1 Transforming Sequences

```python
class Encoder(nn.Module):
    def __init__(self, input_dim, emb_dim, hid_dim):
        super().__init__()
        self.hid_dim = hid_dim
        self.embedding = nn.Embedding(input_dim, emb_dim)
        self.rnn = nn.LSTM(emb_dim, hid_dim)

    def forward(self, src):
        # my code here
        embedded = self.embedding(src)
        outputs, (hidden, cell) = self.rnn(embedded)
        return hidden, cell
```

Figure 1: Encoder Forward Snippet

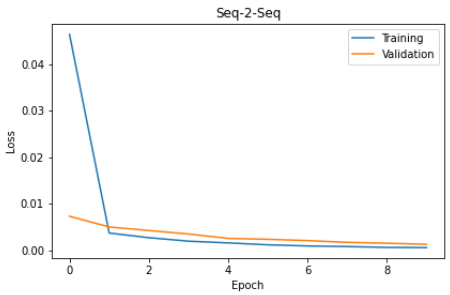

Figure 2: Loss Curve for Seq-2-Seq training

Decoded texts:
1. 'answer the following'
2. 'why is the order of the output reversed'
3. 'what is the point of teacher forcing'
Ans 2. Reversing the output makes the LSTM learn better, this phenomena may be because it introduces short term dependencies to the dataset.
Ans 3. Teacher forcing use output from prior time steps as input. It addresses slow convergence and instability.

Sequence Lengths
With longer chunks, the performance is poor. This could be because the model output does not correspond to words in the vocabulary.