

读书报告

09118119 黄一凡

2020 年 9 月 17 日

1 别人提出的问题

1. 为什么 MT 会导致一些不好的结果？

对于 nerual generation 来说，其目标是产生一个合理的句子，所以不太受形式限制，但是对于机器翻译这种对于句子内容比较有针对性的模式来说，可能完全根据概率来选择不是很好，还需要考虑到对应的翻译文本和其产生的前文，才能产生比较好的翻译。

2. Beam search 上一段中最后一句话说这个方法“invariant doesn't hold”具体指什么意思吗？

应该是说动态规划的条件不成立，因为类似维持比的算法应该有很多的限制，但是对于这种句子长度都不太能确定的应用可能不是很好使用。

3. Attentio 机制与直接上下文之间的比较，Attention 是否存在缺点？

因为是对序列的所有元素并行处理的，所以无法考虑输入序列的元素顺序，这在自然语言处理任务中比较糟糕。而在自然语言中，语言的顺序是包含了十分多的信息的，如果缺失了该部分的信息，则得到的结果往往会大大折扣。

4. attention 机制是否就是指通过给予编码器和解码器一定权重从而使得机器可以更有针对性的学习？

attetion 其本质其实和我们人类进行翻译的过程是比较类似的，因为两种文本一般有一种对齐的关系，某些部分会对应某些部分，如果使用的是之前那种上下文向量的话，其实是不符合人的思考习惯的。attetion 的意义在

于其进行了一种对齐操作，比如说根据我现在翻译的信息，我可以通过向量积的形式找到对于接下来一步翻译最重要的原文部分，这样就为每一步翻译有效的提供了其需要的上下文，而且信息缺失也很少。

5. Self-Attention and Transformer Networks 这两个概念是指什么？是否是 Attention 的延申。

应该可以看做 attention 的延伸。self-attention 显然是 attention 机制的一种。上面所讲的 attention 是输入对输出的权重，例如在上文中，是 I am a student 对学生的权重。self-attention 则是自己对自己的权重，例如 I am a student 分别对 am 的权重、对 student 的权重。之所以这样做，是为了充分考虑句子之间不同词语之间的语义及语法联系。

2 读书计划

本周所读： 《Speech and Language Processing》 Chapter 10

下周计划： 《Speech and Language Processing》 Chapter 11