Answering Reading Comprehension Tests: N-gram and Multi-view Regression

Abstract

Create a system that answers reading comprehension tests, specifically cloze questions, using n-gram and multi-view regression approach. Evaluate and compare predictions from the two approaches.

Background & Motivation

Can machines understand English?
- Understanding English can be interpreted in two ways

Exhibiting facts (tuples) out of passage
- Example: (Bob, birthday, 6/23/1968)
- Utilization of knowledge database
- TrueKnowledge Ltd., Wolfram Alpha, Cyc, etc
- Costly, no context capturing

Answering reading comprehension (“cloze”) questions
- Utilization of context capturing; fact extraction useless
- Local/immediate context
  - Hidden Markov Models
  - N-gram
- Document-level context (“topic”)
  - Latent semantic indexing
  - Latent Dirichlet allocation
- Local & broader context
  - Multi-view regression (“MVR”)
- Lexile: Company that publishes cloze questions
  - 8 different levels; around 50 questions per level

Cloze question example
Anna and her mother changed the ___ of the yarn.

Hypothesis: Because MVR captures both the local and broader context, it answers cloze questions better than n-gram, which only captures the local context

System Design

1st approach: N-gram

- N-gram: Given N-1 words, what is the frequency of Nth word?
- Google n-gram: N-grams extracted from the web by Google

2nd approach: Multi-view regression

- Assume past and future words are conditionally independent at the present state that produces the target word.
- Each word can be fully explained by two views: words preceding and following the target word.

Results

- MVR yields higher percentage of correct answers than n-gram in general – better results in higher level tests suggest that MVR captures broader context while N-gram does not
- As level of tests gets higher, it generally requires capturing more context to answer the cloze question.
- Despite the relatively smaller size of MVR training corpus compared to Google n-gram, MVR performs as well as, if not better than, n-gram
- MVR captures more context than N-gram, given a corpus

Conclusion

- Faculty Advisor: Lyle Ungar