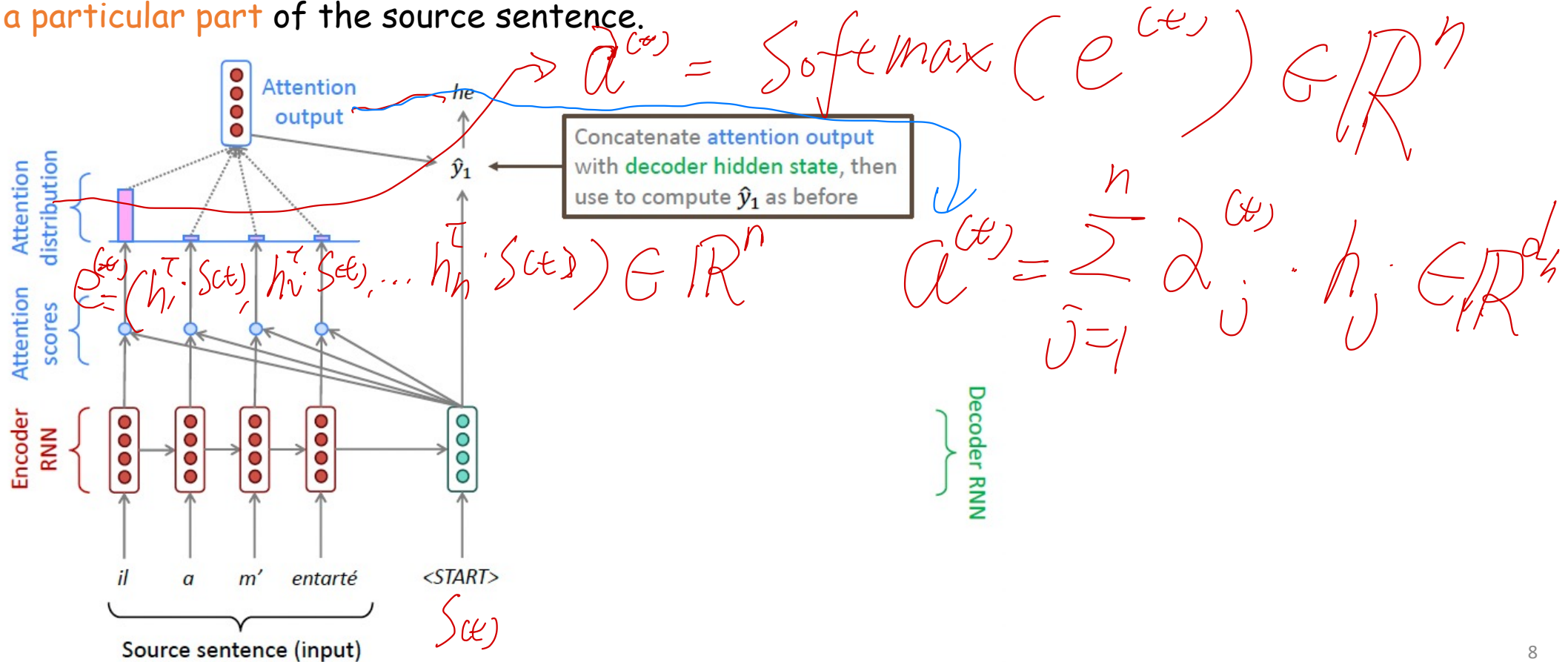


Attention Mechanism in Seq2Seq

Reference: Stanford CS224N, Lecture 8: <https://web.stanford.edu/class/cs224n/slides/cs224n-2024-lecture08-transformers.pdf>

- Idea of attention: On each step of the decoder, use **direct connection to the encoder** to **focus on a particular part** of the source sentence.



h_1, h_2, \dots, h_n