COMP6248: Lab Exercise 7

David Jones (dsj1n15@soton.ac.uk)

Task: Transforming Sequences

1 Exercise 1

Exercise 1.1: Training a Sequence-to-Sequence Model

```
def forward(self, src):
    embedded = self.embedding(src)
    _, (hidden, cell) = self.rnn(embedded)
    return (hidden, cell)
```

Listing 1: Encode `forward` code snippet.

Exercise 1.2: Decoding Test

answer the following

- **why is the order of the output reversed**
  Input sentences are read in reverse because it introduces short term dependencies that make the optimisation problem easier. This results in reversed outputs.
- **what is the point of teacher forcing**
  Teacher forcing is where the ground truth is used as an input to the model for $t+1$, as opposed to the generated value at t. This is done to speed up convergence and improve model skill.

Exercise 1.3: Effect of Sequence Length

Added options in `decode` for configuring `span` (the number of spaces before splitting into a chunk), and `maxlen` (passed to `Seq2Seq`). If only the `span` is changed, only the last encoded letter is returned (`"..-. .-"` [`fa`] returns `a`). Increasing `maxlen` to 3 corrects this, however this adds SOSs (`^`) for shorter inputs. Therefore `maxlen` should be 1 + `span` and the input should be divisible by `span`.

Setting `maxlen` and `span` appropriately such that the full string can be decoded as a single chunk: inputting the first item from the dataset (`".--.  .-.  .  ..-.  .-"` [`prefa`]) will decode successfully. Appending an extra character `".-"` [`a`] causes an unsuccessful decode (`rpaswa`). Notably, it is not just the length of string that allows `prefa` to decode, an input of the same length not present in the dataset (`aaaaa`) will fail (returns `raaaa`). This indicates the LSTM has learnt the dataset.