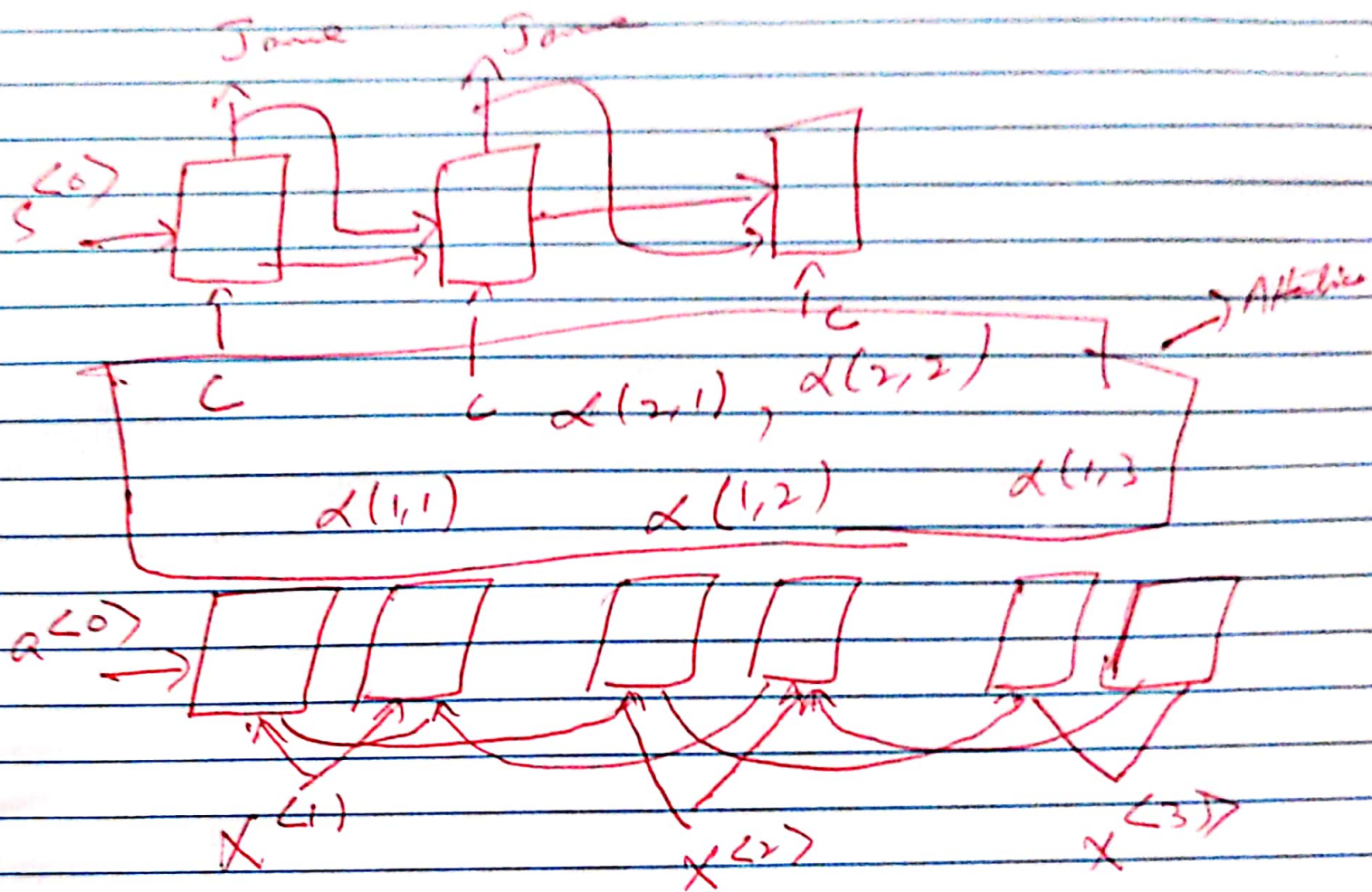


* Inspired by Human Brain, gives weights to the important information.

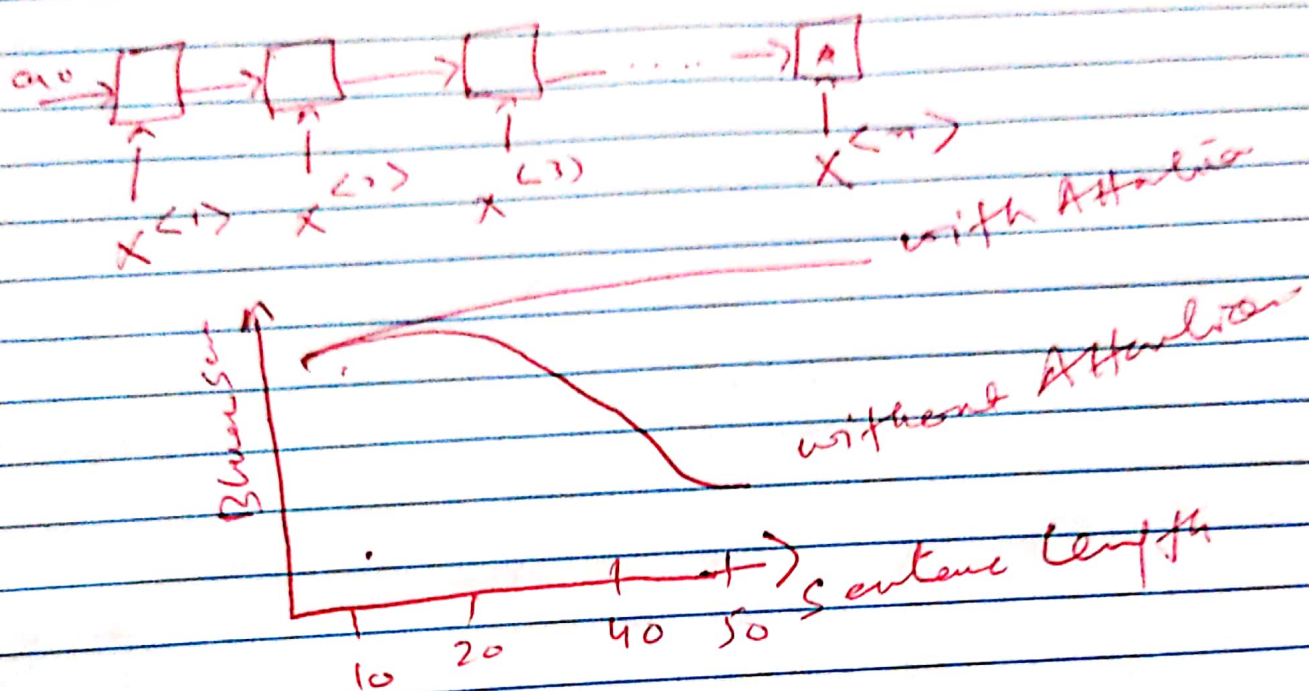


* I/P the Complete Sentence.

* Bidirectional RNN \rightarrow I/P the whole Sentence

$$\sum_{t=1}^T \alpha^{(1,t)} = 1$$

$$c^{(1)} = \sum_{t=1}^T \alpha^{(1,t')} h^{(1,t')}$$



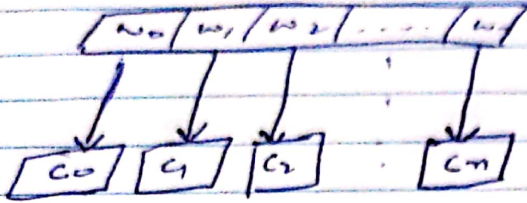
BLEU score: Performance metric for Comparison of machine translation with Human.

- * 1 means Complete Match
- * 0 means Complete Mismatch.

$$P = \frac{n}{w_t} \rightarrow \begin{array}{l} n \rightarrow \text{number of words in Language 1} \\ w_t \rightarrow \text{Correctly Translated by the NN.} \end{array}$$

- * NLTK Library is used.

Attention Mechanism

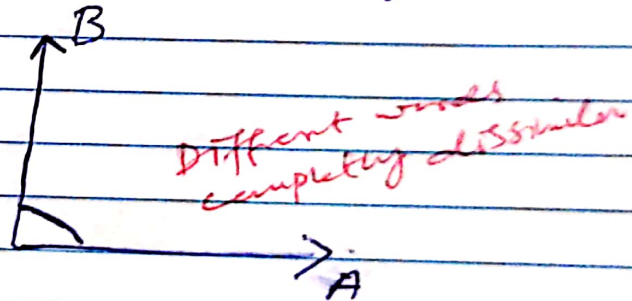


* We want to Quantify how Similar a Word "Query" c_k to each of this word.

* How we gonna do this?

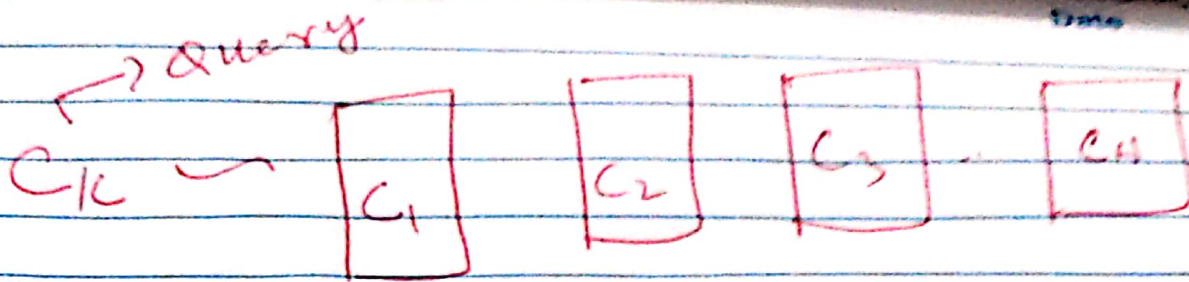
"Inner product" or dot product

* After that we can Quantify this relationship b/w the Query word and the Sequence.



* We can thought of two words A and B, totally different from each other if the dot product is zero.

* The words are Similar exactly, if the dot Product is one, the Values b/w zero and one determined the degree of relationship b/w two words.



$$C_K = r_{K1} \times C_1 + r_{K2} \times C_2 + r_{K3} \times C_3 + \dots + r_{Kn} \times C_N$$

$$C_K = \sum_{n=1}^N r_{Kn} \times C_n = 1 \quad \rightarrow \text{Attention mechanism}$$

\rightarrow Value

$r_{K1} \Rightarrow$ relative relationship b/w word 1 and query K

High values means words are related.

\rightarrow How much Attention to word 1.

If relationship is high, Pay high relationship.

C_K is assumed to C_K .