Sequence to Sequence Semantic Parsing with Attention

Qianying Liu
School of Mathematical Science
Peking University
1400010684@pku.edu.cn

Haoran Zhang
School of Mathematical Science
Peking University
1400010667@pku.edu.cn

Abstract

This document contains the report of our second project of Peking Universitys EMNLP course. As we all known, Semantic Parsing has been increasingly important in NLP (Natural Language Processing) field, which are often used to understand the meaning of natural language and transform the original input of natural language into what a machine can understand. Recently neural models have reached high performance on this problem, so we decided to imply neural networks. Our results show that neural models are powerful and this architecture is useful.

1 Credits

This work is mainly done by Qianying Liu and Haoran Zhang. Qianying Liu designed the architecture of the model and trained the embedding. Haoran Zhang wrote the preprocessing part. We wrote the core code of the neural model together.

2 Introduction

Semantic parsing has been an important NLP (Natural Language Processing) task which benefits many downstream tasks. This task translates text to a logical or structured form. In order to predict the correct logical form for a given utterance, most previous systems rely on predefined templates. Recently machine learning methods have developed great interest, which rely on manually designed features, which often render the parsing model domain or representation specific.

Recent years, there are several researches on this field. Some of them are traditionally rule based(Berant and Liang, 2014). Lambda Calculus(Wong and J.Mooney, 2007) and statistic machine translation(Wong, 2007) seem to be useful. And as the development of neural network, it's convenient to find the rule hidden behind the question and answer. Among them, sequence-to-sequence model(van Noord and Bos, 2017) will be easy to reach, but in recent research ,sequence-to-tree model perform a little better(Dong and Lapata, 2016).

We wish to build a system that relies on minimal external knowledge and hand craft systems. So we decided to build a sequence to sequence based system.

3 Model

The whole architecture is based on a sequence to sequence idea. We propose a neural model with attention for this task. For the encoder, we use two layers of Bidirectional GRU and use dropout to avoid overfitting. For the decoder, we use one layer of one directional GRU with attention. Here we propose a simple linear layer for attention computing and then dot the vectors with the encoding input. We use teacher forcing while training, which is a novel technique of machine translation. A ratio is used to control the percent of teacher forcing usage. We randomly chose whether to use teacher forcing or not. We use a hidden size of 100 for the three layers. Dropout is set to 0.5 . The batch size is 16. Limited to our laptops, we only ran for 40 epochs. The learning rate is initially set to 0.0001 and we use a gradient clip of 10.

3.1 Problem Formulation

Our aim is to learn a model which maps natural language input $q = x_1 \cdots x_{|q|}$ to a logical form representation of its meaning $a = y_1 \cdots y_{|a|}$. The conditional probability $p(a|q)$ is decomposed as:

$$p(a|q) = \prod_{t=1}^{|a|} p(y_t|y_{<t,q}) \qquad (1)$$

where $y_{<t} = y_1 \cdots y_{t-1}$.

3.2 seq2seq model

This model regards both input q and output a as sequences. The encoder and decoder are two different neural networks with GRU units and they do not share parameters.

Let $h_t^l \in R^n$ denote the hidden vector at step t and layer l. So that h_t^l can be computed by:

$$h_t^l = GRU(h_{t-1}^l, h_t^{l-1}) \tag{2}$$

3.3 Model Training

Object function is:

$$min \ -\sum_{q,a} \log p(a|q) \tag{3}$$

in order to maximize the likelihood of the generated logical form.

3.4 Inference

The Prediction form is:

$$\hat{a} = \arg\max_x p(a'|q) \tag{4}$$

We can use formula (1) for decomposition and then use greedy search for quicker prediction.

4 Results

Our system is seriously limited by the device. Our laptops cannot present very reliable results.

References

Jonathan Berant and Percy Liang. 2014. Semantic parsing via paraphrasing. *ACL2014*, pages 1415–1425.

Li Dong and Mirella Lapata. 2016. Language to logical form with neural attention. *ACL2016*.

Rik van Noord and Johan Bos. 2017. The meaning factory at semeval-2017 task 9 producing amrs with neural semantic. *arXiv:1704.02156*.

Yuk Wah Wong. 2007. Learning for semantic parsing and natural language generation using statistical machine translation techniques. *Dissertations and Theses - Gradworks 87.87(2007)*, pages 1–6.

Yuk Wah Wong and Raymong J.Mooney. 2007. Learning synchronous grammars for semantic parsing with lambda calculus. *ACL2007*, pages 960–967.