotional states and views about specific entities) from source materials, usually represented in textual form (Feldman, 2013; Westerski, 2007). Related disciplines are "behavior computing" – or "behavioral informatics" (Cao and Yu, 2012) – and "affective computing" (Ahn, 2010). Most common, practical sentiments analysis applications are in the area of reviews of consumer products and services.

The research tools used in the sentiment analysis domain consist normally of computational linguistics and text mining techniques that perform some sort of "surface" analysis of the original sources in order, e.g., to determine the "negative/positive polarity" of words or sentences, recognizing the presence of words or expressions within specific sentiment lexica, detecting sentences that contain comparative opinions, etc. In this paper, we suggest that these surface techniques, often strongly statistically-oriented, could be usefully complemented by "deep" conceptual analysis tools aiming at describing, in sufficient detail, the behaviors (according to the most general meaning of this term) and the mutual relationships of the (human and non-human) characters that appear in the original natural language documents. To this end, this paper focuses on the conceptual representation tools proper to a (wholly implemented) knowledge representation language and computer system environment, NKRL, the Narrative Knowledge Representation Language (Zarri, 2009).

In a nutshell, the term *narrative* denotes a general unifying framework used for relating real-life or fictional stories (novels, tales…) involving the *common relationships* between concrete or imaginary characters. Narratives deal then, among other things, with those opinions, beliefs, emotional states and viewpoints about specific entities that, as already stated, represent the *basic, raw material* used to perform the sentiment analysis operations. Narratives are normally conveyed by NL supports as, in a *non-fictional* context, news stories, corporate memory documents (memos, reports, minutes…), normative and legal texts, medical records, etc. However, they can also be represented by multimedia documents like audio records, surveillance videos, actuality photos for newspapers and magazines, etc. A photo representing President Obama addressing the Congress, or a short video showing three nice girls chattering on a beach, must be considered as "narrative" documents even if they are not, of course, NL documents. A narrative is a *highly-dynamic entity*, since it can be synthetically defined as a *sequence of logically structured and temporally and spatially bounded "elementary events" (a non-linear 'stream' of elementary events)*. An "elementary event" corresponds in turn to the conceptual representation of the bundle of mutual relationships among characters associated with a *single* "generalized predicate" isolated within the natural language formulation of the whole stream. Generalized predicates correspond not only to the usual tensed/untensed "verbs", but also to "adjectives" ("…worth several dollars…", "…a dormant volcano…"), nouns ("…Jane's amble along the park…", "…a possible attack…), etc., when they have a predicative function.

* Tel.: +33 1 40463288.
E-mail addresses: zarri@noos.fr, gian-piero.zarri@u-pec.fr.

http://dx.doi.org/10.1016/j.neunet.2014.05.010
0893-6080/

To justify the use within the sentiment analysis/opinion mining domain of formal tools created for the analysis of "narrative" documents, let us examine briefly other "*conceptual*" – as opposed to pure statistical – approaches used in this domain. For example, the so-called "*sentiment (or opinion) lexica*" can be defined in general as lists of *words and expressions* used to denote people's subjective feelings and sentiments/opinions ("negative" or "positive" *prior polarities*). The term "expressions" is used here to denote not just individual words, but also phrases and idioms. These lexica can be built up according to three main ways, a manual approach (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011), a corpus-based approach that relies on the detection of syntactic patterns in large corpora (Ding, Liu, & Yu, 2008; Kaji & Kitsuregawa, 2007) and a dictionary-based approach. Lexica pertaining to this last category are often developed by making use of WordNet's synsets and hierarchies to acquire opinion words, see in this context, e.g., WordNet-Affect (Strapparava & Valitutti, 2004) and SentiWordNet (Esuli & Sebastiani, 2006). WordNet-Affect was developed through the selection and labeling, using the terms included in a specific hierarchy of "affective domain labels", of the WordNet synsets representing affective concepts. SentiWordNet is a version of WordNet where the independent values "positive", "negative", and "objective" are associated with 117,660 WordNet's synsets. Each of the three values ranges from 0.0 to 1.0, and their sum is 1.0 for every synset.

In a "sentiment lexica" context, one of the most well-known and advanced approaches is represented by SenticNet. This system exists in three versions of increasing complexity, SenticNet 1 (Cambria, Speer, Havasi, & Hussain, 2010), SenticNet 2 (Cambria, Havasi, & Hussain, 2012) and SenticNet 3 (Cambria, Olsher, & Rajagopal, 2014). Partially inspired from WordNet-Affect and SentiWordNet, SenticNet makes use of the so-called "*sentic computing*" *approach*. This is a new paradigm that exploits both AI and Semantic Web techniques to recognize, interpret, and process natural language opinions going beyond a simple "syntactic" strategy. In its version 2 for example, it provides the semantics and the "*sentic information*" – i.e., the *cognitive* and *affective* information – that concern over 14,000 concepts. Unlike SentiWordNet, SenticNet discards concepts with neutral or almost neutral polarity, i.e., concepts with polarity magnitude close to zero. Moreover, while SentiWordNet stores three values for each synset, SenticNet associates each concept $\boldsymbol{c}$ with just one value $\boldsymbol{p_c}$, i.e., a float in the range $[-1, 1]$ representing its polarity. This choice allows SenticNet to avoid redundancy and facilitates its representation as a (ConceptNet, see below) semantic network. In SenticNet, eventually, concepts like **make good impression**, **look attractive**, **show appreciation** or **good deal** are likely to have a $\boldsymbol{p_c}$ very close to 1 while concepts such as **being fired**, **leave behind** or **lose control** are likely to have $\boldsymbol{p_c} \approx -1$ (Cambria et al., 2010: 16). An important, common characteristic of the three SenticNet versions concerns the fact that their 'basic knowledge' derives from ConceptNet (Liu & Singh, 2004; Speer & Havasi, 2012), a semantic network built up from *nodes* representing *concepts* in the form of *words or short phrases in natural language* and from *labeled relationships between them*. The relationships (21, including the standard **IsA**) are in the form of, e.g., **CreatedBy**, **PartOf**, **UsedFor**, **PrerequisiteOf**, **DefinedAs**, **LocatedNear**. Thus, ConceptNet knowledge is mainly associated with *general compound concepts* instead of single words/concepts. The compound concepts are represented in semi-structured English by composing, using the labeled relationships, a verb/concept with a noun phrase/concept or a prepositional phrase/concept. (Recursive) compound concepts can then be, e.g., "**[*wake up in the morning*] PrerequisiteOf [*eat breakfast*]**", "**[*kitchen table*] UsedFor [*eat breakfast*]**", "**[*chair*] LocatedNear [*kitchen table*]**", etc. (Liu & Singh, 2004: 213).

Independently from the formal semantic/syntactic details, the *knowledge* included in all the systems mentioned above has in common the fact of being, basically, a sort of *terminological/definitional knowledge*. It denotes, then, some *stable, self-contained, a priori and basic notions/concepts* that can be considered, *at least in the short term*, as '*a-temporal*' (or '*static*') and '*universal*'. This means that their definitions *are not subject to change*, at least within the framework of a given application, even if they can evolve *in the long term* as a consequence, e.g., of the progress of our knowledge or of criticisms/comparisons with different approaches. These static notions can be very general, see concepts like **human being**, **color** or **chair** that are proper to several application domains, or linked to more specific contexts as **make person happy**, **feel guilty** or **shed tears** in a sentiment analysis environment.[1]

The self-contained and stable character of this terminological/definitional knowledge (where, as stated above, the temporal phenomena can be ignored) justifies the use of a relatively simple *formal model* for its conceptual representation/definition. This can be limited to the description of some main properties—sometimes, only the use of the genus/species **IsA** relationships is actually required. This formal model can then correspond to the usual binary one, where properties are simply expressed as a *binary* (i.e., accepting only two arguments) relationship linking two individuals or an individual and a value. And this independently from the fact that these binary relationships are organized into, e.g., frame format as in the original Protégé software (Noy, Fergerson, & Musen, 2000) or take the form of a set of "property" statements used to define a "class" (a "concept") in some W3C language. In a sentiment analysis/opinion mining framework we can note that, accordingly, WordNet 3 is now represented in (*binary*) RDF/W3C format; RDF is also used in a ConceptNet 5 environment and to encode the nodes of the SenticNet network.

In the context of the "*narrative information*" analysis evoked above and of similar applications, the *main* knowledge to be dealt with corresponds, on the contrary, to a sort of particularly *complex and "structured"* information. This type of knowledge denotes, in fact, the *dynamic, interpersonal, often accidental and unpredictable, spatio-temporal characterized behavior* proper to *specific subsets* of the *terminological/definitional entities* examined above. Examples of this sort of dynamic/structured knowledge that can be of interest in a sentiment analysis/opinion mining environment correspond, e.g., to the description of "*elementary events*" in the style of "On November 17, 2003, in an unspecified location in Afghanistan, an armed group of people shot a woman dead", "Yesterday, John gave a book to Mary for her birthday", "Peter has recently bought his first iPhone in the Carrousel Apple Store of Paris", "On November 20, 1999, in Sulu province, the family of the kidnapped journalist was asked for a ransom", "On August 8, 2012, at Beta Bank's premises, Mary Collins fired John Smith", "Tom returned his new Ultrabook yesterday", etc. In a "structured/dynamic" context, then, some *static, terminological/definitional entities* ("John", "Mary", "woman", "present"',

1 We can note that this terminological/definitional knowledge coincides largely with the "*common knowledge*" as defined, e.g., in Cambria, Olsher, et al. (2014) and Cambria and White (2014). More precisely, Cambria and his colleagues make a distinction between "*common knowledge*" and "*common-sense knowledge*". The first corresponds to general knowledge about the world, e.g., "a chair is a type of furniture". On the other hand, common-sense knowledge denotes "…accepted things that people normally know about the world but which are usually left unstated in discourse, e.g., that *things fall downwards* (*and not upwards*) and *people smile when they are happy*" (Cambria & White, 2014: 51). The two types can be both classified as static, a priori, a-temporal knowledge as the terminological/definitional knowledge introduced above. In an NKRL context we prefer, however, to think about the common-sense knowledge as that "*operational knowledge*" definitely needed for setting up useful *inference rules*, see Section 4 below.

"journalist", "ransom", "Ultrabook", etc.), *a priori and independently defined*, are involved within structured networks of *mutual and dynamic relationships*. These networks are modeled by broad conceptual structures, spatio-temporally constrained, which correspond to scenarios/events/situations/circumstances, etc. The necessity of making use of:

- "*conceptual predicates*" that correspond to surface verbs like "give", "buy", "ask", "fire", etc. in the previous examples (or to those "generalized predicates" mentioned above) to specify the *basic type* of scenario, situation, etc. described in each elementary event;
- the notion of "*functional role*" (Zarri, 2011) to denote the logical and semantic function of the terminological entities involved in the different events: in "John gave a book to Mary...", the "individual" (instance of a concept) **JOHN_** is the **SUBJ**(ect) of the action of "giving", **BOOK_1** the **OBJ**(ect) and **MARY_** the **BEN**(e)**F**(iciary);
- an adequate, specific formalism to denote the "*temporal and location information*" and its relationships with the global representation of the whole elementary event;
- a way of "*reifying*" this global representation to be able to use it within larger, complex scenarios of the type "an armed group of people shot a woman dead *because* she was an employee of UNHCR (the UN Agency for the Refugees)";

makes it difficult to use the standard binary approach to represent correctly (and effectively) this sort of structured/dynamic information—which is surely of interest within a sentiment analysis/opinion mining framework.

As we will see in more detail in Section 2.1, a complex type of *n*-ary representation like that used in an NKRL context is then *required to represent correctly this structured/dynamic knowledge*. Note, moreover, that the knowledge representation issues addressed above are not only proper to a specific "narrative" context. They concern, in general, other up-to-date and economically important Computer Science applications like the representation and management of many-sided events, of crisis and terrorism situations, of corporate memory records, cultural heritage, eChronicles, the encoding of complex temporal information, etc. Note also, moreover, that the partition between "terminological/definitional" and "structured/dynamic" knowledge has been evoked several times, implicitly or explicitly, in the ontological and cognitive literature. However, from a '*concrete*' representation point of view, this difference has been frequently neglected and a *unique*, often very poor, (binary) *knowledge representation schema* has been used for describing the two sorts of knowledge.

In the following, Section 2 will supply first a short overview of the main representational features of NKRL referring the reader, for more information, to the rich literature on the subject. Section 3 represents the core of the paper, and consists of a series of examples intended to emphasize those *behavioral* properties of NKRL that could prove particularly useful in the context of a full understanding of the sentiment analysis domain. This section will be concluded by the analysis of a detailed example of NKRL encoding concerning a complex sentiment analysis scenario. Section 4 supplies some information about NKRL's inference procedures. Section 5 will mention the related work; Section 6 is the "Conclusion".

2. Basic facts about NKRL

2.1. Ontology of concepts and ontology of events

NKRL innovates with respect to the current ontological paradigms by *adding* an (*n*-ary) "*ontology of elementary events*" to the usual (binary) "*ontology of concepts*". In an NKRL context, this last is called HClass, hierarchy of classes, and includes at present (May 2014) more than 7500 concepts. The ontology of elementary events is *a new sort of hierarchical organization where the nodes correspond to n-ary structures called "templates", represented schematically according to the syntax of Eq. (1) below*. This ontology is then denoted as HTemp (hierarchy of templates) in NKRL. Templates can be conceived as the canonical, *formal representation of generic classes of elementary events* like "evaluating an artefact", "having a specific attitude towards someone/something", "producing a service", "threatening someone with violence", etc. The HTemp hierarchy can then be equated, in a sense, to a *dynamic* generalization of the usual "*sentiment lexica*" mentioned in the previous section.

$$(\mathbf{L}_i(\mathbf{P}_j(\mathbf{R}_1\ \mathbf{a}_1)(\mathbf{R}_2\ \mathbf{a}_2)\cdots(\mathbf{R}_n\ \mathbf{a}_n))). \qquad (1)$$

In Eq. (1), $\mathbf{L}_i$ is the "*symbolic label*" identifying the particular *n-ary* structure corresponding to a specific template, $\mathbf{P}_j$ is a "*conceptual predicate*", $\mathbf{R}_k$ is a generic "*functional role*" and $\mathbf{a}_k$ the corresponding "*predicate arguments*". Note that the conceptual predicates are 'translations', at deep/conceptual level, of the surface/linguistic level "generalized predicates" like "give", "buy", "ask", "fire", but also "amble", "dormant", "worth", etc., introduced in the previous section and used to recognize the presence of the elementary events.

When a template following the syntax of Eq. (1) is "*instantiated*" to provide the representation of a simple elementary event like "John gives a book to Mary", the predicate $\mathbf{P}_j$ (of the **MOVE** type in NKRL) will introduce its three arguments $\mathbf{a}_k$ (*"individuals", i.e., instances of standard HClass concepts*) **JOHN_**, **MARY_** and **BOOK_1** *through three functional relationships* ($\mathbf{R}_k$ *roles*) as **SUBJ**(ect), **BEN**(e)**F**(iciary) and **OBJ**(ect). The global construction is then *reified making use of the symbolic label* $\mathbf{L}_i$ *and necessarily managed as a coherent block at the same time*. The instantiations of structures in the style of Eq. (1) are called "*predicative occurrences*" and correspond then to the NKRL representation of specific elementary events. The "*reifying*" function proper to the $\mathbf{L}_i$ conceptual labels allows us to create *new objects* (new formalized elementary events) out of *pre-existing conceptual entities* (predicates, roles, concept instances, etc.) and to deal with these new objects *without making reference to the original entities*. This function is then *absolutely essential to guarantee the n-ary nature of the templates/predicative occurrences in NKRL*. Moreover, it provides the possibility of *formalizing large scenarios/complex events* under the form of structured associations of the *constitutive elementary events*, see below Sections 2.2 and 3.3.

To avoid the ambiguities of natural language and the possible "combinatorial explosion" problems, both the (unique) conceptual predicate of Eq. (1) and the associated functional roles are "*primitives*". Predicates $\mathbf{P}_j$ pertain then to the set {**BEHAVE**, **EXIST**, **EXPERIENCE**, **MOVE**, **OWN**, **PRODUCE**, **RECEIVE**}, and the functional roles $\mathbf{R}_k$ to the set {**SUBJ**(ect), **OBJ**(ect), **SOURCE**, **BEN**(e)**F**(iciary), **MODAL**(ity), **TOPIC**, **CONTEXT**}.

The single arguments $\mathbf{a}_k$ of a template/predicative occurrence, or the templates/occurrences as a whole, may be characterized by "*determiners*" (attributes) that (i) introduce further details/precisions about the 'meaning' of these arguments or templates/occurrences, but that (ii) are never strictly necessary for their basic semantic interpretation in NKRL terms. Determiners are represented mainly by:

- "*locations*", denoted in general by lists of concepts or individuals associated with the arguments $\mathbf{a}_k$ of templates/occurrences through the "colon" operator, "**:**", see Table 1;
- "*modulators*", which apply to a full template or occurrence to particularize their meaning according to the modulators used: they pertain to three categories, temporal, deontic and modal modulators, see the examples in Tables 4, 7, 10 and 11;

Table 1
Deriving a predicative occurrence from a template.

name: **Produce:NegativeCondition/Result**
father: **Produce:CreateCondition/Result**
position: **6.42**
natural language description: "Creation of Explicit Negative Conditions, Situations or Results"

```
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   =   negative_condition, negative_relationship, negative_result
var4   =   human_being_or_social_body
var6   =   human_being_or_social_body
var8   =   sector_specific_activity, intellectual_activity, process_, temporal_sequence
var9   =   sortal_concept, symbolic_label
var10  =   situation_, symbolic_label
var2, var5, var7 = geographical_location

cred5.c20)  PRODUCE   SUBJ      MARY_COLLINS: (BETA_BANK_PREMISES)
                      OBJ       dismissal_
                      BENF      JOHN_SMITH
                      CONTEXT   (SPECIF BETA_BANK PRIORITY_PROJECT)
                      date-1:   8-august-2012
                      date-2:
```

- the "*temporal determiners*" **date-1** and **date-2**, used only in association with predicative occurrences in order to introduce the temporal information associated with an elementary event, see again Table 1, etc.

Fig. 1 reproduces a fragment of the HTemp hierarchy that displays, in particular, the *conceptual labels* of some off-springs of the **Behave:** sub-hierarchy.

As it appears from this figure, HTemp is structured into *seven branches*, where each one includes *only* the templates organized, following the syntax of Eq. (1), around one of the seven predicates ($\mathbf{P_j}$) admitted by the NKRL language. Note also that *polarity-like values* ("favorable", "negative" in Fig. 1) are associated with the templates (and their associated predicative occurrences) through the templates' symbolic labels see, e.g., the **Experience:** templates in Section 3.2. HTemp includes at present (May 2014) more than 150 templates, very easy to specialize and customize according to the application domain.

Several predicative occurrences (elementary events) – denoted by their symbolic labels $\mathbf{L_i}$ – can be associated within the scope of *second order structures* called "*binding occurrences*", denoting then *complex narrative scenarios*. The binding occurrences are labeled lists made up of a *binding operator* ***Bn*** like **CAUSE**, **GOAL**, **COND**(ition), etc. and its *arguments*. The binding operators are then used to formalize those "*connectivity phenomena*" denoting the logico-semantic links that assure the *mutual coherence* of the basic building blocks (elementary events) of a narrative. A binding occurrence can be expressed as:

$$(\boldsymbol{Bn_k}\ \mathbf{arg_1}\ \mathbf{arg_2} \ldots \mathbf{arg_n}). \quad (2)$$

The arguments $\mathbf{arg_i}$ of Eq. (2) can correspond directly to $\mathbf{L_i}$ labels: in this case, they denote simply the presence of particular elementary events represented formally as predicative occurrences. However, they can also correspond recursively to sets of labeled lists in Eq. (2) format, i.e., to complex combinations

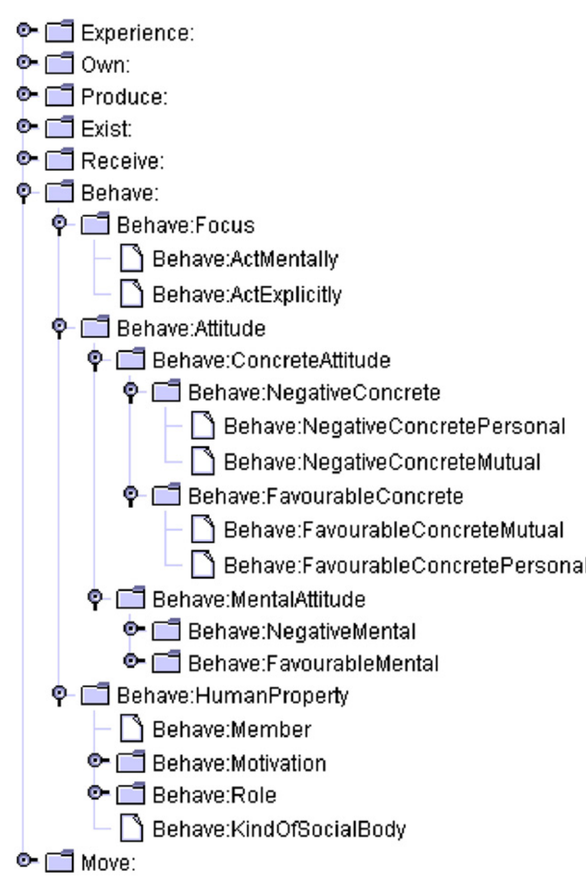

Fig. 1. A (severely reduced) image of HTemp, where the **Behave:** branch has been partially unfolded.

Table 2
Binding and predicative occurrences.

cred5.c20)	**PRODUCE**	**SUBJ**	**MARY_COLLINS: (BETA_BANK_PREMISES)**
		OBJ	**dismissal_**
		BENF	**JOHN_SMITH**
		CONTEXT	**(SPECIF BETA_BANK PRORITY_PROJECT)**
		date-1:	**8-august-2012**
		date-2:	
Produce:NegativeCondition/Result (6.42)			
cred5.c21)	**EXPERIENCE**	**SUBJ**	**JOHN_SMITH**
		OBJ	**(SPECIF difficulty_ strong_)**
		TOPIC	**(SPECIF programming_activity JOHN_SMITH)**
		CONTEXT	**(SPECIF BETA_BANK PRORITY_PROJECT)**
		date-1:	**3-january-2012**
		date-2:	**8-august-2012**
Experience:NegativeHuman/Social (3.211)			
cred5.c22)	**(CAUSE cred5.c20 cred5.21)**		

of **CAUSE**, **GOAL**, **COND**, etc. clauses. Note that, according to Eq. (2), the *upper level of the formal NKRL representation of a complex narrative scenario* is always denoted by a binding occurrence, see also the detailed example of Section 3.3.

2.2. A first example

As already stated, when a specific elementary event pertaining to one of the general classes of events denoted by templates must be represented, the corresponding template is instantiated, then giving rise to one of those *predicative occurrences* (formal descriptions of elementary events) introduced before. Let us suppose we want to represent in NKRL format a (simplified) elementary event, significant from a sentiment analysis point of view, like: "On August 8, 2012, at Beta Bank's premises, Mary Collins fired John Smith in the context of the activities related to the bank's Priority project". This is an example of the information dealt with in a current NKRL application about credentials, permissions and security levels in an industrial context; the names of the characters involved and of the specific Bank have been, of course, changed. Additional information about the roles of the two characters and their hierarchical relationships has been removed for sake of simplicity.

To represent this event we must select first, in the HTemp hierarchy, the template corresponding to "create a negative condition/situation/result with respect to someone", represented in the upper part of Table 1 (see also Fig. 1). This template is a specialization of the particular **PRODUCE** template corresponding to "create a condition/situation/result".

In a template, the arguments of the predicate (the $\mathbf{a}_k$ terms in Eq. (1)) are concretely represented by *variables with associated constraints*. The constraints are expressed as HClass concepts or combinations of concepts, i.e., the two ontologies, HTemp and HClass (the 'standard' ontology of concepts), are *strictly intermingled*. When creating a predicative occurrence like **cred5.c20** in Table 1, the role fillers in this occurrence must conform to the constraints of the father-template. For example, **MARY_COLLINS** and **JOHN_SMITH** are individuals, instances of the HClass concept **individual_person**, specialization in turn of the concept **human_being_or_social_body**, see the constraints on the fillers of the **SUBJ**(ect) and **BEN**(e)**F**(iciary) roles; **dismissal_** is a concept, specialization of **negative_relationship**, a specific term of **mutual_relationship**. In the "*location*" determiner associated with the SUBJ's filler, **BETA_BANK_PREMISES** is an individual, instance of **commercial_premises** that is a specific term of **premises_**, specific term in turn of **geographical_location**, etc. Note that, in NKRL, the HClass concepts are conventionally represented in lower case, while their *instances* (i.e., the *individuals*) are in upper case.

The meaning of the expression "**CONTEXT (SPECIF BETA_BANK PRIORITY_PROJECT)**" in **cred5.c20** is: the dismissal occurs in the context of a particular project, **PRIORITY_PROJECT**, which is proper to the **BETA_BANK**. The "attributive operator", **SPECIF**(ication) is, like **ALTERN**(ative) or **COORD**(ination), one of the operators used for the set up of *structured arguments* of the predicates, see Zarri (2009: 68–70). An example of use of the **ALTERN** operator is given in Table 10 below. The terms included within square parentheses, "**[ ]**", are 'possible/optional', which means that they can be found or not in the corresponding occurrences when the original template is instantiated see, in Table 1, the roles **SOURCE**, **MODAL** and **TOPIC**, or the different location's vectors. In **cred5.c20**, the temporal interval **date-1**, **date-2** used to denote the duration of the elementary event is reduced to the indication of the "point date" of the dismissal, conventionally associated with the **date-1** determiner; see Zarri (2009: 80–86, 194–201) for a description of the representation and management techniques of temporal information in an NKRL context.[2]

To supply now an at least intuitive idea of how a *narrative scenario* (see Eq. (2) above) can be represented in NKRL, and returning to the example in Table 1, let us suppose we would now state that: "Mary Collins fired John Smith … because of John's inadequacy as computer programmer". In this (very simple) example, see Table 2, the specific event corresponding to the firing action is still represented by occurrence **cred5.c20** in Table 1.

To encode correctly the new information, we must introduce first an additional predicative occurrence labeled as **cred5.c21**, see Table 2, meaning that: "During his employment at Beta Bank (January 3rd, 2012–August 8th, 2012), John Smith has been confronted with strong difficulties with respect to his job as computer programmer". Additional predicative occurrences could also be used to denote John's hiring on January 3rd, the details of his employment contract, etc. We will then add a binding occurrence **cred5.c22** labeled with a **CAUSE** ***Bn*** operator,

[2] Computational Linguistics techniques are utilized to produce, in a semi-automatic way, full NKRL-like formal representations from the original NL texts. The procedures employed to implement this NL/NKRL 'translation' are based mainly on the use of generalized "if-then" rules to "trigger" NKRL structures from the results of a previous syntactic analysis of the texts. All of them derive from the algorithms developed in the eighties in the framework of the RESEDA (in French, *Reseau Sémantique Documentaire*) project, an ancestor of NKRL, see Zarri (1983). A recent prototype in this style, created in the context of an "assisted living" application, is described in Ayari, Chibani, and Amirat (2013).

Table 3
The **HumanProperty** sub-domain of **Behave:**.

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

used to link **cred5.c20** (the result) to **cred5.c21** (the triggering factor). The global meaning of **cred5.c22** can be verbalized as: "The activity described in **cred5.c20** is the consequence (**CAUSE**) of the situation illustrated in **cred5.c21**". Note the use, in the **TOPIC**'s filler of **cred5.c21**, of "**(SPECIF programming_activity JOHN_SMITH)**" to denote "his own programming activity"; the **TOPIC** role has the general meaning of "apropos of". **difficulty_**, **programming_activity** and **strong_** are HClass concepts. **strong_** is, in particular, one of the (several) specific terms of the concept **general_characterising_property** of HClass, specific term in turn of **qualifier_** that represents one of the main sub-tree of the **non_sortal_concept** branch of HClass. "Non sortal" are all the concepts (including, e.g., the specific terms of **substance_** or **color_**), which *cannot be endowed with direct instances* (individuals).

3. Sentiment analysis and the NKRL templates

We are now ready to examine some important "sentiment analysis" features that characterize the HTemp templates. Given the space limitations, this description will be necessarily quite limited: the interested reader will find many additional details in Zarri (2009: 149–177).

3.1. "*Behave:*" templates

The **Behave:** templates appear, obviously, as particularly appropriate as vectors of "*sentiment analysis*" features, even if the presence in a template of a particular conceptual predicate like **BEHAVE** (or **MOVE**, **EXIST**, etc.) does not imply, by itself, the assertion of any particular 'conceptual meaning'. A *full meaning* can only arise, in fact, when the four main elements of Eq. (1) *are all present together*.

The templates of the **Behave:** branch of HTemp, see Fig. 1, *can be gathered in two main groups according to the mandatory/forbidden presence of the* ***OBJ****(ect) role*.

Filling the **OBJ**(ect) role is *forbidden*, **+(OBJ)**, in the predicative occurrences derived from templates pertaining to the two branches **Behave:HumanProperty** and **Behave:Focus** of Fig. 1. The **Behave:HumanProperty** templates are used, in general, in situations where *one or more characters perform according to a specific, proper 'function', 'task' or 'role'*. Hence, the most important among them are those represented by **Behave:Role** and its specializations. The **Behave:Focus** templates are employed when a character or group of characters *would like, concretely or as a desire, intention, etc., to make a given situation happen*. In the two cases, the presence in the derived occurrences of a "direct object", **OBJ**(ect), of the **SUBJ**(ect)'s *behavior*, is *logically inconsistent*.

A second group of **Behave:** templates corresponds to the **Behave:Attitude** branch of Fig. 1. They are used to model situations where a **SUBJ**(ect) *manifests directly a given behavior*, real or purely speculative, *in favor of or against a person, a social body, a situation/activity, etc.* In the derived occurrences, filling the **OBJ** role is *mandatory*. The **BEN**(e)**F**(iciary) role is now 'forbidden', **+(BENF)**, given that the "direct object" of the **SUBJ**'s 'attitude' corresponds here to the **OBJ**'s filler.

The Behave:HumanProperty templates. Their general schema is shown in Table 3; the "/" symbol indicates the presence of syntactic alternatives, "**+**" means "forbidden". The constraints on the variables have been suppressed for sake of simplicity; see, however, the template in Table 4 below.

Filling the **MODAL** role is then *mandatory* in all the predicative occurrences derived from all the **Behave:HumanProperty** templates. In Table 3, this is signaled by the absence, for this role, of the "possible/optional" code ("**[ ]**", see the previous section). For the sake of generality, the possibility of filling the **BEN**(e)**F**(iciary) role has been preserved, see again Table 3, for some of these templates, like the two **Behave:Motivation** templates (willing/unwilling about the execution of a given task) of Fig. 1.

Filling the **BEN**(e)**F**(iciary) role is, on the contrary, *strictly forbidden*, **+(BENF)**, in the **Behave:Role** predicative occurrences (the **Behave:Role** template is reproduced in full in Table 4).

We can note, moreover, that filling the **TOPIC** role is *possible/optional* for the occurrences derived from the generic **Behave:Role** template. On the contrary, it is *necessarily required* in the occurrences derived from all the specializations of this template, like **Behave:User** or **Behave:Believer**. In general, it is required when, in agreement with the function of the **TOPIC** role, it is necessary to give *additional precisions about a specific function or task*. The lower part of Table 4 reproduces two examples of use of **Behave:Role**, where the first does not imply the use of **TOPIC**, which is needed in the second.

In the two predicative occurrences of Table 4, the "temporal modulator" (see Section 2.1 above) "**obs**(erve)" is used to denote the fact that the date associated with **date-1** represents only a specific point within the temporal interval associated with the event, without giving any information about the lower and upper limits of this interval.

Acting to obtain a given result. The **Behave:Focus** templates are used to translate the general idea of *acting to obtain a given result* according to the following modalities:

- A predicative occurrence, which must necessarily be an instance of a **Behave:Focus** template (see, e.g., the template **Behave:ActExplicitly** in Table 5 below) is used to express the *'acting' component*. This occurrence allows us, then, to identify the **SUBJ**(ect) of the action, the temporal information, possibly the **MODAL**(ity) or the instigator (**SOURCE**) of this component, etc. In this occurrence, the **OBJ**(ect) role is *'empty'*, in conformity with what stated at the beginning of Section 3.1.
- A second occurrence, which can be a single predicative occurrence or a binding occurrence denoting several predicative occurrences, is used to express the *'intended result' component*. This occurrence, which happens 'in the future' with respect to the first, i.e., the **Behave:Focus** one, must necessarily be marked as *hypothetical*. This implies *adding* to the second occurrence, if this is a predicative one, an *uncertainty validity code* "*****", see occurrence **brit.c27** in Table 5 below. If the second occurrence is a binding one, all the *included predicative occurrences* are characterized by the addition of this code.

Table 4
Examples of the use of **Behave:Role**.

name: **Behave:Role**
father: **Behave:HumanProperty**
position: **1.33**
NL description: 'A Human Being or a Social Body Acts in a Particular Role'

BEHAVE	**SUBJ**	***var1***: [(***var2***)]
	+(OBJ)	
	[SOURCE	***var3***: [(***var4***)]]
	+(BENF)	
	MODAL	***var5***
	[TOPIC	***var6***]
	[CONTEXT	***var7***]
	{ [modulators], ≠(abs) }	

var1 = **human_being_or_social_body**
var3 = **human_being_or_social_body**
var5 = **role_**
var6 = **entity_**
var7 = **situation_, symbolic_label**
var2, var4 = **geographical_location**

mod33.c9)	**BEHAVE**	**SUBJ**	**ARIEL_BROWN**
		MODAL	**journalist_**
		{ obs }	
		date-1:	**13-june-1999**
		date-2:	

Behave:Role (1.33)

On June 13, 1999, we can remark, temporal modulator **obs***(erve), that Ariel Brown is a journalist (***journalist_** *is an HClass specific term of* **professional_role***).*

sent4.c17)	**BEHAVE**	**SUBJ**	**JOHN_KERRY**
		MODAL	**chairman_**
		TOPIC	**(SPECIF foreign_relations_committee US_SENATE)**
		{ obs }	
		date-1:	**9-march-2009**
		date-2:	

Behave:Role (1.11)

On March 9, 2009, we can note that John Kerry is the chairman of the Foreign Relations Committee of the Senate of the USA.

Table 5
Binding and predicative occurrences.

brit.c26)	**BEHAVE**	**SUBJ**	**BRITISH_TELECOM**
		MODAL	**planning_**
		date1:	**march-1998**
		date2:	

Behave:ActExplicitly (1.12)

***brit.c27)**	**MOVE**	**SUBJ**	**BRITISH_TELECOM**
		OBJ	**payg_internet_service**
		BENF	**(SPECIF customer_ BRITISH_TELECOM)**
		date-1:	**after-1-september-1998**
		date-2:	

Move:TransferOfServiceToSomeone (4.11)

brit.c28) **(GOAL brit.c26 brit.c27)**

- A third occurrence, *a "binding" one*, *which makes necessarily use of a* ***GOAL*** *operator*, is then used to link the previous two, see again Table 5.

The general syntax of the NKRL expressions used to code the "acting to obtain a given result" situations is then given by:

(c_α) BEHAVE SUBJ <human_being_or_social_body>
(*c_β) <predicative occurrence(s), with any syntax>
(c_γ) (**GOAL c_α c_β**)

The example of Table 5 translates then the following situation: "We can note that, on March 2008, British Telecom *plans to offer* to its customers, in autumn 1998, a pay-as-you-go (**payg_**) Internet service", where the template **Behave:ActExplicitly** is used to express the notion of "planning". In the predicative occurrence **brit.c26** of Table 5, **march-1998** is interpreted as a "point data". Note that the **GOAL** machinery exemplified by the binding occurrence **brit.c28** is, in a sense, the mirror image of that used, in a **CAUSE** context, in the occurrence **cred5.c22** of Table 2 above.

Table 6
The "attitude" sub-domain of **Behave:**.

[Behave:Attitude]	BEHAVE	SUBJ	*var1*: [(*var2*)]
		OBJ	*var3*: [(*var4*)]
		[SOURCE	*var5*: [(*var6*)]]
		+(BENF)	
		MODAL	*var7*
		[TOPIC	*var8*]
		[CONTEXT	*var9*]
		{ for / against }	

Table 7
Example of use of **Behave:FavourableConcretePersonal**.

cob1.c1) BEHAVE	SUBJ	(SPECIF GOVERNMENT_1 (SPECIF CARLO_AZEGLIO_CIAMPI prime_minister)): ROME_
	OBJ	(SPECIF sale_ CRDI_)
	MODAL	COMMITMENT_1
	CONTEXT	italian_privatisation_programme
	{ for }	
	date-1:	(1-august-1999), 7-september-1999
	date-2:	

Behave:FavourableConcretePersonal (1.2122)

(Before September 7, 1999), the government of Carlo Azeglio Ciampi has pledged to sell the Credito Italiano SpA (CRDI) bank as part of Italy's privatization program.

Note that the possible addition of a **ment**(al) "modal" modulator to the **BEHAVE** occurrence, $\mathbf{c}_\alpha$, that introduces an "acting to obtain a result" construction should imply that *no concrete initiative* has actually been taken by the **SUBJ** of **BEHAVE** in order to fulfill the result. To return to the British Telecom example, this would be the case if, e.g., the British Telecom's initiative represented by occurrence **brit.c27** of Table 5 were only a project. With the addition of **ment**, the 'result', $*\mathbf{c}_\beta$, reflects then only the *planned intentions of the **SUBJ**(ect)*; note that, in this last case, the template to be used for $\mathbf{c}_\alpha$ should be now **Behave:ActMentally** instead of **Behave:ActExplicitly**, see Fig. 1. For further details about the representation of the general "*motivational attitudes*" domain (goals, wants, desires, preferences, wishes, choices, intentions, commitments, behaviors, plans), see Zarri (2009: 153–155). An example of use of the modal modulator "**wish**" is given in Table 10.

The Behave:Attitude templates. The templates corresponding to the **Behave:Attitude** branch of Fig. 1 follow the general schema of Table 6. As already stated at the beginning of Section 3.1, filling the **OBJ** role is now *strictly mandatory* in the derived occurrences; for this class of templates, also the **MODAL** role is mandatory. Moreover, unlike the **Behave:Focus** occurrences discussed in the previous sub-section, these predicative occurrences cannot be included within binding occurrences of the **GOAL** type.

A simple example of *'positive' attitude* is represented in Table 7. As shown in Table 6, the predicative occurrences derived from templates of the **Behave:Attitude** type *must* necessarily be associated with one of the two modal modulators "**for/against**". The *global* meaning of occurrence **cob1.c1** is then: the government has a *specific attitude* about the sale of the bank, which is defined as *favorable* thanks to the association of the modal modulator "**for**" to the *whole occurrence*. In the **date-1** temporal attribute, the first element of the temporal interval is a "reconstructed date".

3.2. "Behavioral" aspects in the templates of the residual HTemp branches

The Exist: templates. They can be classified in two main categories:

- Templates that represent specializations of **Exist:BePresent**, to be used to denote *situations where a given entity, human or not, is present at a given location*, including virtual ones. They *all require as mandatory* the presence, in the derived occurrences, of the *location of the **SUBJ**(ect)*. Moreover, **OBJ**(ect) is normally *forbidden* in these occurrences.
- Templates that represent specializations of **Exist:OriginOrDeath** and that are used to model *the 'birth' or the 'final end' or a given entity*, human or not. They can be employed, e.g., to represent *the creation or the dismantling* of a social body, company, political party, university, etc.

The Experience: templates. These templates are mainly used to represent events where a given entity, human or not, *is exposed to some sort of 'experience'* (illness, richness, economic growth, starvation, success, racism, violence…). This experience can be further specialized making use of a 'positive' (**Experience:PositiveSituation**), 'negative' (**Experience:NegativeSituation**) or 'neutral' (**Experience:GenericSituation**) polarity. The particular experience undergone is represented, in the derived predicative occurrences, by the filler of the **OBJ**(ect) role: this role is then *mandatory* for all the **Experience:** templates. On the contrary, given that all the experiences are considered as 'personal', the **BEN**(e)**F**(iciary) role is forbidden, **+(BENF)**. An example of negative experience is represented by occurrence **cred5.c21** of Table 2, which is an instance of the **Experience:NegativeHuman/Social** template. A further specialization of this last template is, e.g., **Experience:HumanBeingInjuring**. An example of template pertaining to the **Experience:GenericSituation** sub-hierarchy is **Experience:BeAged**.

The Move: templates. They are distributed into four branches, **Move:TransferToSomeone**, **Move:ForcedChange**, **Move:TransmitInformation** and **Move:AutonomousDisplacement**. These templates present some interesting *syntactic variants* linked, at least partly, *to the different possible layouts of the "location" items*. As already stated, these are associated with the arguments of the predicate making use of the (external) "colon" operator, "**:**".

For example, an 'autonomous movement' of the **SUBJ** (**Move:AutonomousDisplacement** templates) is always interpreted as: "The **SUBJ** moves herself/himself/itself as an **OBJ**". The location associated with the filler of the **SUBJ** role, l_1, is then interpreted as the 'initial location', and the location associated with the **OBJ**,

Table 8
A simple example of use of the **Move:ForcedChangeOfState** template.

```
aal1.c5) MOVE   SUBJ      HOME_CONTROL_SYSTEM_1
                OBJ       (SPECIF lighting_apparatus BATHROOM_1): (switch_off, switch_on)
                date-1:   11- april-2011/9:16
                date-2:

Move:ForcedChangeOfState (4.12)
```

On April 4, 2011, at 9h16, the system has turned the bathroom lights on.

Table 9
An example from the **Produce:Violence** domain.

```
name: Produce:HumanBeingKilling
father: Produce:Violence
position: 6. 492
NL description: 'Killing the Filler (Human Being) of the BEN(e)F(iciary) Role'

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   =  human_being_killing
var4   =  human_being_or_social_body
var6   =  human_being
var8   =  criminality/violence_related_tool, machine_tool, small_portable_equipment, violence_, weapon_
var9   =  h_class
var10  =  situation_, symbolic_label
var2, var5, var7 = geographical_location

afga0404.c6)   PRODUCE   SUBJ      (SPECIF INDIVIDUAL_PERSON_84 weapon_wearing
                                     (SPECIF cardinality_ several_)): (GEOGRAPHICAL_LOCATION_6)
                         OBJ       human_being_killing
                         BENF      WOMAN_2
                         CONTEXT   terrorist_shooting_attack
                         date-1:   17/11/2003
                         date-2:

Produce:HumanBeingKilling (6.351)
```

On November 17, 2003, in an unspecified location, an armed group of people shot a woman dead.

l_2, as the 'arrival location', $l_1 \neq l_2$, see occurrence **sent1.c7** in Table 11 below. Locations are represented in general as lists. For example, the **Move:ForcedChange** templates are used whenever an agent (**SUBJ**), which is located in its proper, known or unknown, location, moves an entity (**OBJ** = physical object, animate entity, process, state...) from the 'initial' location l_1 to l_2. In this case, l_1 and l_2 correspond, respectively, to the initial and final terms of the *location list* (l_1, l_2) associated with the filler of the **OBJ** role. A simple example of use of the **Move:ForcedChangeOfState** template in the framework of an on-going NKRL application in the "assisted living" domain is represented in Table 8.

An important class of **Move:** templates is represented by the specializations of **Move:TransmitInformation**. In these templates, the filler of the (mandatory) **OBJ**(ect) role can be simply a specific HClass concept (or an instance) like, e.g., **query_** or **message_** see, e.g., occurrence **sent1.c4** in Table 11 below. However, this filler can also designate a "conceptual label" in the style of $\mathbf{L}_i$ in Eq. (1) or $\boldsymbol{Bn}_k$ in Eq. (2). In this second case, thanks to the mechanism of the "*completive construction*", see 3.3 below and Zarri (2009: 87–91), the transmitted message is a *complex/structured one*.

The Own: templates. They are mainly used to represent the different nuances of the notion of *"possessing some sort of entity"*. Under the **Own:Property** form, they are employed *to specify the 'properties' of NKRL inanimate entities* making use of the **TOPIC** role, see Table 12 below. Note that the 'properties' of human beings and social bodies must be described, in contrast, making use of the **Behave:** templates.

The Produce: templates. This class of templates is particularly large. The meaning of the **Produce:Entity** templates (e.g., **Produce:Hardware**) is self-evident. Examples of the **Produce:PerformTask/Activity** templates are, e.g., **Produce:Buy** and **Produce:Sell** (in the predicative occurrences derived from these two templates, the filler of the **OBJ** role is necessarily **purchase_** or **sale_**, or specializations/instances of these concepts).

A **Produce:PerformTask/Activity** template especially important from a sentiment analysis point of view is, e.g., **Produce:Violence**, which involves several specializations. Table 9 concerns, in particular, the subsumed **Produce:HumanBeingKilling** template and explains its use to encode one of the elementary event examples introduced informally in Section 1. Other **Produce:** templates interesting in a sentiment analysis context are, e.g., **Produce:Acceptance/Refusal** and **Produce:CreateCondition/Result**. A specialization of this last template,

Table 10
An example of **Receive:DesiredAdvice** predicative occurrence.

skin1.c8)	**RECEIVE**	**SUBJ**	**CRYSTAL_EYES: (KANATA_ON)**
		OBJ	**advice_**
		SOURCE	**BEAUTYNET_COMMUNITY**
		TOPIC	**(ALTERN (SPECIF use_ NOVA_UNDEREYE_CREAM) (SPECIF use_ baby_oil))**
		CONTEXT	**(SPECIF therapy_ (SPECIF undereye_dry_skin CRYSTAL_EYES))**
		{ wish }	
		date-1:	**28-december-2001**
		date-2:	

Receive:DesiredAdvice (7.311)

On December 28, 2001, Crystal Eyes wishes to receive an advice from the Beauty Net community about the utilization of baby oil or of a product like Nova Undereye Cream in the context of her eye dry skin problem.

Produce:NegativeCondition/Result, has been used to create the predicative occurrence **cred5.c20** of Table 1; the template **Produce:PositiveCondition/Result** can be used to express, e.g., an official approval with respect to a given action/situation. The specializations of **Produce:RelationInvolvement** like **Produce:MutualCommittment** are used to represent *all forms of 'agreement' among several participants*. These are mentioned in a **COORD**(ination) list that fills the **SUBJ**(ect) role and that is duplicated as filler of the **BEN**(e)**F**(iciary) role. Predicative occurrences derived from the **Produce:Growth** specialization of **Produce:Increment/Decrement** can be employed to represent the increase/acceleration/intensification/amplification, etc. of a given process/action.

The Receive: templates. **Receive:DesiredAdvice** is a **Receive:** template that can be particularly important in a sentiment analysis context. In its derived predicative occurrences, the use of the modal modulator **wish** is *mandatory*. An example, pertaining to an NKRL application in the "beauty care" domain, is supplied in Table 10.

Note, in the filler of the **TOPIC** role of **skin1.c8**, the use of the so-called "priority rule" in order to formulate correctly the alternative between "the use of the Nova Undereye cream" and "the use of a baby oil" by stating that **ALTERN**(ative) has priority over **SPECIF**(ication).

3.3. A global example in a sentiment analysis context

Table 11 reproduces the NKRL conceptual encoding of a typical news story, slightly modified for the sake of intelligibility and conciseness: "On November 10, 2008, Barack Obama and George W. Bush talked briefly at the White House about Iraq, and in the end President Obama asked President Bush how he was enjoying his new endeavor and whether the Bush family had moved into the new house. Barack Obama is the current USA president and George W. Bush the former one (November 10, 2008)".

The text of this news story has been extracted from the "*Sentiment-annotated quotation corpus*" that includes a set of 1590 quotations from news in English language prepared at the European Joint Research Centre of Ispra (Italy) for facilitating sentiment analysis research about news. The full corpus can be retrieved at http://ipsc.jrc.ec.europa.eu/?id=61; a related paper is Balahur et al. (2010).

As already stated in 2.1, the (*mandatory*) starting point for the creation of any NKRL full representation of complex stories/narratives/scenarios is the set up of a *binding occurrence listing the main topics dealt with in that context*. This 'upper level' occurrence corresponds frequently, as in the present case (**sent1.c1**), to a **COORD**(ination) binding occurrence. We have assumed here, then, that the story was formed of *five independent but strictly connected items*. The first (**sent1.c2**) relates the occurrence of a meeting between George W. Bush and Barack Obama and the second (**sent1.c3**) introduces the main themes of the conversation. The third (**sent1.c4/#sent1.c5**) concerns President Obama's 'last query' and its topics, and the last two, **sent1.c8** and **sent1.c9**, are the descriptions of the actual 'jobs' of the two involved people. Note that, within **sent1.c1**, the conceptual label **#sent1.c5** is there only to allow the NKRL software to carry out some coherence controls. In the global code of Table 11, in fact, the two occurrences **sent1.c4/#sent1.c5** perform in practice as a *unique conceptual unit* by virtue of the mechanism of the "*completive construction*" already mentioned, see below for additional details. The presence of **#sent1.c5** in **sent1.c1** does not alter at all, then, the cardinality (five) of the **COORD**'s arguments in this top binding occurrence.

After having set up the top level of the global scenario, the different information blocks listed in this binding occurrence are *successively encoded in NKRL terms* in order to model, according to the NKRL's specifications, the single elementary events included in the narrative. Let us consider, e.g., the logical/semantic structure of the 'last query' block introduced by occurrence **sent1.c4**.

This occurrence is an instance of the **Move:GenericInformation** template, to be used when the 'message' sent by the **SUBJ**(ect)'s filler to the **BEN**(e)**F**(ciary)'s filler and represented by the **OBJ**(ect)'s filler is denoted by a *simple* HClass term, **query_** in the present case, or by a conjunction/alternative of simple HClass terms. The **query_** concept pertains to the **entity_** branch of HClass and is a specific term, through an intermediary concept **word_content**, of **information_content**, the default filler of **OBJ** in the **Move:GenericInformation** template. When the message to be transmitted is a '*complex*' one and is represented then by a full predicative occurrence or a set of predicative occurrences, the **Move:StructuredInformation** template and the "*completive construction*" mechanism must be used. The "*completive construction*" represents an alternative modality, with respect to the binding occurrences mechanism, of linking together predicative occurrences. In a **Move:StructuredInformation** template context, the predicative occurrences derived from this template are then characterized by the presence, as **OBJ**(ect)'s filler, of an instance of **#symbolic_label**. As already stated above, the symbol "**#**" is there only to assure the correct functioning of the NKRL software; note, moreover, that **symbolic_label** is a regular concept of HClass having as *instances* all the *concrete labels* used to denote the predicative/binding occurrences defined within a specific NKRL application. The particular symbolic_label used in the concrete case of transmission of a complex message corresponds then to the symbolic name of a single predicative occurrence c_i or to the name of a binding occurrence b_j (and refers then to a set of predicative occurrences). In the **Move:StructuredInformation** case, c_i or b_j denote, eventually, the 'structured' message to be transmitted.

Even if particularly frequent in an NKRL context, the use of **#symbolic_label** as filler of the **OBJ**(ect) functional role within predicative occurrences derived from **Move:Structured**

Table 11
NKRL representation of a full story/scenario.

```
sent1.c1)  (COORD  sent1.c2  sent1.c3  sent1.c4  #sent1.c5  sent1.c8  sent1.c9)
```

The NKRL representation of the sent001 *story is formed of several components.*

```
sent1.c2)  BEHAVE   SUBJ      (COORD1 BARACK_OBAMA GEORGE_W_BUSH): (WHITE_HOUSE)
                    MODAL     participant_
                    CONTEXT   CONVERSATION_1
                    date-1:   10/11/2008
                    date-2:
Behave:Participant (1.114)
```

On November 10th, 2008, Barack Obama and George W. Bush had a talk in the White House.

```
sent1.c3)  OWN      SUBJ     CONVERSATION_1
                    OBJ      property_
                    TOPIC    (COORD1 (SPECIF is_about IRAQ_) (SPECIF duration_ short_))
                    date-1:  10/11/2008
                    date-2:
Own:CompoundProperty (5.42)
```

We can remark that the talk is about Iraq and that its duration is short.

```
sent1.c4)  MOVE     SUBJ      BARACK_OBAMA: (WHITE_HOUSE)
                    OBJ       (SPECIF query_ last_)
                    BENF      GEORGE_W_BUSH
                    TOPIC     #sent1.c5
                    CONTEXT   CONVERSATION_1
                    date-1:   10/11/2008
                    date-2:
Move:GenericInformation (4.41)
```

On November 10th, 2008, President Obama asks President Bush about what is described in occurrence sent1.c5.

```
sent1.c5)  (COORD  sent1.c6  sent1.c7)
```

The two occurrences sent1.c6 *and* sent1.c7 *mentioned in the binding occurrence* sent1.c5 *are logically linked.*

```
*sent1.c6) EXPERIENCE   SUBJ     GEORGE_W_BUSH
                        OBJ      positive_feeling
                        TOPIC    (SPECIF personal_activity new_ GEORGE_W_BUSH)
                        date-1:  10/11/2008
                        date-2:
Experience:PositiveHuman/Social (3.221)
```

The first part of the question is whether President Bush has a positive feeling about his new type of activity.

```
*sent1.c7) MOVE   SUBJ     (COORD1 GEORGE_W_BUSH (SPECIF family_ GEORGE_W_BUSH)): (WHITE_HOUSE)
                  OBJ      (COORD1 GEORGE_W_BUSH (SPECIF family_ GEORGE_W_BUSH)): (HOME_1)
                  date-1:  10/11/2008
                  date-2:
Move:AutonomousPersonDisplacement (4.31)
```

The second part of the question is whether President Bush and his family have moved to a (different) house.

```
sent1.c8)  BEHAVE   SUBJ     BARACK_OBAMA
                    MODAL    current_republic_president
                    TOPIC    USA_
                    { obs }
                    date-1:  10/11/2008
                    date-2:
Behave:Role (1.11)
```

On November 10th, 2008, we can note that Barak Obama is the current USA president.

```
sent1.c9)  BEHAVE   SUBJ     GEORGE_W_BUSH
                    MODAL    former_republic_president
                    TOPIC    USA_
                    { obs }
                    date-1:  10/11/2008
                    date-2:
Behave:Role (1.11)
```

On November 10th, 2008, we can note that George W. Bush is the former USA president.

Information does not represent the only possibility of use of the completive construction mechanism. This particular filler can also be associated, in fact, with the **OBJ**, **MODAL**, **TOPIC** and **CONTEXT** functional roles in the context of predicative occurrences derived from several different templates. In the example of Table 11, the completive construction mechanism is used in **sent1.c4** to denote the *subject matter* (introduced by the **TOPIC** functional role) of the query. In **sent1.c4**, the **TOPIC**'s filler denotes then the binding occurrence **sent1.c5** that introduces the two themes of the query. The corresponding predicative occurrences are both marked as "uncertain" because, at the moment of the query, it cannot be verified if they correspond to any concrete reality.

The predicative occurrence **sent1.c7** represents an instance of one of the most interesting conceptual structures of NKRL, used

Table 12
An example of multi-consequent transformation.

"explicit hostility" transformation:

antecedent

EXIST SUBJ *var1*: (*var2*)
TOPIC (SPECIF negative_relationship *var3*)

***var1* = mass_demonstration**
***var2* = country_**
***var3* = country_**
var2* ≠ *var3

consequent1

OWN SUBJ *var4*
OBJ property_
TOPIC (SPECIF located_in *var2*)

***var4* = city_**

consequent2

PRODUCE SUBJ (SPECIF human_being (SPECIF cardinality_ several_)): (*var4*)
OBJ *var5*
BENF *var3*

***var5* = violence_**

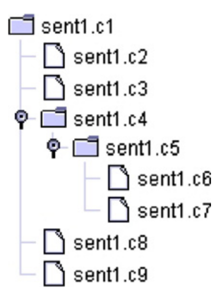

Fig. 2. Tree structure corresponding to the conceptual representation of Table 11.

to represent the autonomous movement of animate entities. As already mentioned in 3.2, a characteristic of this structure concerns the fact that the person(s) mentioned as filler(s) of the **SUBJ**(ect) role, in this case President Bush and his family, are conceived as *moving herself/himself/themselves as an **OBJ**(ect)*. The "location of the **SUBJ**(ect)", which appears after the "**:**" code, corresponds then to the starting point of the movement, the "location of the **OBJ**(ect)" to the arrival point. Eventually, we can note the use of the "temporal modulator" **obs**(erve) in **sent1.c8** and **sent1.c9**. The association of this modulator with the timestamp **10/11/2008** filling the temporal attribute **date-1** means that the situation described in the two occurrences is true at this particular date (in our case, the date of the meeting between the two Presidents). For lack of interest, lack of information or for sake of conformity with the original wording of the narrative we do not care, then, about the real duration of this situation, which surely extends in time before and after the given date.

We will conclude this example by noticing that this sort of conceptual representation, in NKRL terms, of complex stories/scenarios/narratives, can always be represented as a tree structure, see Fig. 2. This remark is not new, and can be considered as valid in general independently of the type of formalization employed see, e.g., the "story trees" of Mani and Pustejovsky (2004).

4. The query/inference aspects

A detailed representation of phenomena in the "sentiment analysis" style would be of scarce utility without some means of *automatically exploiting the power of this representation*. In this section, we will supply some (basic) information about the query/inference aspects of NKRL, referring the interested readers to Zarri (2005, 2009: 183–243, 2013) for more details.

Querying/reasoning in NKRL ranges from the *direct questioning* of a knowledge base (*KB*) of information in NKRL format to *high-level inference procedures*. Making use of a powerful ***InferenceEngine***, these utilize *mainly* two classes of rules, "transformations" and "hypotheses"; all the NKRL "rules" are founded on the use of that "operational/common sense" knowledge already mentioned in Note 1 above. The '*reasoning steps*' of these rules are represented under the form of *partially instantiated, standard NKRL templates* (including then explicit variables under $\boldsymbol{var_i}$ form, see Table 12). According to the specific values associated with the variables, the reasoning steps are converted by the different versions of ***InferenceEngine*** into "*search patterns*" $\boldsymbol{p_i}$ (formal queries) that (try to) unify information in the base thanks to a powerful *Filtering Unification Module* (***Fum***). In the "*direct questioning*" mode, users can also build up 'by hand' the search patterns $\boldsymbol{p_i}$ needed to retrieve directly (using ***Fum***) information from the *KB*.

Transformation rules try to '*adapt*', from a semantic point of view, a search pattern $\boldsymbol{p_i}$ that '*failed*' (that was unable to find a unification within the knowledge base) to the real contents of this base making use of a sort of *analogical reasoning*. These rules attempt then to automatically 'transform' $\boldsymbol{p_i}$ into one or more *different* $\boldsymbol{p_1}, \boldsymbol{p_2} \ldots \boldsymbol{p_n}$ that *are not strictly 'equivalent' but only 'semantically close'* to the original one. A transformation rule can be conceived as made up of a left-hand side, the "*antecedent*" and of one or more right-hand sides, the "*consequent(s)*". The antecedent corresponds to the formulation of the 'query' to be transformed, the consequent(s) to the representation(s) of one or more search patterns to be substituted for the given one. Denoting with $\boldsymbol{A}$ the antecedent and with $\boldsymbol{Cs_i}$ the possible consequents, these rules can then be expressed as:

$$\boldsymbol{A}(\boldsymbol{var_i}) \Rightarrow \boldsymbol{Cs_i}(\boldsymbol{var}_j), \quad \boldsymbol{var_i} \subseteq \boldsymbol{var_j}. \qquad (3)$$

The restriction $\boldsymbol{var}_i \subseteq \boldsymbol{var}_j$ (all the variables declared in the antecedent $\boldsymbol{A}$ must also appear in $\boldsymbol{Cs_i}$) assures the logical congruence of the rules.

Let us consider a concrete, quite simple example. This concerns a recent NKRL application about the 'intelligent' management of

Table 13
Details about an anti-western manifestation in Afghanistan.

afga0314.c8)	PRODUCE	SUBJ	(SPECIF INDIVIDUAL_PERSON_90 (SPECIF cardinality_ several_) (SPECIF approximate_amount 10000)): (MEHTARLAN_)
		OBJ	flag_burning
		BENF	(COORD1 UK_ USA_ ALLIED_COALITION)
		date-1:	24/3/2003
		date-2:	
Produce:Violence (6.35)			
The protesters burned flags of the United Kingdom, of the US and of the countries members of the Western coalition.			

news in the context of the Afghanistan conflict and could also present an interest from a "sentiment analysis" point of view. Let us suppose then we would like to ask whether, in 2001–2003, anti-US demonstrations took place in Afghanistan; we suppose also that the corresponding 'manual' search pattern $\boldsymbol{p_i}$ fails. Under these conditions, it will be then necessary to resort to transformation rules like that reproduced in Table 12 to try to *infer* an *indirect answer* to the original question. One of the possible answers will be represented by the information that, in the Afghan city of Mehtarlan, US flags were burned by demonstrators at a date congruent with the original temporal constraints. Note that the transformation of Table 12 is "*multi-consequent*". This means that the two reasoning steps must be *simultaneously satisfied* to produce a valid implicit answer.

After unification of the antecedent with the original query, variable ***var2*** is bound to **AFGHANISTAN_** and variable ***var3*** to **USA_**. ***consequent1*** is used to retrieve, in the Afghanistan conflict knowledge base, all the Afghan cities (***var4***) known by the system. **located_in** is a concept included in the **binary_relational_property** branch of the **relational_property** sub-hierarchy of HClass; an example of **multiple_relational_property** concept is, e.g., **between_**. As already stated in Section 3.2, the **Own:Property** templates are used *to specify the 'properties' of NKRL inanimate entities* making use of the **TOPIC** role. ***consequent2*** allows us to select, among all the possible cities bound to ***var4*** a city, Mehtarlan where, in agreement with the temporal limits imposed, we can observe the existence of some form of mass protest against USA. The predicative occurrence **afga0314.c8** in Table 13 tells us in fact that, on March 24, 2003, in Mehtarlan, a number of protesters set fire to the US flag. The search pattern $\boldsymbol{p_i}$ built up from ***consequent2*** after having bound the value **violence_** to ***var5*** is actually *congruent* with the **afga0314.c8**'s format and can then *unify* this last occurrence. Note that **MEHTARLAN_** is among the possible values associated with ***var4*** and that **USA_** is already bound to ***var3***.

With respect now to the *hypothesis rules*, these allow us to build up automatically a sort of '*causal explanation*' for an elementary event retrieved within a NKRL knowledge base. These rules can be expressed as *biconditionals* of the type:

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

where the 'head' $\boldsymbol{X}$ of the rule corresponds to a predicative occurrence $\boldsymbol{c_i}$ and the 'reasoning steps' $\boldsymbol{Y_i}$ (called 'condition schemata' in a hypothesis context) *must all be satisfied.* This means that, for each of them, at least one '*successful*' search pattern $\boldsymbol{p_i}$ must be derived. 'Successful' means that this pattern must be able to find, using the ***Fum*** module, a *positive unification* with some predicative occurrences within the knowledge base. *In this case, the set of* $\mathbf{c_1}$, $\mathbf{c_2} \ldots \mathbf{c_n}$ *predicative occurrences retrieved by the* $\boldsymbol{Y_i}$ *thanks to their conversion into* $\boldsymbol{p_i}$ *can be interpreted as a context/causal explanation of the original occurrence* $\mathbf{c_i}$ (i.e., of the head $\boldsymbol{X}$ of the rule).

As an informal example, and to continue with the NKRL application about the Afghanistan conflict, let us suppose we want to find a 'plausible explanation' for the narrative corresponding to the occurrence **afga0404.c6** above, see Table 9, i.e., "On November 17, 2003, in an unspecified location, an armed group of people shot a woman dead". We can make use for this of two 'parallel' hypothesis rules, where the 'common sense knowledge' underpinning the first says that: "The Taliban (we are in an Afghanistan context) do not like to see women working" and the second says: "The Taliban dislike people working for the UN agencies". In the two cases, the first reasoning step consists then in verifying that the Taliban are effectively at the origin of the murder. After this, in the first hypothesis we must prove, directly (or indirectly through transformations) that the Taliban are against the employment of women. A third step will, therefore, consist in proving that the killed woman was a regular employee of an UN agency: more precisely, we will discover that she was an employee of UNHCR (the UN Agency for the Refugees). In the second case, the intermediate reasoning step consists in verifying (once again directly or indirectly) that the Taliban detest the UN agencies like UNHCR. The last step is then identical to that to be verified for the first hypothesis.

A recent development of NKRL concerns the possibility of making use of the two above modalities of inference *in an 'integrated' way* (Zarri, 2005). More exactly, *it is possible to make use of "transformations" when working within the "hypothesis" inference environment*. This means that, whenever a search pattern $\boldsymbol{p_i}$ is derived from a condition schema $\boldsymbol{Y_i}$ of a hypothesis to implement the corresponding reasoning step, we can use it '*as it is*', in accordance then with its 'father' condition schema $\boldsymbol{Y_i}$ as this last had been originally encoded. However, we can also make use of a new pattern $\boldsymbol{p_j}$ obtained by transformation of the original $\boldsymbol{p_i}$ *if the appropriate transformation rules exist within the system*. The advantages of this approach are essentially of two types:

- From a 'practical' point of view, a hypothesis that was deemed to fail because of the impossibility of deriving a 'successful' $\boldsymbol{p_i}$ from one of its condition schemata *can now continue* if a transformed $\boldsymbol{p_j}$ can unify some information within the *KB*, getting then new values for the hypothesis variables.
- From a more general point of view, this strategy allows us to explore systematically all the possible *implicit* relationships among the data in the base. One of its modalities consists, in fact, *of transforming all the possible* $\boldsymbol{p_i}$ *derived from the condition schemata of a hypothesis also in case of successful unifications with information in the base*. This permits, e.g., to confirm in many different ways the existence of relationships between people/entities.

Table 14 supplies the informal description of the reasoning steps ("condition schemata") to be validated, making use of the hypothesis tools, in order to prove that, in the Southern Philippines' terrorism context, a generic "kidnapping" corresponds in reality to a more precise "kidnapping for ransom".

Making use of transformations, the hypothesis of Table 14 becomes then *potentially equivalent* to that represented in Table 15.

Table 14
Inference steps for the 'kidnapping for ransom' hypothesis.

(***cond1***)	The kidnappers are part of a separatist movement or of a terrorist organization.
(***cond2***)	This separatist movement or terrorist organization currently practices ransom kidnapping of specific categories of people.
(***cond3***)	In particular, executives or assimilated categories are concerned.
(***cond4***)	It can be proven that the kidnapped is really a businessperson or assimilated.

Table 15
'Kidnapping' hypothesis in the presence of transformations.

(***cond1***) The kidnappers are part of a separatist movement or of a terrorist organization.
- (**Rule *t3*, *Consequent1***) *Try to verify whether a given separatist movement or terrorist organization is in strict control of a specific sub-group and, in this case,*
- (**Rule *t3*, *Consequent2***) *check if the kidnappers are members of this sub-group. We will then assimilate the kidnappers to 'members' of the movement or organization.*

(***cond2***) The movement/organization performs ransom kidnapping of specific categories of people.
- (**Rule *t2*, *Consequent***) *The family of the kidnapped has received a ransom request from the separatist movement or terrorist organization.*
- (**Rule *t4*, *Consequent1***) *The family of the kidnapped has received a ransom request from a group or an individual person, and*
- (**Rule *t4*, *Consequent2***) *this second group or individual person is part of the separatist movement or terrorist organization.*
- (**Rule *t5*, *Consequent1***) *Try to verify if a particular sub-group of the separatist movement or terrorist organization exists, and*
- (**Rule *t5*, *Consequent2***) *check whether this particular sub-group practices ransom kidnapping of specific categories of people.*
- …

(***cond3***) In particular, executives or assimilated categories are concerned.
- (**Rule *t0*, *Consequent1***) *In a 'ransom kidnapping' context, we can check whether the kidnapped person has a strict kinship relationship with a second person, and*
- (**Rule *t0*, *Consequent2***) *we can then check if this second person is a businessperson or assimilated.*

(***cond4***) It can be proven that the kidnapped person is really an executive or assimilated.
- (**Rule *t6*, *Consequent***) *In a 'ransom kidnapping' context, 'personalities' like physicians, journalists, artists etc. can be assimilated to businesspersons.*

For example, the proof that the kidnappers are part of a terrorist group or separatist organization (***cond1*** of Table 14) can be now obtained indirectly, transformation ***t3*** of Table 15, by checking whether they are members of a specific subset of this group/organization. Note that transformations ***t2*** and ***t6*** of Table 15 imply only one step of reasoning (only one consequent), whereas all the residual transformations are "multi-consequent". We can see, in particular, that there is a *whole family of transformations* corresponding to the condition schemata ***cond2*** of Table 15. They all correspond to variants of this general scheme: the separatist movement or terrorist organization, or some group or single person affiliated with them, have requested/received money for the hostage's ransom.

5. Related work

As already stated in the "Introduction", the sentiment analysis domain normally makes use of "*surface*", statistically-oriented computational linguistics and text mining techniques. See, in this context, the broad bibliography put together by Jan Wiebe: http://www.cs.pitt.edu/~wiebe/subjectivityBib.html. These techniques are conceptually very different, then, from the "*deep analysis*" tools used in an NKRL context.

Likening, however, NKRL's templates to some sort of highly expressive "*semantic*" *patterns*, some equivalence can be found between the NKRL's approach and the work described in Riloff and Wiebe (2003). This work aims, in fact, at deriving "*extraction patterns*" for recognizing "subjective sentences", even if these extraction patterns concern mainly the syntactic/surface level. A related work is Cardie, Wiebe, Wilson, and Litman (2004), which evokes, among other things, the "elementary events/full narratives and scenarios" dichotomy in NKRL. It relies on a set of "*opinion-oriented template relations*" to identify the different expressions of opinion in a text along with their source, type and strength. Once recognized, these low-level relations can be combined to create an *opinion-based scenario*, i.e., a summary representation of the opinions expressed in a document, a group of documents, or an arbitrary text span. Detection of "opinionated expressions" making use of relational features derived from grammatical and semantic role structures is described in Johansson and Moschitti (2010). From an "inference/reasoning" point of view, Cambria, Song, Wang, and Howard (2014) illustrate how the "Semantic Multi-Dimensional Scaling" approach can be used to implement reasoning techniques able to infer general conceptual and affective information from a large common-sense knowledge base.

However, it is indubitable that, to find valid analogies between NKRL and comparable tools, it is necessary to look at that class of *pure AI-based systems* that inherit from both the "*semantic network*" tradition (Lehmann, 1992) and from the "*conceptual approaches*" originally popularized by Schank (1973). These systems include, e.g., CYC (Lenat & Guha, 1990), Conceptual Graphs (Sowa, 1984, 2000) and Topic Maps (Pepper, 2000).

CYC is an important system that utilizes a huge knowledge base containing about a million of hand-entered "logical assertions". The base includes both simple statements of facts and rules about what conclusions can be inferred if certain statements of facts are satisfied. The "*upper level*" of the CYC ontology (OpenCyc) is now freely accessible on the Web at http://www.cyc.com/cyc/opencyc. The *n*-ary knowledge representation language utilized by CYC is called CycL. A criticism often addressed to this language concerns its uniform use of the same representation model (substantially, a frame system rewritten in logical form) to represent phenomena conceptually very different (the "*uniqueness syndrome*"). In NKRL, on the contrary, concepts are represented in the (usual) binary way, elementary events/situations (and general classes of events/situations) like *n*-ary predicate/roles-based structures, the connectivity phenomena as labeled lists with reified arguments. Moreover, special conceptual structures have been conceived to take the temporal phenomena into account. With respect to Topic Maps, and given their (at least apparent) simplicity, they are often only considered (surely wrongly) as a sort of *downgraded version* of Semantic Networks, strongly influenced as well by other conceptual models like CGs.

CGs are based on a powerful *graph-based representation scheme* that can be used to represent *n*-ary relationships between complex objects in a global system. For example, a CG corresponding to the narrative "John is going to Boston by bus" is represented by a conceptual structure where a "concept node", "Go" (*functionally equivalent* to an NKRL *primitive* predicate) is associated with three "relation nodes" (roles) like **Agnt**, **Dest** and **Inst**. These 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"). The resemblance to *HTemp and to the NKRL representation of elementary events* is evident. Moreover, for any CGs system, it is assumed that there exists a *pre-defined type hierarchy of "concept-types"*, different according to the domain to formalize, similar then to HClass.

Notable differences between the NKRL and the CGs approach however exist. They concern, among other things, the nature of the predicates (*primitives in NKRL and free surface terms in CGs*, see "Go" in the previous "John is going..." example), the choice and definition of the roles/relation nodes, etc. A remark often raised about CGs concerns the use of the "*canonical graphs*". In a CGs context, canonical graphs are general conceptual structures, similar in principle to NKRL templates, which could be used for describing complex, dynamic and semantically-rich phenomena like narratives/behaviors. However, unlike what has been done for the NKRL templates, *an exhaustive and authoritative list of these graphs*, equivalent then to the HTemp hierarchy illustrated in Fig. 1, *does not exist*. Note also that, because of the "free" nature of the CGs' predicates, the set up of a *universally acceptable* list of canonical graphs would probably be impossible. The practical consequence of this state of affairs is often the need, whenever a concrete application of CGs must be implemented, of *defining anew for this application a distinct list of canonical graphs*. This contrasts with a fundamental characteristic of NKRL, where its catalog of "basic templates" (HTemp) is, in practice, part and parcel of the definition of the language.

A system of Computational Linguistics origin that, at least in principle, could be used to represent complex narratives/behaviors is the "Text Meaning Representation" model, TRM (Nirenburg & Raskin, 2004). Some comments and criticisms about TRM can be found in Zarri (2009: 146–149).

6. Conclusion

The "sentiment analysis" (opinion mining) domain concerns all those applications that try to identify and extract "*subjective information*" (opinions, beliefs, emotional states and views about specific entities) from source materials. In this paper, we have focused our attention on the fundamental aspects of *production and fruition of subjective information* that are associated with the everyday "*behavior*" (in the most general meaning of this term) of human and non-human characters.

We have then supplied an extremely condensed overview of a knowledge representation language, NKRL, which deals with this behavior when it is expressed in the form of non-fictional and real-time "*narratives*". These are logically and temporally structured sequences of elementary events, similar then to other high-level, spatio-temporally constrained information structures like "scenarios", "situations", "complex events", etc. After having supplied the basic syntactic and semantic principles of NKRL, we have discussed in depth the *templates* (the **Behave:** templates in particular) that represent the main interest of the language from a sentiment analysis point of view. The inference procedures that can be used in an NKRL context have also been briefly discussed.

In general, we will refer now to the recent IEEE Computational Intelligence Magazine review on NLP (Natural Language Processing) research (Cambria & White, 2014) and to progress in this domain that, according to the Authors of the review, could be associated with the possibility of "*jumping the NLP curves*". We can note, in this context, that NKRL is certainly located at the frontier between the "*Semantics*" and "*Pragmatics*" curves. The former field is characterized by the use of a "*bag-of-concepts*" model: NKRL makes, in fact, an intensive usage of (an accurately structured) bag-of-concepts under the form of the HClass hierarchy. However, NKRL has also more than a foot in the pragmatic curve field. According to Cambria and White (2014: 52), this last is characterized by the use of a "*bag-of-narratives*" model, where "...each piece of text will be represented by ministories or interconnected episodes, leading to a more detailed level of text comprehension and sensible computation". It is easy then to associate the "ministories or interconnected episodes" with NKRL's "elementary events" and with their interconnecting conceptual tools (binding occurrences, completive construction). According to the Authors of the review, jumping from the Semantics Curve to the Pragmatics Curve is *mandatory* given that this move will enable NLP "...to be more adaptive and, hence, open-domain, context-aware, and intent-driven. Intent, in particular, will be key for tasks such as sentiment analysis ...". NKRL seems then to be (at least in principle) well placed to play an important role not only with respect to the progress of the sentiment analysis/opinion mining domain but also of NLP in general.

We can conclude the paper by noticing that, apart from being a knowledge representation language, NKRL is also a *fully operational computer science environment*, implemented in Java and built up thanks to several European projects. A (relatively in-depth) description of the NKRL software can be found in Zarri (2009, Appendix A). The environment exists in two versions, a (standard) SQL-based version and a (lighter) file-oriented one, to be used mainly as a "demonstration" version. The environment includes also powerful "inference engines" able to carry out complex inference procedures based, e.g., on "analogical" and "causal" reasoning principles, see Section 4 above.

References

Ahn, H.-il (2010). *Modeling and analysis of affective influences on human experience, prediction, decision making, and behavior* (Ph.D. thesis), Cambridge, MA: MIT.

Ayari, N., Chibani, A., & Amirat, Y. (2013). Semantic management of human–robot interaction in ambient intelligence environments using n-ary ontologies. In *Proceedings of the 2013 IEEE international conference on robotics and automation* (pp. 1164–1171). Piscataway, NJ: IEEEXplore.

Balahur, A., Steinberger, R., Kabadjov, M., Zavarella, V., van der Goot, E., Halkia, M., et al. (2010). Sentiment analysis in the news. In *Proceedings of the 7th conference on language resources and evaluation* (pp. 2216–2220). Paris: European Language Resources Association (ELRA).

Cambria, E., Havasi, C., & Hussain, A. (2012). SenticNet 2: a semantic and affective resource for opinion mining and sentiment analysis. In *Proceedings of the twenty-fifth international Florida artificial intelligence research society conference* (pp. 202–207). Menlo Park, CA: AAAI Press.

Cambria, E., Olsher, D., & Rajagopal, D. (2014) SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In *Proceedings of AAAI 2014*. Available at: http://sentic.net/senticnet-3.pdf, Accessed 16.05.14.

Cambria, E., Song, Y., Wang, H., & Howard, N. (2014). Semantic multi-dimensional scaling for open-domain sentiment analysis. *IEEE Intelligent Systems*, *29*(2), 2–9.
Cambria, E., Speer, R., Havasi, C., & Hussain, A. (2010). SenticNet: a publicly available semantic resource for opinion mining. In *Papers from the AAAI fall symposium on commonsense knowledge (FS-10-02)* (pp. 14–18). Menlo Park, CA: AAAI Press.
Cambria, E., & White, B. (2014). Jumping NLP curves: a review of natural language processing research. *IEEE Computational Intelligence Magazine*, *9*(2), 48–57.
Cao, L., & Yu, P. S. (Eds.) (2012). *Behavior computing: modeling, analysis, mining and decision*. Berlin: Springer.
Cardie, C., Wiebe, J., Wilson, T., & Litman, D. (2004). Low-level annotations and summary representations of opinions for multi-perspective question answering. In *New directions in question answering* (pp. 87–98). Menlo Park, CA: AAAI Press/MIT Press.
Ding, X., Liu, B., & Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. In *Proceedings of the 2008 ACM international conference on web search and data mining* (pp. 231–239). New York: ACM.
Esuli, A., & Sebastiani, F. (2006). SentiWordNet: a publicly available lexical resource for opinion mining. In *Proceedings of the 5th conference on language resources and evaluation* (pp. 417–422). Paris: European Language Resources Association (ELRA).
Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89.
Johansson, R., & Moschitti, A. (2010). Syntactic and semantic structure for opinion expression detection. In *Proceedings of the fourteenth conference on computational natural language learning* (pp. 67–76). Stroudsburg, PA: ACL.
Kaji, N., & Kitsuregawa, M. (2007). Building lexicon for sentiment analysis from massive collection of HTML documents. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 1075–1083). Stroudsburg, PA: ACL.
Lehmann, F. (1992). *Semantic networks in artificial intelligence*. Oxford: Pergamon Press.
Lenat, D. B., & Guha, R. V. (1990). *Building large knowledge based systems*. Reading, MA: Addison-Wesley.
Liu, H., & Singh, P. (2004). ConceptNet—a practical commonsense reasoning tool-kit. *BT Technology Journal*, *22*(4), 211–226.
Mani, I., & Pustejovsky, J. (2004). Temporal discourse models for narrative structure. In *Proceedings of the ACL workshop on discourse annotation* (pp. 57–64). Stroudsburg, PA: ACL.
Nirenburg, S., & Raskin, V. (2004). *Ontological semantics*. Cambridge, MA: The MIT Press.
Noy, F. N., Fergerson, R. W., & Musen, M. A. (2000). The knowledge model of Protégé-2000: combining interoperability and flexibility. In *LNCS*: *Vol. 1937*. *Knowledge acquisition, modeling, and management—Proceedings of EKAW'2000* (pp. 17–32). Berlin: Springer.
Pepper, S. (2000). *The TAO of topic maps: finding the way in the age of Infoglut*. Oslo: Ontopia, AS, Available at: http://www.ontopia.net/topicmaps/materials/tao.html, Accessed 16.05.14.
Riloff, E., & Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 105–112). Stroudsburg, PA: ACL SIGDAT.
Schank, R. C. (1973). Identification of conceptualizations underlying natural language. In *Computer models of thought and language* (pp. 187–247). San Francisco: W.H. Freeman and Co.
Sowa, J. F. (1984). *Conceptual structures: information processing in mind and machine*. Reading, MA: Addison-Wesley.
Sowa, J. F. (2000). *Knowledge representation—logical, philosophical, and computational foundations*. Pacific Grove, CA: Brooks/Cole.
Speer, R., & Havasi, C. (2012). Representing general relational knowledge in ConceptNet 5. In *Proceedings of the 8th conference on language resources and evaluation* (pp. 3679–3686). Paris: European Language Resources Association (ELRA).
Strapparava, C., & Valitutti, A. (2004). Wordnet-affect: an affective extension of WordNet. In *Proceedings of the 4th conference on language resources and evaluation* (pp. 1083–1086). Paris: European Language Resources Association (ELRA).
Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, *37*, 267–307.
Westerski, A. (2007). Sentiment analysis: introduction and the state of the art overview. *Text*, 211–218. Available at: http://www.adamwesterski.com/wp-content/files/docsCursos/sentimentA_doc_TLAW.pdf, Accessed 16.05.14.
Zarri, G. P. (1983). Automatic representation of the semantic relationships corresponding to a French surface expression. In *Proceedings of the first international conference on applied natural language processing* (pp. 143–147). Stroudsburg, PA: ACL.
Zarri, G. P. (2005). Integrating the two main inference modes of NKRL, transformations and hypotheses. *Journal on Data Semantics (JoDS)*, *4*, 304–340.
Zarri, G. P. (2009). *Representing and managing of narrative information—theoretical principles and implementation*. London: Springer.
Zarri, G. P. (2011). Differentiating between "functional" and "semantic" roles in a high-level conceptual data modeling language. In *Proceedings of the 24th international Florida AI research society conference* (pp. 75–80). Menlo Park, CA: AAAI Press.
Zarri, G. P. (2013). Advanced computational reasoning based on the NKRL conceptual model. *Expert Systems with Applications (ESWA)*, *40*, 2872–2888.