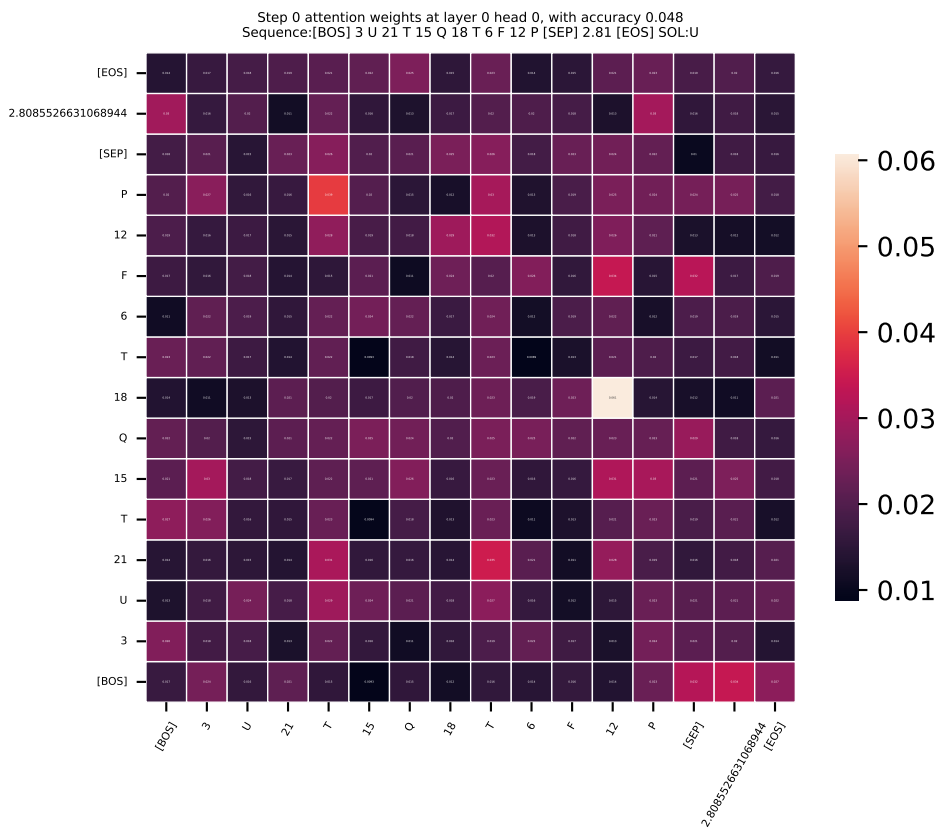
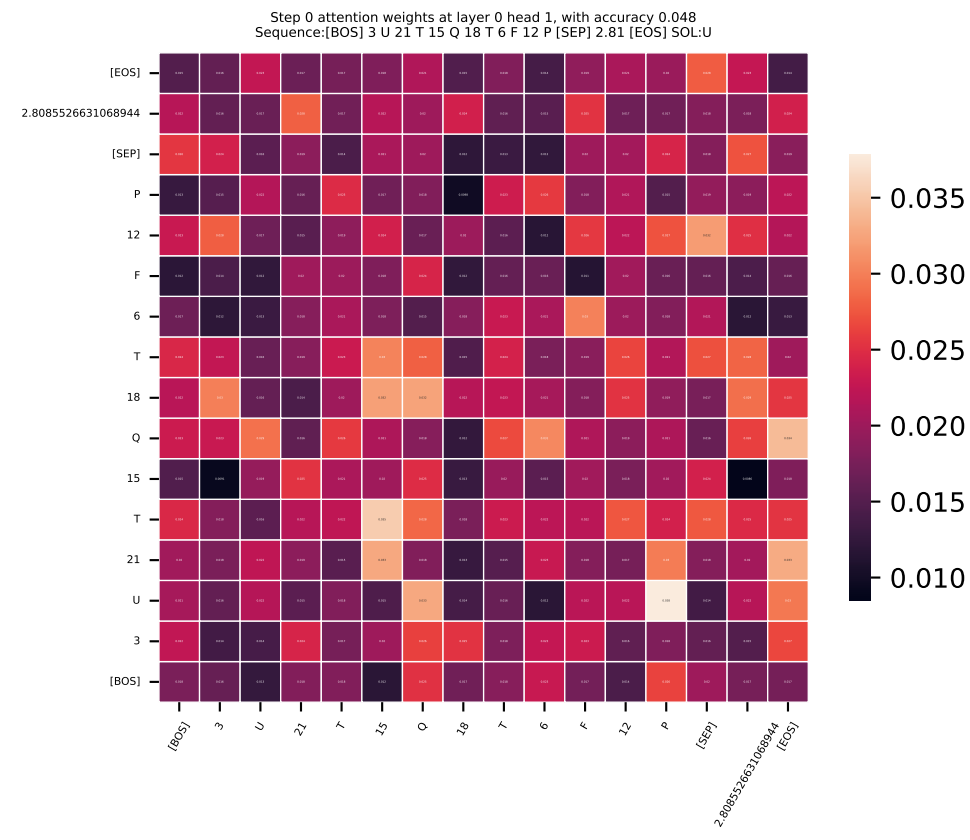
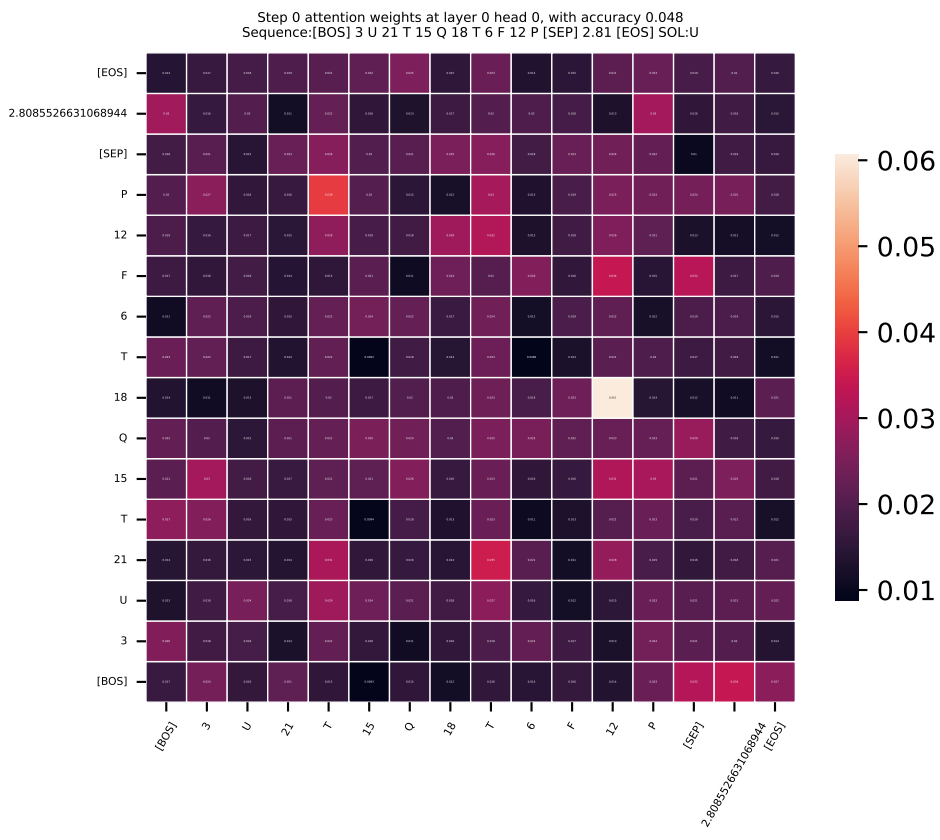
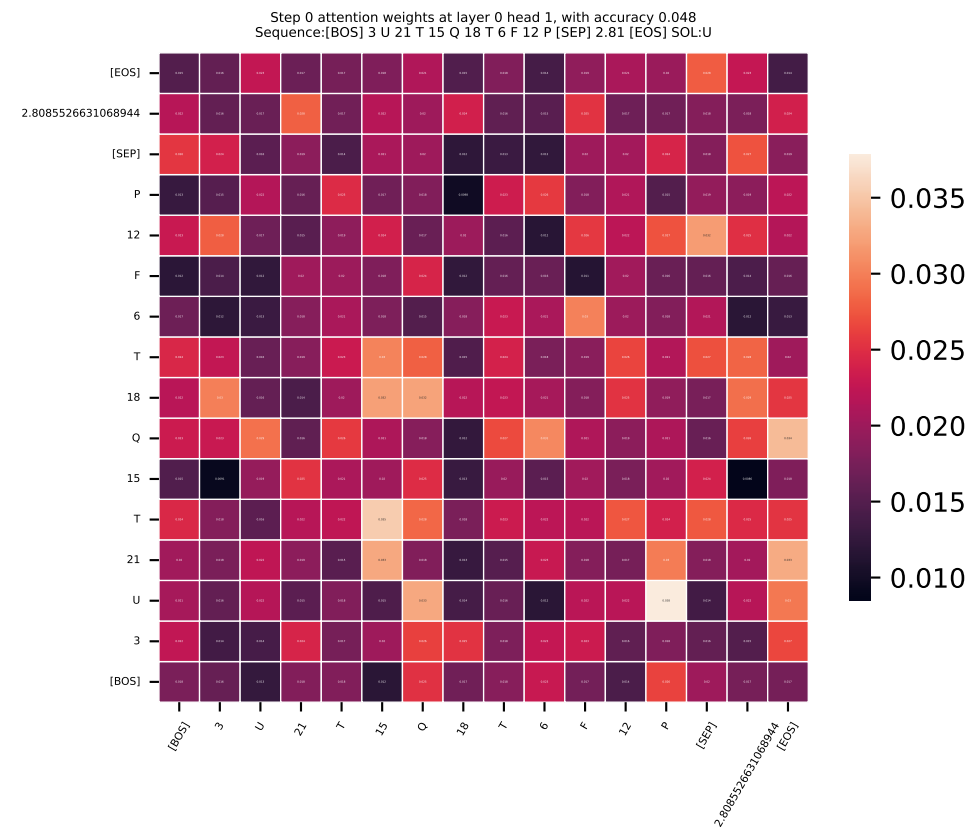
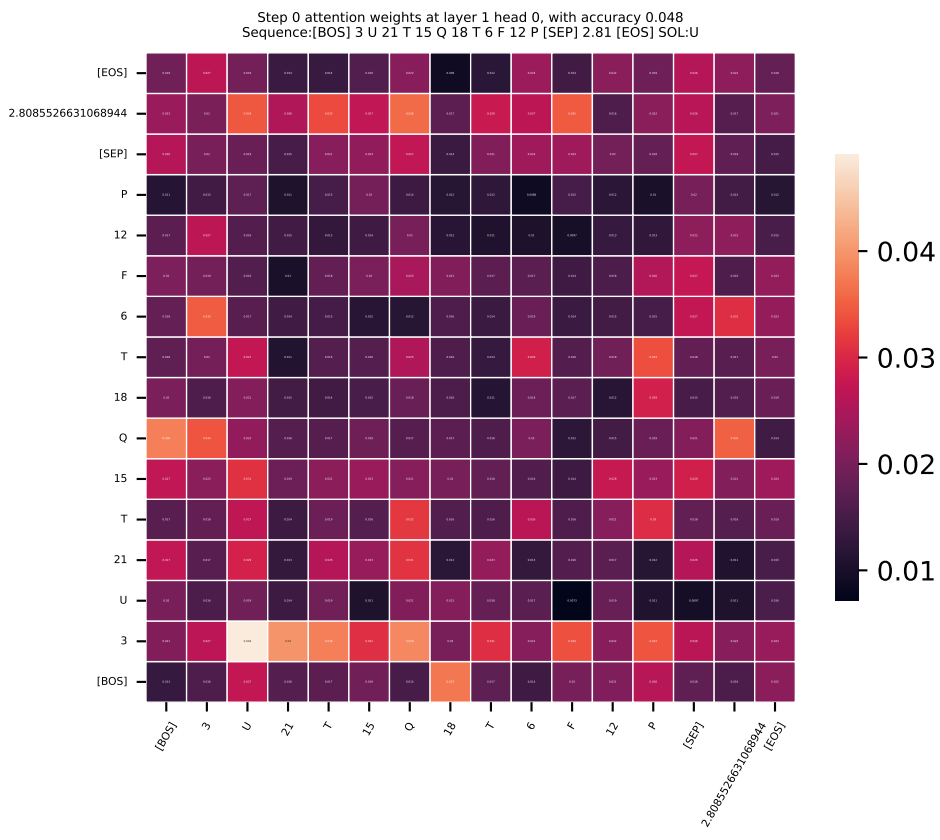
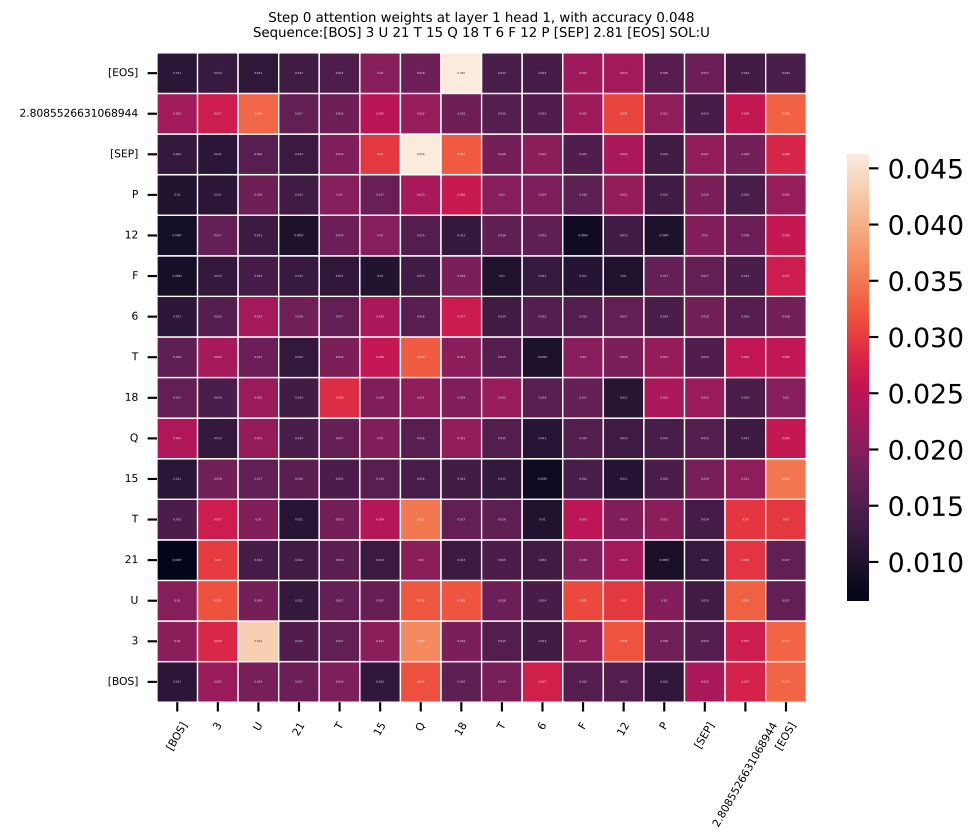

Step 0 attention weights at layer 0 head 0, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.06
0.05
0.04
0.03
0.02
0.01

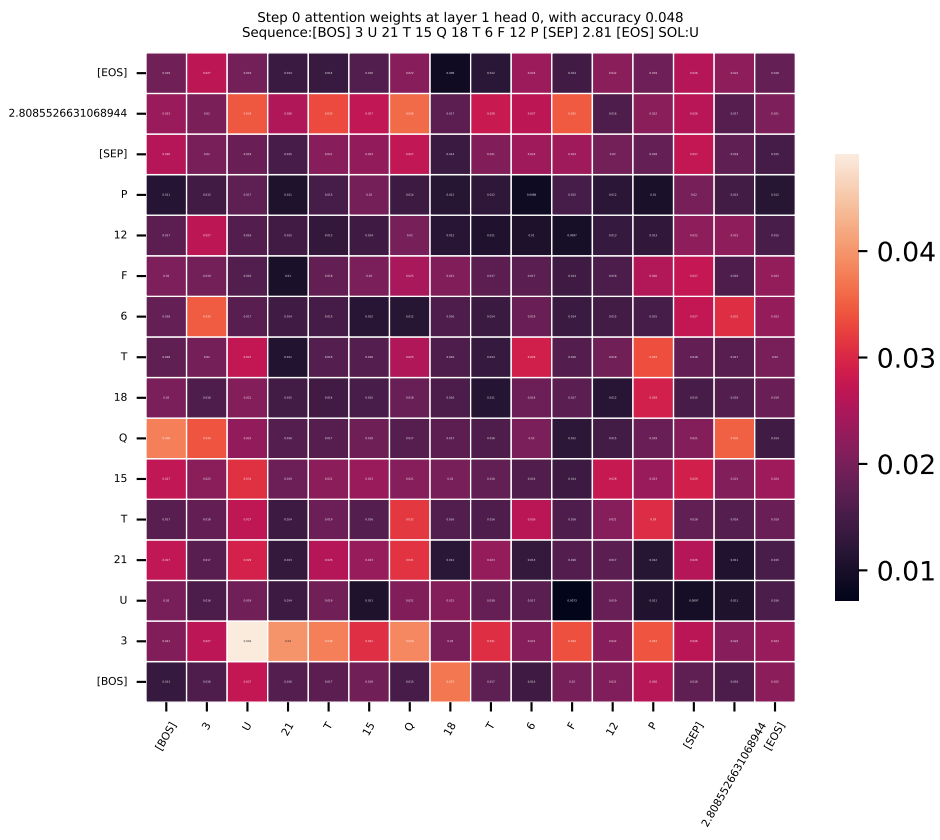

Step 0 attention weights at layer 0 head 1, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.035
0.030
0.025
0.020
0.015
0.010

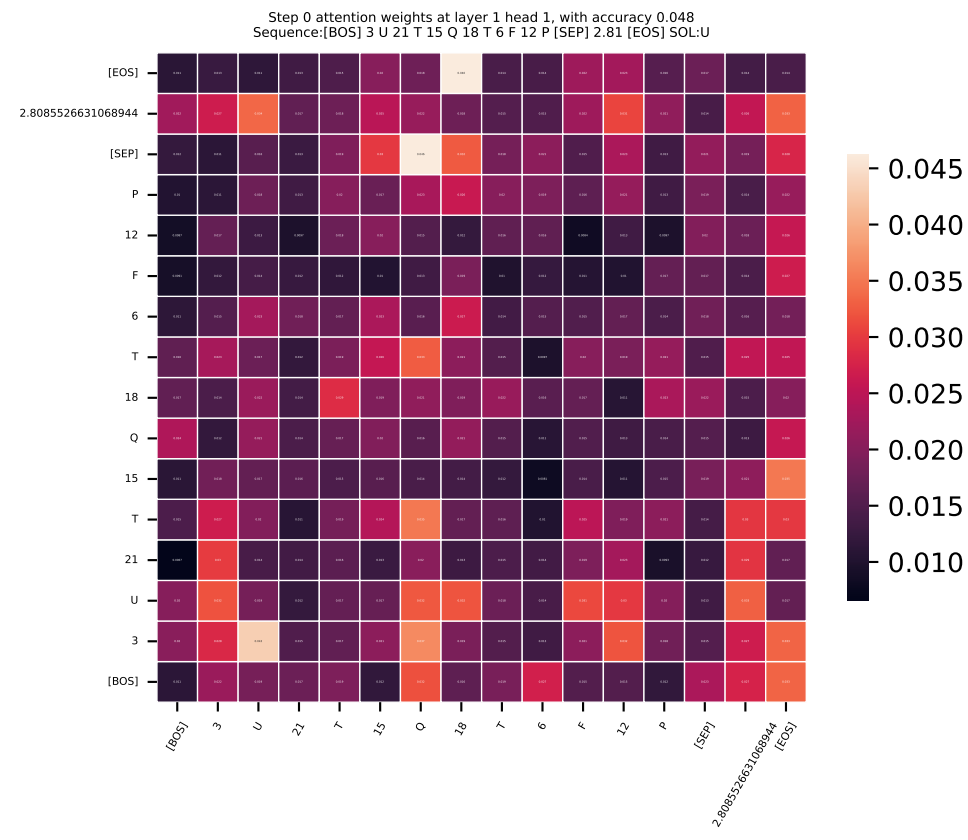

Step 0 attention weights at layer 0 head 0, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.06
0.05
0.04
0.03
0.02
0.01


Step 0 attention weights at layer 0 head 1, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.035
0.030
0.025
0.020
0.015
0.010


Step 0 attention weights at layer 1 head 0, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.04
0.03
0.02
0.01


Step 0 attention weights at layer 1 head 1, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010


Step 0 attention weights at layer 1 head 0, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.04
0.03
0.02
0.01


Step 0 attention weights at layer 1 head 1, with accuracy 0.048
Sequence:[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.81 [EOS] SOL:U
[EOS]
2.8085526631068944
[SEP]
P
12
F
6
T
18
Q
15
T
21
U
3
[BOS]
[BOS] 3 U 21 T 15 Q 18 T 6 F 12 P [SEP] 2.8085526631068944 [EOS]
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010