本周进度

- 复习英语 / 图论助教；
- 论文阅读 -- Encoder-Agnostic Adaptation for Conditional Language Generation。

---

1. 论文阅读

- 提出一种`pseudo self attention`的方法
- **核心问题：**
  - how to adapt a pretrained decoder to effectively utilize arbitrary source information
- **Experiments：**
  - `预训练模型：`Gpt2
  - `Tasks:`
    - class-conditional generation
    - document summarization
    - story generation
    - image paragraph captioning
  - `Encoder-agnostic variants considered`

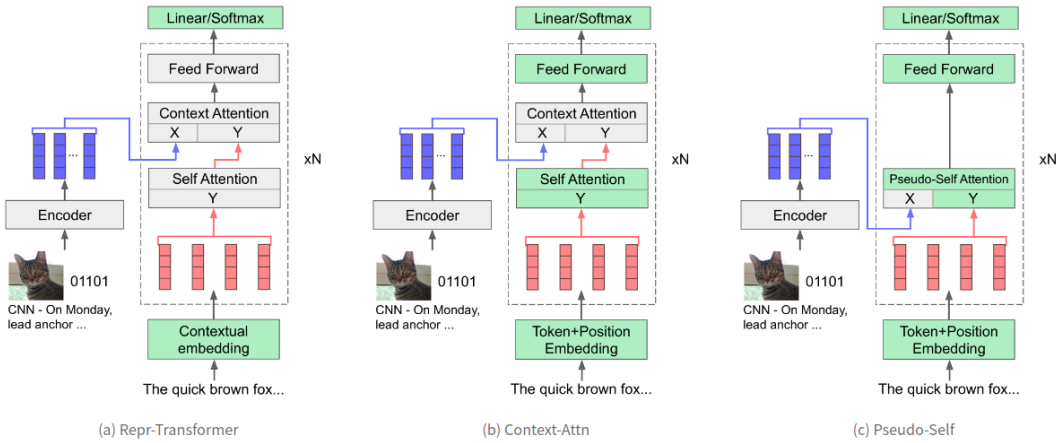

Figure 1: Encoder-agnostic variants considered. All methods utilize a problem-specific source encoder, but vary in which parts of the decoder are pretrained and which are randomly initialized. Repr-Transformer trains a new full transformer decoder, Context-Attn trains a new context attention layer, Pseudo-Self attention only modifies part of the self attention layer. Residual connections and layernorm have been omitted for clarity. Green indicates that parameters are initialized with pretrained weights, gray indicates random initialization. Red vectors indicate the target activations at each layer, Blue vectors indicate the source features at the output of the encoder. xN indicates the section within the dotted lines is stacked N times.

  - `实验结果`
    - 还没看完。。。

下周任务

- 整理一下条件语言模型的论文；
- 继续看论文，然后尝试训练一个条件语言模型。