Reinforced Extractive Summarization with Question-Focused Rewards

Kristjan Arumae, Fei Liu

Computer Science Department, University of Central Florida, Orlando, FL 32816

ACL 2018

Motivation

- We investigate a new training paradigm for extractive summarization.
- Existing supervised approaches to extractive summarization frequently use human abstracts to create annotations for extraction units.
- However, the labels are often inaccurate, because human abstracts and source documents cannot be easily aligned at the word level.
- In this paper we convert human abstracts to Cloze-style comprehension questions and encourage system summaries to preserve salient source content useful for answering these questions.

Proposed Approach

Source Document
The first doses of the Ebola vaccine were on a commercial flight to West Africa and were expected to arrive on Friday, according to a spokesperson from GlaxoSmithKline (GSK) one of the companies that has created the vaccine with the National Institutes of Health.

Another vaccine from Merck and NewLink will also be tested.

"Shipping the vaccine today is a major achievement and shows that we remain on track with the accelerated development of our candidate Ebola vaccine," Dr. Moncef Slaoui, chairman of global vaccines at GSK said in a company release...

Abstract
The first vials of an Ebola vaccine should land in Liberia Friday

Questions
Q: The first vials of an ____ vaccine should land in Liberia Friday
Q: The first vials of an Ebola vaccine should ____ in Liberia Friday
Q: The first vials of an Ebola vaccine should land in ____ Friday

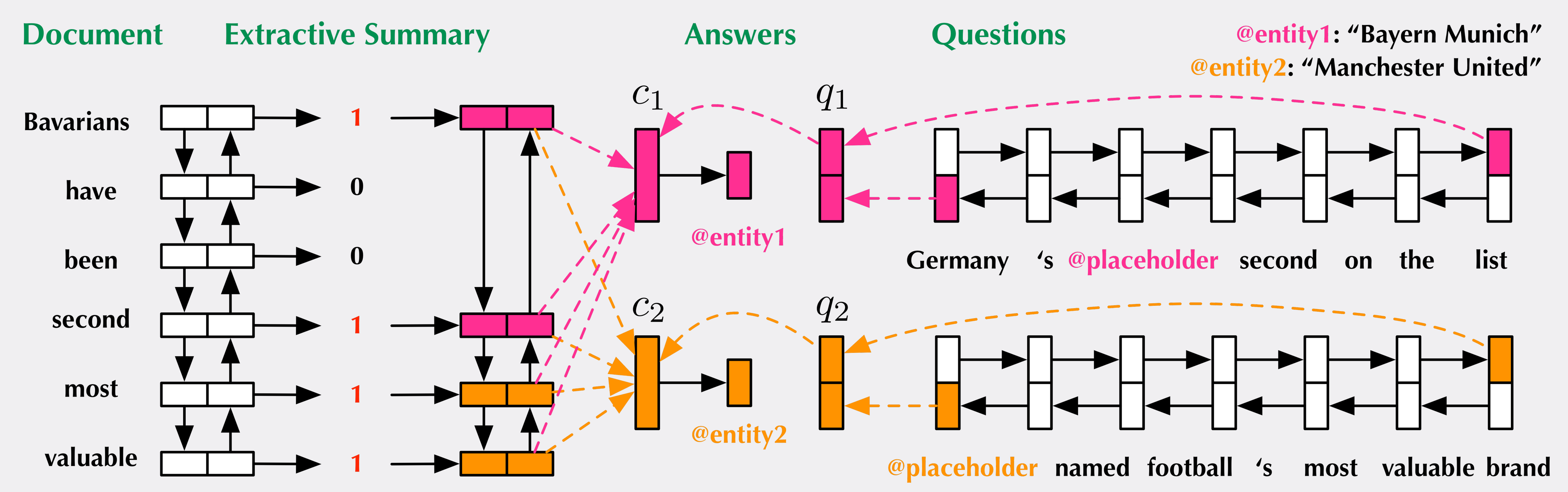
- We convert human abstracts to a set of Cloze-style comprehension questions, where the question body is a sentence of the abstract with a blank, and the answer is an entity or a keyword.
- The system summaries are encouraged to preserve salient source content that is relevant to the questions (≈ human abstract) such that the summaries can work as a document surrogate to predict correct answers.
- We use reinforcement learning to explore the space of extractive summaries.
- An attention mechanism is used to locate segments of a summary that are relevant to the given question, hence a summary can be used to answer multiple questions.

Conclusion

- We introduce a novel framework promoting extractive summaries that are concise, fluent, and adequate for answering questions. Results show that our approach is effective, surpassing state-of-the-art systems.

Kristjan Arumae gratefully acknowledges a travel grant provided by the National Science Foundation.

Model Architecture



- A bidirectional LSTM encodes the question into a vector.
- The same Bi-LSTM encodes the [sampled] summary to a sequence of vectors.
- A context vector **(2)** is constructed as a weighted sum of all summary words relevant to the k-th question, and it is used to predict the answer **(3)**.

$$\alpha_{k,i} = \frac{\exp(\mathbf{q}_k \mathbf{W}^a \mathbf{h}_i^S)}{\sum_{i=1}^{|S|} \exp(\mathbf{q}_k \mathbf{W}^a \mathbf{h}_i^S)} \quad (1)$$

$$\mathbf{c}_k = \sum_{i=1}^{|S|} \alpha_{k,i} \mathbf{h}_i^S \quad (2)$$

$$P(e_k | Y, Q_k) = \mathrm{softmax}(\mathbf{W}^c \mathbf{c}_k) \quad (3)$$

$$\mathcal{R}_a(Y) = \frac{1}{K} \sum_{k=1}^{K} \log P(e_k^* | Y, Q_k) \quad (4)$$

Reinforcement Learning

$$\mathcal{R}(Y) = \mathcal{R}_a(Y) + \gamma \mathcal{R}_b(Y) - \alpha \mathcal{R}_f(Y) - \beta \mathcal{R}_s(Y) \quad (5)$$

$$\mathcal{R}_s(Y) = \left| \frac{1}{|Y|} \sum_{t=1}^{|Y|} y_t - \delta \right| \quad (6)$$

$$\mathcal{R}_f(Y) = \sum_{t=2}^{|Y|} |y_t - y_{t-1}| \quad (7)$$

Our objective consists of **4 components**: the question answering reward **(4)**; the bigram overlap of system and reference summaries ($\mathcal{R}_b$); fluency **(7)**, to promote consecutive summaries; a summary length threshold **(6)** defined as a percentage of the input document.

Sampling Summary

$$P(y_t | \hat{y}_{1:t-1}, X) = \sigma(\mathbf{W}^h [\mathbf{h}_t^D || \mathbf{s}_{t-1}] + b^h) \quad (8)$$

$$\mathbf{s}_t = \mathrm{LSTM}([\mathbf{h}_t^D || \hat{y}_t], \mathbf{s}_{t-1}) \quad (9)$$

$$P(\hat{Y}|X) = \prod_{t=1}^{|Y|} P(\hat{y}_t | \hat{y}_{1:t-1}, X) \quad (10)$$

Whether to include the *t*-th source word in the summary can be decided based on contextual embeddings of the input Bi-LSTM. However, we also want to accommodate the previous *t-1* sampling decisions to improve the fluency of the extractive summary. We introduce a single-direction LSTM **(9)** whose hidden state tracks the sampling decisions up to time step *t*.

Results

CNN Test Set (ROUGE)

System	R-1	R-2	R-L
Distraction-M3 (Chen et al., 2016b)	27.1	8.2	18.7
Seq2Seq w/ Attn. (See et al., 2017)	25.0	7.7	18.8
Pointer-Gen w/ Cov. (See et al., 2017)	29.9	10.9	21.1
Graph-based Attn (Tan et al., 2017)	30.3	9.8	20.0
Extr+EntityQ (this paper)	**31.4**	**11.5**	**21.7**
Extr+KeywordQ (this paper)	**31.7**	**11.6**	**21.5**

We experiment with two variants of our approach. "**EntityQ**" uses QA pairs whose answers are named entities. "**KeywordQ**" uses pairs whose answers are sentence root words. Both methods are superior to the baseline systems on the benchmark dataset.

Question Answering Results

	K1	K2	K3	K4
Uniq Entities	23.7K	37.0K	46.1K	50.3K
Train Acc	46.1	37.2	34.2	33.6
Valid Acc	12.8	14.0	14.7	15.7
Valid R-2	11.2	11.1	11.2	11.1

We vary the number of QA pairs used per article in the reward function. When more QA pairs are used, we observe that the number of answer tokens has increased and almost doubled. This enlarged answer space has an impact on QA accuracies.