Attention probabilities for head 11.4 swapping positions C2 <-> C3

