Attention probabilities for head 11.4 swapping positions C1 <-> C3

