“Narrative” Information and the NKRL Solution

Gian Piero Zarri
LaLIC, University Paris 4-Sorbonne, France

INTRODUCTION

In a companion article of this Encyclopaedia: ‘Narrative’ Information, the Problem, we have introduced the problem of finding a complete and computationally efficient system for representing and managing ‘*nonfictional narrative information*’. We have stressed there the *important economic value of this multimedia type of information* – that concerns, e.g., corporate memory documents, news stories, normative and legal texts, medical records, intelligence messages, surveillance videos or visitor logs, actuality photos, eLearning and Cultural Heritage material, etc. We have also emphasised that the usual Computer Science tools – including those pertaining to the now very popular ‘Semantic Web’ domain, see (Bechhofer *et al.*, 2004, Beckett, 2004) – *are not really suitable* for dealing with this type of information.

BACKGROUND

In this article, we will present an Artificial Intelligence tool, NKRL (Narrative Knowledge Representation Language) that has been especially developed for dealing in an ‘intelligent’ way with the nonfictional narrative information. NKRL is, at the same time:

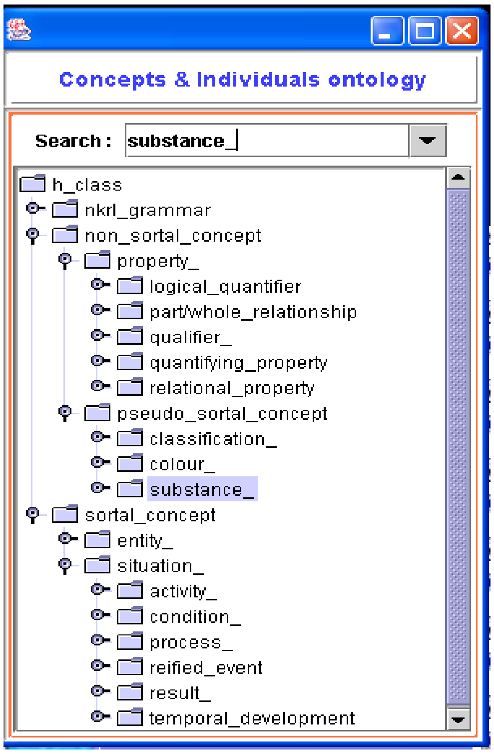
- a *knowledge representation system* for describing in the best possible detail the essential content (the ‘meaning’) of complex nonfictional ‘narratives’;
- a *system of reasoning (inference) procedures* that, thanks to the richness of the representation system, is able to automatically establish ‘interesting’ relationships among the represented data;
- an *implemented software environment* that allows the user to encode the original narratives in terms of the representation language to create ‘NKRL knowledge bases’ in a specific application domain and to exploit ‘intelligently’ these bases.

The main innovation introduced by NKRL with respect to the usual ontological paradigms concerns the addition to the traditional *ontology of concepts* – called HClass, ‘hierarchy of classes’ in the NKRL’s jargon – an *ontology of events*, i.e., a new sort of hierarchical organization where the nodes correspond to *n*-ary structures called ‘*templates*’ (HTemp, ‘hierarchy of templates’). A partial image of the ‘upper level’ of HClass – that follows then the standard Protégé approach, see (Noy *et al.*, 2000) – is given in Figure 1; for HTemp, see Table 1 and Figure 2 below.

A SHORT DESCRIPTION OF NKRL

Instead of using the traditional (binary) *attribute/value* organization, the templates are generated from the

Figure 1. A partial representation of the ‘upper level’ of HClass, the NKRL ‘traditional’ ontology of concepts.



n-ary combination of quadruples connecting together the *symbolic name* of the template, a *predicate*, and the *arguments* of the predicate introduced by named relations, the *roles*. The quadruples have in common the *name* and *predicate* components. Denoting then with L_i the generic symbolic label identifying a given template, with P_j the predicate used in the template, with R_k the generic role and with a_k the corresponding argument, the core data structure for templates has the following general format (see also the companion article, 'Narrative' Information, the Problem):

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

Predicates pertain to the set {BEHAVE, EXIST, EXPERIENCE, MOVE, OWN, PRODUCE, RECEIVE}, and roles to the set {SUBJ(ect), OBJ(ect), SOURCE, BEN(e)F(iciary), MODAL(ity), TOPIC, CONTEXT}. An argument of the predicate can consist of a simple 'concept' or of a structured association ('expansion') of several concepts. Templates can be conceived as the *formal representation* of *generic classes of elementary events* like "move a physical object", "be present in a place", "produce a service", "send/receive a message", etc. When a particular event pertaining to one of these general classes must be represented, the corresponding template is *instantiated* to produce a *predicative occurrence*.

To represent then a simple narrative like: "On November 20, 1999, in an unspecified village, an armed group of people has kidnapped Robustiniano Hablo", we must then select firstly in the HTemp hierarchy the template corresponding to "execution of violent actions", see Figure 2 and Table 1 below – this example refers to a recent application of NKRL in a 'terrorism' context in the framework of an European project see, e.g., (Zarri, 2005).

As it appears from Table 1a, the arguments of the predicate (the a_k terms in (1)) are represented by *variables with associated constraints* expressed as HClass concepts or combinations of concepts. When deriving a predicative occurrence (an instance of a template) like mod3.c5 in Table 1b, the role fillers in this occurrence must conform to the constraints of the father-template. For example, ROBUSTINIANO_HABLO (the 'BEN(e)F(iciary)' of the action of kidnapping) and INDIVIDUAL_PERSON_20 (the unknown 'SUBJECT', actor, initiator etc. of this action) are both 'individuals', instances of the HClass concept individual_person. The constituents – as SOURCE in Table 1a – included in square brackets are optional. A 'conceptual label' like mod3.c5 is the symbolic name used to identify the NKRL code corresponding to a specific predicative occurrence.

The 'attributive operator', SPECIF(ication), is one of the four operators used in NKRL for the construction of 'structured arguments' ('complex fillers' or 'expansions') see, e.g., (Zarri, 2003). The SPECIF lists, with syntax (SPECIF $e_i\ p_1 \ldots p_n$), are used to represent the properties or attributes that can be asserted about the first element e_i, concept or individual, of the list – e.g., in the SUBJ filler of mod3.c5, Table 1b, the attributes weapon_wearing and (SPECIF cardinality_ several_)) are both associated with INDIVIDUAL_PERSON_20.

The 'location attributes', represented in the predicative occurrences as lists, are linked with the arguments of the predicate by using the colon operator, ':', see the individual VILLAGE_1 in Table 1b. In the occurrences, the two operators date-1, date-2 materialize the temporal interval normally associated with narrative events, see (Zarri, 1998) – and, more in general, (Allen, 1981, Ferro *et al*., 2005).

150 templates are permanently inserted into HTemp; Figure 2 reproduces the 'external' organization of the PRODUCE branch of HTemp. This branch includes the Produce:Violence template used in Table 1. HTemp corresponds then to a sort of 'catalogue' of narrative formal structures, that are very easy to 'customize' to derive the new templates that could be needed for a particular application.

What expounded until now illustrates the NKRL solutions to the problem of representing 'elementary' (simple) events. To deal now with those 'connectivity phenomena' that arise when several elementary events are connected through causality, goal, indirect speech etc. links – see also (Mani and Pustejovsky, 2004) – the basic NKRL knowledge representation tools have been complemented by more complex mechanisms that make use of second order structures, see (Zarri, 2003). For example, the *binding occurrences* consist of lists of symbolic labels (c_i) of predicative occurrences; the lists are differentiated using specific binding operators like GOAL, CONDITION and CAUSE. Let us suppose that, in Table 1, we state now that: "…an armed group of people has kidnapped Robustiniano Hablo *in order to* ask his family for a ransom", where the new elementary event: "the unknown individuals will ask for a ransom" corresponds to a new predicative occurrence, e.g., mod3.

Related Content