

Towards a mechanical analysis of French tense forms in texts.¹⁾

Christian Rohrer

Universität Stuttgart

In this paper we want to present a system that analyzes tense forms and temporal adverbs in texts.²⁾ For reasons of exposition we restrict the analysis to narrative texts. The system comprises the following components:

- a) A parser which assigns a tree structure to sentences. The fragment which the parser can analyse corresponds roughly to the fragment of English described by R. Montague in 'The proper treatment of quantification in ordinary English'. Our fragment contains however more tense forms, temporal adverbs, and temporal conjunctions. In particular we treat the notoriously difficult pair of tense forms passé simple and imparfait, frame adverbials like aujourd'hui, hier, demain, ce jour-là, un jour plus tard, etc., as well as the conjunctions quand, pendant que, depuis que, après que.
- b) The syntactic analysis constitutes the input to the rules that derive the corresponding discourse representations. The syntactic structure determines the discourse representation.

Let $D = S_1 \dots \dots \dots, S_n$ be a discourse of L .

A D(iscourse) R(epresentation) S(tructure) for D is constructed by reducing the sentences $S_1 \dots \dots \dots, S_n$ in the order in which they occur in D . The reduction of a sentence S_i proceeds through the application of certain DRS-construction rules which operate on the syntactic analysis of the sentence and work so to speak 'from the top down' (thus the syntactic analysis imposes a partial order on how the rules are to be applied).

What DRS-construction rule is to be applied depends on the particular syntactic formation rule that was used to form the syntactic compound that the application of the rule is to reduce. Where one of the immediate components of the compound is a singular term, moreover, the choice of construction rule will be determined by the type of this term; thus, for instance there is a different rule for proper names, for indefinite descriptions and for pronouns, respectively. Similarly there are different rules for each of the tenses considered here.

- c) In ordinary truth-conditional semantics truth conditions are stated for individual sentences: A discourse consisting of a set of sentences is true iff it consists of nothing but true sentences. The order in which the individual sentences occur in the discourse is irrelevant for the truth conditions. In this system truth conditions are formulated for a sequence of sentences $D = S_1, \dots, S_n$ in such a way that the order in which the sentences occur in the text plays a crucial role. The truth of a discourse D in a model M is defined as the existence of a proper embedding into M of the representation of D .

A major advantage of the present system lies in its ability to deal with cases of intersentential and intrasentential anaphora in exactly the same fashion. Tense forms and temporal adverbs behave in many respects like anaphoric pronouns. The DRS-construction rules handle both temporal and pronominal anaphora. The system extracts all the temporal information which is contained in the input text. At a later stage the information extracted from the text will form the basis for a question - answering system. An implementation of the system on a VAX 780/11 is in progress.

- 1) The work reported in this paper was supported by the Deutsche Forschungsgemeinschaft grant Ro 245/12.
- 2) The system is based on the following papers: Kamp, H. (1980) 'A theory of truth and semantic representation', in: Groenendijk, J., Janssen, T., Stokhof, M. (eds.), Formal methods in the study of language, Amsterdam, p. 277 - 322.
Kamp, H. (1981), 'Evènements, représentations discursives et référence temporelle', Langages, p. 39 - 64.
Kamp, H., Rohrer C. (1981) 'Tense in texts', Ms. Stuttgart