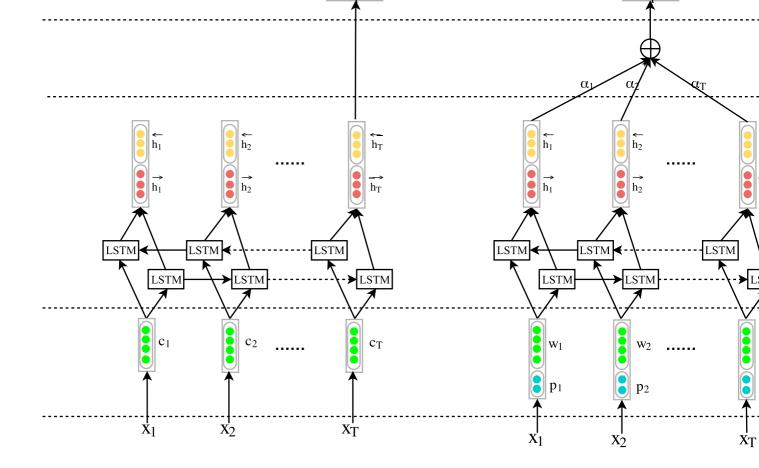
output



	attention layer			
← h _T				
$\overrightarrow{h_{\mathrm{T}}}$	LSTM layer			
STM				
w_{T}	Embedding layer			
p _T				
	Input layer			