个人总结040

通过人工智能导论这门课程，我们进一步知道人工智能在很多领域得到了发展，在我们的日常生活和学习中发挥了重要的作用。如在语言工程中的应用之一：机器翻译。机器翻译是利用计算机把一种自然语言转变成另一种自然语言的过程，用以完成这一过程的软件系统叫做机器翻译系统，利用这些机器翻译系统我们可以很方便的完成一些语言翻译工作。在本次大作业中，我们实现了基于Seq2Seq结构和注意力机制的英文-中文机器翻译系统，在这个过程中，我研究了序列模型以及知识推理算法，对神经网络机器翻译领域的应用知识以及人工智能有了进一步的学习认识和收获，其中包括：

1. 知识推理是知识利用的基础。神经机器翻译(NMT)是一种端到端自动翻译学习方法，它的优势在于它直接学习从输入文本到相关输出文本的映射，因此其知识推理是通过神经网络的学习训练实现的，即充分利用样本性知识。
2. 对机器学习有了进一步的认识。机器翻译的发展实际上始于基于规则的机器翻译，可以对一组规则进行编码，但不同的语言有太多的例外情况，模型也变得越来越复杂。机器学习出现后，被引入到机器翻译中，算法可以通过学习双语语料得到提升，而不是通过预先的规则设定。
3. 神经网络的seq2seq学习，即序列对序列的学习。对于深度学习而言，如果要学习一个序列，一个重要的困难就是这个序列的长度是变化的，而深度学习的输入和输出的维度一般是固定的，不过，有了RNN结构，这个问题就可以解决了，因此在我们的模型中encoder和decoder使用的是GRU结构。
4. 在学习研究模型结构的过程中，由于Seq2Seq其实是循环神经网络（RNN）的变种，为充分理解掌握该结构，需要学习RNN作为先验知识，而LSTM、GRU是为了解决长期记忆和反向传播中的梯度等问题而提出来的，即RNN到Seq2Seq到LSTM到GRU 不断变化进而弥补之前缺陷，因而需要将他们串联起来进行学习，方便更好地理解模型。
5. 关于使用注意力机制的原因，Attention机制是对神经网络中一些重点结构或节点加以关注，例如学习一个权重分布，该权重可以对网络节点加权地赋予不同的重要程度。我认为Attention机制并不是一个架构或模型，更类似于一种思想。它来源于人类的视觉注意力，最简单的例子，比如看一个图片，会有特别显眼的场景率先吸引注意力，因为大脑中对这类东西很敏感。机器学习模型架构中注意力机制的引入，进一步体现了人工智能的概念。
6. 我发现神经网络翻译存在两点问题和挑战：模型把原文句子整体读进去以后形成了一个向量，然后再对这个向量进行解码。如果翻译模型认为有些词不应该产生，从而就会漏掉了译文；相比于统计机器翻译，数据稀疏问题对神经网络翻译而言，更严重，实验表明，神经网络对于数据量更敏感。
7. 对于翻译评价指标，我们使用的是BLEU score。在我们的模型上进行评价分析，发现我们的模型虽然简单，但在性能上还有一定的提升空间，需要后续的进一步寻求改善。