Encoder | **Decoder** | Output Probabilities
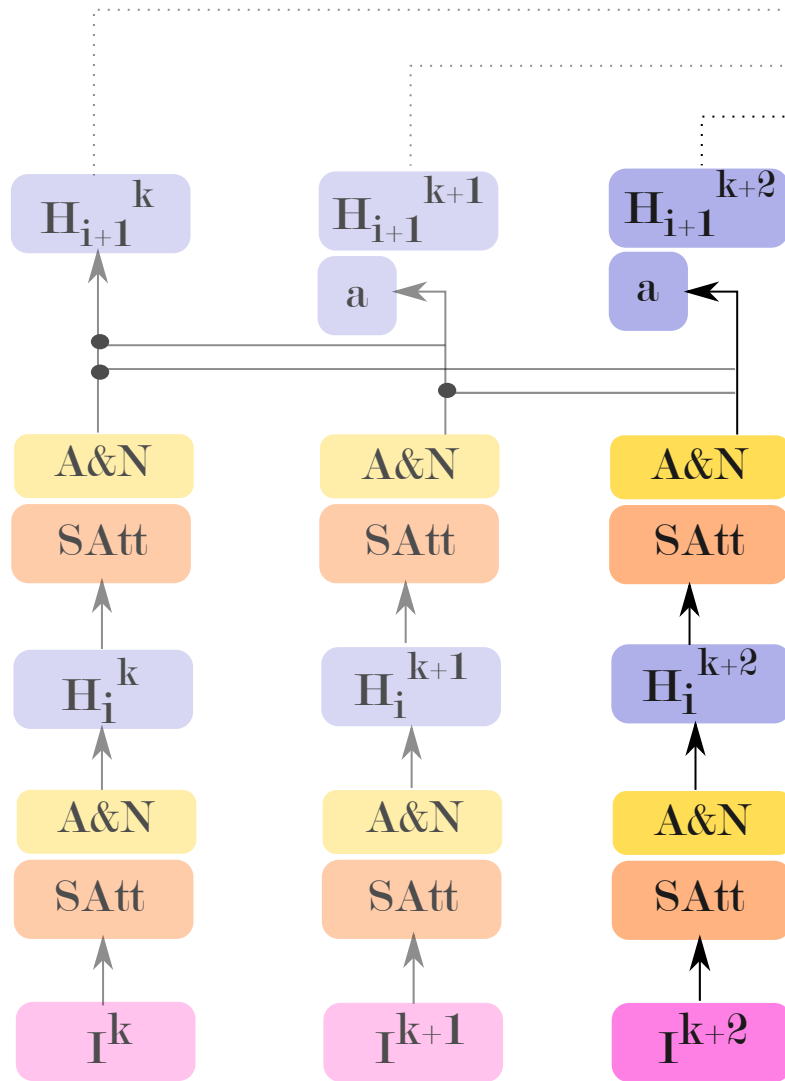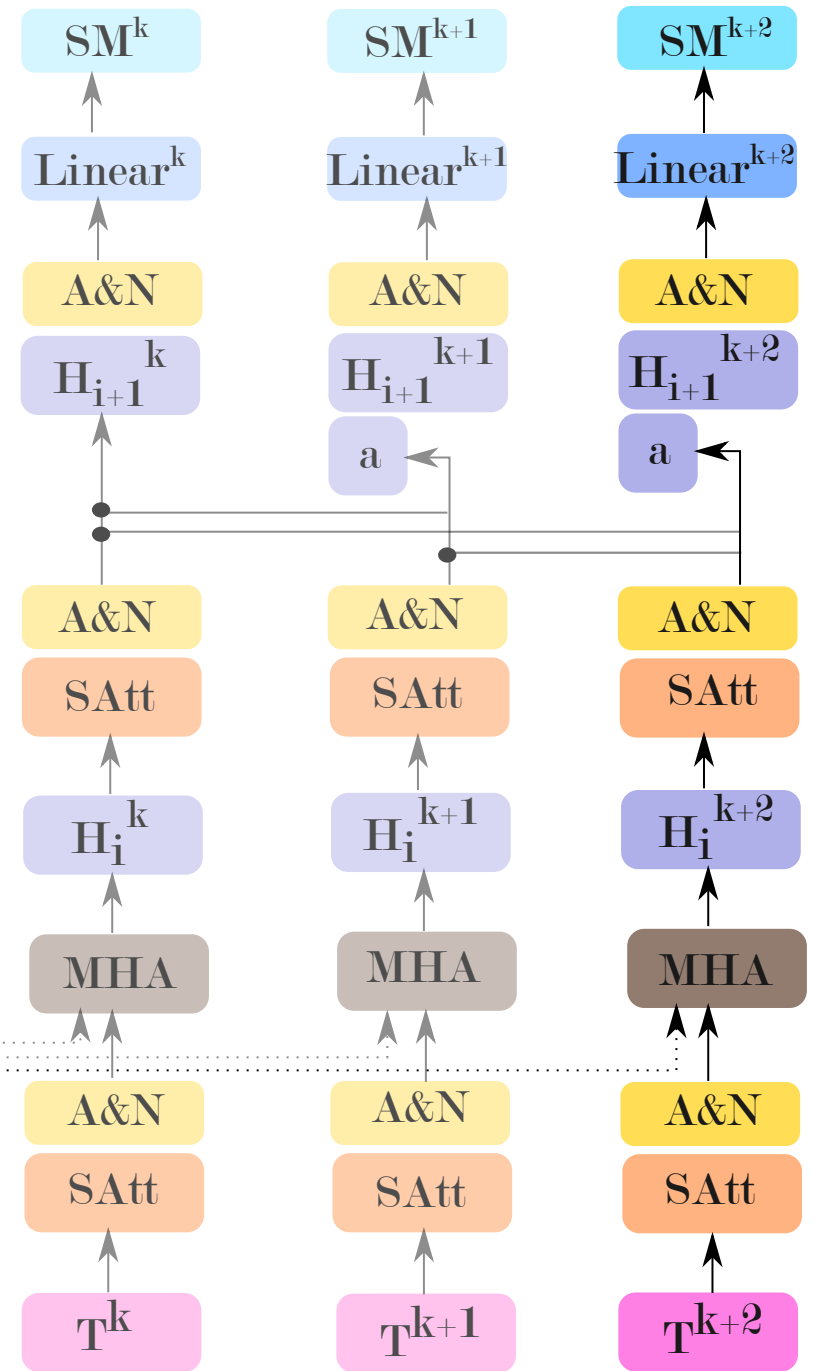
I - Input Sequences : Transformed to Embeddings
T - Target Sequences : Transformed to Embeddings
SAtt - Self Attention Mechanism
A&N - Add and Normalize
a - Adapter Function
H - Hidden Layer Network

SM - Softmax Layer to output the probabilities
Transfer of Hidden Layer State
Representations from Encoder
to Decoder