Temporal Event Knowledge Acquisition via Identifying Narratives

Wenlin Yao, Ruihong Huang
Department of Computer Science and Engineering, Texas A&M University
{wenlinyao, huangrh}@tamu.edu

Introduction

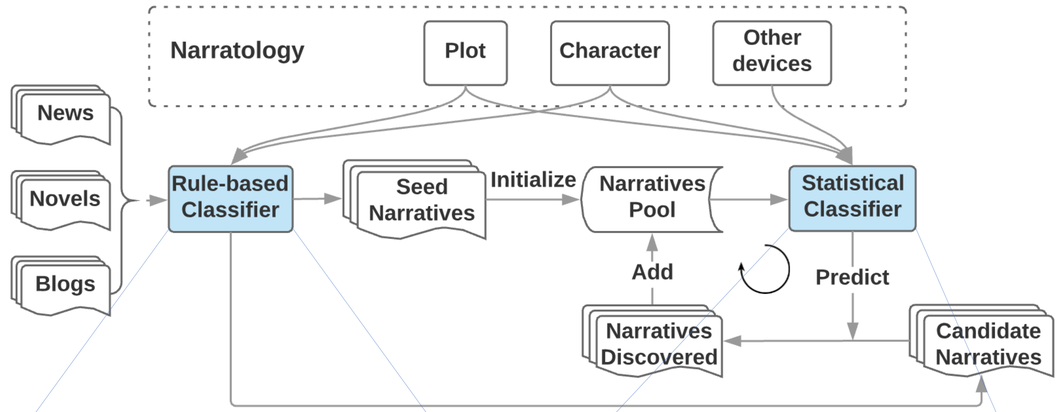
- Chronological telling: In narratives, events are governed by a **double temporality** - the chronology of the events and their presentation in the text.
- Double temporality enables us to acquire temporal "before/after" event knowledge **across sentences**.
- We explored narratology principles and built a **bootstrapping approach** that identifies 287k narrative paragraphs from three large text corpora.
- **Distilled knowledge** is useful to improve temporal relation classification and outperform neural network models on the narrative cloze task.

Key Elements of Narrative

Michael Kennedy graduated with a bachelor's degree from Harvard University in 1980. He married his wife, Victoria, in 1981 and attended law school at the University of Virginia. After receiving his law degree, he briefly worked for a private law firm before joining Citizens Energy Corp. He took over management of the corporation, a non-profit firm that delivered heating fuel to the poor, from his brother Joseph in 1988. Kennedy expanded the organization goals and increased fund raising.

Beth paid the taxi driver. She jumped out of the taxi and headed towards the door of her small cottage. She reached into her purse for keys. Beth entered her cottage and got undressed. Beth quickly showered deciding a bath would take too long. She changed into a pair of jeans, a tee shirt, and a sweater. Then, she grabbed her bag and left the cottage.

- **Plot**: A plot event is likely to have the actantial syntax "NP VP" (Greimas, 1971) with the main verb in the past tense.
- **Character**: A narrative usually describes events caused or experienced by actors. Therefore, a narrative story often has one or two main characters.
- **Other Devices**: Time, place, the emotional and psychological states of characters, which do not advance the plot but provide essential information for event interpretation (Pentland, 1999).



Phase I: Narrative Identification

Grammar Rules to Identify Plot Events: use context-free grammar production rules to identify sentences that describe an event in an actantial syntax structure (S → NP VP).

- Also consider more complex sentence structures that are derived from the basic structure. For example, "S → NP ADVP VP", "S → S CC S".
- The headword of the VP should be in the past tense.
- The NP referring to the character should have a simple structure.

Character Rules: a protagonist character appears in multiple sentences and ties a sequence of events.

- The normalized length of the longest entity chain. # entity mentions / # sentences in the paragraph

Features for Identifying Plot Events:

- Production rules of all sentences in a paragraph.
- Perplexity score of a verb bigram language model based on learned narratives.

Features for the Protagonist Characters

- Top 3 normalized length of entity chains in the paragraph.
- Top 3 normalized length from the previous and following paragraph.

Other Writing Style Features:

- Linguistic Inquiry and Word Count (LIWC): words denoting relativity (e.g., motion, time, space) and referring to emotion and cognitive.
- Parts-of-Speech (POS) tag frequencies

Phase II: Distill Temporal Knowledge

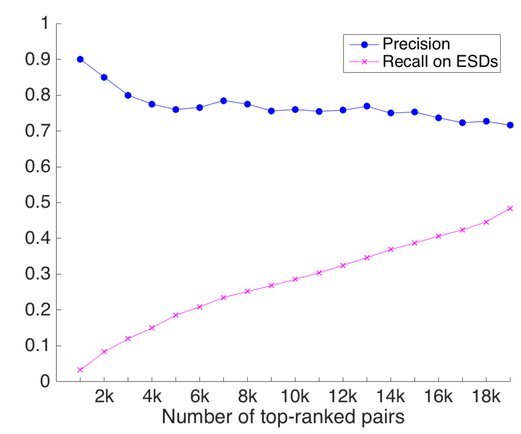
$$cp(e_i, e_j) = pmi(e_i, e_j) + log\frac{P(e_i \rightarrow e_j)}{P(e_j \rightarrow e_i)} \qquad CP(e_1, e_2, \cdots, e_n) = \frac{\sum_{d=1}^{3}\sum_{j=1}^{n-d}\frac{CP(e_j, e_{j+d})}{d}}{n-1}$$

Result

	0 (Seeds)	1	2	3	4	Total
News	20k	40k	12k	5k	1k	78k
Novels	75k	82k	24k	6k	2k	189k
Blogs	6k	10k	3k	1k	-	20k
Sum	101k	132k	39k	12k	3k	287k

pairs	graduate → teach (5.7), meet → marry (5.3) pick up → carry (6.3), park → get out (7.3) turn around → face (6.5), dial → ring (6.3)
chains	drive → park → get out (7.8) toss → fly → land (5.9) grow up → attend → graduate → marry (6.9) contact → call → invite → accept (4.2) knock → open → reach → pull out → hold (6.0)

Evaluation



Models	Acc.(%)
Choubey and Huang (2017)	51.2
+ CP score	**52.3**

Results on TimeBank

Method	Acc.(%)
(Chambers and Jurafsky, 2008)	30.92
(Granroth-Wilding and Clark, 2016)	43.28
(Pichotta and Mooney, 2016)	43.17
(Wang et al., 2017)	46.67
Our Results	**48.83**

Results on MCNC task

Data & model: http://nlp.cs.tamu.edu/resources.html