A Novel Framework for Extraction and Representation of Event-Time Relation in Natural language Text

ABSTRACT

A Novel framework for automatic extraction and representation of event-time information is an essential area of research in Computational Linguistics and Artificial Intelligence. In linguistic terms, events are occurrences, situations, or happenings that utter with a period of timestamp. In natural language processing (NLP), time is a kind of ontology. Time exists in quantitative or qualitative form, and it can be used for temporal reasoning to increase temporal awareness. The properties associated with events change over time. Processing of event and time information addresses few challenges in NLP and clues towards the applications like question answering, information retrieval, and text summarization.

Several models exist in the literature for extracting and representing the events and time from natural language text. Existing models are useful in obtaining lexical, syntactic features of events and time. Very few models addressed semantic features of event and time information, but most of them are constrained to a particular domain.

The thesis aims to build a frame work for the extraction and representation of events, times, and event-time relationship from natural language text. Various techniques NLP tools such as Conditional Random field, Semantic Role labels, and WordNet are used to achieve the task. Extracted event-time information and relations are further applied to a temporal question answering system to check the working of the developed model.

The significant contribution of the thesis focused on extracting event and time expression/ with lexical, syntactic, and semantic features. Events are noticeable if that falls within time, but in the majority of the cases, they may not fall within a sequence or in a timeline. This thesis work presents a frame

work that addresses the challenges related to time-dependent events information. The information obtained is further used in the temporal question answering system to answer the time-sensitive queries that occur in this dynamic world.

Message Understanding Conference (MUC), Semantic Evaluation (SEMEVAL), Text Retrieval Conference (TREC), Forum for Information retrieval (FIRE), News articles and Wikipedia articles are collected to evaluate the model. The standard evaluation metrics are used to assess the results of the proposed framework. The obtained precision values are for event extraction 91%, for time extraction is 85% and for event time relations applying on temporal question answering got 84% of precision. The results have shown better performance when compared with other models.