图神经网络应用——机器翻译

2017141531099 汤龙宇

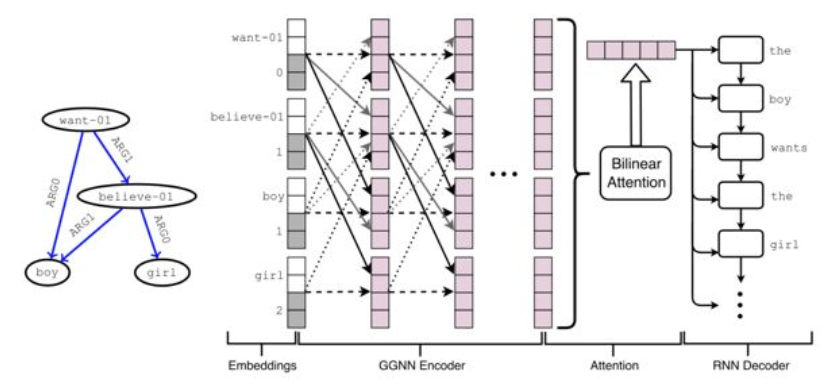
当图神经网络应用于其数据结构为非图结构的场景，例如文本。在此类场景中，GNN探索此类数据中隐含的图结构，如文本的句法树结构。

图神经网络在机器翻译中应用一般是考虑结合句法树，因为树本身就是图的特例。而融合句法信息一般认为对机器翻译是有帮助的。现在机器翻译的范式已经变成神经机器翻译（NMT），而NMT的通用架构是Encoder-Decoder,图神经网络一般出现在Encoder端，目的也就是尽量多地融合句法信息。

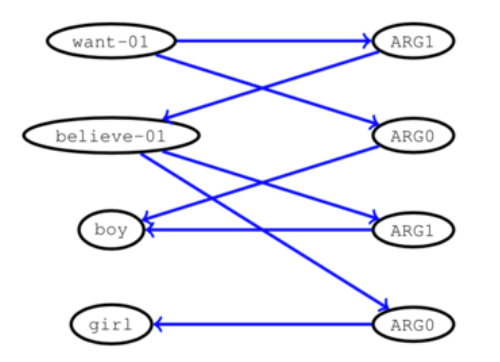
阅读了一篇关于图神经网络应用于机器翻译的论文<ACL-18> Beck D, Haffari G, Cohn T.Graph-to-sequence learning using gated graph neural networks。通过将 sourcelanguage 的句法结构图输入 GNN 进行 encode，继而附加 Attention layer 和 decoder 层，输出 target language 序列。

这篇文章以“Graph tosequence”为要点，实验部分包括两个任务，其中一个是 AMR 图输出序列，另一个做 syntax-aware 的机器翻译任务。其中，AMR Graph 为从句子的句法表示中抽象出来的语义表示语言，具体相似语义的句子有相同的 AMR Graph。

模型在 Seq2seq 模型的基础上，以 AMR graph 为输入，通过 GGNN 网络进行 encode，经过 Attention 层和 RNNdecoder 输出序列。如下图所示。



随着 Graph 中边类型的增多，很容易引起参数爆炸的问题，因此，本文提出了一种融合 Graph 中 edge 信息的方式，即将 edge 转化为附加的节点(no labelled edges)。具体地，通过将 Graph 转化为对应的二分图，由此graph中的边的类型只有一种。如下图所示。



在阅读的另一篇文章Graph Convolutional Encoders for Syntax-aware Neural Machine Translation中，讲句法结构整合到基于神经注意的机器翻译编译码模型中，GCN使用源语句的预测语法依赖树来生成对它们的语法领域敏感的单词表示，即编码器的隐藏状态。GCN将单词表示作为输入，并生成单词表示作为输出，因此他们可以很容易地作为层合并到标准编码器中，例如双向RNN或CNN之上。

原则上，GCNs具有足够的灵活性，可以容纳任何语言结构，只要它们可以表示为图。例如与递归神经网络相比，GCNs不要求图是树，只关注依赖语法，将更多一般性研究留给未来的工作。GCNs在句法依赖树之上运行，在语义角色标记上下文中显示出来有效性，由于语法GCNs在单词级产生表示，在给予注意力的编解码器框架中使用它们作为编码器是很简单的。GCNs课以将Word embeddings作为输入，但如果将其作为递归神经网络或卷积神经网络编码器之上的层，使用语法信息丰富它们的状态时，会更加有效。