Mohamed Sameh Saafan

Joint Learning of Alignment and Translation in Neural Machine Translation

**Introduction**

- Traditional statistical machine translation has many sub-components; neural machine translation (NMT) unifies them into a single model.  
- Basic encoder–decoder NMT suffers from the limitation of encoding the entire sentence into a fixed-length vector, which reduces performance for long sentences.  
- This paper introduces a model that learns to align and translate simultaneously using a soft attention mechanism.

**Model Architecture**

- Encoder uses a bidirectional RNN to produce annotation vectors for each word.  
- Decoder uses a soft attention mechanism to compute context vectors dynamically for each output word.  
- Context vectors are weighted combinations of encoder annotations based on an alignment model.

**Attention Mechanism**

- The alignment model computes attention weights αij using a feedforward network.  
- The decoder generates each target word based on its previous state, the last word, and the dynamic context vector.  
- Soft alignment improves translation of long sentences and manages word reordering and unknown word handling better.

**Experiments**

- Trained on WMT’14 English–French data using RNNsearch and RNNencdec variants.  
- BLEU scores significantly higher for RNNsearch, especially for long sentences.  
- RNNsearch matches or outperforms phrase-based systems on sentences with known words.

**Results and Analysis**

- RNNsearch more robust with long sentences; avoids performance drop seen in RNNencdec.  
- Visualized alignments show accurate and interpretable attention across words.  
- Qualitative samples show better preservation of meaning in long and complex sentences.

**Conclusion**

- Jointly learning to align and translate addresses key limitations in earlier NMT architectures.  
- The soft attention mechanism leads to better alignment, especially for longer sentences.  
- The model performs comparably to traditional systems and offers a foundation for future NMT advancements.