A Hierarchical Encoder-Decoder for Paragraph Summarization

Farzaneh Mahdisoltani

Department of Computer Science
University of Toronto

CSC2541, Fall 2016

Paragraph Summarization

Storyline

After a serial killer named Red John murdered Patrick Jane's wife and daughter, Jane dedicated his life to hunting down and killing Red John. To that end he gave up his lucrative pretense of being a psychic and joined the California Bureau of Investigation (CBI) as a consultant to the team responsible for investigating the Red John case, led by Senior Agent Teresa Lisbon. Using Jane's exceptional gift for observation and his mentalist tric able to close an unprecedented number of cases, but Jane's unconventional and often outright illegal methods also bring much censure down on Lisbon's head, making his assistance both a blessing and a curse. Meanwhile, the hunt for Red John continues...

Summary

A famous "psychic" outs himself as a fake and starts working as a consultant for the California Bureau of Investigation so he can find "Red John," the madman who killed his wife and daughter.
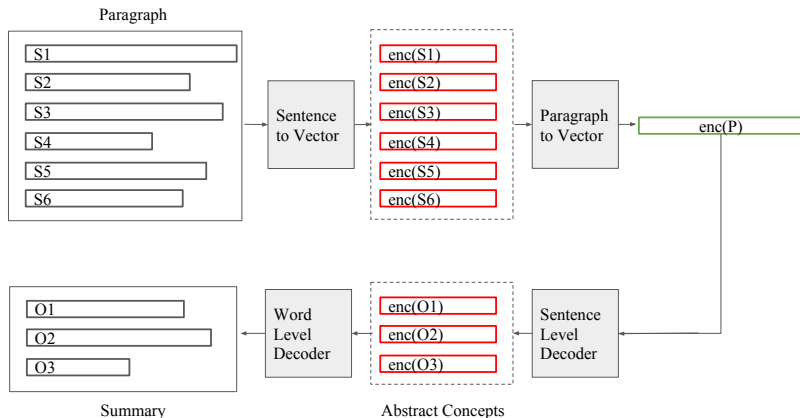
Plot Summary | Plot Synopsis

Summarization Phenomena:

- Deletion
- Paraphrase
- Generalization

Related Work

- **Extractive Summarization:** Stitching words and phrases from the source
  - Dorr et al. (2003), Cohn et al. (2008), Woodsend et al. (2010)
- **Abstractive Summarization:** Generating richer summaries based on understanding the input
  - **Attention based:** Rush et al. (2015)
  - **Statistical machine translation for headline generation:** Banko et. al.(2000)
  - **RNN with attention:** Chopra et al.(2016)
  - **Encoder-Decoder:** Hu et al.(2015)

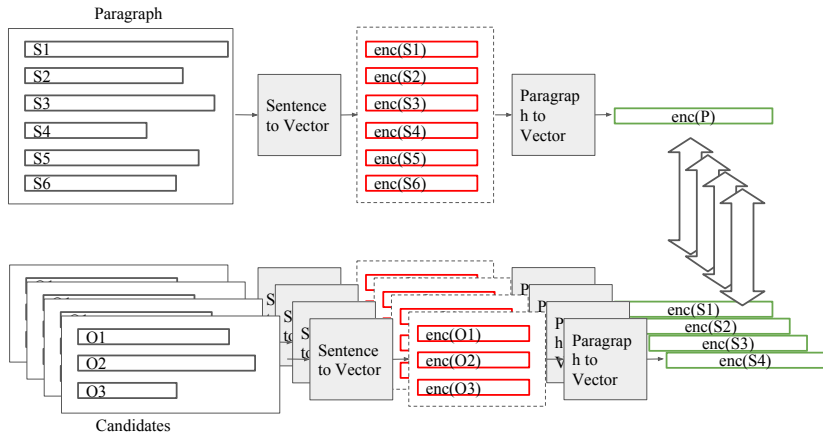# Hierarchichal Encoder-Decoder Model for Summarization

Hierarchichal Encoder-Decoder Model for Summarization

Objective Function: $\sum_t P(o_w^t | o_w^{<t}, o_s, h_p)$
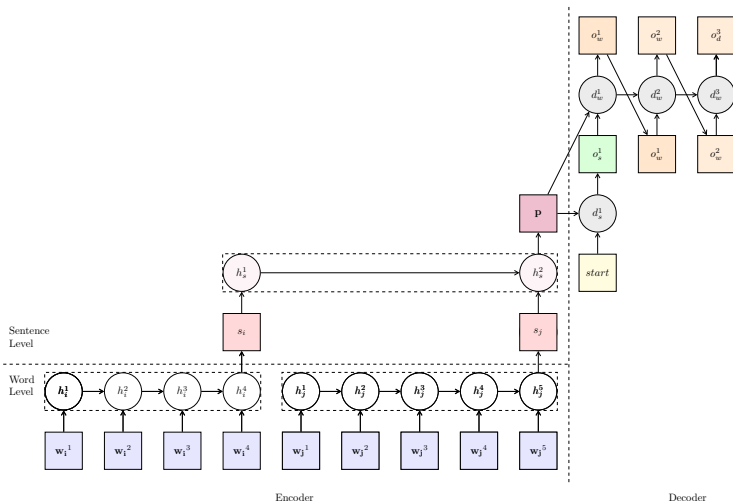
Selecting the Correct Summary from Candidates

Preliminary Results

Paragraph Representation Method	Accuracy
Random	25%
Word2Vec	65%
Skip-Thought Vectors	62%
Bag of Words	67%
RNN Encoder	57%
Hierarchichal Encoder	54%

Table: Training the encoder for multiple choice question

Thank You