1. Record Nr.   TD20015123

Autore   CASELLI, TOMMASO

Titolo   Time, events and temporal relations: an empirical model for temporal processing of Italian texts [Tesi di dottorato]

Editore   Pisa University, 2009-04-26

Lingua di pubblicazione   Italiano

Formato   Tesi di dottorato

Livello bibliografico   Monografia

Note   diritti: info:eu-repo/semantics/openAccess diritti: Copyright information available at source archive

Sommario   The aim of this work is the elaboration a computational model for the identification of temporal relations in text/discourse to be used as a component in more complex systems for Open-Domain Question-Answers, Information Extraction and Summarization. More specifically, the thesis will concentrate on the relationships between the various elements which signal temporal relations in Italian texts/discourses, on their roles and how they can be exploited. Time is a pervasive element of human life. It is the primary element thanks to which we are able to observe, describe and reason about what surrounds us and the world. The absence of a correct identification of the temporal ordering of what is narrated and/or described may result in a bad comprehension, which can lead to a misunderstanding. Normally, texts/discourses present situations standing in a particular temporal ordering. Whether these situations precede, or overlap or are included one within the other is inferred during the general process of reading and understanding. Nevertheless, to perform this seemingly easy task, we are taking into account a set of complex information involving different linguistic entities and sources of knowledge. A wide variety of devices is used in natural languages to convey temporal information. Verb tense, temporal prepositions, subordinate conjunctions, adjectival phrases are some of the most obvious. Nevertheless even these obvious

devices have different degrees of temporal transparency, which may sometimes be not so obvious as it can appear at a quick and superficial analysis. One of the main shortcomings of previous research on temporal relations is represented by the fact that they concentrated only on a particular discourse segment, namely narrative discourse, disregarding the fact that a text/discourse is composed by different types of discourse segments and relations. A good theory or framework for temporal analysis must take into account all of them. In this work, we have concentrated on the elaboration of a framework which could be applied to all text/discourse segments, without paying too much attention to their type, since we claim that temporal relations can be recovered in every kind of discourse segments and not only in narrative ones. The model we propose is obtained by mixing together theoretical assumptions and empirical data, collected by means of two tests submitted to a total of 35 subjects with different backgrounds. The main results we have obtained from these empirical studies are: (i.) a general evaluation of the difficulty of the task of recovering temporal relations; (ii.) information on the level of granularity of temporal relations; (iii.) a saliency-based order of application of the linguistic devices used to express the temporal relations between two eventualities; (iv.) the proposal of tense temporal polysemy, as a device to identify the set of preferences which can assign unique values to possibly multiple temporal relations. On the basis of the empirical data, we propose to enlarge the set of classical finely grained interval relations (Allen, 1983) by including also coarse-grained temporal relations (Freska, 1992). Moreover, there could be cases in which we are not able to state in a reliable way if there exists a temporal relation or what the particular relation between two entities is. To overcome this issue we have adopted the proposal by Mani (2007) which allows the system to have differentiated levels of temporal representation on the basis of the temporal granularity associated with each discourse segment. The lack of an annotated corpus for eventualities, temporal expressions and temporal relations in Italian represents the biggest shortcomings of this work which has prevented the implementation of the model and its evaluation. Nevertheless, we have been able to conduct a series of experiments for the validation of procedures for the further realization of a working prototype. In addition to this, we have been able to implement and validate a working prototype for the spotting of temporal expressions in texts/discourses.

Localizzazioni e accesso	http://memoria.depositolegale.it/*/http://etd.adm.unipi.it/theses/available/etd-04242009-113147/