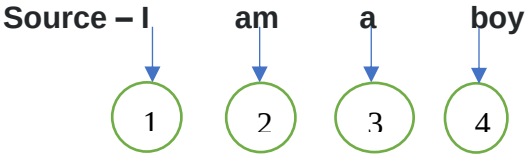
Exercise 5.3: It has frequently been observed that feeding the source language into the seq2seq encoder backward (but leaving the decoder working forward) improves MT performance by a slight but constant amount. Make up a plausible story for why this could be the case.

Answer:

Usually Seq2Seq model maps a source sequence to the target sequence. In this process of concatenating the source sentence with the target sentence, each word in the source sentence is far from its corresponding word in the target sentence.

For example, let's take "**I am a boy**" being machine translated to Japanese as "**Watashi wa otokonokodesu".**



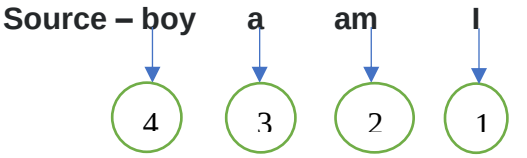
Target – Watashi wa otokonokodesu

Here the problem is that it has a large "minimal time lag".

The paper "**Sequence to Sequence Learning with Neural Networks**" by Ilya Sutskever and others says that "reversing the order of the words in all source sentences (but not target sentences) improved the performance markedly, because doing so introduced many short term dependencies between the source and the target sentence which made the optimization problem easier".

While reading the paper much further, it has been understood that by reversing the words in the source sentence, the average distance between corresponding words in the source and target language is unchanged. However, the first few words in the source language are now very close to the first few words in the target language, so the problem's minimal time lag is greatly reduced.

After reversing the source,



Here we take words in the sequence order of 1, 2 and so on. While predicting the target sentence, the average distance has not changed. You can now pretty well predict the target word.

Thus, backpropagation has an easier time "**establishing communication**" between the source sentence and the target sentence, which in turn results in substantially improved overall performance.

References

Sutskever I., Vinyals O., & V. Le Q. (2014). Sequence to Sequence Learning with Neural Networks. New York, NY: Cornell University.