

Chapter 8.11

Representation Languages for Narrative Documents

Gian Piero Zarri

University of Paris IV/Sorbonne, France

INTRODUCTION

A big amount of important, “economically relevant” information, is buried into unstructured “narrative” information resources: This is true, for example, for most of the corporate knowledge documents (memos, policy statements, reports, minutes, etc.), for the news stories, the normative and legal texts, the medical records, many intelligence messages as well as for a huge fraction of the information stored on the Web. In these “narrative documents,” or “narratives,” the main part of the information content consists in the description of “events” that relate the real or intended behavior of some “actors” (characters, personages, etc.)—the term “event” is taken here in its more general meaning, also covering strictly related notions like fact, action, state, and situation. These actors try to attain a specific result, experience particular situations, manipulate some

(concrete or abstract) materials, send or receive messages, buy, sell, deliver, and so forth. Note that in these narratives, the actors or personages are not necessarily human beings; we can have narrative documents concerning, for example, the vicissitudes in the journey of a nuclear submarine (the “actor,” “subject,” or “personage”) or the various avatars in the life of a commercial product. Note also that even if a large amount of narrative documents concerns natural language (NL) texts, this is not necessarily true. A photo representing a situation that verbalized could be expressed as “Three nice girls are lying on the beach” is not of course an NL text, yet it is still a narrative document.

Because of the ubiquity of these “narrative” resources, being able to represent in a general, accurate, and effective way their semantic content—that is, their key “meaning”—is then both conceptually relevant and economically impor-

tant: Narratives form, in fact, a huge underutilized component of organizational knowledge. This type of explicit yet unstructured knowledge can be, of course, indexed and searched in a variety of ways, but it requires, however, an approach for formal analysis and effective utilization that is neatly different from the “traditional” ones.

BACKGROUND

Usual ontologies—both in their “traditional” and “Semantic Web” versions (see the “Knowledge Representation” and “RDF and OWL” articles in this Encyclopedia)—are not very suitable for dealing with narratives. Basically, ontologies organize the “concepts”—that we can identify here with the important notions to be represented in a given application domain – into a hierarchical structure, able to supply an elementary form of definition of these concepts through their mutual generic/specific relationships (“IsA” links). A more detailed definition of the concepts is obtained by associating them with a set of binary relationships of the “property/value” type (e.g., a “frame”). The combination of these two representational principles is largely sufficient to provide a static, a priori definition of the concepts and of their properties.

Unfortunately, this is no more true when we consider the dynamic behavior of the concepts, that is, we want to describe their mutual relationships when they take part in some concrete action or situation (“events”). First of all, representing an event implies that the notion of “role” must be added to the traditional generic/specific and property/value representational principles. If we want to represent adequately a narrative fragment like “NMTV (an European media company) ... will develop a lap top computer system ...,” besides asserting that NMTV_ is an instance of the concept company_ and that we also must introduce an instance of a concept like lap_top_pc, we have to create a sort of

“threefold” relationship; this relationship includes a “predicate” (like DEVELOP or PRODUCE), the two instances, and a third fundamental component, the “roles” (like SUBJECT or AGENT for NMTV_ and OBJECT or PATIENT for the new lap top system) used to specify the exact function of these two instances within the formal description of the event. Moreover, in an event context, we also must deal with those “connectivity phenomena” like causality, goal, indirect speech, co-ordination, and subordination that link together the basic “elementary events.” It is very likely, in fact, that dealing with the sale of a company, the global information to represent is something like: “Company X has sold its subsidiary Y to Z because the profits of Y have fallen dangerously these last years due to a lack of investments,” or, returning to our previous example, that “NMTV will develop a lap top computer system to put controlled circulation magazines out of business,” or, that dealing with the relationships between companies in the biotechnology domain, “X made a milestone payment to Y because they decided to pursue an in vivo evaluation of the candidate compound identified by X.” In computational linguistics terms, we are here in the domain of the “Discourse Analysis” which deals, in short, with the two following problems: (1) determining the nature of the information that in a sequence of statements goes beyond the simple addition of the information conveyed by a single statement; (2) determining the influence of the context in which a statement is used on the meaning of this individual statement or part of it.

It is now easy to imagine the awkward proliferation of binary relationships that sticking to the traditional ontological paradigm it would be necessary to introduce to approximate high-level notions like those of “role” and “connectivity phenomena.”

Solutions for representing narratives in computer-usable ways that could move beyond a strict “binary” framework have, therefore, already

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/representation-languages-narrative-documents/25343

Related Content

Secure Knowledge Management for Healthcare Organizations

Darren Mundy and David W. Chadwick (2008). *Knowledge Management: Concepts, Methodologies, Tools, and Applications* (pp. 2168-2181).

www.irma-international.org/chapter/secure-knowledge-management-healthcare-organizations/25250

Conflict of Interest: What It Is, Its Causes and Consequences

Yanli Zhang, Jeffrey Hsu, Ruben Xing and Wencang Zhou (2015). *International Journal of Knowledge-Based Organizations* (pp. 1-18).

www.irma-international.org/article/conflict-of-interest/125582

Structural Capital and Innovation Capabilities: Theory and Empirical Evidence from Spain

Elsa Mercedes Alama Salazar (2010). *Intellectual Capital and Technological Innovation: Knowledge-Based Theory and Practice* (pp. 194-208).

www.irma-international.org/chapter/structural-capital-innovation-capabilities/44323

Assessing Travel Websites Based on Service Quality Attributes Under Intuitionistic Environment

Abhishek Tandon, Himanshu Sharma and Anu Gupta Aggarwal (2019). *International Journal of Knowledge-Based Organizations* (pp. 66-75).

www.irma-international.org/article/assessing-travel-websites-based-on-service-quality-attributes-under-intuitionistic-environment/216841

How to Accumulate Empirical Engineering Knowledge in the Complex Problem-Solving Process for Novice Engineers

Lijun Liu and Zuhua Jiang (2021). *International Journal of Knowledge Management* (pp. 1-20).

www.irma-international.org/article/how-to-accumulate-empirical-engineering-knowledge-in-the-complex-problem-solving-process-for-novice-engineers/269383