10. Conclusion

The thesis attempts to provide a contribution to the problem of temporal reasoning in Natural Language Processing (NLP). The key achievements concern both the theoretical framework presented in the first part of the work and its application to automated multi-document text summarization problem discussed in the second part.

With respect to the automatic reasoning framework, representation of duration information is provided as an important extension to the TIMEML representation language. Duration information of non-instantaneous events influences the positioning of the events on a timeline and therefore provides the order of the events to be presented in the text. The theoretical framework provides automatic extraction, representation and reasoning with time expressions of events in natural language text. Furthermore the results of the general Constraint Satisfaction Problem (CSP) research are extended to reason with the events occurrences inorder to infer underspecified and incomplete information. Algorithms of polynomial complexity are devised for resolution of event occurrences and chronological ordering of events on a timeline. Even when the exact time and duration of events occurrences is unknown, the relative ordering of the end-points and durations are utilized to impose an ordering of the events on the timeline.
A short contribution of the thesis is concerned with the task of multi-document summarization. The thesis contributes an additional level of refinement by extracting the temporal characteristics of the events, segmenting text and relating segments of multiple documents based on the events described by them. Relating event temporal characteristics relative to a set of one or more related events in the same as well as multiple documents imposed a graphical representation on the documents.

The notion of salience of the events, and coherence of the summary generated are not easy to evaluate. Nonetheless, the generated summaries of proposed model have been tested with respect to F-measure and compared the results against the dataset used for contest on text summarization.

A particular rewarding aspect of the research was that, the F-measure of proposed approach was much better than semantic based approaches which require linguistic analysis for the text understanding.

Future work of this thesis ranges to specific tasks, enhancing the framework with spatial aspects to wider and more open problems related to text understanding such as question answering system.

Within the first direction there is the possibility of extracting the spatial information of the events to reason out time and place of event occurrences enhancing the text understanding capability of the model.
A second enhancement of the model involves the applicability of the model to question answering system. This task however requires integration of language based mechanisms for deep understanding of the text.