AttentionMerge	input:	[Data Encoder,Question Encoder, Hop 1]
	output:	[n_memories,n_neurons]

TimeDistributedDense	input:	n_memories, n_neurons
	output:	n_memories, n_neurons

AttentionRecurrent (Hop 2)	input:	n_memories, n_neurons
	output:	n_neurons