

Attention probabilities for head 8.11 swapping positions C1 <-> C2

