

# Narratives Plot Comparison

Sharath Srivatsa & Srinath Srinivasa

Web Sciences Lab, IIIT Bangalore 19th Dec, 2016

## 1 Introduction

Narratives are extremely versatile way of telling imaginary or fictional and true or empirical incidents whereas expositions are simple and concise documentation based on true and well researched content [1]. Writing narratives is not bounded by any style, it is limited by the authors intention to entertain, his experience and effort to compose. A similar message can be conveyed in varying grades of style and illustrative cases and hence comparing two narratives and scoring their similarity is non-trivial. Narratives have two aspects the flow of events called the Fabula and the expression style called Discourse, both aspects affect the reading experience and the impact of the intention or message to be conveyed by the author. Our hypothesis is that two narratives can be compared by matching the verbs and nouns of events of each subject. Similarity scoring is done by parsing the text and for a given subject extracting all the events in which they are referred to and computing the WordNet word sense similarity of verbs and nouns in the events. Event sets of each subject is extracted using the co-reference annotations of Stanford NLP library.

## 2 Related Work

The primary task to compare narratives involves parsing and annotating key informative elements of narratives text and arriving at a model to score relatedness. Story Intention Graph (SIG) [3] has been used as formal representation of narratives. At current state of SIG, the annotation process is manual using Scheherazade [4] and there are no automated processes. To arrive at a computational model of narratives through automated process we explored Abstract Meaning Representation (AMR). AMR representations has been adopted for many other research works where the AMRs are parsed in specific techniques to achieve the intended goals. This representation is supported by automated parsing tool JAMR [5]. The parsing tool converts a document with one sentence in each line to simple and human readable PENMAN notation. PENMAN notation is a way of representing a directed graph in a simple tree-like form. We were unable to find a con-

vincing method to compare PENMAN notations of sentences. One of the key disadvantages we noticed was that, the PENMAN notations of a sentence and its complex rephrased version were completely different and counter-intuitive to compare.

## 3 Proposed Research Work

Our approach to compare narratives is by extracting the set of events of each subject of which they are part of and compute the word sense LIN match score between the verbs and nouns of events. Events in each set of a subject of first narrative is compared with all sets of events of all subjects in the second narrative. From the resulting matrix max-knapsack value is computed which would be the similarity score between the narratives. We have implemented the approach and have observed good results with movie plot narratives from IMDB.

We are working on precise way of interpreting the score as match or mismatch. We are also compiling 20 sets of narratives of movie plots from IMDB of 3 each where 2 plots are from remakes and the third plot completely different. The current implementation is consuming large computation cycles to arrive at the results and we would refine the implementation to reduce time complexity.

## References

- [1] Computational Modeling of Narrative (Synthesis Lectures on Human Language Technologies) , Dr Inderjeet Mani
- [2] "Abstract Meaning Representation (AMR)" URL: <http://amr.isi.edu/index.html>
- [3] David K. Elson. *Detecting Story Analogies from Annotations of Time, Action and Agency* Columbia University, New York City
- [4] "Scheherazade tool" URL: <http://www.cs.columbia.edu/~delson/software.shtml>
- [5] "AMR Parser" URL: <https://github.com/jflanigan/jamr>