Neural Text Generation in Stories Using Entity Representations as Context

Presenter: Longxu Dou

Paper Authors: Elizabeth Clark, Yangfeng Ji, Noah A. Smith
University of Washington

NAACL2018 Outstanding paper

May 24, 2018

Overview

Motivation

- Problem: Automatically generating narrative text with neural models
- Challenges:
  - Character-centered generation
  - Combination of different context representation
- Applications:
  - Personalized education
  - Assistive tools for human authors
  - Social conversational agents

Centering Theory

Context	All of a sudden, [*Emily*]$_1$ walked towards [*the dragon*]$_2$.
Current Sentence	[*Seth*]$_3$ yelled at [*her*]$_1$ to get back but ________________

Figure: An example of entity-labeled story data. The goal is to continue the story in a coherent way. The actual story reads, "Seth yelled at her to get back but she ignored him."

Entities are an important element of narrative text. Centering Theory places entities at the center of explaining what makes text coherent.

Overview

Model Description

ENGEN (entity-based generation model) combines three different sources of contextual information for text generation

- The content that has already been generated within the current sentence.
- The content that was generated in the previous sentence.
- The current state of the entities mentioned in the document so far.

Context from Previous Sentence

Notation Setup:

- $\boldsymbol{h}_{t,i}$: LSTM hidden states of sentence t at timestep i
- $\boldsymbol{h}_{t-1,j}$: LSTM hidden states of previous sentence $t-1$ at timestep j
- j : ranges over the number of words in the previous sentence

previous sentence's representation $\boldsymbol{p}_{t-1,i}$

$$\boldsymbol{p}_{t-1,i} = \sum_j \alpha_{i,j} \boldsymbol{h}_{t-1,j} \tag{1}$$

$$\alpha_{t-1,j} = \frac{\exp(\boldsymbol{h}_{t-1,j} \boldsymbol{W}_a \boldsymbol{h}_{t,i})}{\sum_{j'} \exp(\boldsymbol{h}_{t-1,j} \boldsymbol{W}_a \boldsymbol{h}_{t,i})} \tag{2}$$

Context from Entities

Notation Setup:

- $\boldsymbol{e}_{i,t}$: Vector representation for entity i at timestep t
- m : the number of entities tracked in the document so far
- $m+1$: is for new, previously unmentioned entity

most salient entity's representation $\boldsymbol{e}_{current}$

The probability that the word is referring to a given entity $i \in \{1, \ldots, m+1\}$ is proportional to:

$$\exp(\boldsymbol{h}_{t-1}^{T} \boldsymbol{W}_{entity} \boldsymbol{e}_{i,t-1} + \boldsymbol{w}_{dist}^{T} \boldsymbol{f}(i)) \tag{3}$$

where $\boldsymbol{W}_{entity}$ is a weight matrix and $\boldsymbol{w}_{dist}^{T} \boldsymbol{f}(i)$ measures the distance between the current and past entity mentions.

Context from Entities

If new entity is selected, then $\boldsymbol{e}_{current}$ is initialized generated from:

$$u \sim \mathcal{N}(\boldsymbol{r}, \sigma^2 \boldsymbol{I}) \tag{4}$$

where $\sigma = 0.01$ and $\boldsymbol{r}$ is a parameter vector that is used to determine whether the next word should refer to an entity

Combining Contexts

Different representations of context:

- $\boldsymbol{h}_{t-1}$: local contextual, hidden state vector of the RNN
- $\boldsymbol{p}_t$: the previous sentence's representation
- $\boldsymbol{e}_{current}$: the most salient entity's representation

Max-pooling function for combined context vector $\boldsymbol{c}_t$

At timestep t, each element of $\boldsymbol{c}_t$ is calculated as follows:
for $k \in \{1, \ldots, |\boldsymbol{c}_t|\}$

$$\boldsymbol{c}_t[k] = \max(h_{t-1}[k], p_t[k], e_{current}[k]) \tag{5}$$

When generating word w_t, class-factored softmax function was proposed for better performance and runtime than standard softmax.

Training objective

The training objective is to maximize the logprobability of $\boldsymbol{X}$:

$$\ell(\theta) = \log P(X; \theta) = \sum_t \log P(X_t; \theta) \tag{6}$$

- $\boldsymbol{\theta}$: model's parameters
- X_t: all decisions at timestep t, by calculating probabilities for each option
  - whether it is part of a entity mention
  - the entity the mention refers to
  - the length of the mention
  - the word itself

Training Methods

- Challenge:
  - Treating the entity-related variables as latent would create a mismatch between training and prediction.
  - Training with latent variables is also expensive.
- Approach: Annotated training data with mention and coreference information (entity clusters).

Overview

Experiment Setup

- Model Variants for ablation study:
  - S2SA: $\boldsymbol{h}_{t-1}$ and $\boldsymbol{p}_t$ and class-factored softmax
  - ENTITYNLM: $\boldsymbol{h}_{t-1}$ and $\boldsymbol{e}_{current}$ and class-factored softmax
  - ENGEN: combined model
- Dataset: 312 adventure books from the Toronto Book Corpus
- Preprocessing:
  - Divide the book into segment-level dataset (33,279 segments for training)
  - Filter infrequency tokens
  - Truncate the mentions with more than three words

Experiment: Mention Generation

- Goal: Investigate each model's capacity to mention an entity in context.
- Method: Given a text and a slot to be filled with an entity mention, a model must choose among all preceding entity mentions and the correct mention.

Context	All of a sudden, [*Emily*]$_1$ walked towards [*the dragon*]$_2$.
Current Sentence	[*Seth*]$_3$ yelled at [*her*]$_1$ to get back but ________________

cluster and mention	cluster only	mention only
[*Emily*]$_1$	*EMILY	Emily
[*the dragon*]$_2$	THE DRAGON	the dragon
[*Seth*]$_3$	SETH	Seth
[*her*]$_1$		her
*[*she*]$_1$		*she

Figure: Candidate lists for each of the mention generation tasks for completing the blank. The asterisk (*) indicates the correct choice.

Result: Mention Generation

- Problem: The order of mention generation influences the evaluation.
- Solution: Report MAP of the correct candidates, where the language model scores are used to rank candidates.

model	cluster and mention	cluster only	mention only
1. Reverse order	0.12	0.38	0.15
2. S2SA	—	—	0.44
3. ENTITYNLM	0.52	0.46	0.54
4. ENGEN	0.53	0.46	0.55

Figure: Reverse order: new baseline, which ranks mentions by recency

Experiment: Pairwise Sentence Selection

- Goal: Given preceding context, select the next sentence between correct sentence and distractor from the same story
- Method: Score these sentences based on words and entity-related information.

Context	All of a sudden, [*Emily*]$_1$ walked towards [*the dragon*]$_2$.
1.	[*Seth*]$_3$ yelled at [*her*]$_1$ to get back but [*she*]$_1$ ignored [*him*]$_3$.
2.	[*She*]$_1$ patted [*its head*]$_4$ and [*it*]$_2$ curled up outside [*the cave*]$_5$.
3.	"[*Emily*]$_1$, how did [*you*]$_1$ keep [*that dragon*]$_2$ from attacking [*us*]$_6$?"

Figure: The first sentence immediately follows the context, while the second and third sentences are 10 lines and 48 lines away from the context, respectively.

Result: Pairwise Sentence Selection

model	mean accuracy	s.d.
1. S2SA	0.546	0.01
2. ENTITYNLM	0.534	0.006
3. ENGEN	*0.566	0.008

* signficantly better than lines 1 and 2 with $p < 0.05$.

Figure: Accuracy in choosing the actual next sentence.

The mean accuracies and standard deviation are calculated across the five rounds of pairwise sentence selection.

Human Evaluation: Sentence Generation

- Goal: Investigate the strengths and weaknesses of model in a downstream application.
- Method: Ask Turkers to decide which sentences they prefer (in a given context) and to explain why.
- Models: ENGEN and S2SA.
- Evaluation Factors:
    - Referring back to entities
    - Introducing new entities
    - Moving the narrative forward
    - Social knowledge

Human Evaluation: Case Study

Context	EnGen	S2SA	#
he says that it was supposed to look random , but he feels it was planned . i was the target . he 's not sure , but he feels that you might have something to do with this , " cassey said sadly . " he ca n't do that ! " manny yelled . " he ca n't accuse me with no justification .	it 's not me . "	he has nothing to do with my life	10
he was wearing brown slacks and a tan button-down shirt , with wool slippers . he looked about sixty , a little paunchy , with balding brown hair and a bushy mustache . ice blue eyes observed alejo keenly , then drifted over to wara ." welcome to my home . " the man 's voice was deep and calm .	" i 'm proud of you , " he said .	" what 's going on ? '	4
bearl looked on the scene , and gasped . this was the white rock of legend , the rock that had lured him to this land . then he stopped . " look , geron . the white rock we saw from the sea . " the struggle was taking place on the white rock . the monster had his back to bearl .	" oh my god ! "	he could not believe his eyes	1

Figure: The last column indicates the number of Turkers who voted for ENGEN's sentence (out of 11).

Overview

Conclusion

Contributions:

- Expand the context representation in RNN-based text generation
- The max pooling combination is simple but effective.
- Provide evidence for the value of entity representations for story generation.