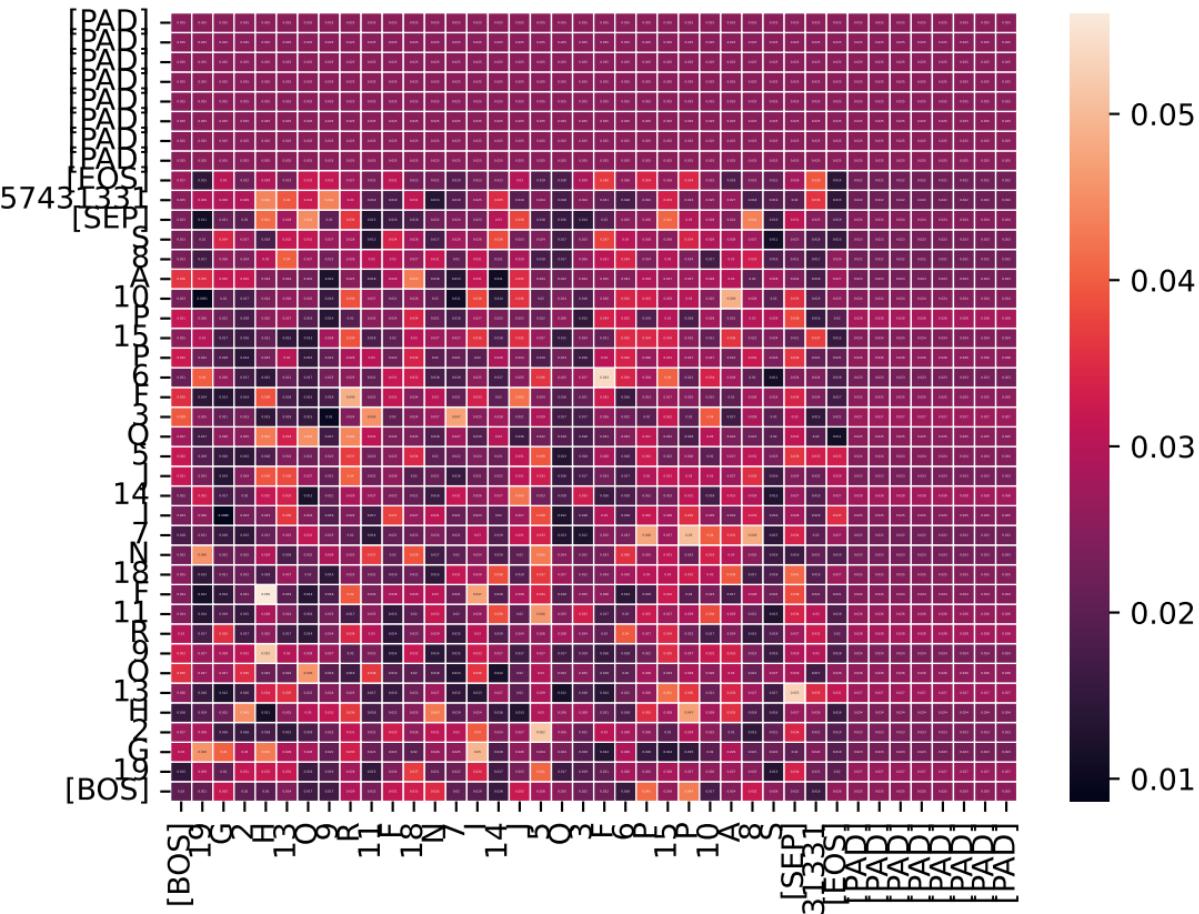


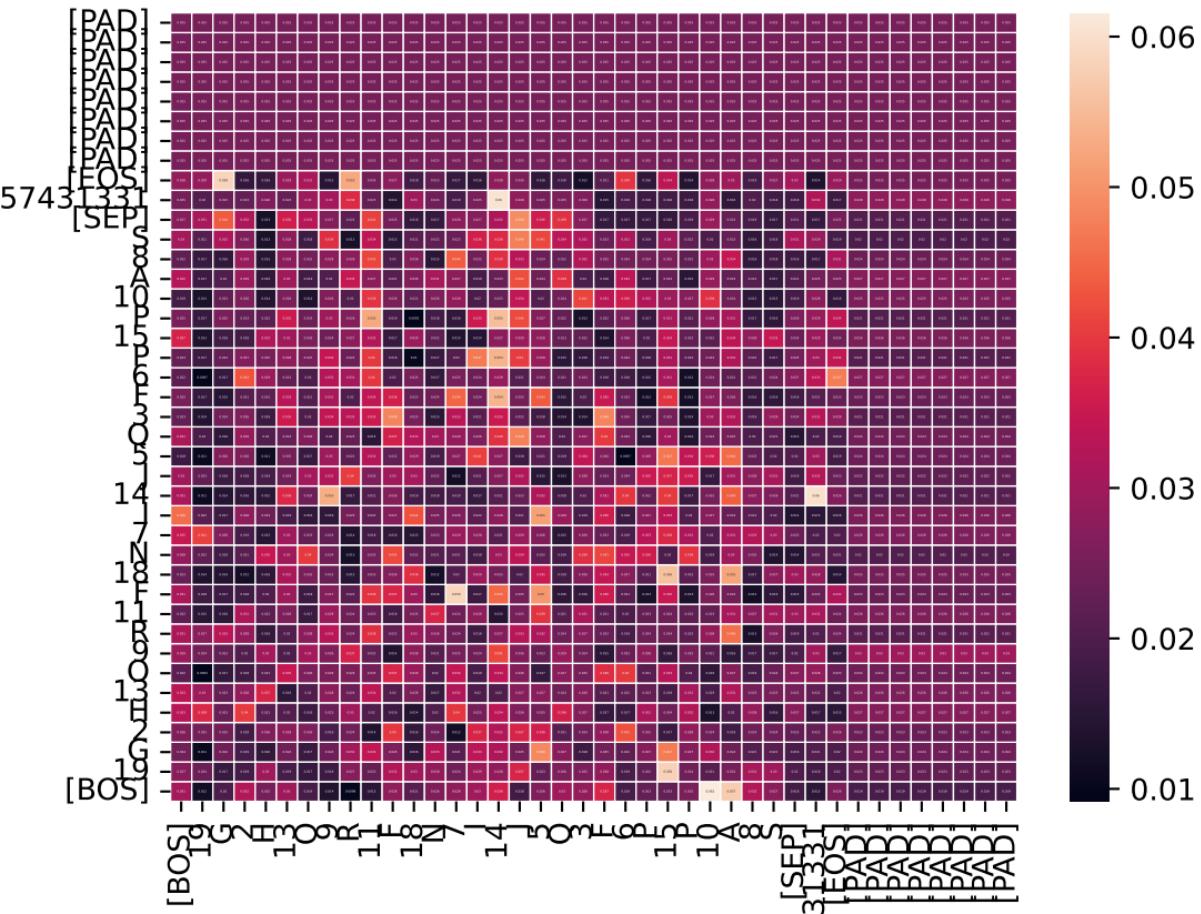
Step 0 attention weights at layer 0 head 0, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 | 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



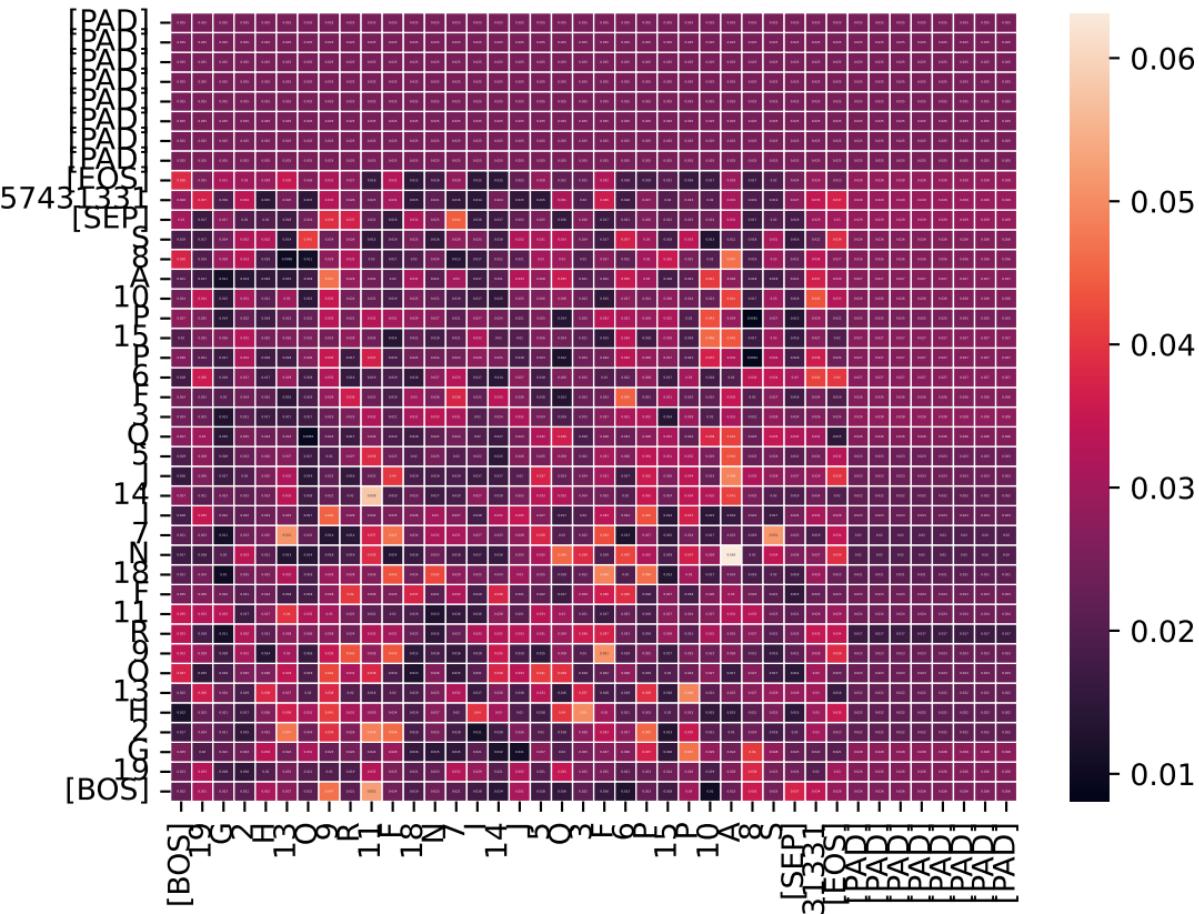
Step 0 attention weights at layer 0 head 1, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 | 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



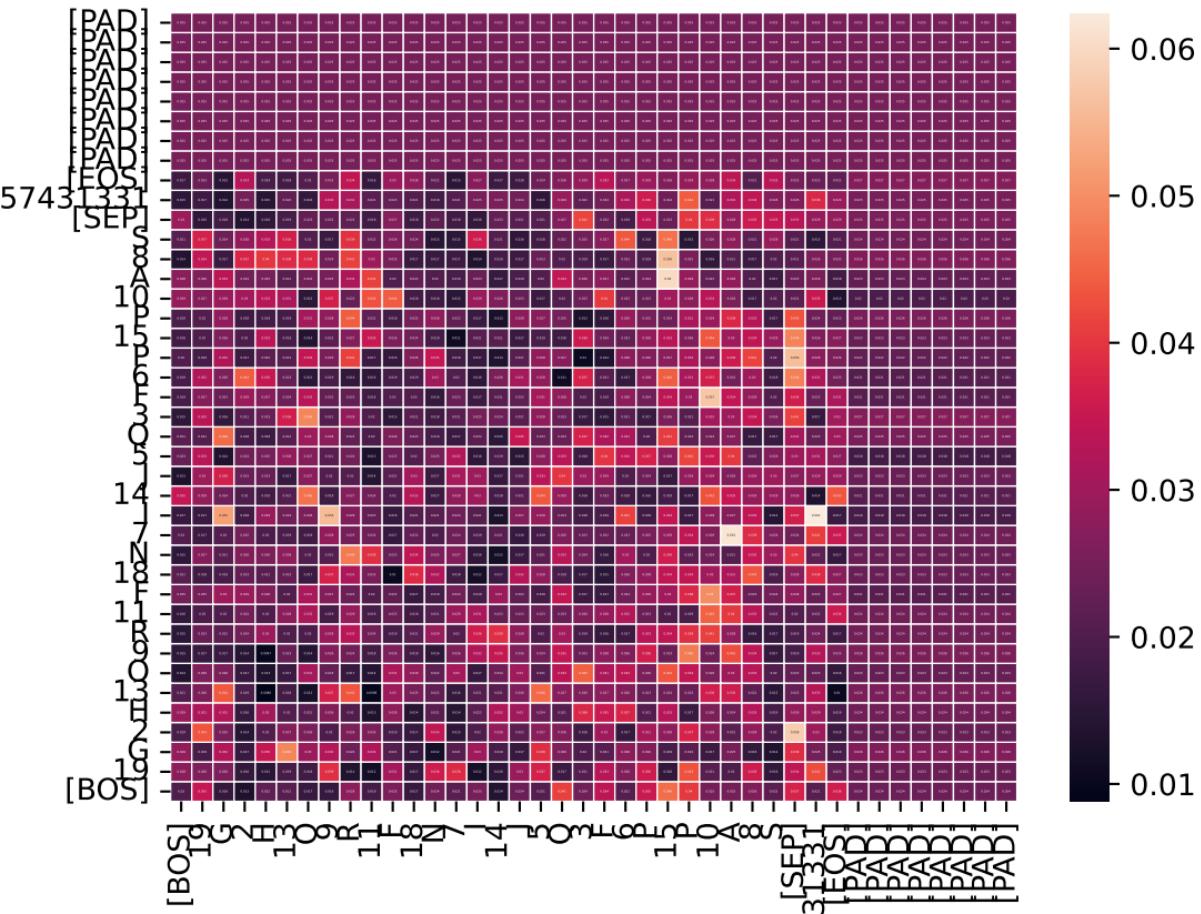
Step 0 attention weights at layer 0 head 2, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 | 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



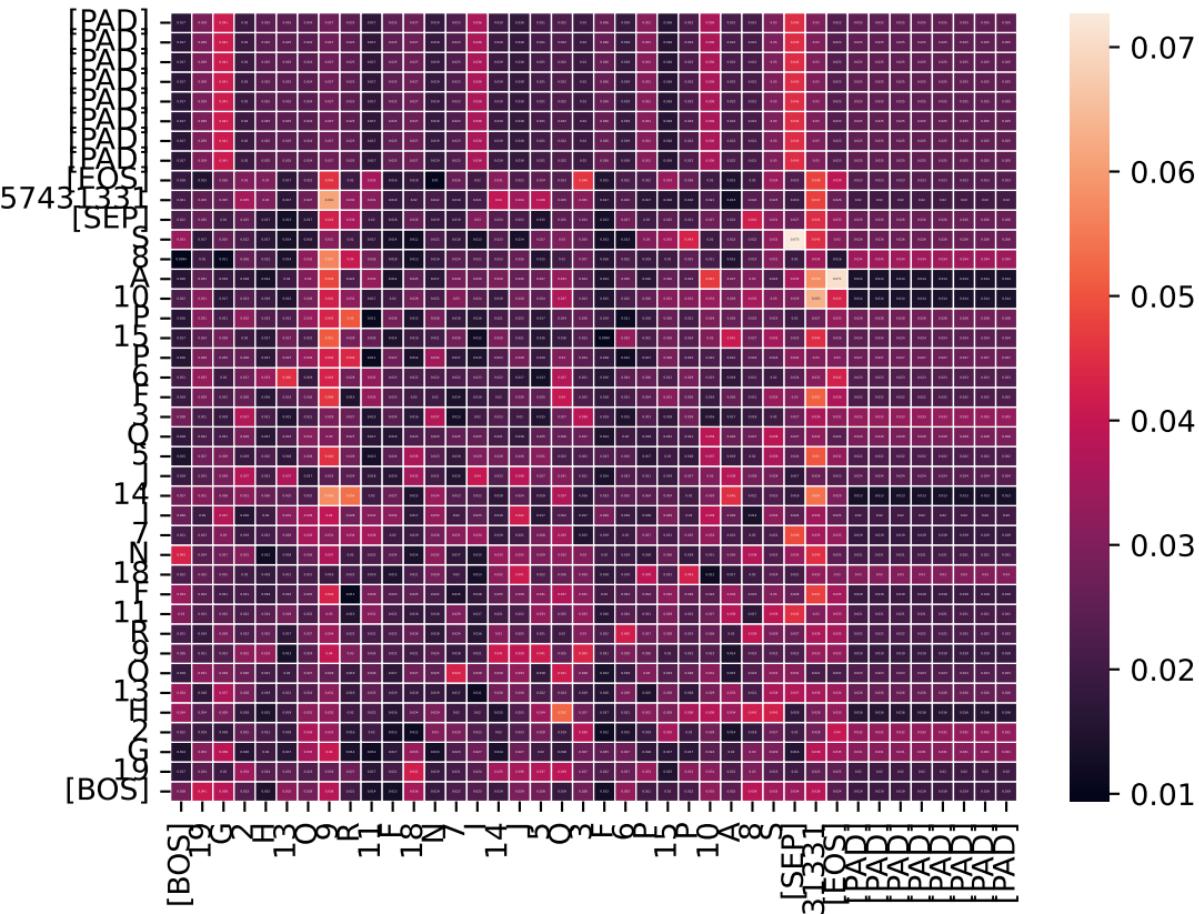
Step 0 attention weights at layer 0 head 3, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 | 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



Step 0 attention weights at layer 1 head 0, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 I 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



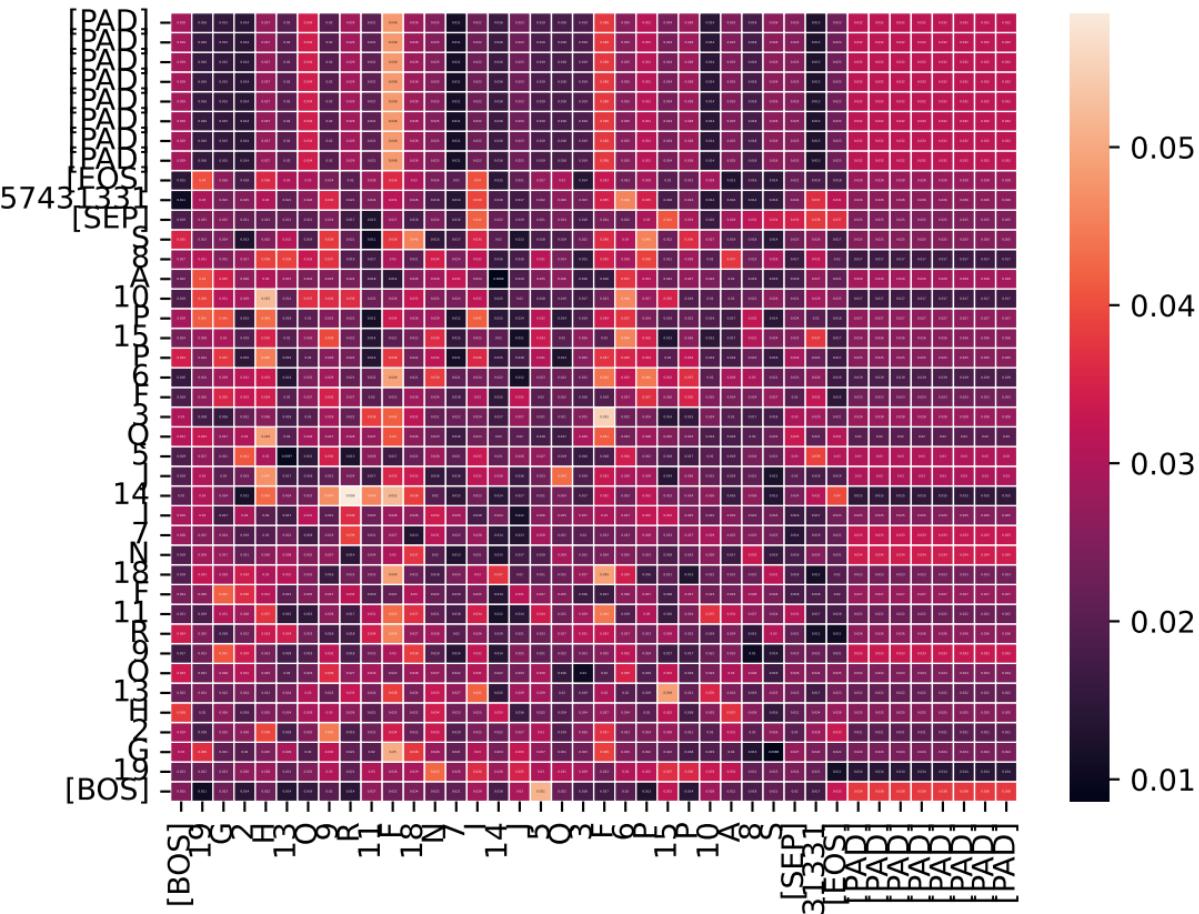
Step 0 attention weights at layer 1 head 1, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 I 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



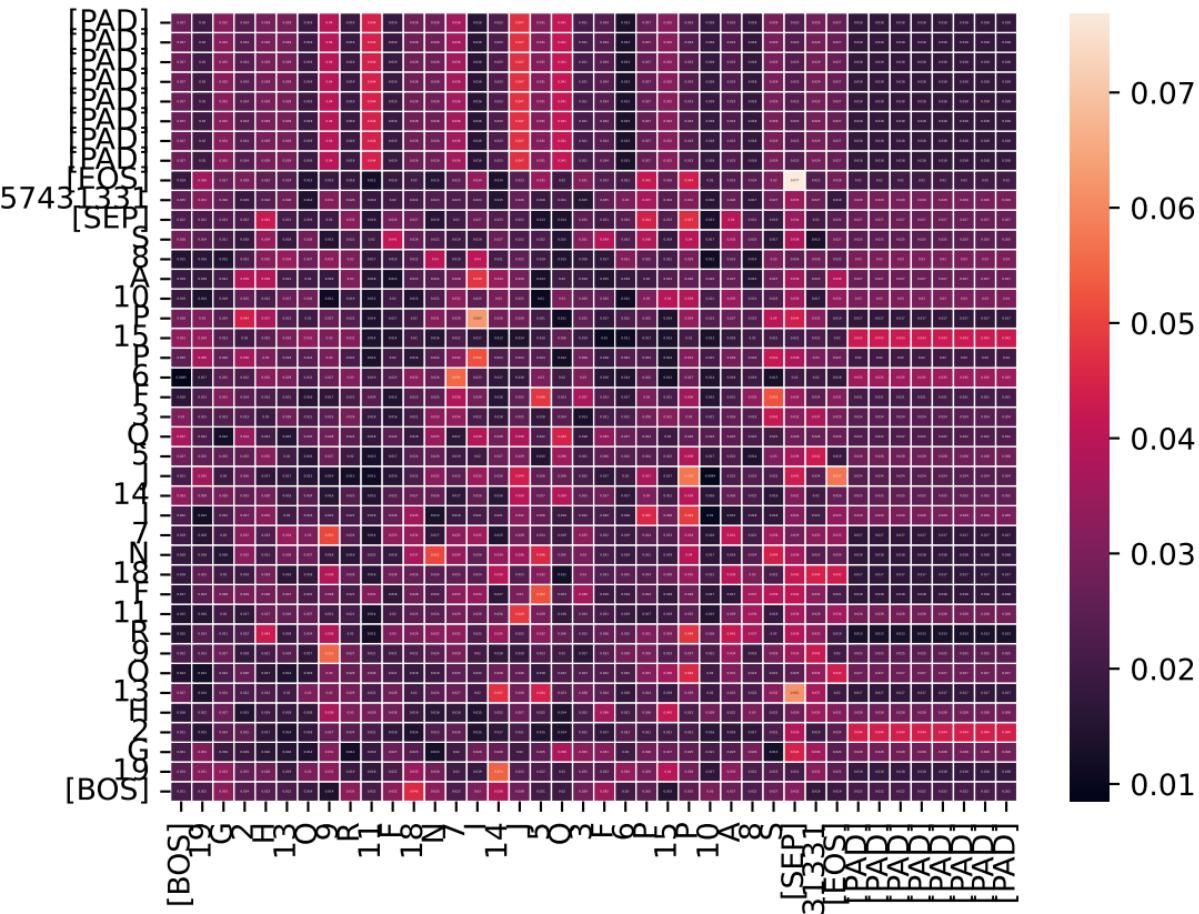
Step 0 attention weights at layer 1 head 2, with accuracy 0.057

Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 | 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R

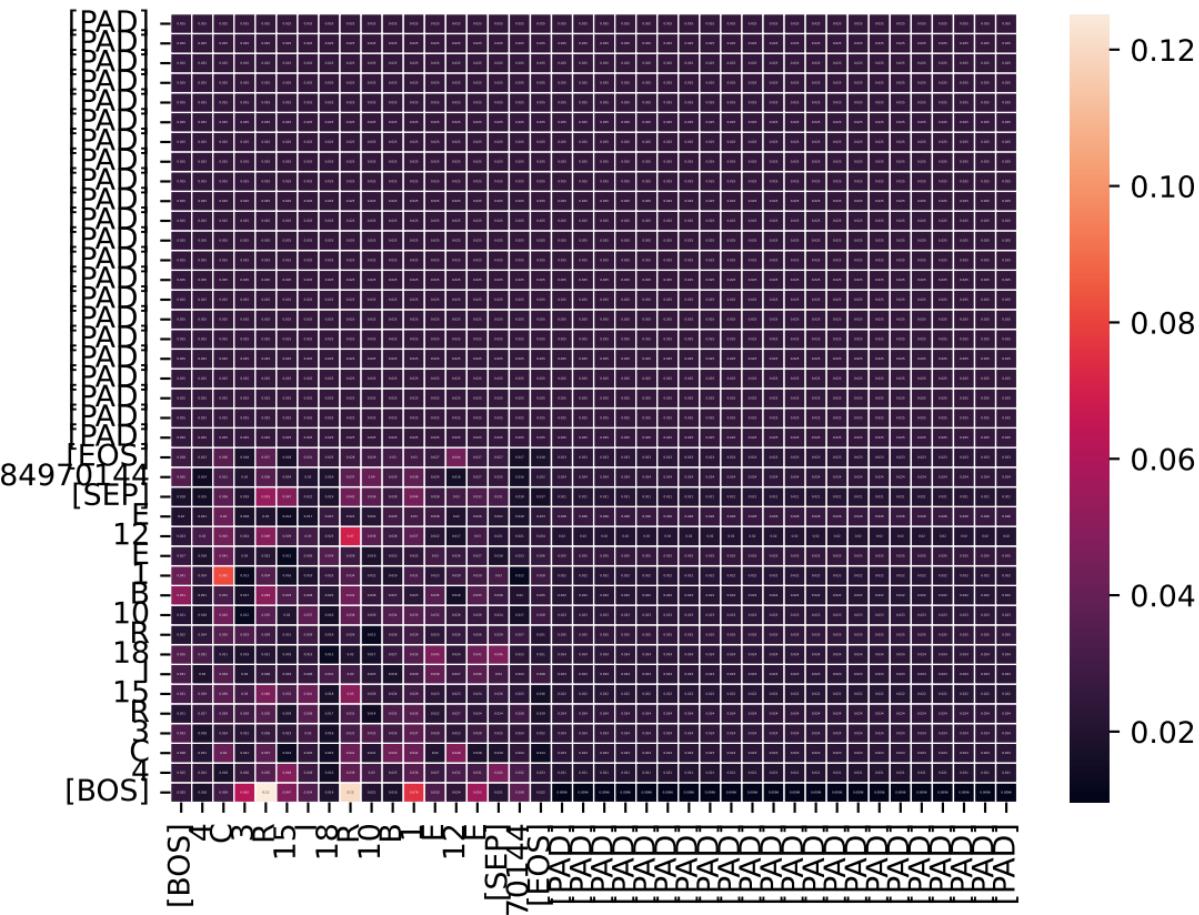


Step 0 attention weights at layer 1 head 3, with accuracy 0.057

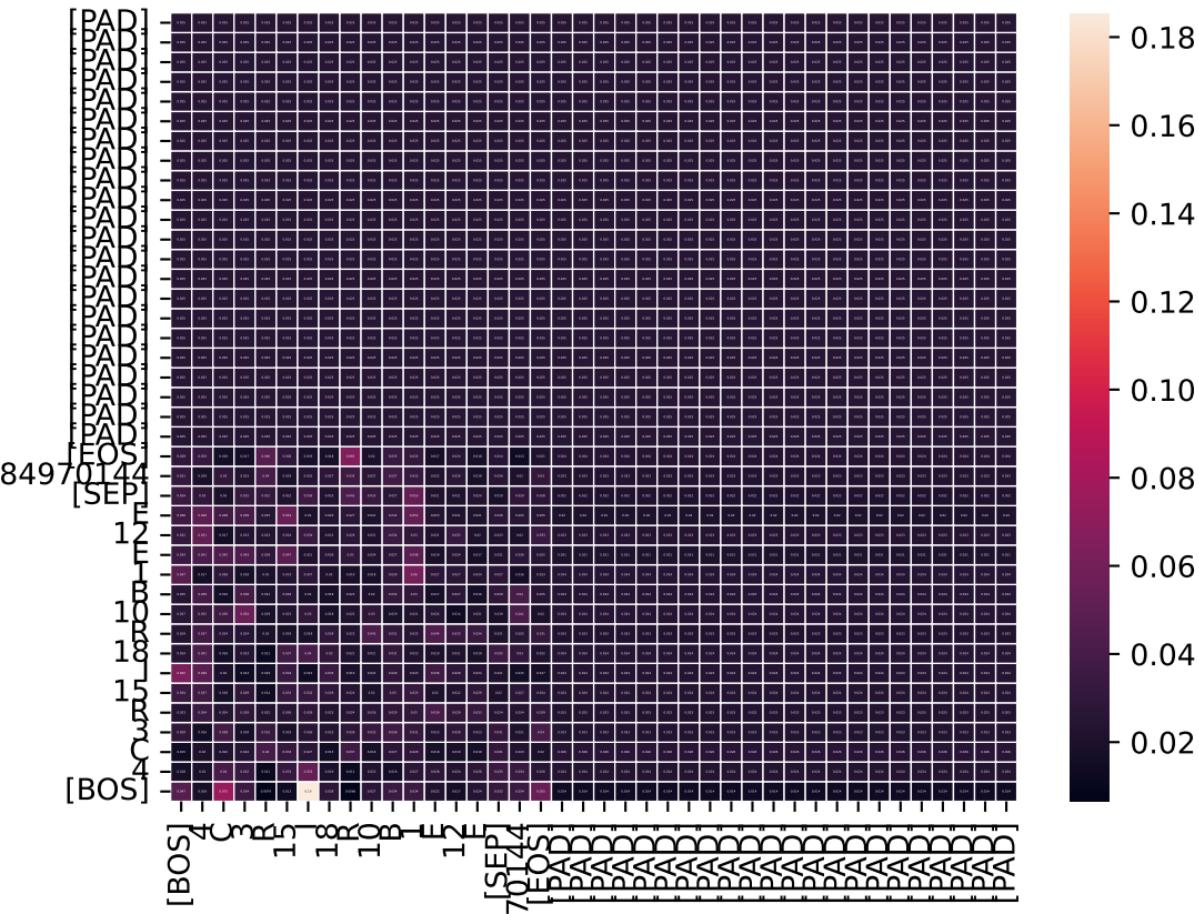
Sequence:[BOS] 19 G 2 H 13 O 9 R 11 F 18 N 7 I 14 J 5 Q 3 F 6 P 15 P 10 A 8 S [SEP] 8.88 [EOS] SOL:R



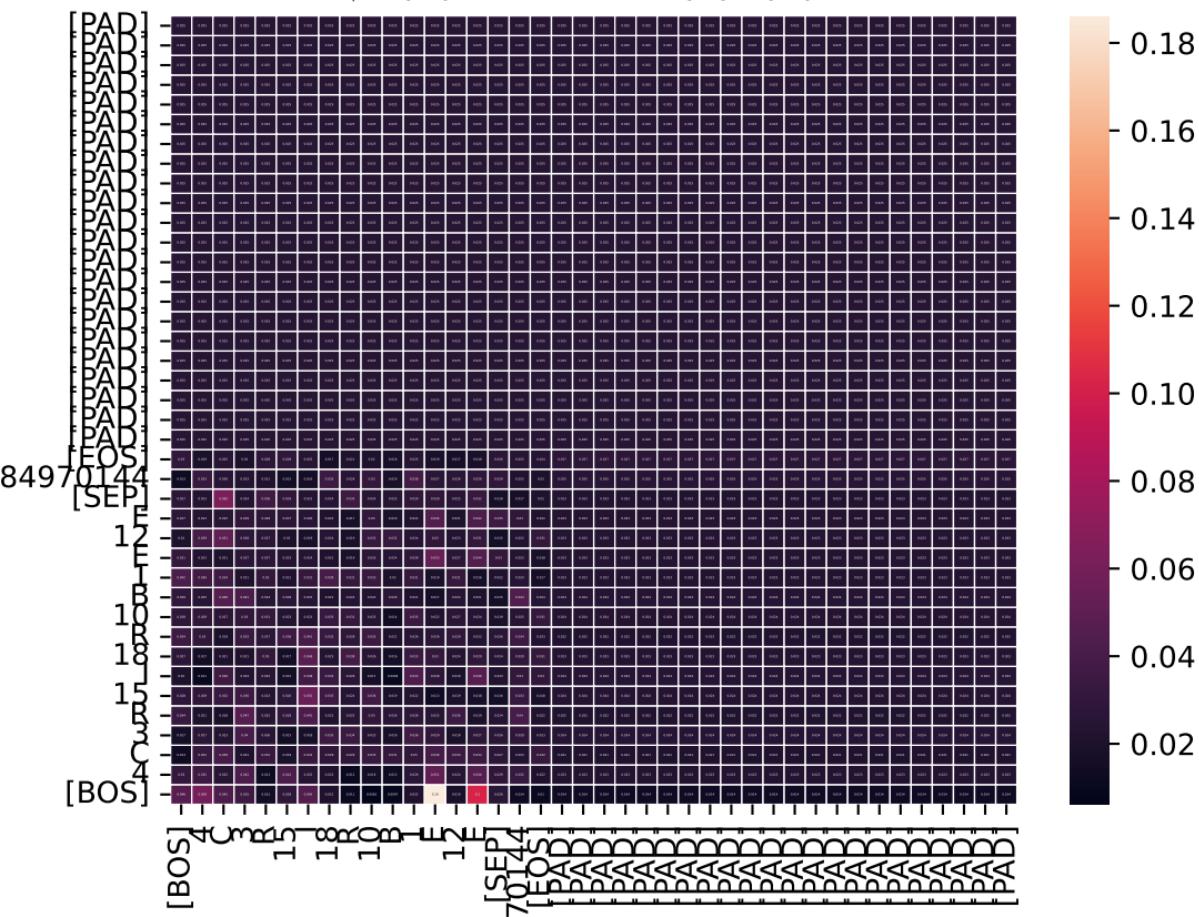
Step 500 attention weights at layer 0 head 0, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



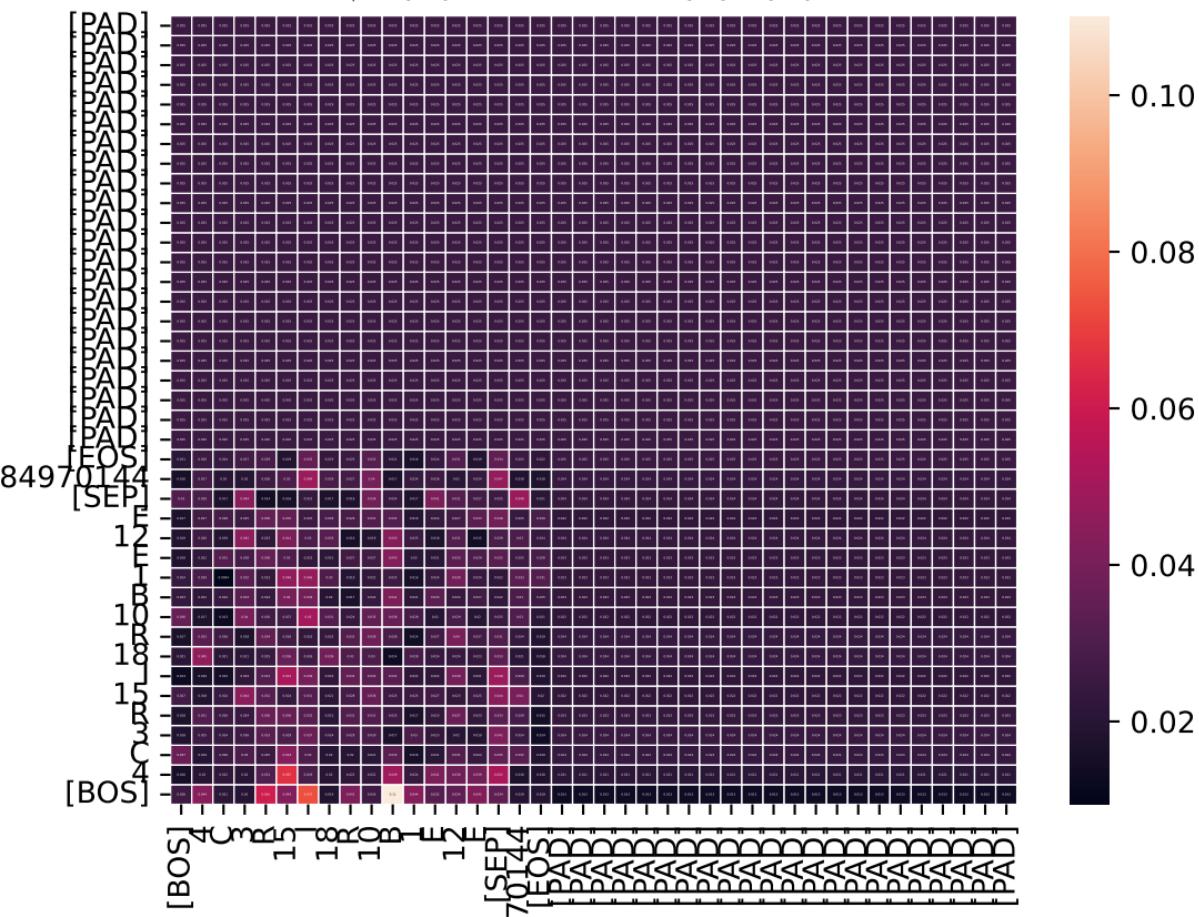
Step 500 attention weights at layer 0 head 1, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



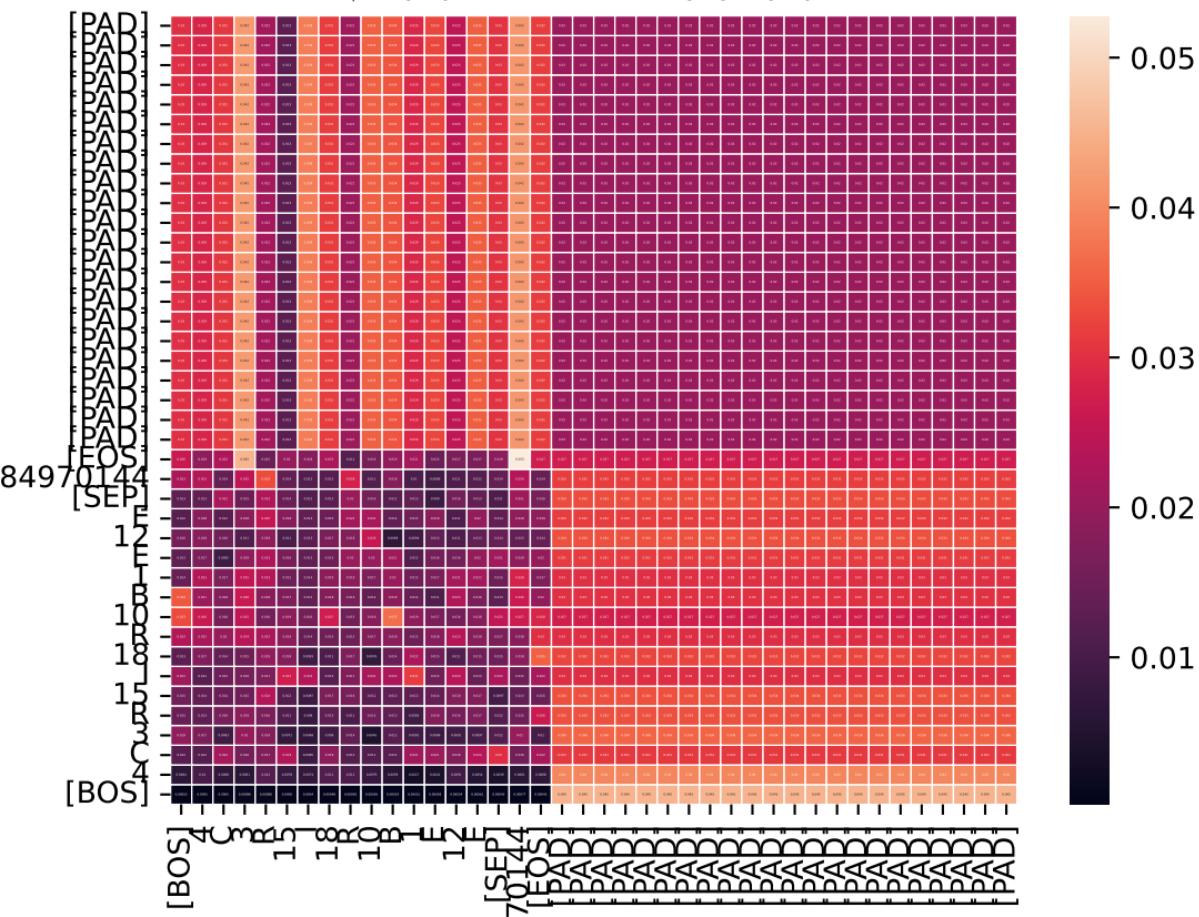
Step 500 attention weights at layer 0 head 2, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



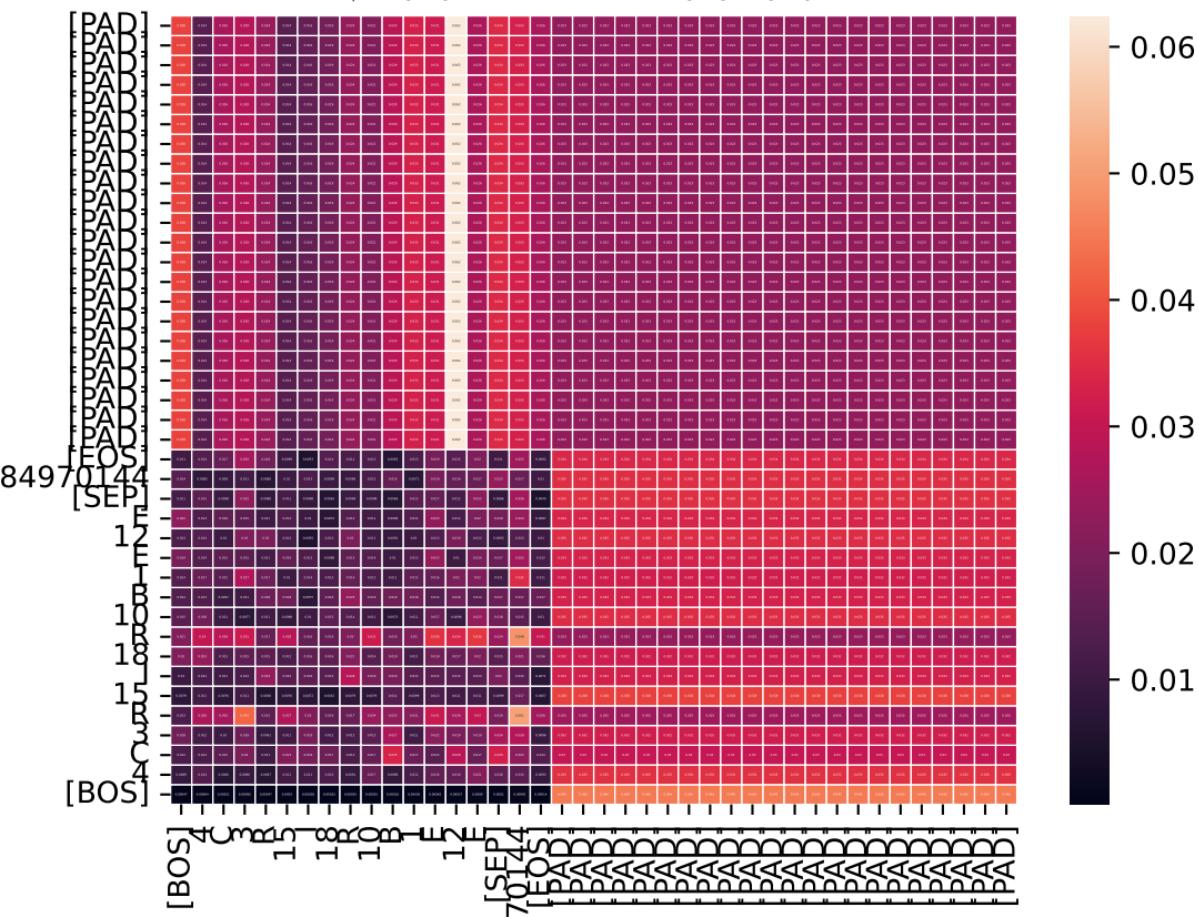
Step 500 attention weights at layer 0 head 3, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



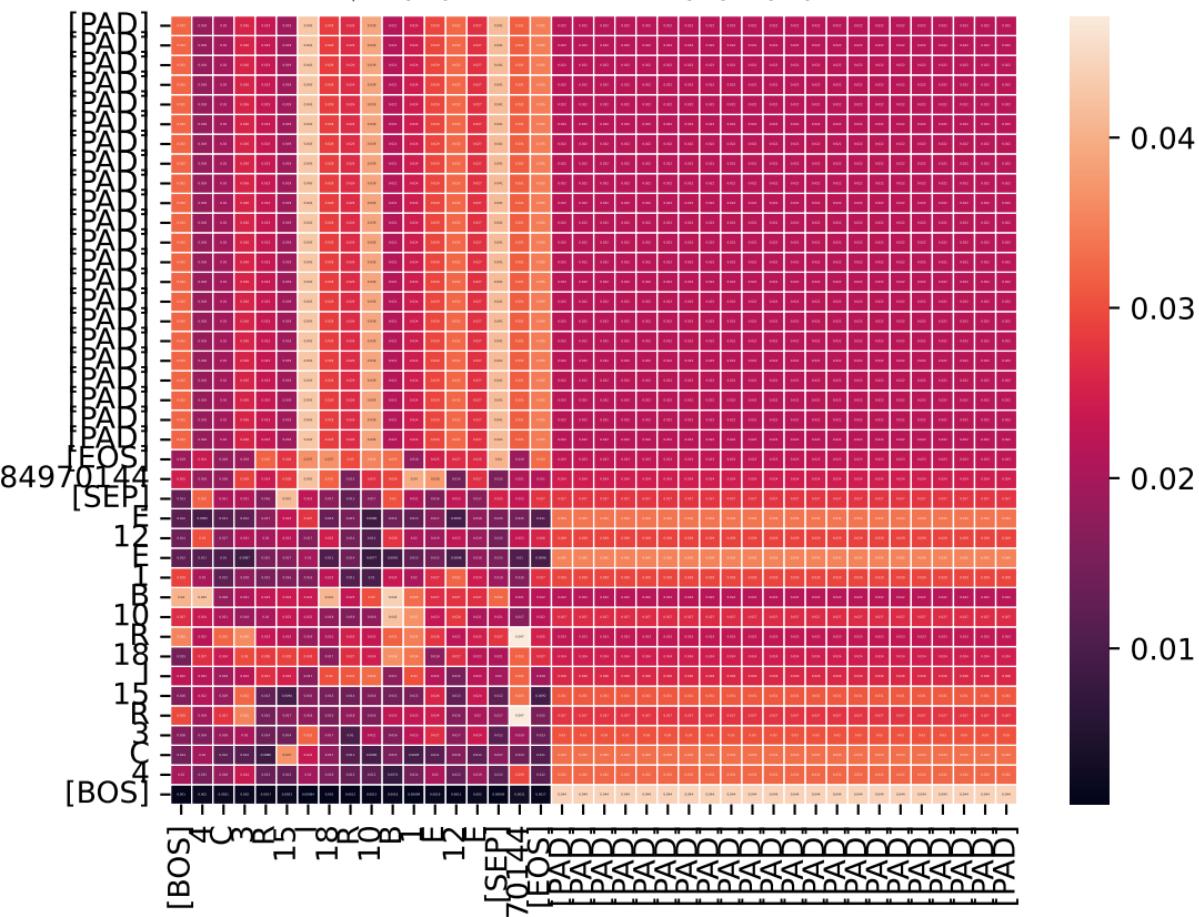
Step 500 attention weights at layer 1 head 0, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



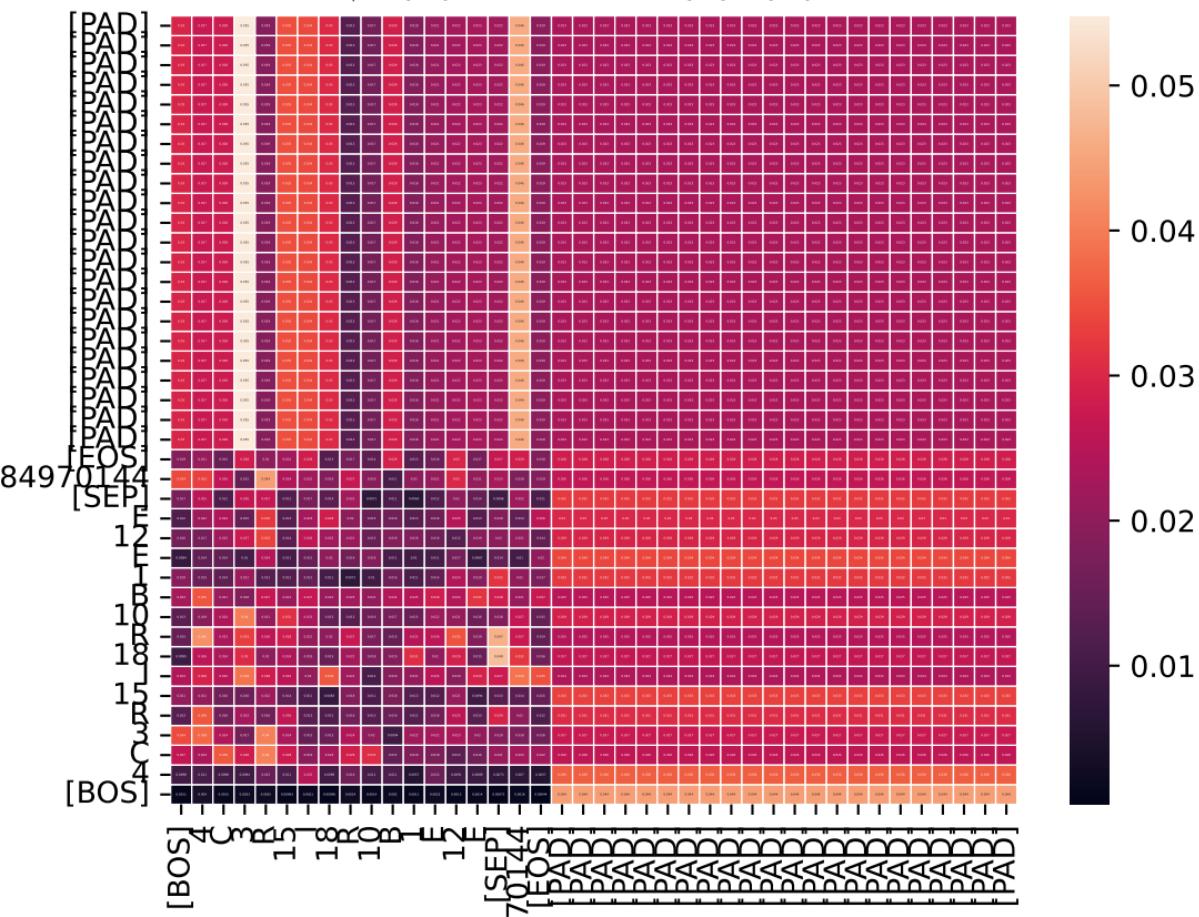
Step 500 attention weights at layer 1 head 1, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R



Step 500 attention weights at layer 1 head 2, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R

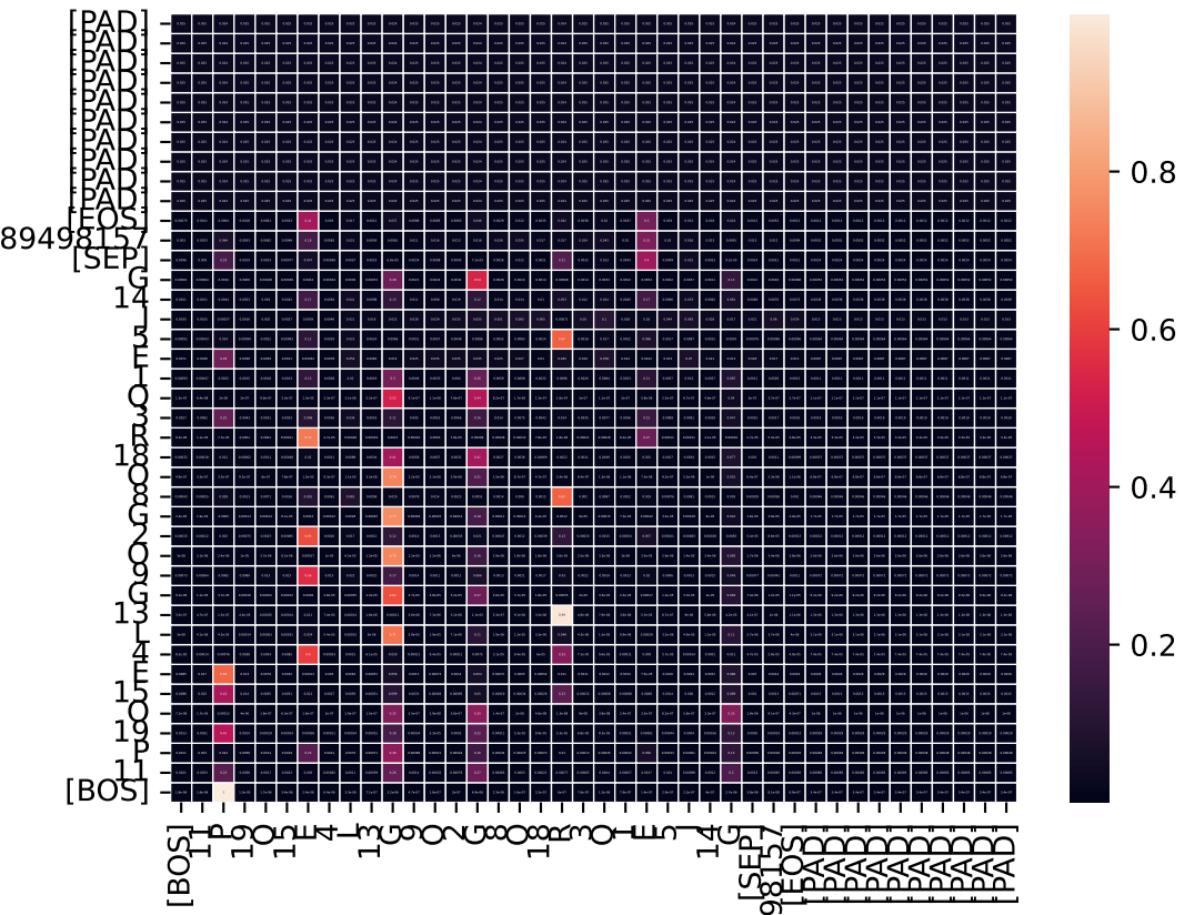


Step 500 attention weights at layer 1 head 3, with accuracy 0.233
Sequence:[BOS] 4 C 3 R 15 I 18 R 10 B 1 E 12 E [SEP] 3.09 [EOS] SOL:R

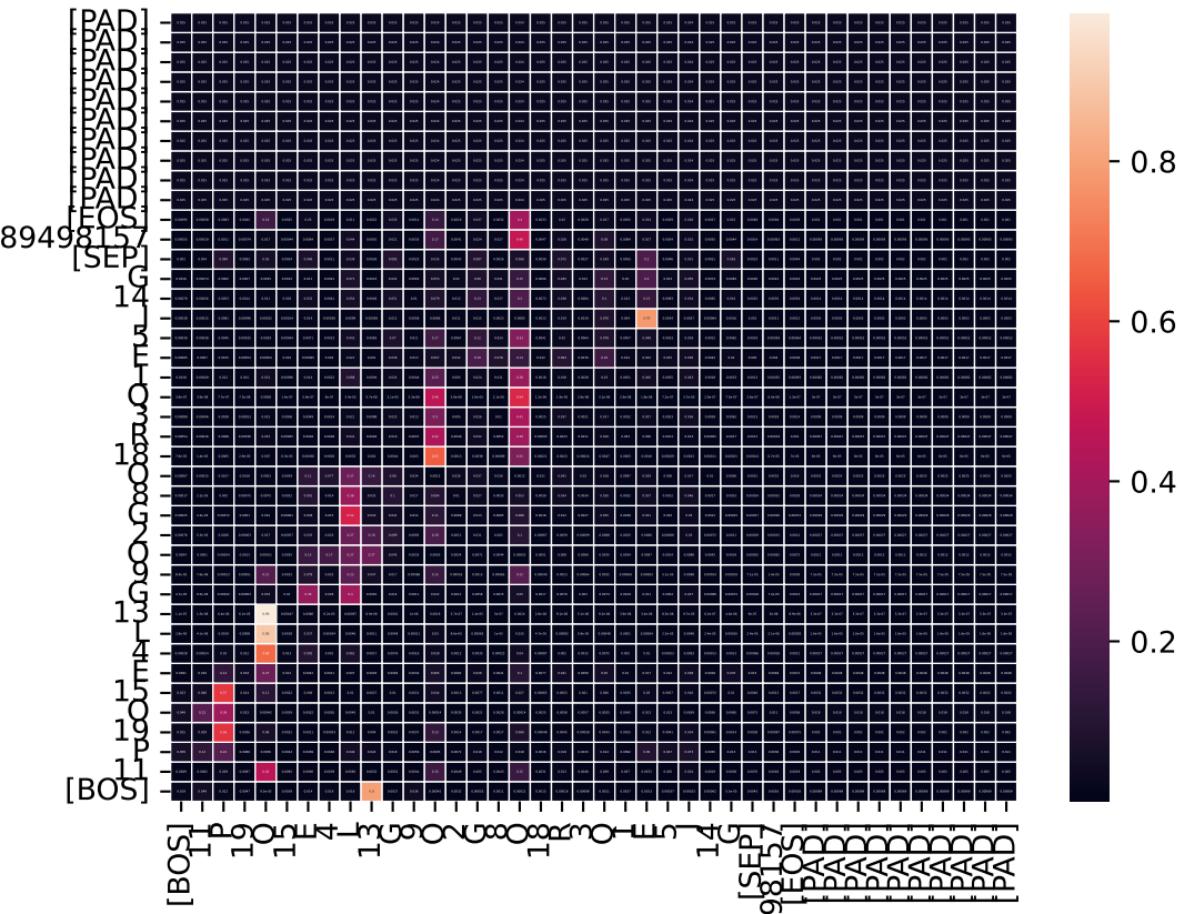


Step 9500 attention weights at layer 0 head 0, with accuracy 0.251

Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SLO:0

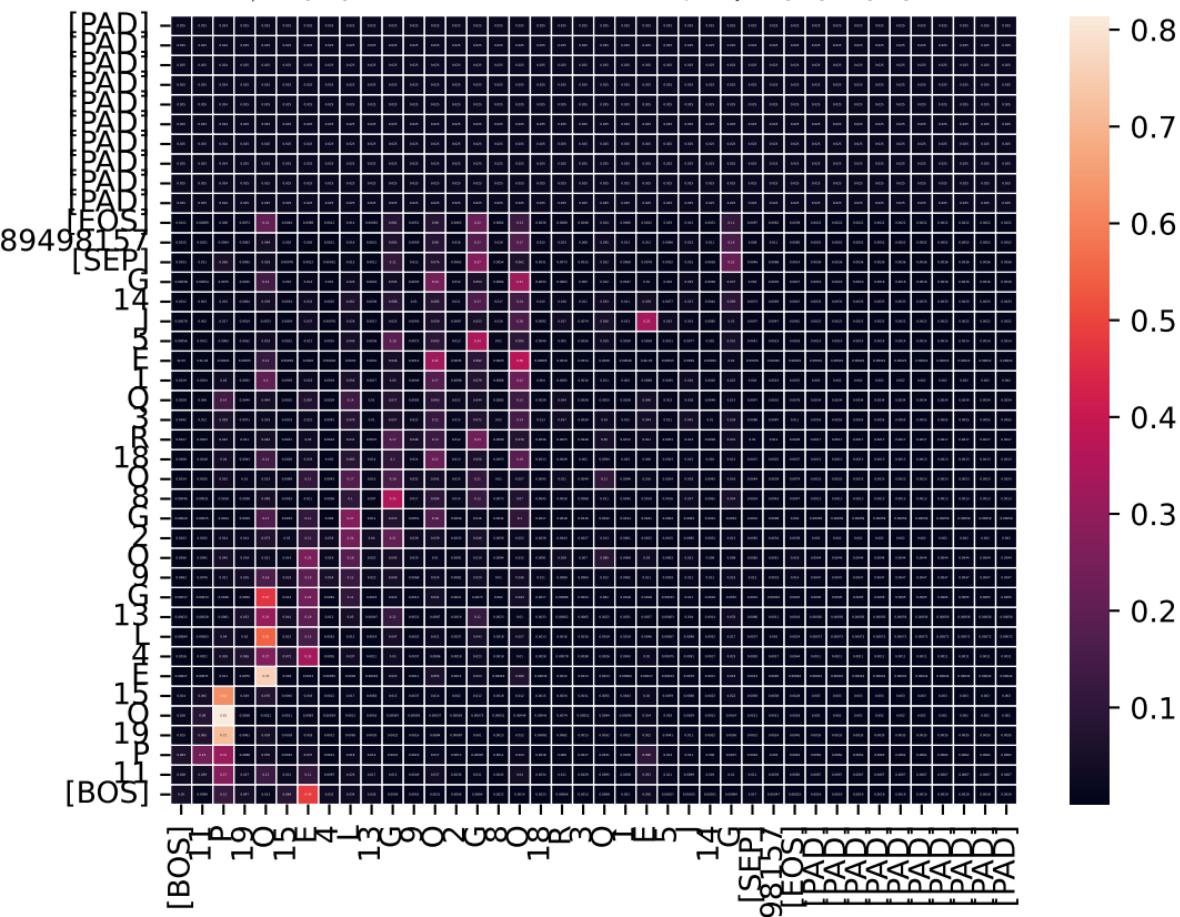


Step 9500 attention weights at layer 0 head 1, with accuracy 0.251
Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SOL:O



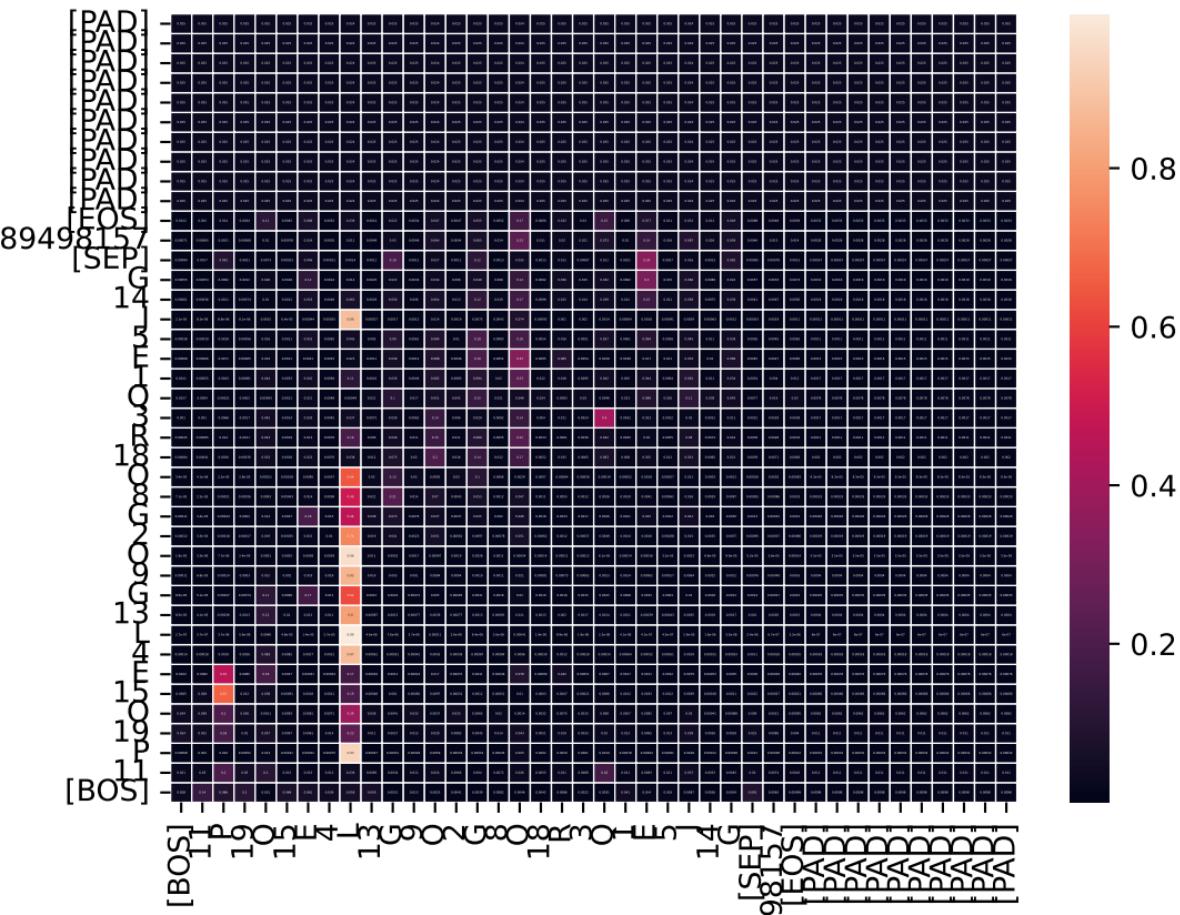
Step 9500 attention weights at layer 0 head 2, with accuracy 0.251

Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] S0L:0



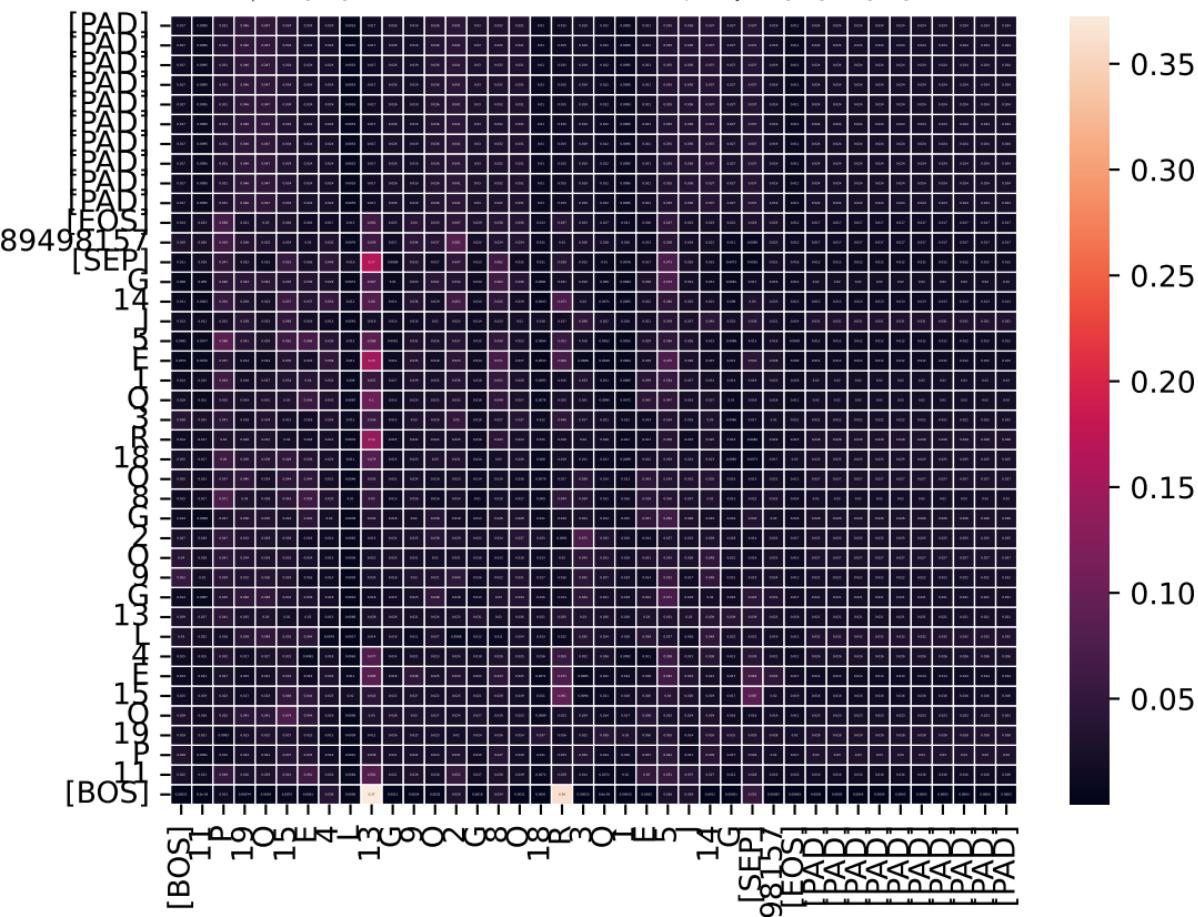
Step 9500 attention weights at layer 0 head 3, with accuracy 0.251

Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SLO:0

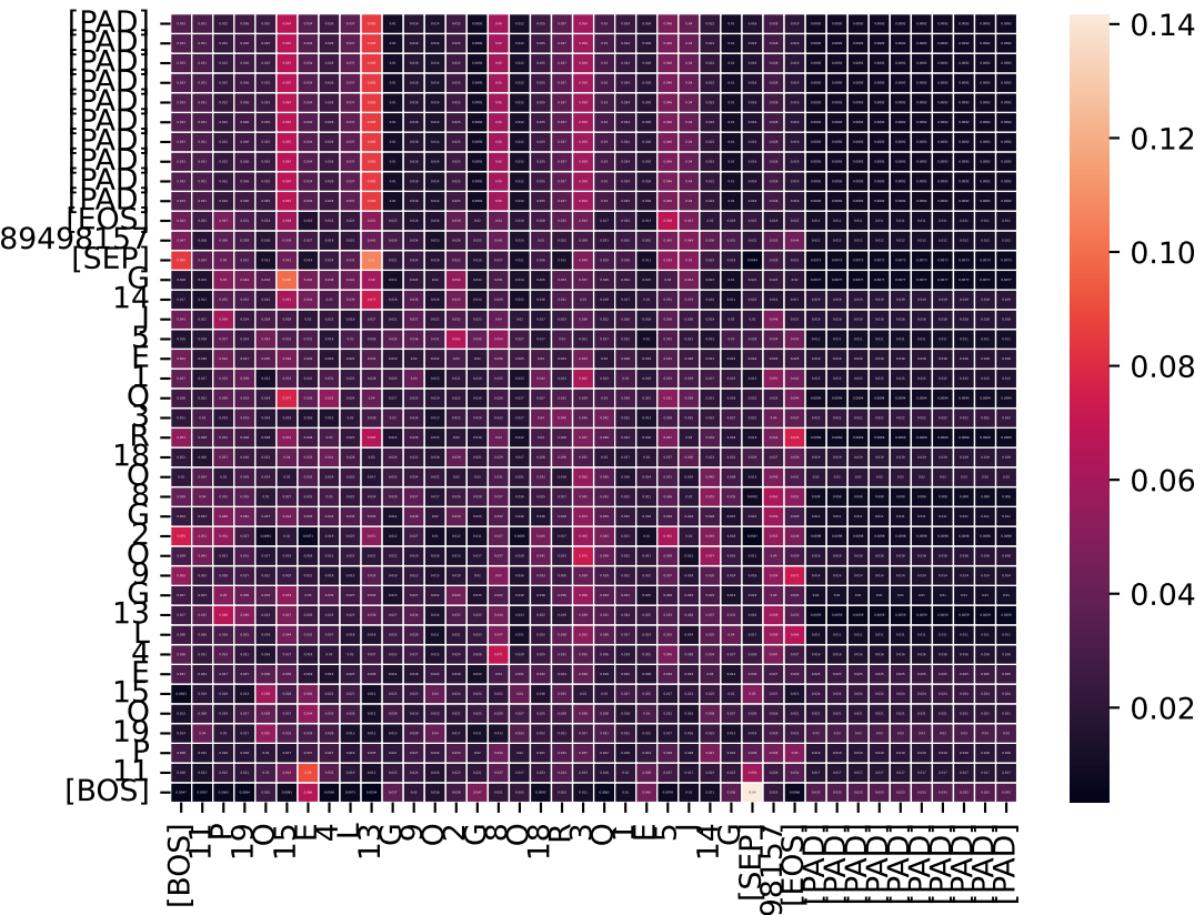


Step 9500 attention weights at layer 1 head 0, with accuracy 0.251

Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SLO:0

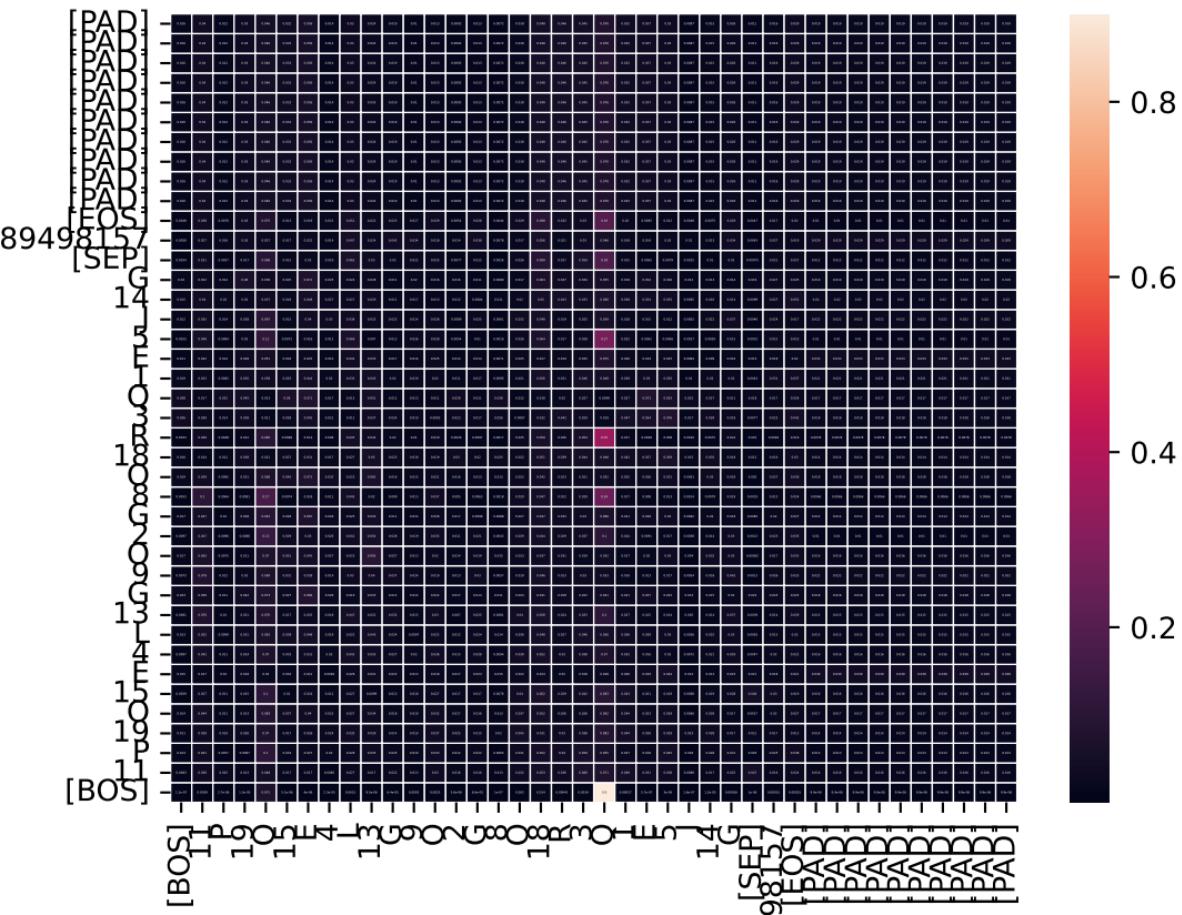


Step 9500 attention weights at layer 1 head 1, with accuracy 0.251
Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5] 14 G [SEP] 9.17 [EOS] SOL:O



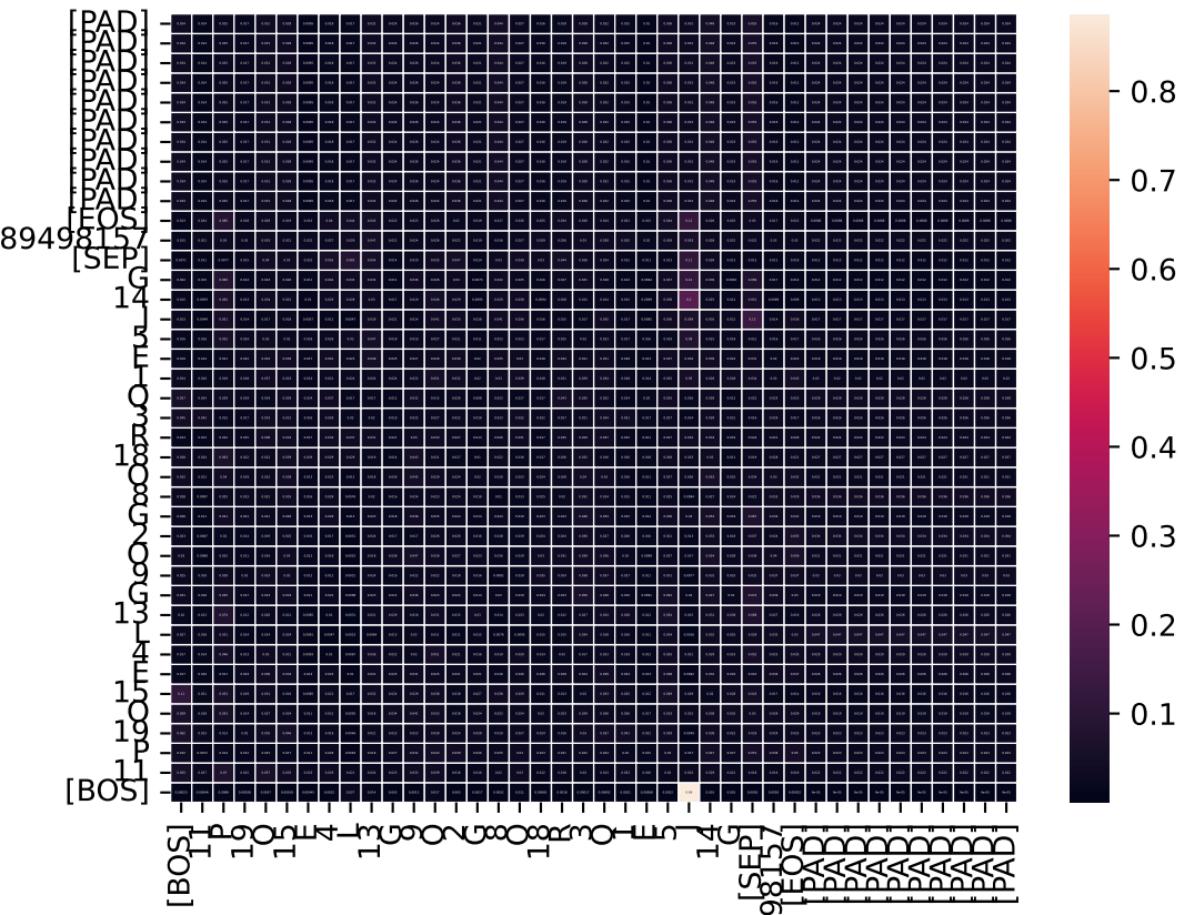
Step 9500 attention weights at layer 1 head 2, with accuracy 0.251

Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SLO:0

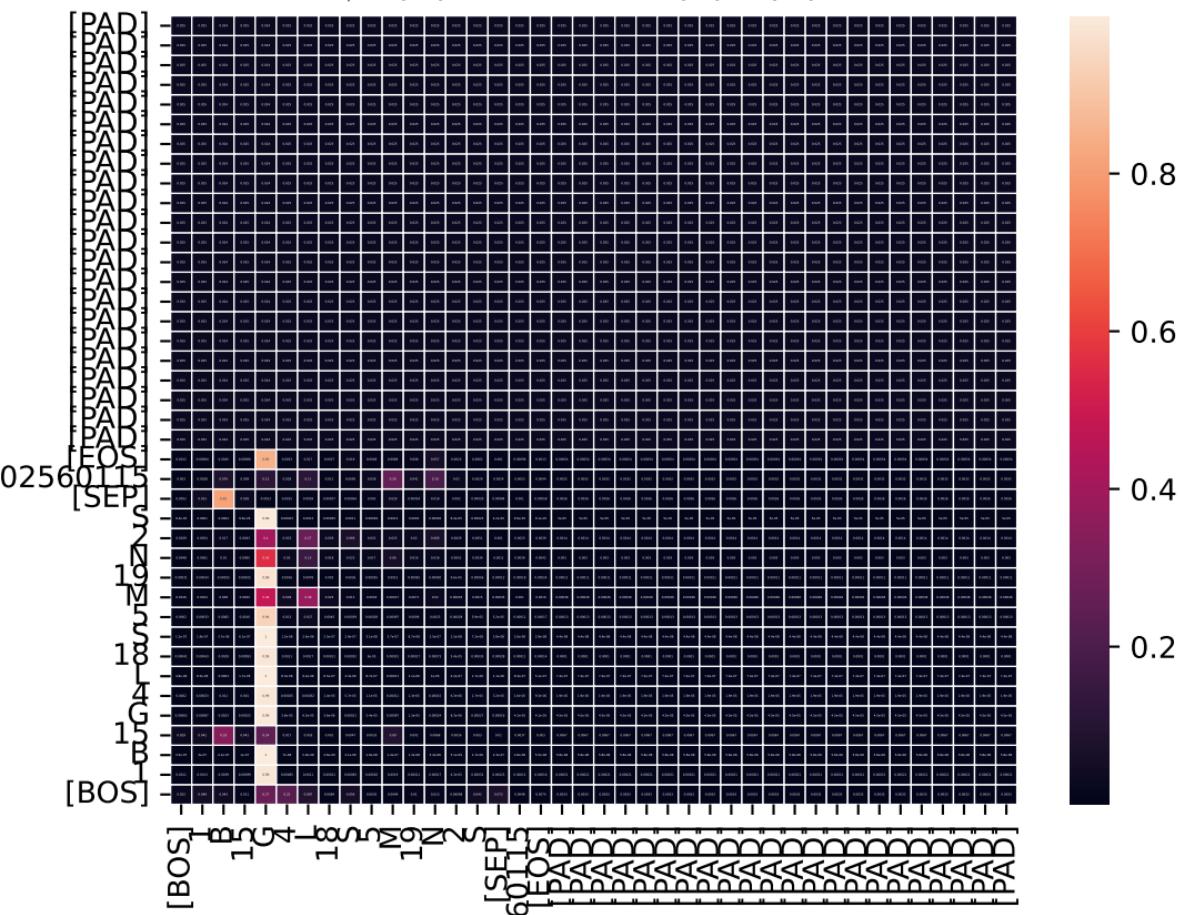


Step 9500 attention weights at layer 1 head 3, with accuracy 0.251

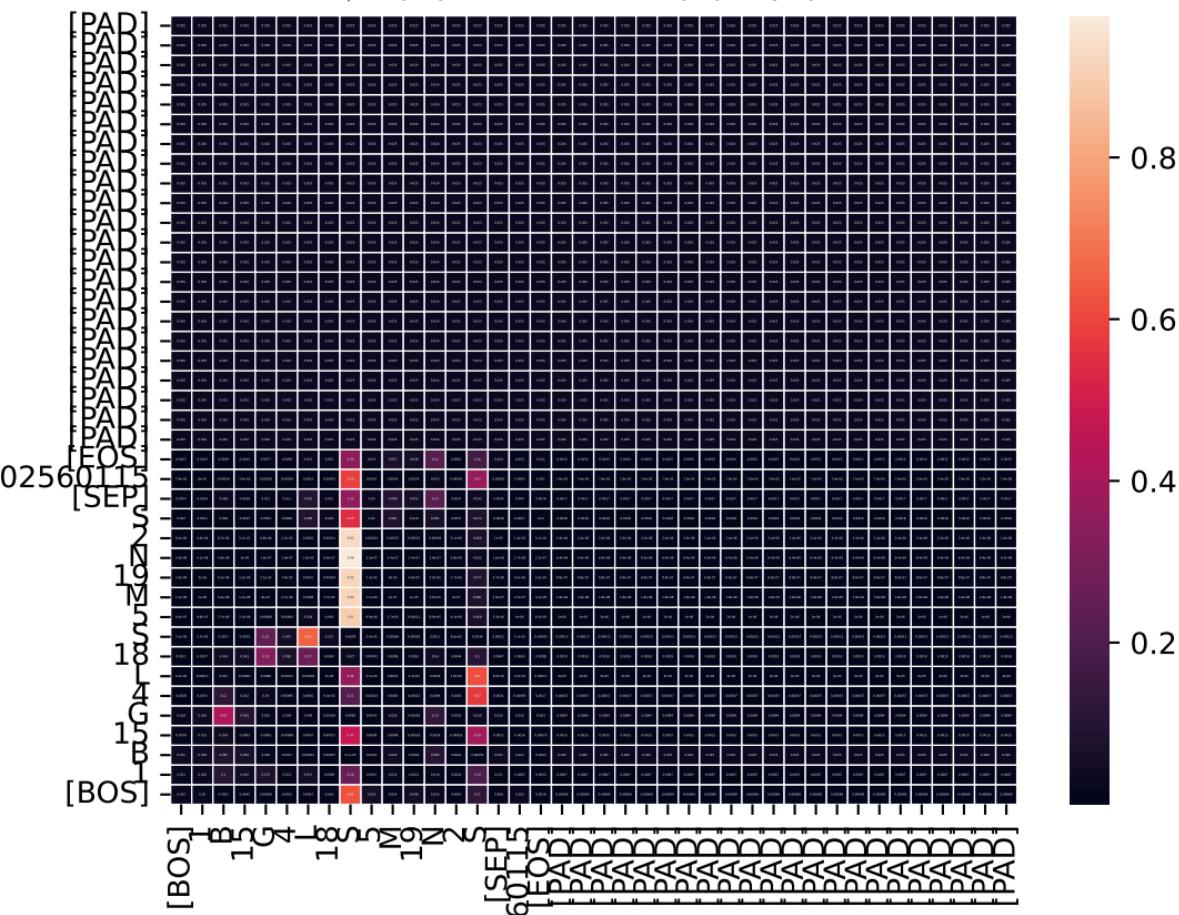
Sequence:[BOS] 11 P 19 O 15 E 4 L 13 G 9 O 2 G 8 O 18 R 3 Q 1 E 5 J 14 G [SEP] 9.17 [EOS] SOL:O



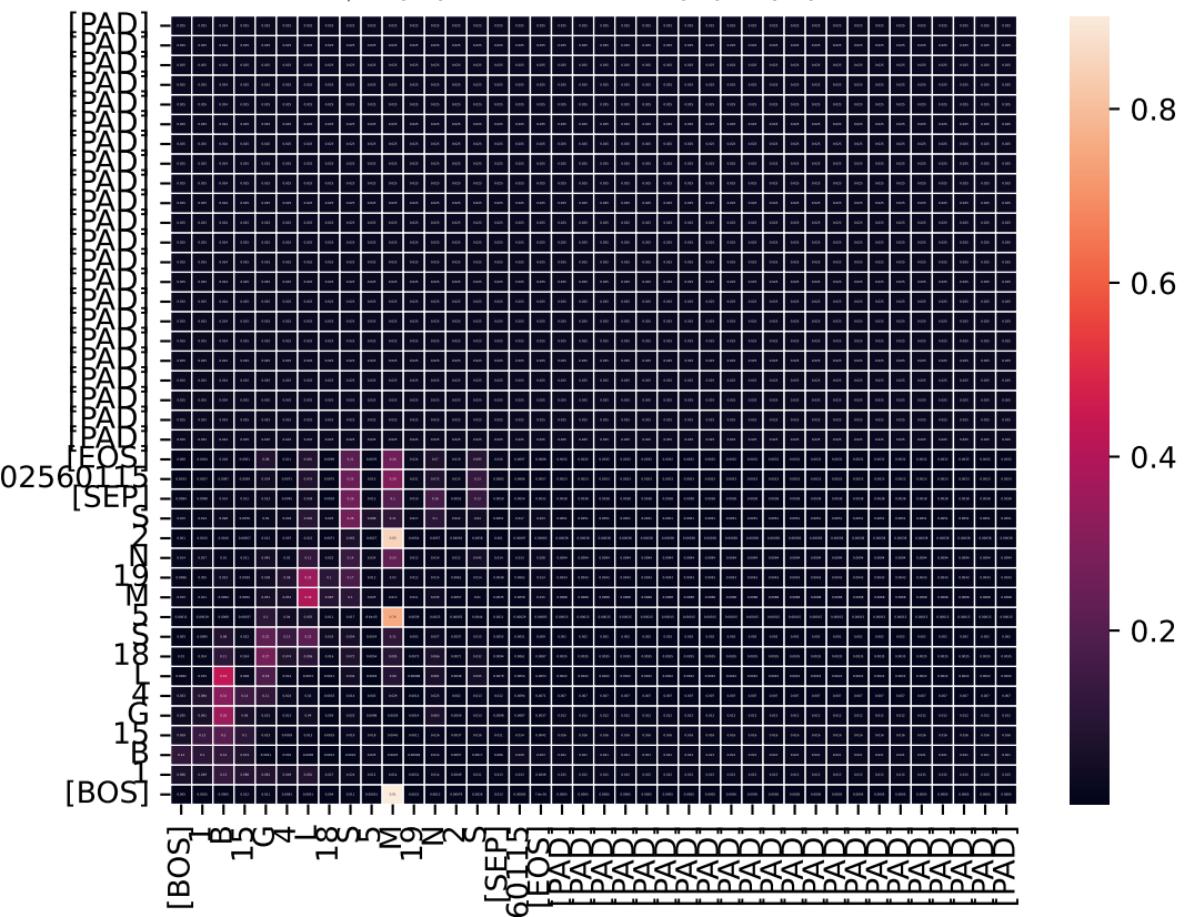
Step 10000 attention weights at layer 0 head 0, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



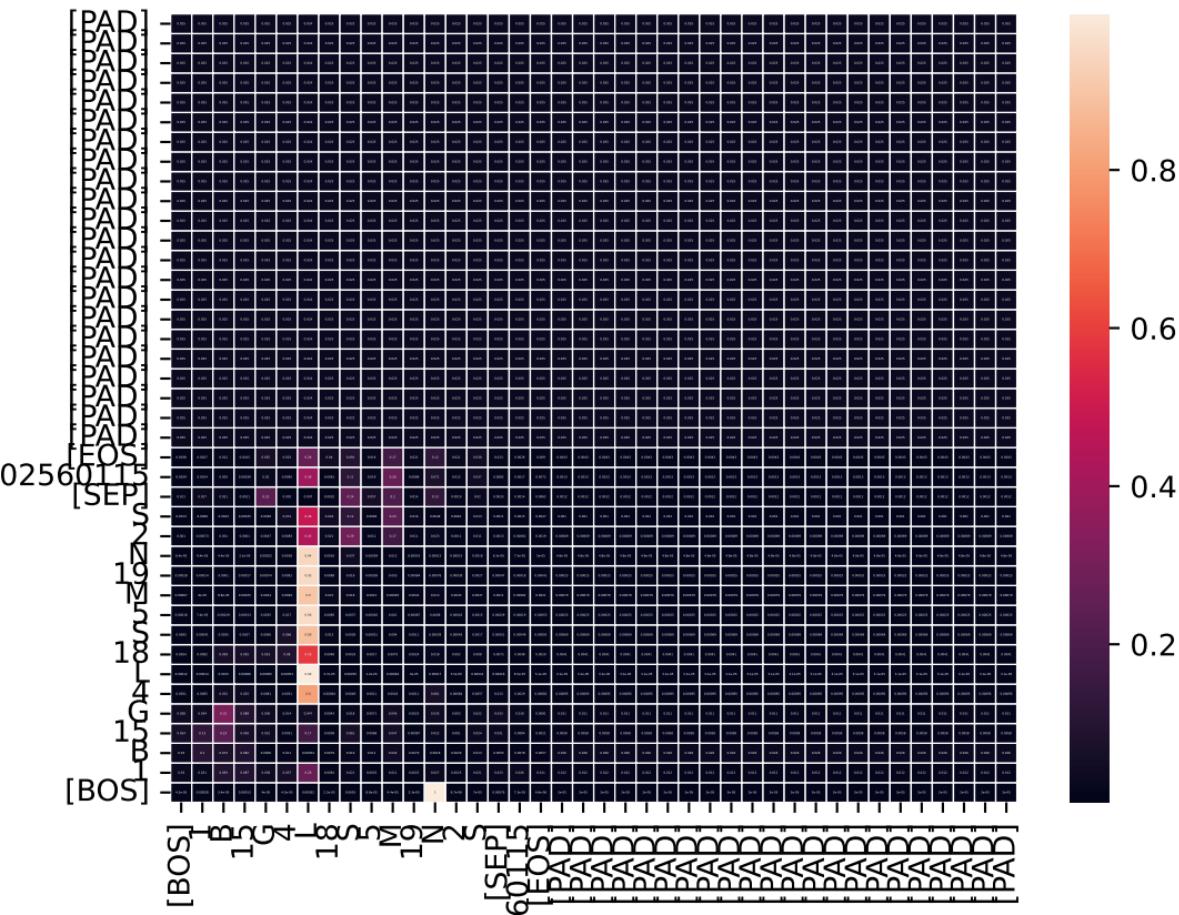
Step 10000 attention weights at layer 0 head 1, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



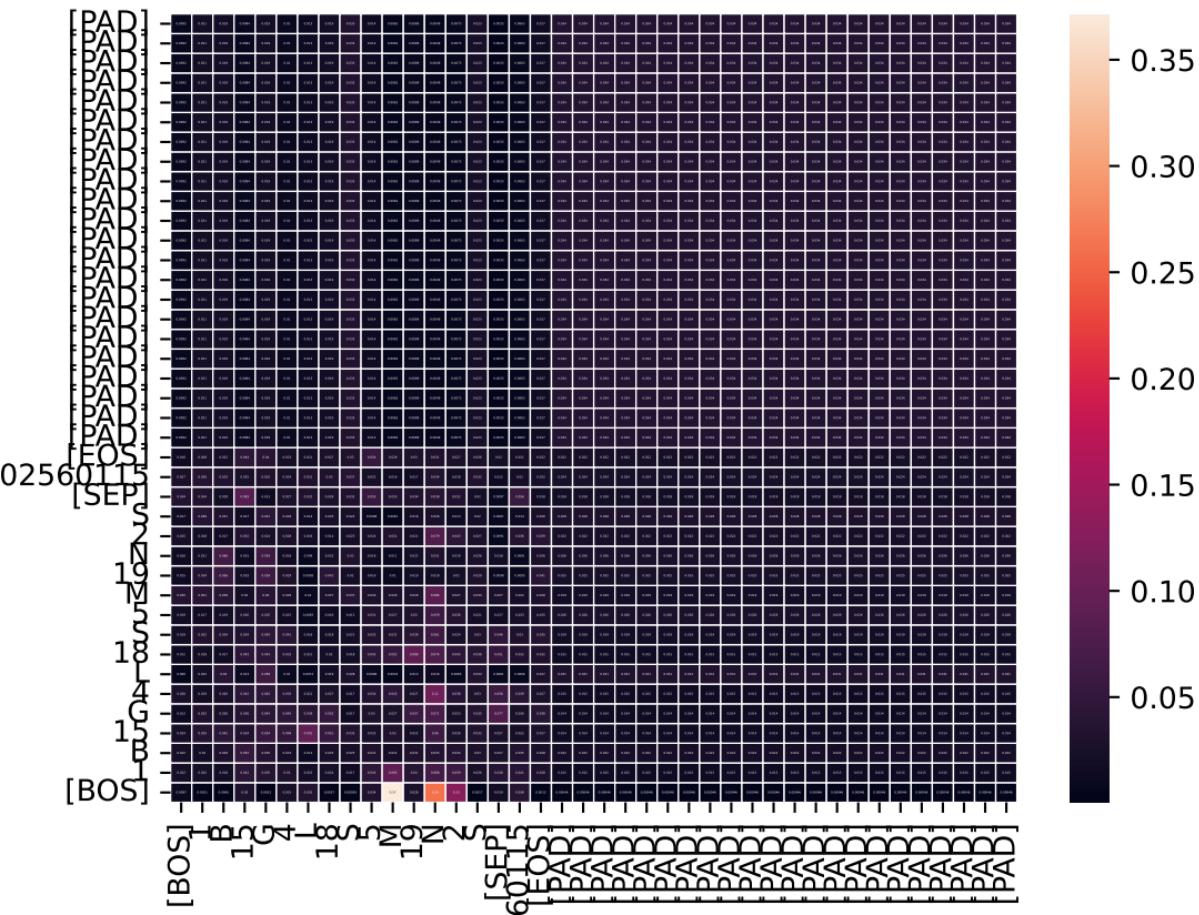
Step 10000 attention weights at layer 0 head 2, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



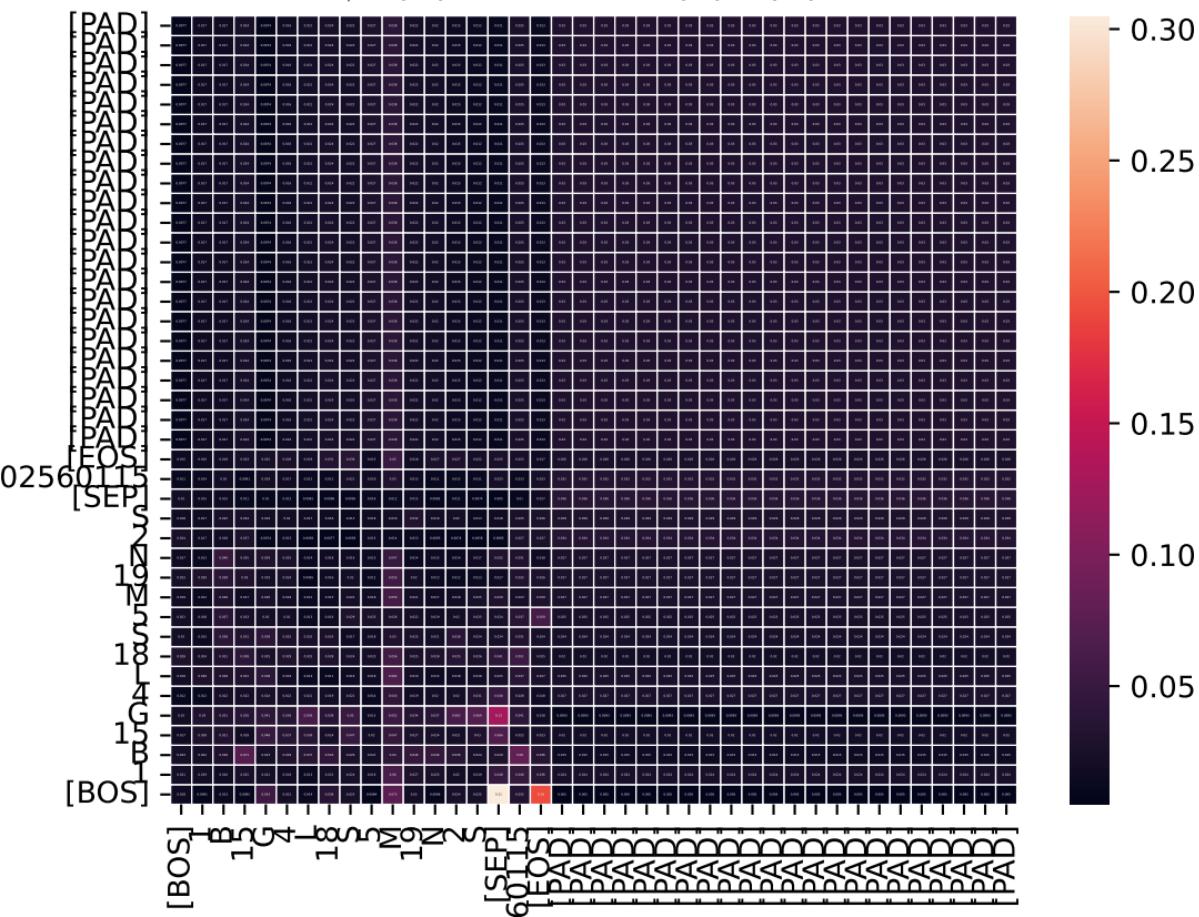
Step 10000 attention weights at layer 0 head 3, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



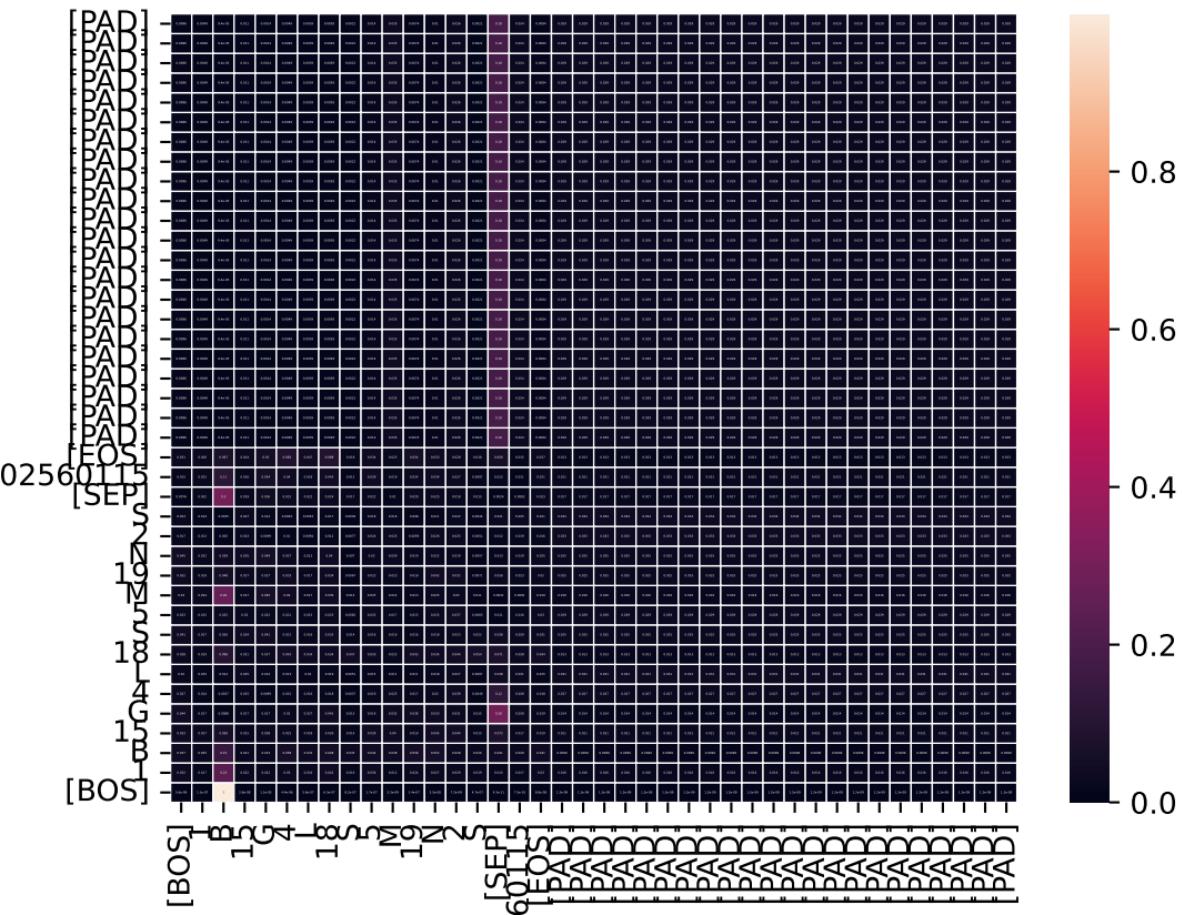
Step 10000 attention weights at layer 1 head 0, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



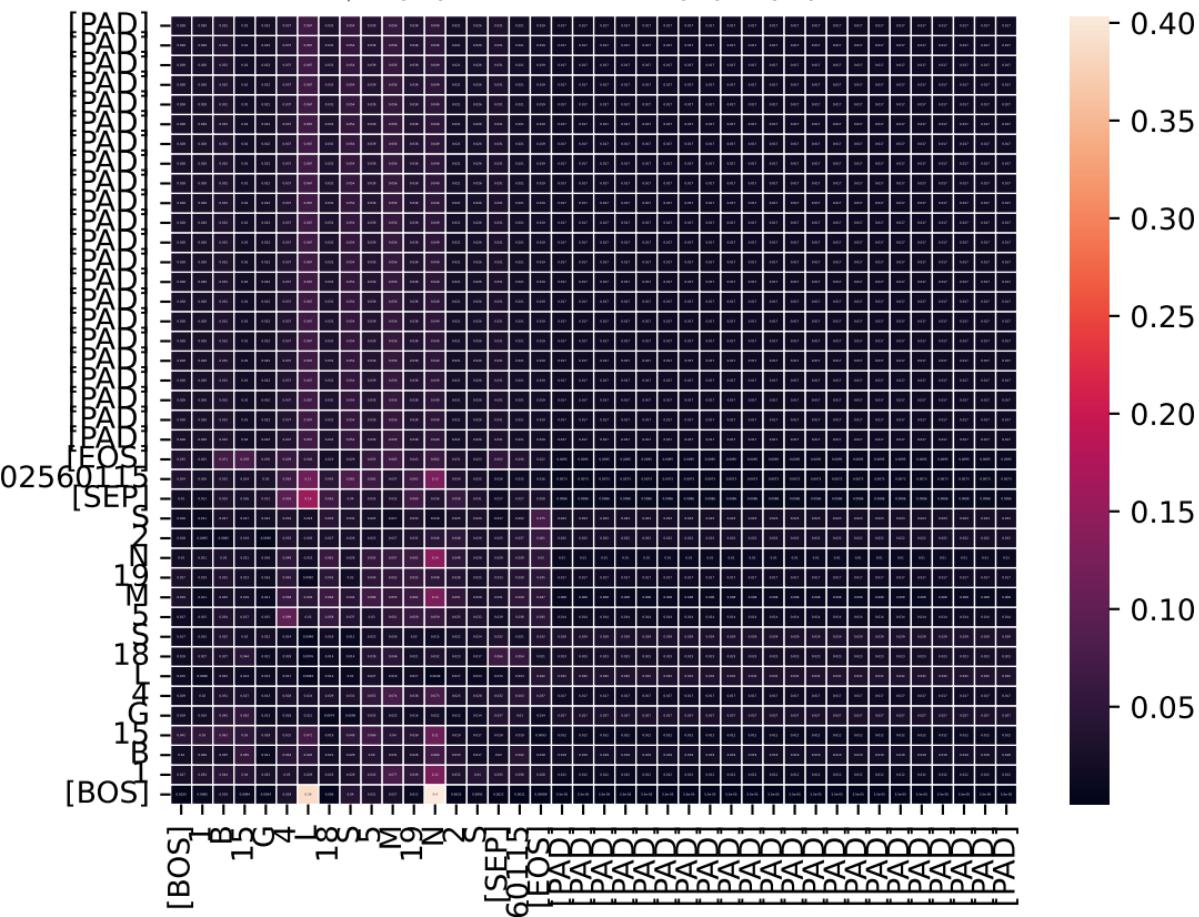
Step 10000 attention weights at layer 1 head 1, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



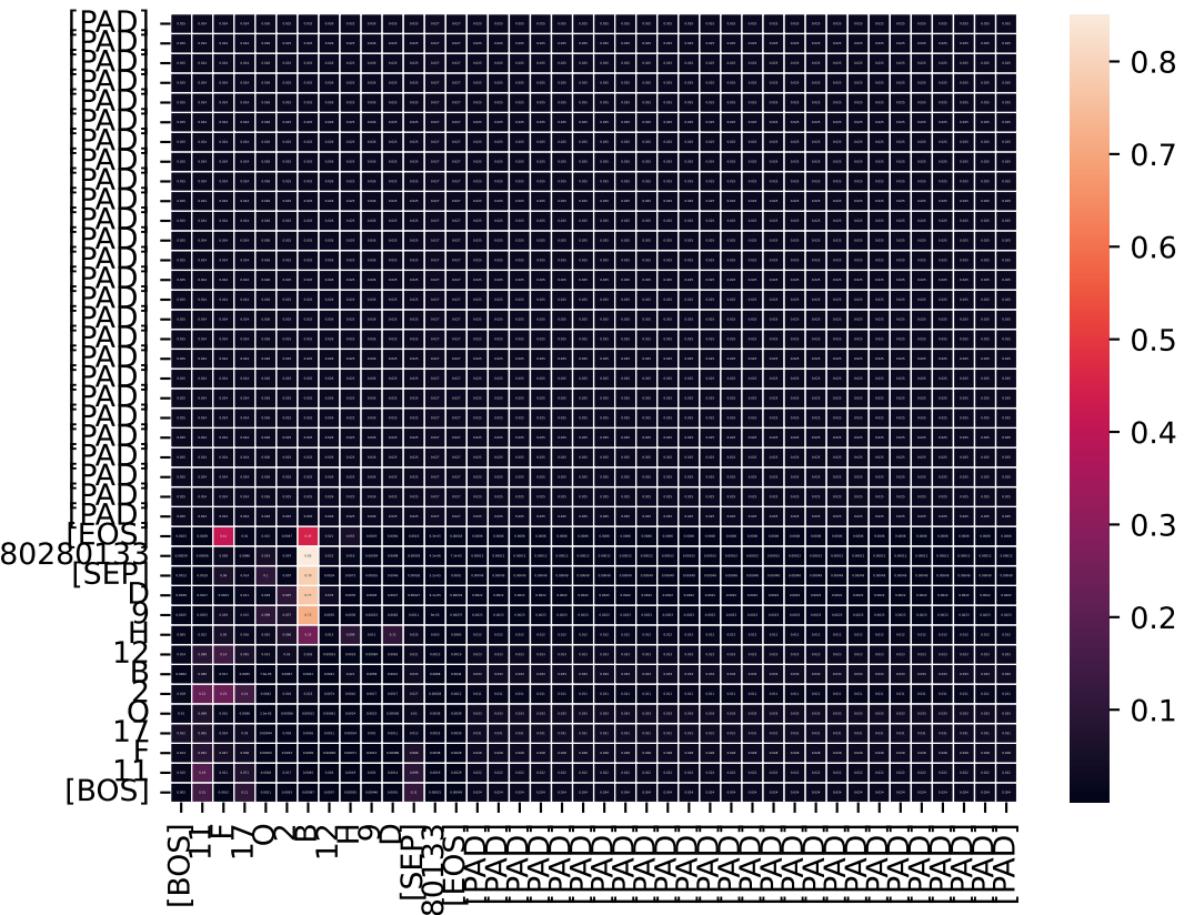
Step 10000 attention weights at layer 1 head 2, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



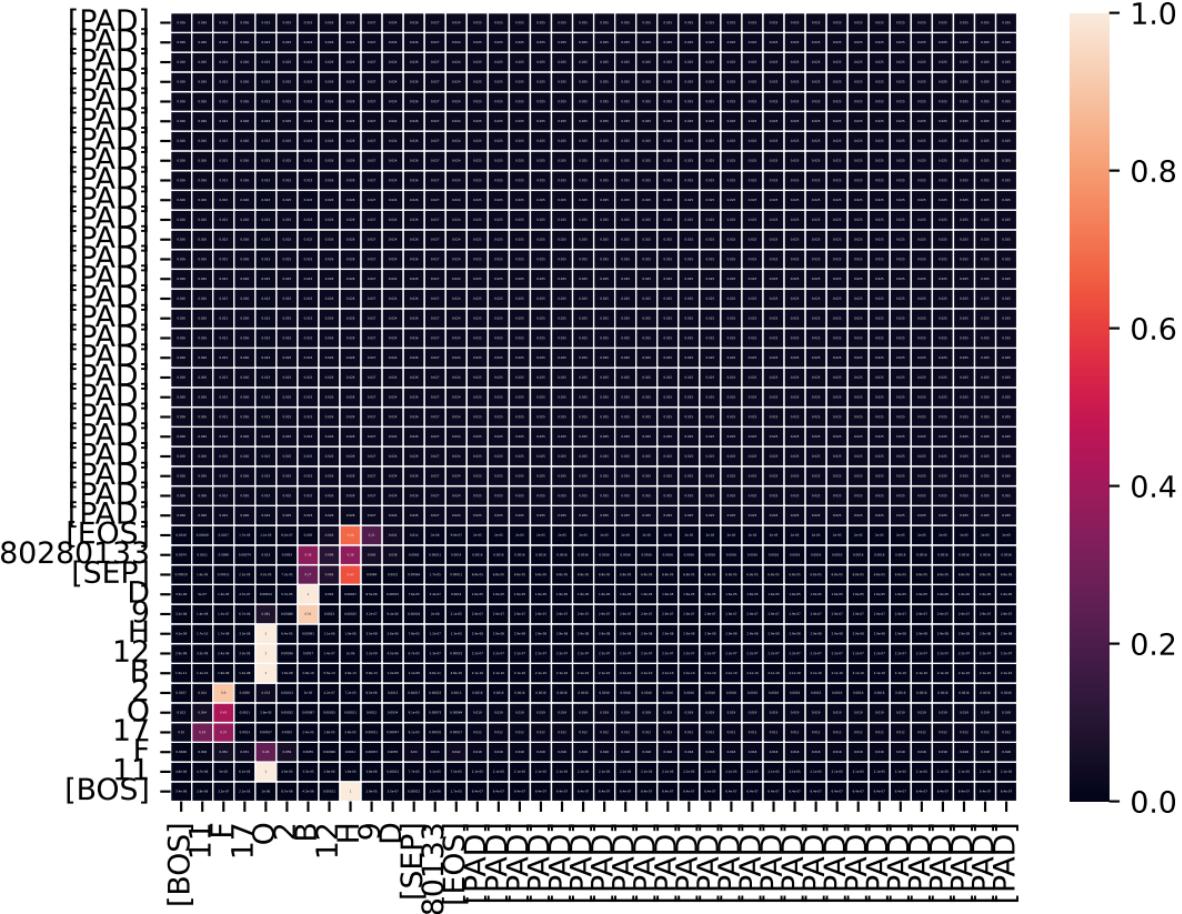
Step 10000 attention weights at layer 1 head 3, with accuracy 0.248
Sequence:[BOS] 1 B 15 G 4 L 18 S 5 M 19 N 2 S [SEP] 14.92 [EOS] SOL:G



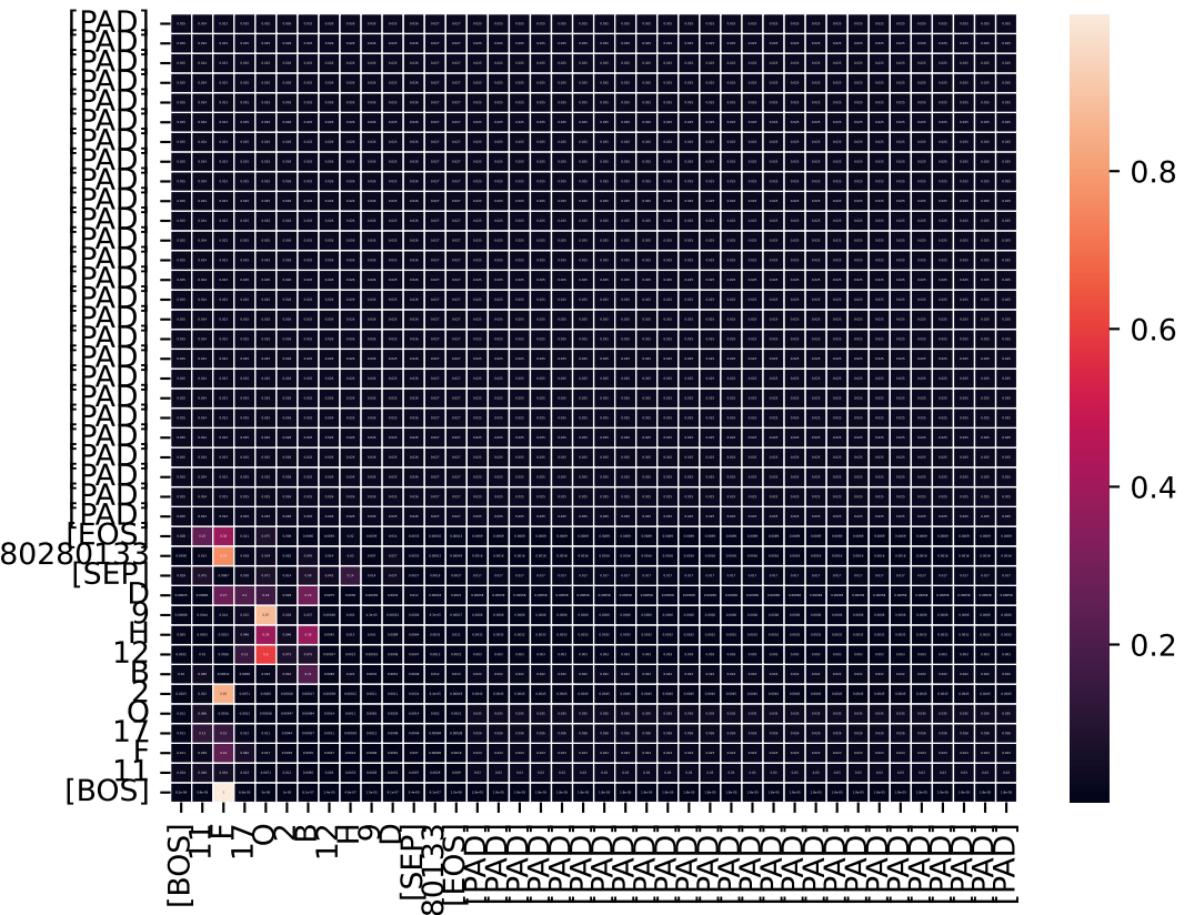
Step 19500 attention weights at layer 0 head 0, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



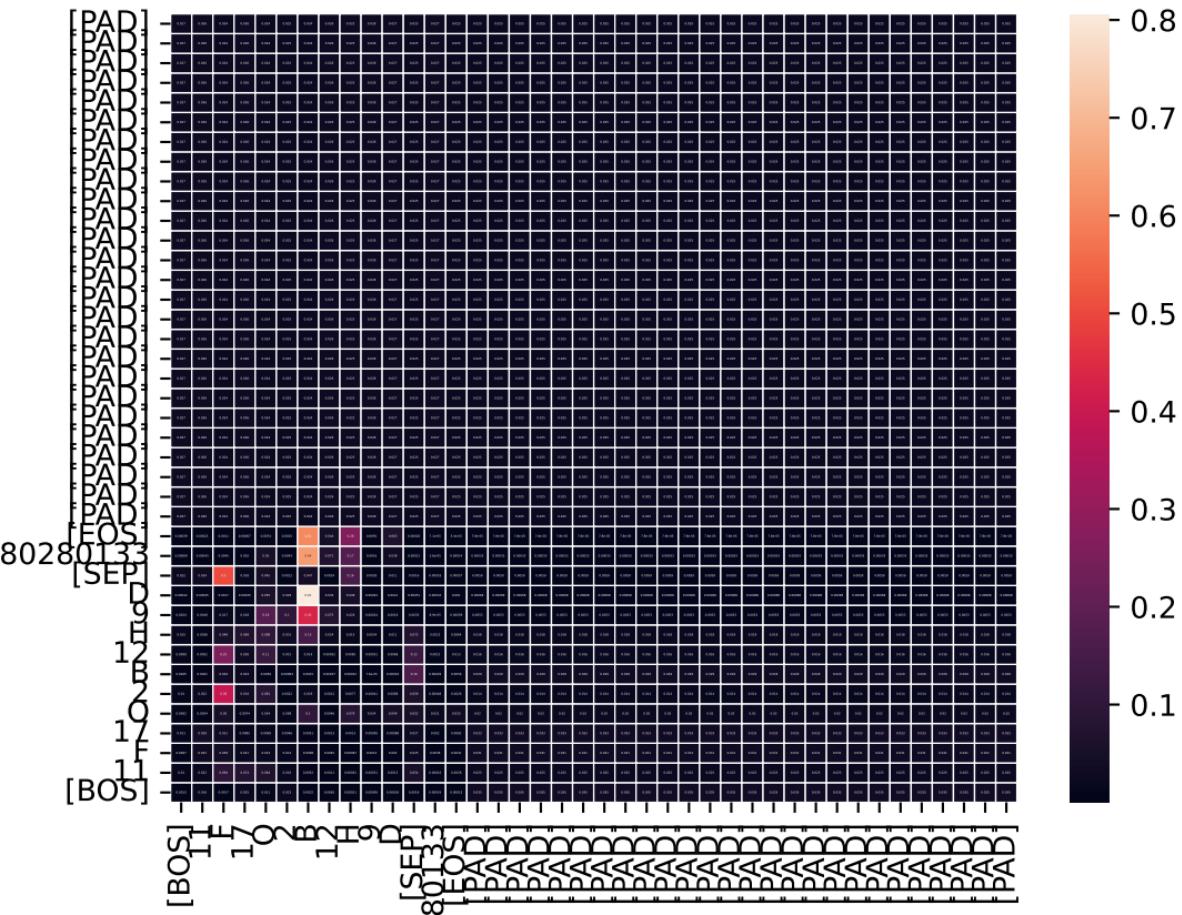
Step 19500 attention weights at layer 0 head 1, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



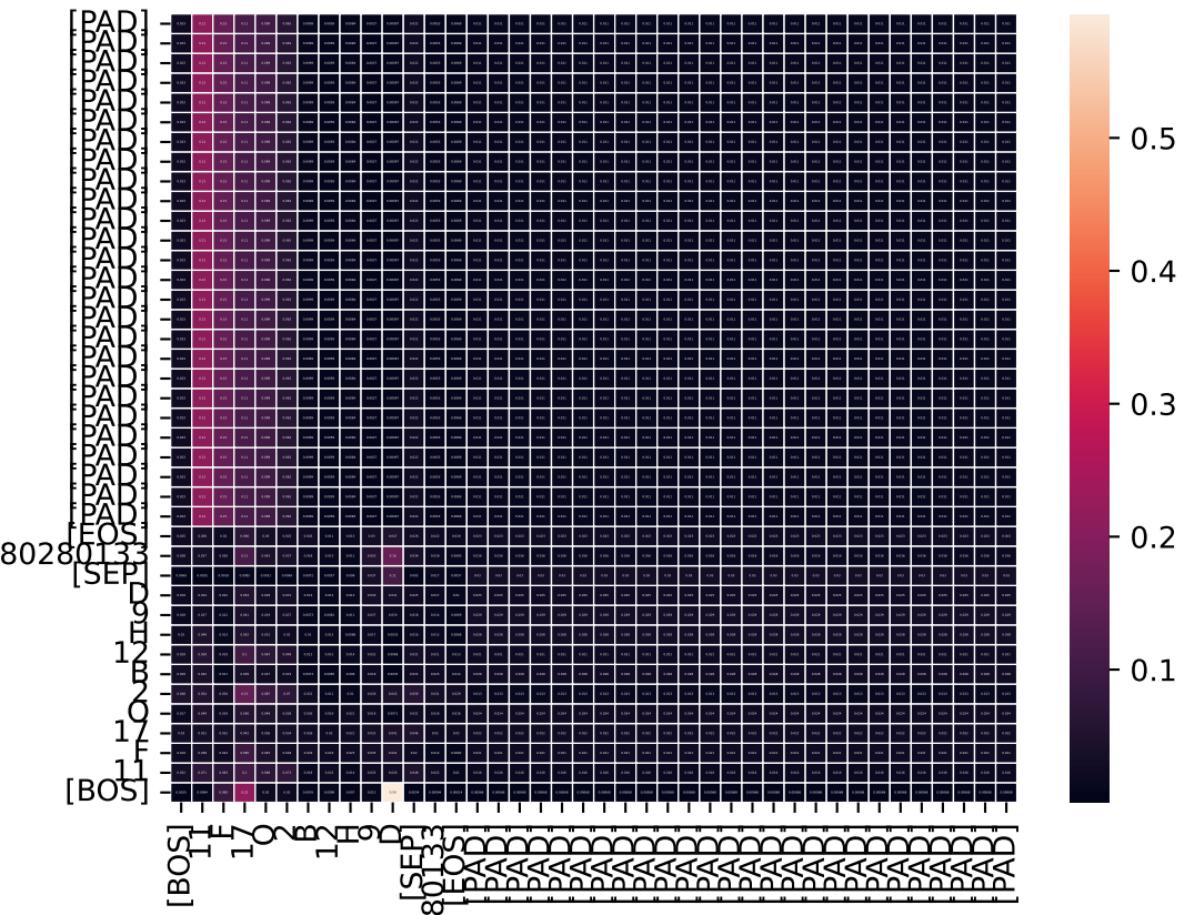
Step 19500 attention weights at layer 0 head 2, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



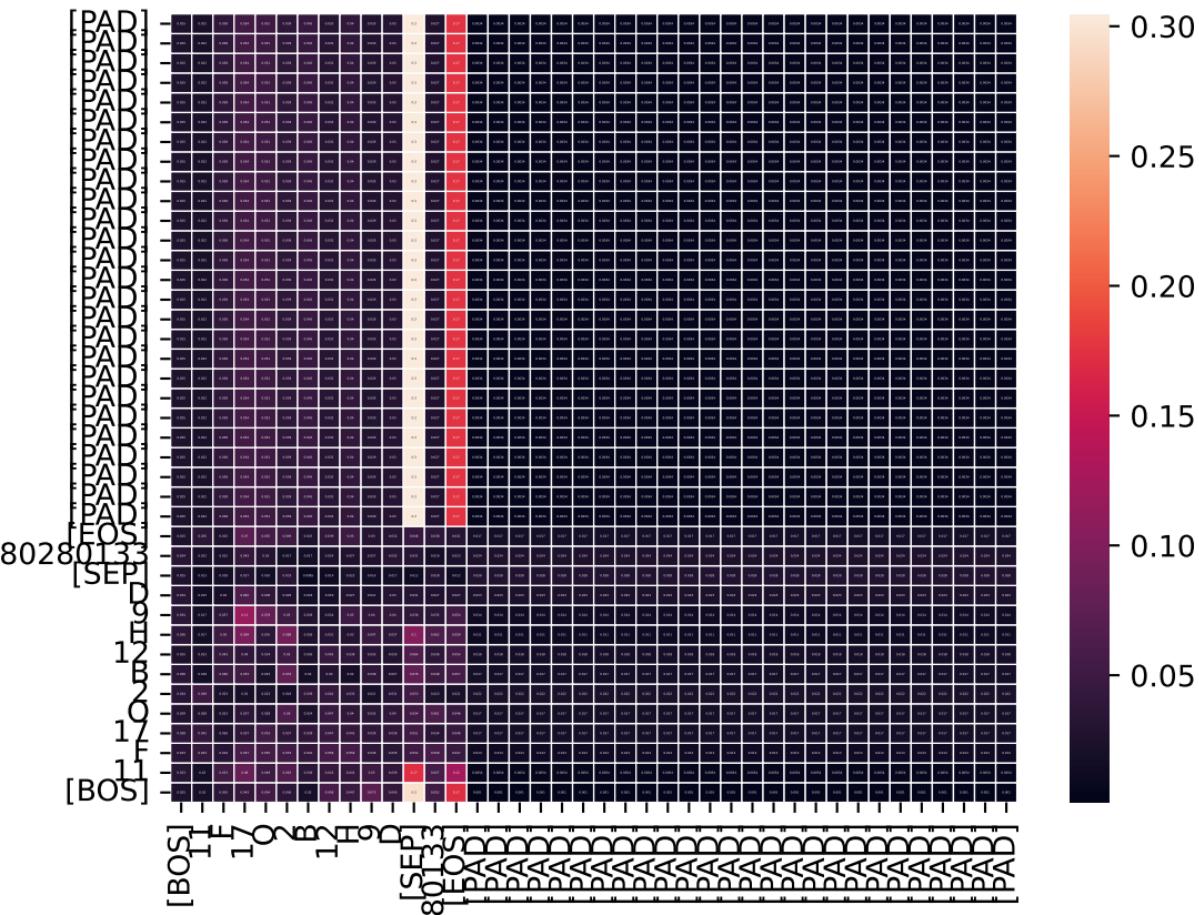
Step 19500 attention weights at layer 0 head 3, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



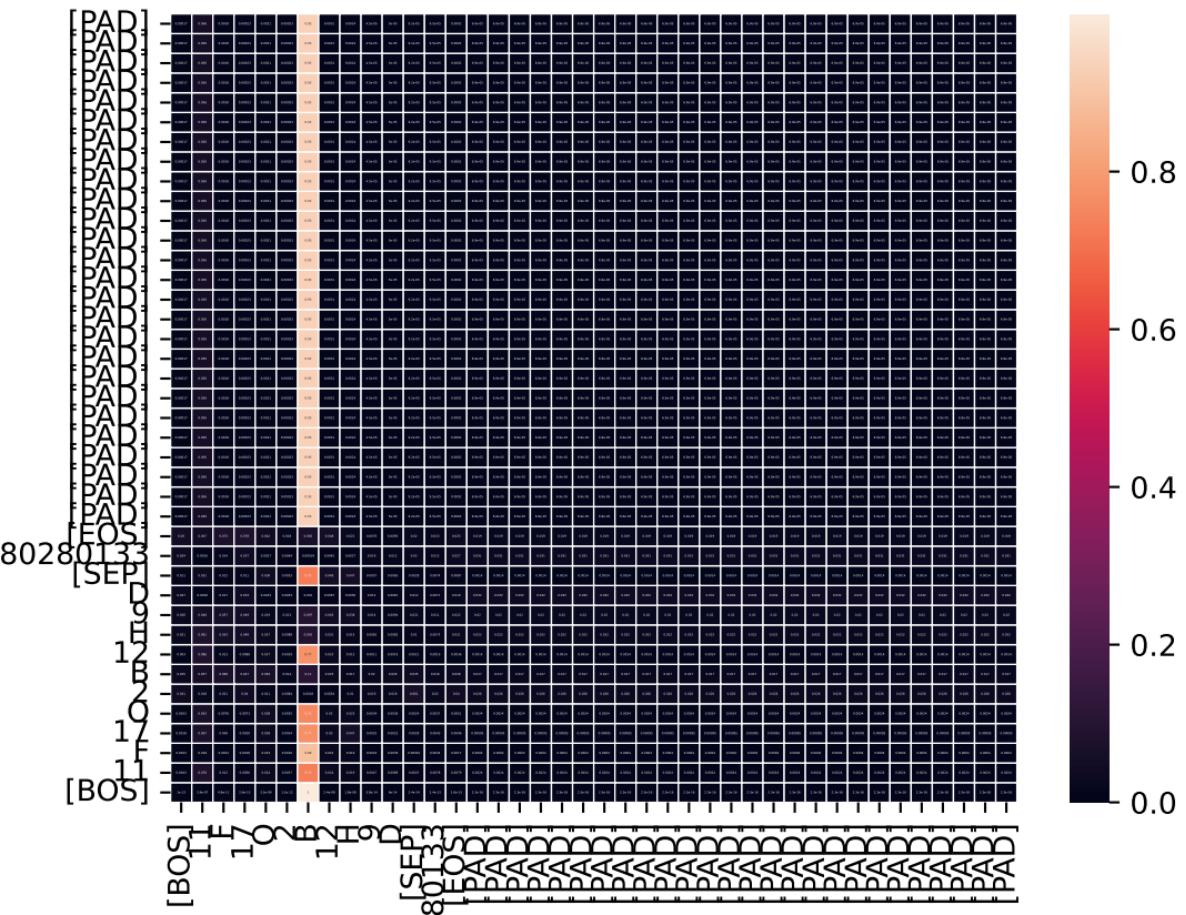
Step 19500 attention weights at layer 1 head 0, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



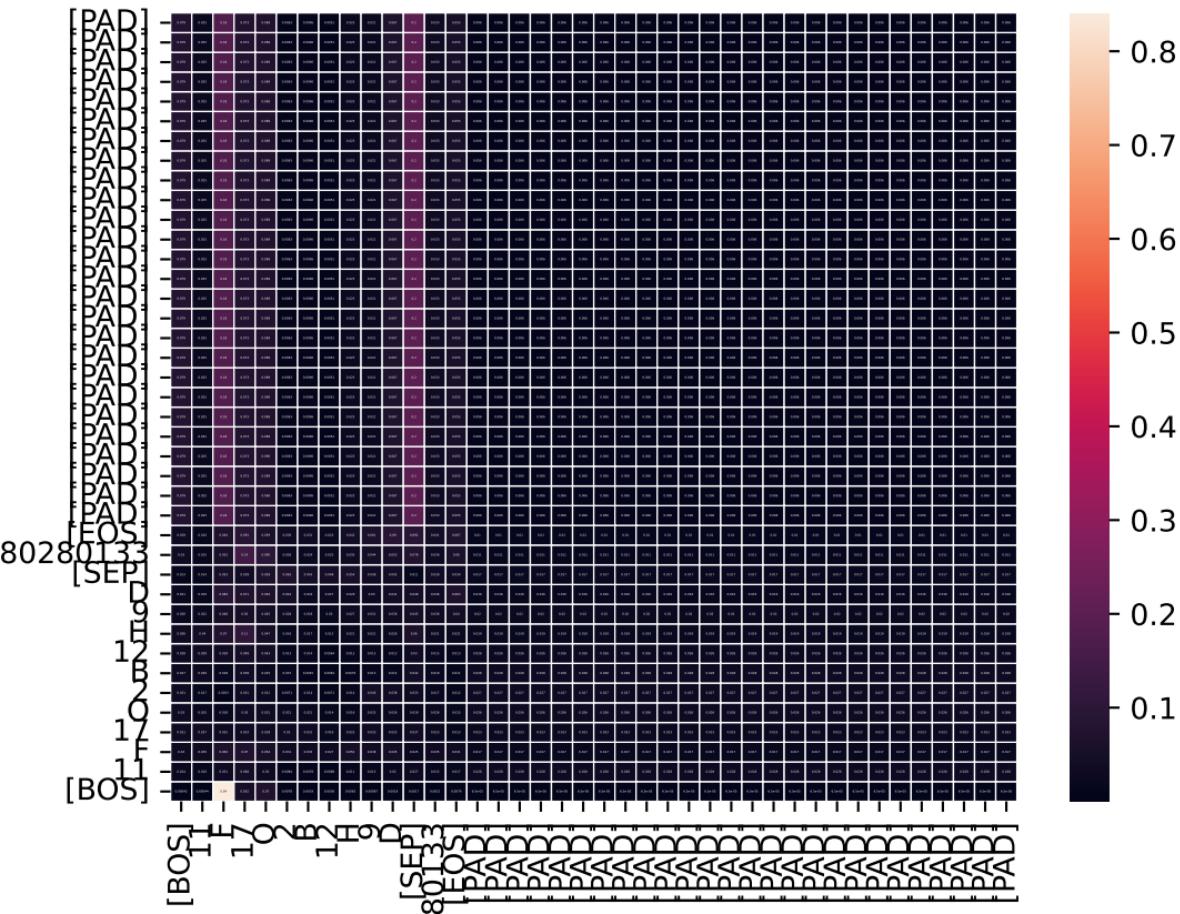
Step 19500 attention weights at layer 1 head 1, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



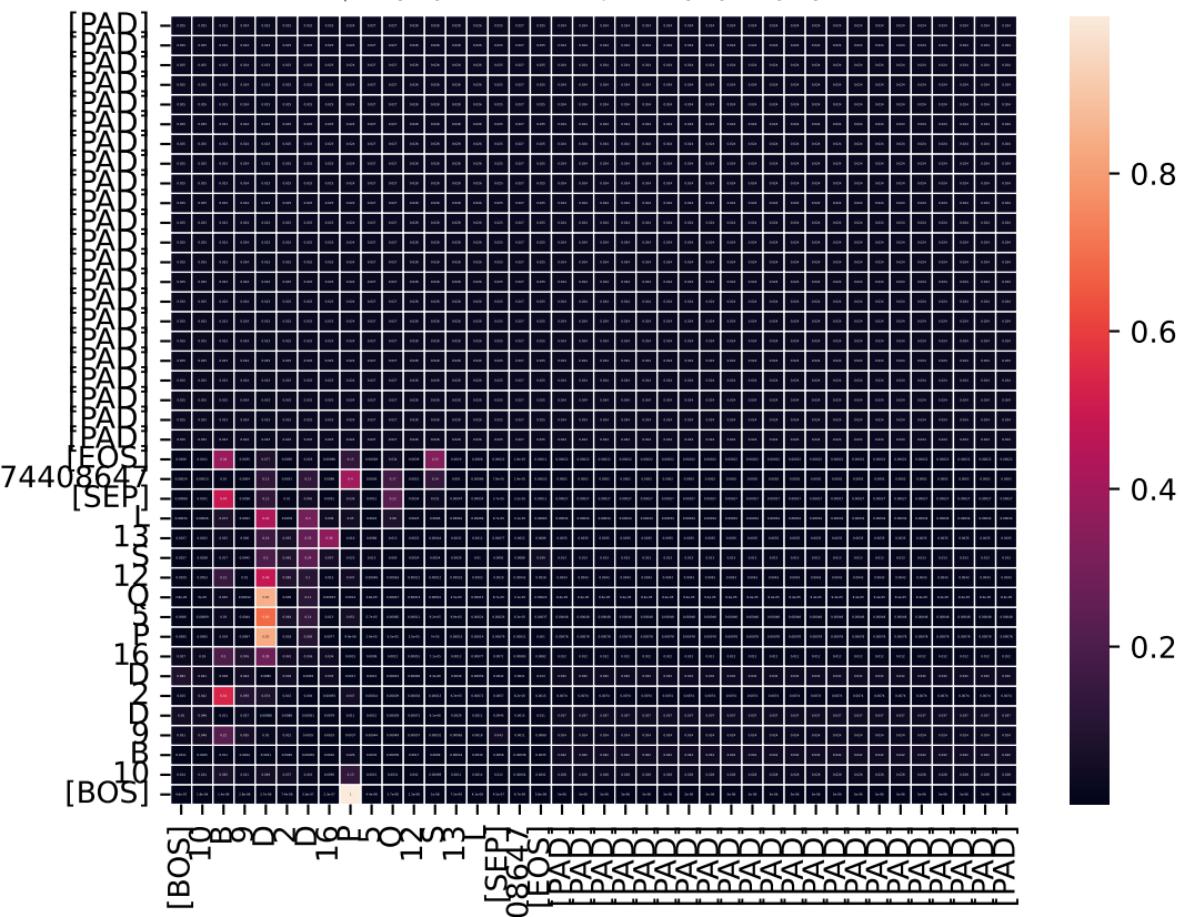
Step 19500 attention weights at layer 1 head 2, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



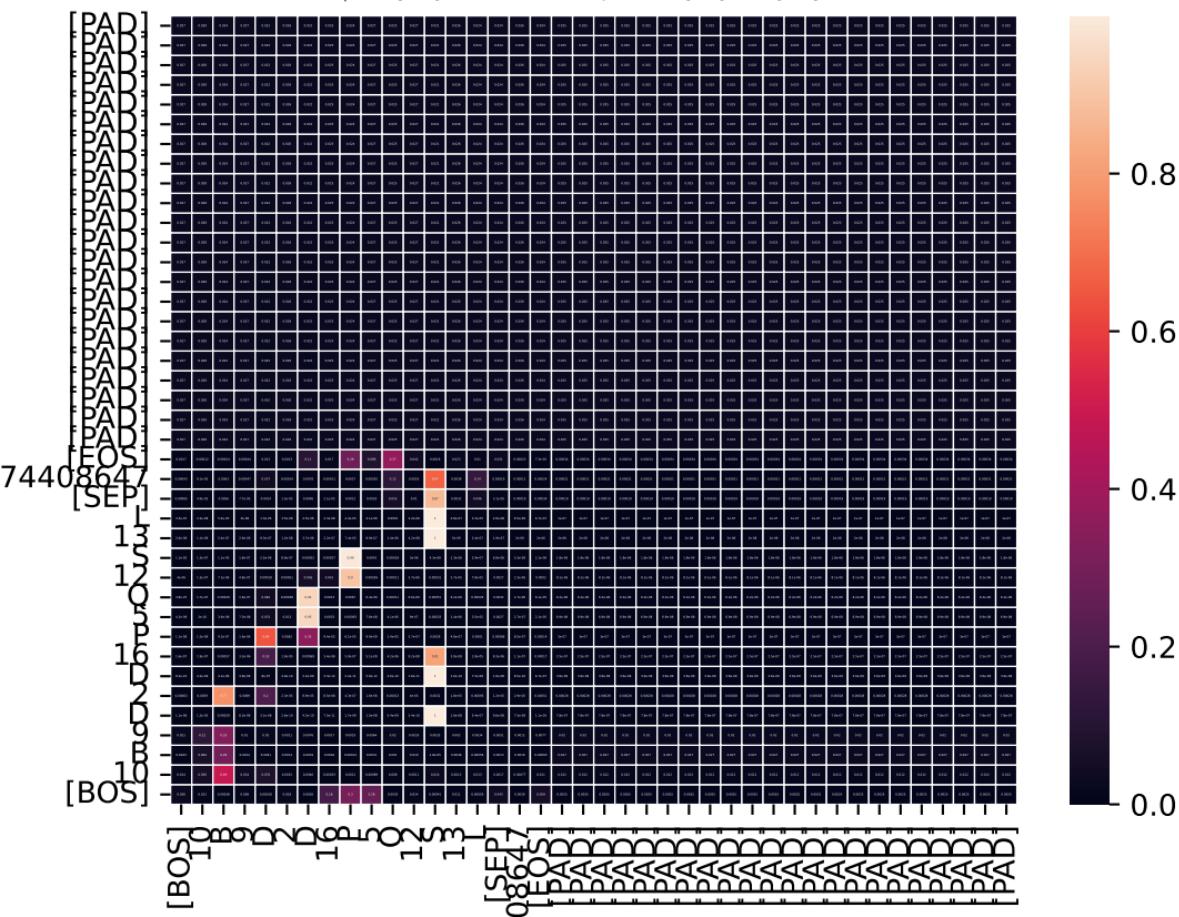
Step 19500 attention weights at layer 1 head 3, with accuracy 0.252
Sequence:[BOS] 11 F 17 O 2 B 12 H 9 D [SEP] 2.16 [EOS] SOL:B



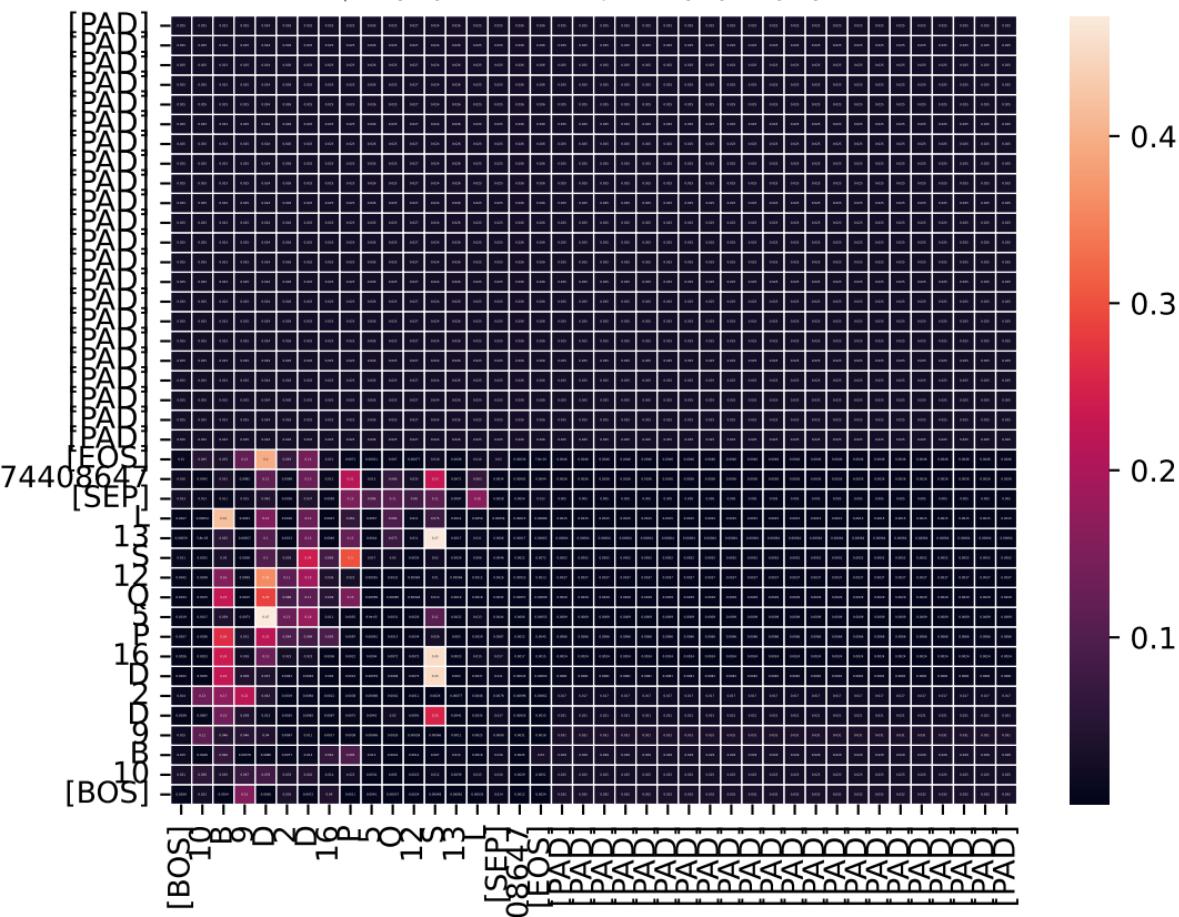
Step 20000 attention weights at layer 0 head 0, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



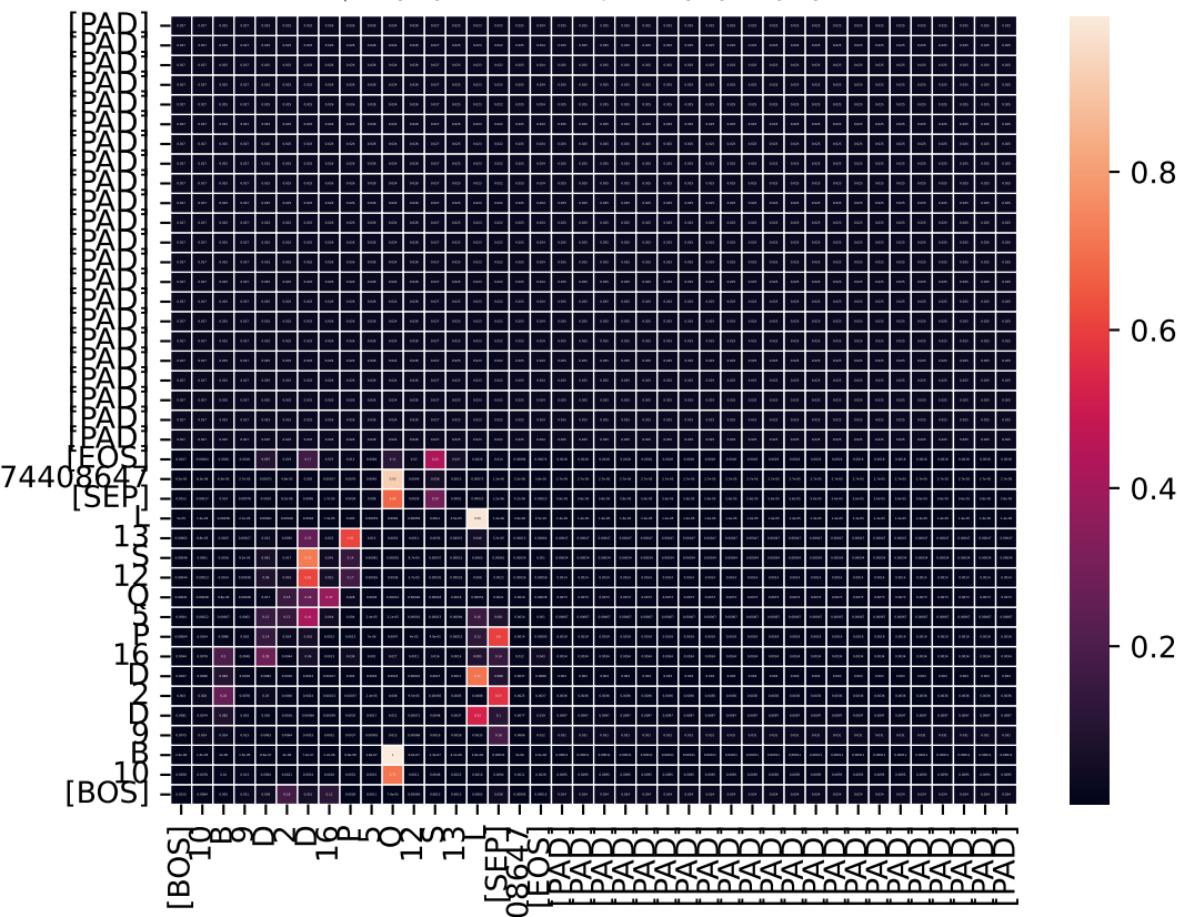
Step 20000 attention weights at layer 0 head 1, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



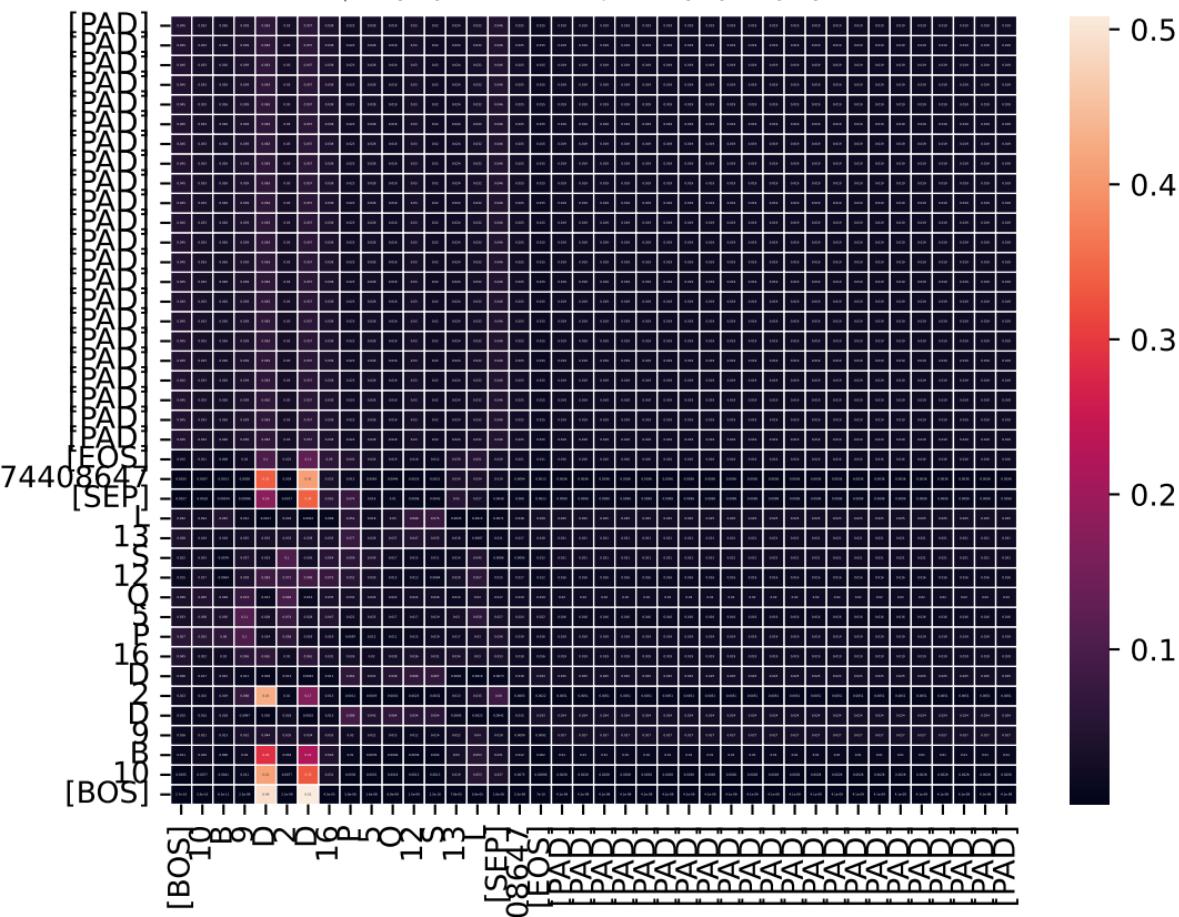
Step 20000 attention weights at layer 0 head 2, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



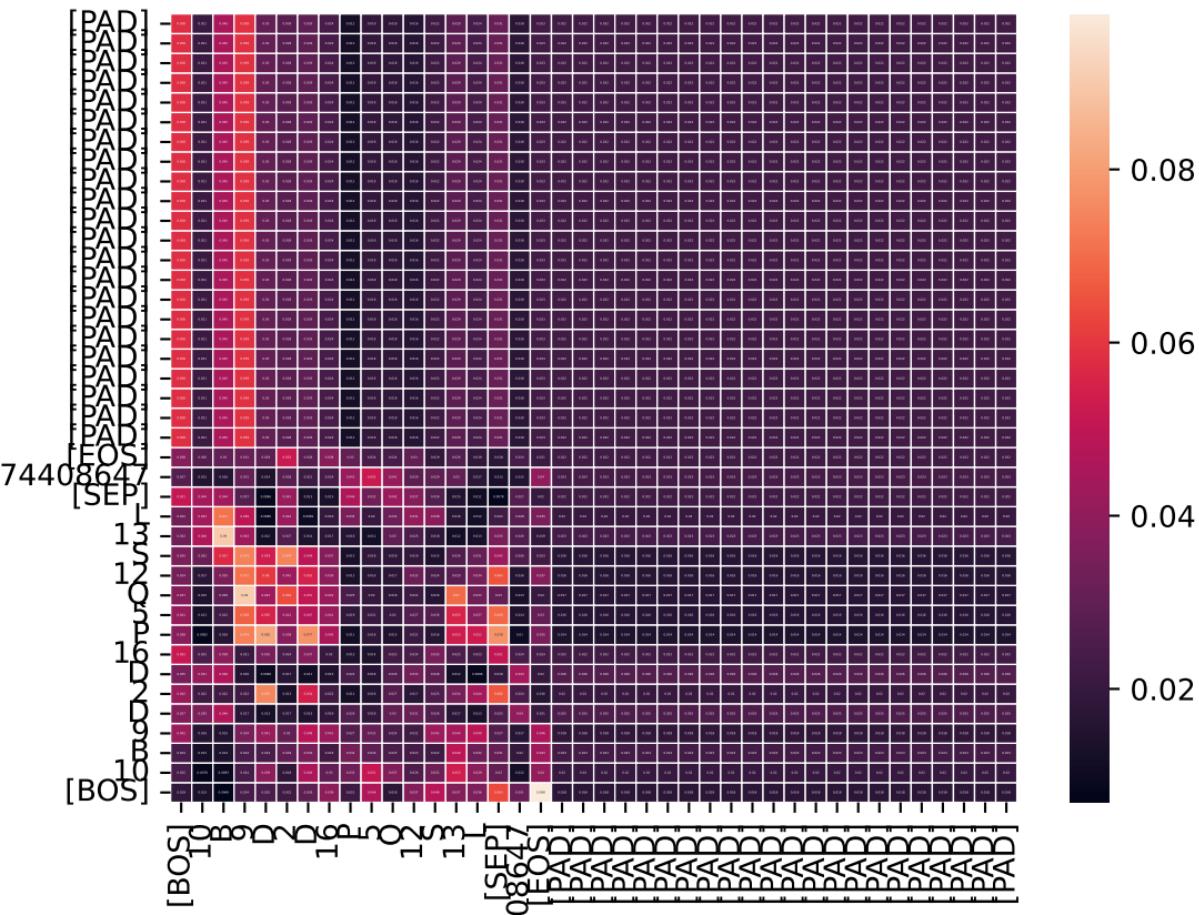
Step 20000 attention weights at layer 0 head 3, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



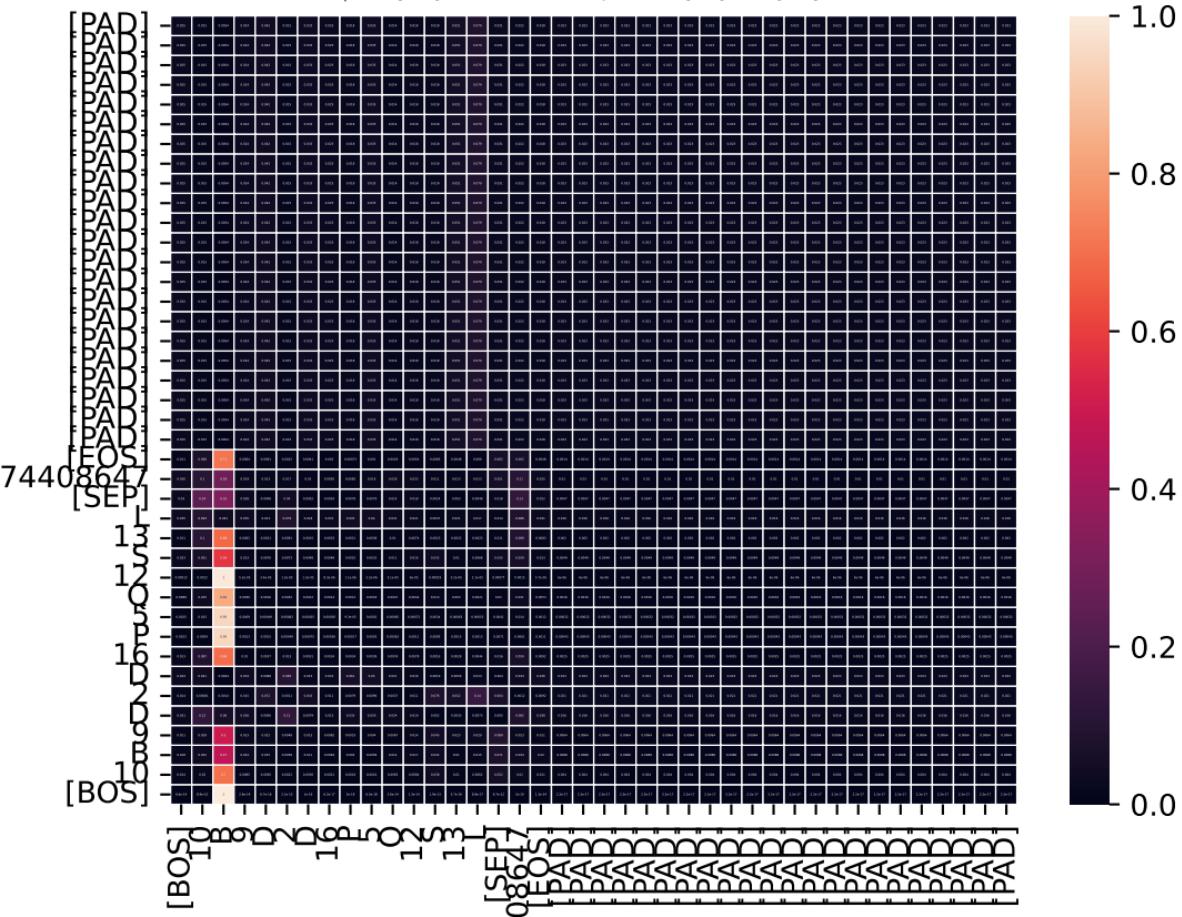
Step 20000 attention weights at layer 1 head 0, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



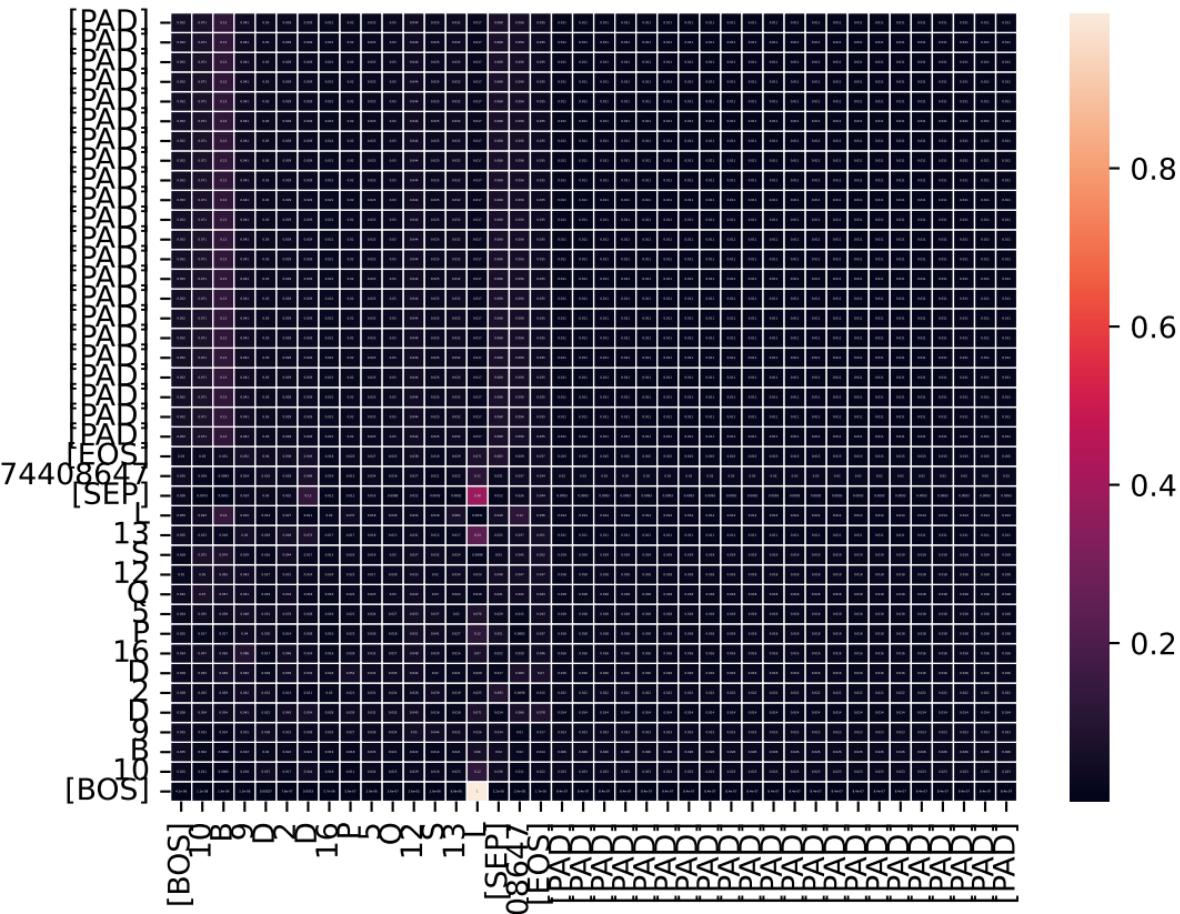
Step 20000 attention weights at layer 1 head 1, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



Step 20000 attention weights at layer 1 head 2, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B

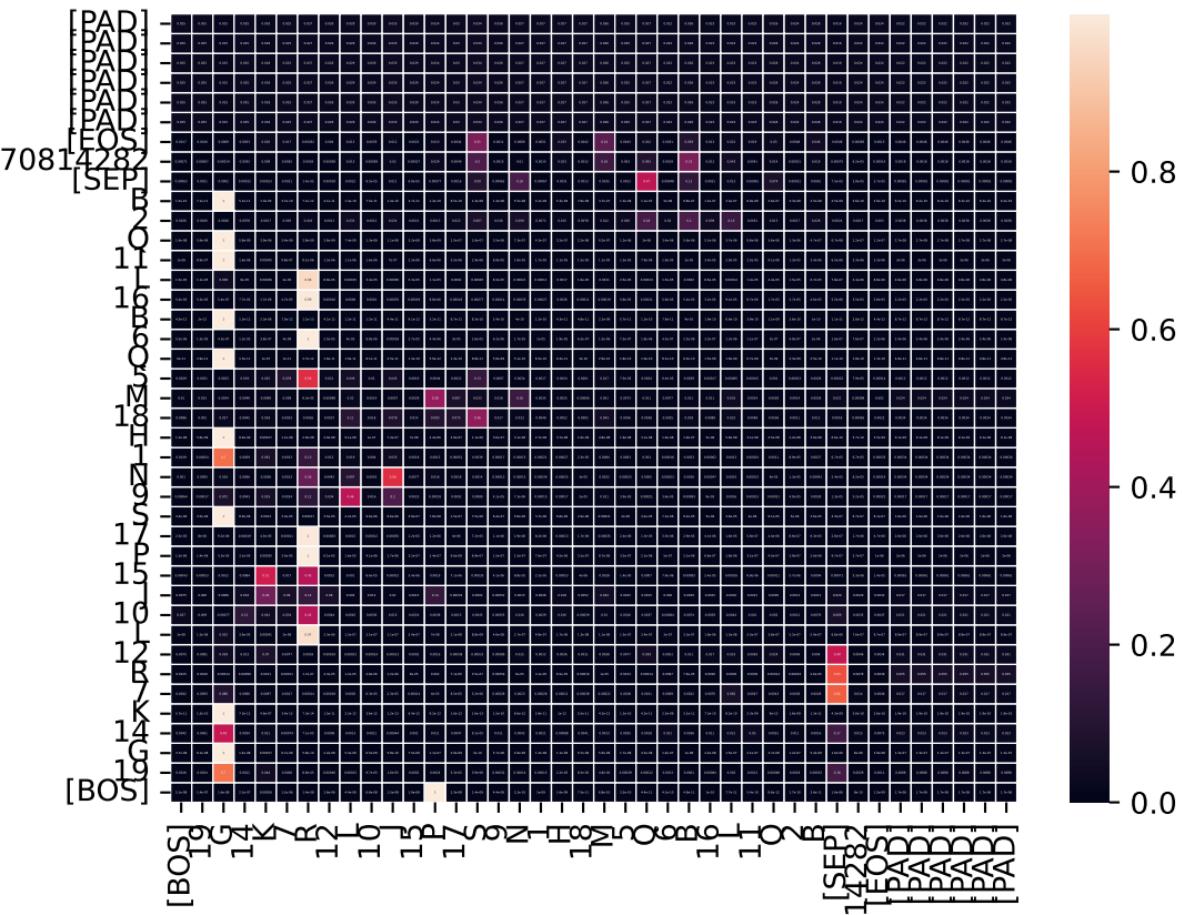


Step 20000 attention weights at layer 1 head 3, with accuracy 0.260
Sequence:[BOS] 10 B 9 D 2 D 16 P 5 Q 12 S 13 L [SEP] 10.05 [EOS] SOL:B



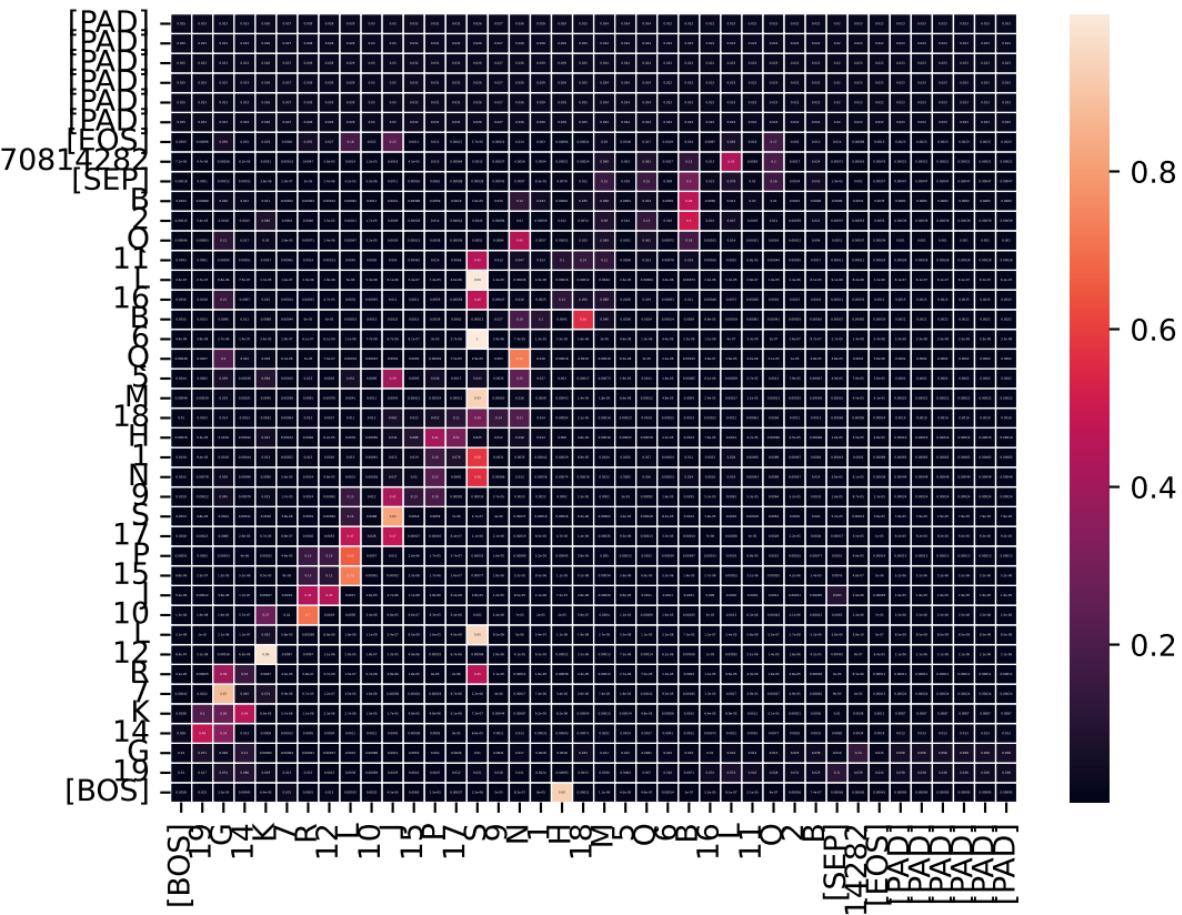
Step 29500 attention weights at layer 0 head 0, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



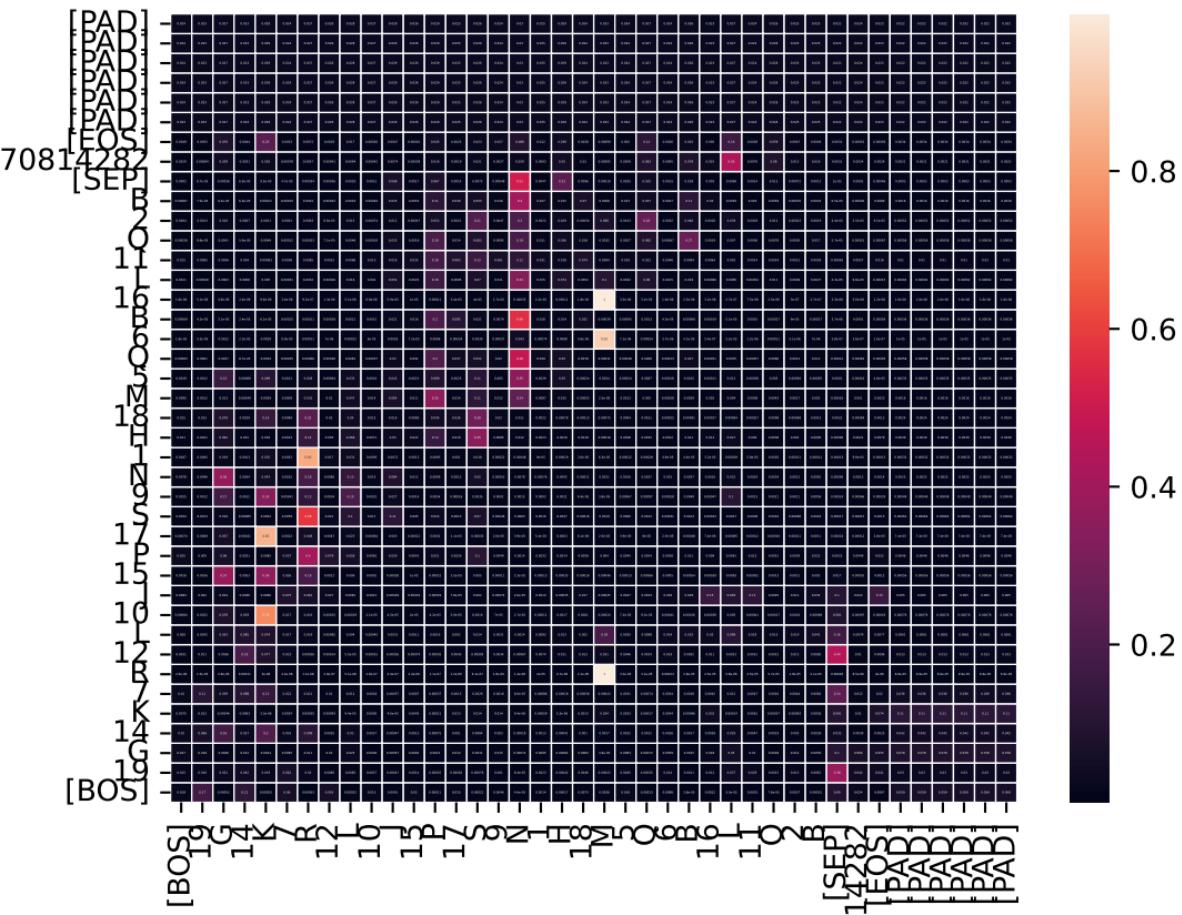
Step 29500 attention weights at layer 0 head 1, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



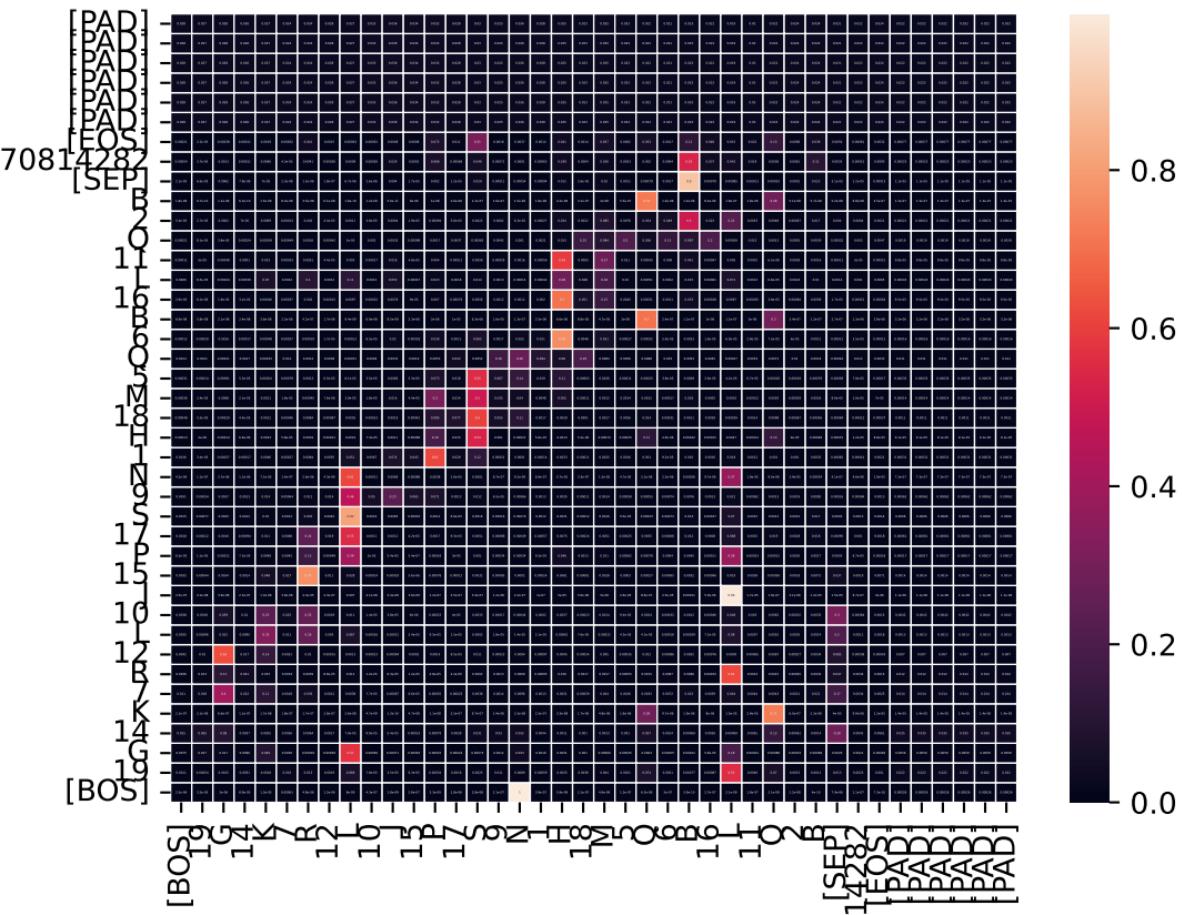
Step 29500 attention weights at layer 0 head 2, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



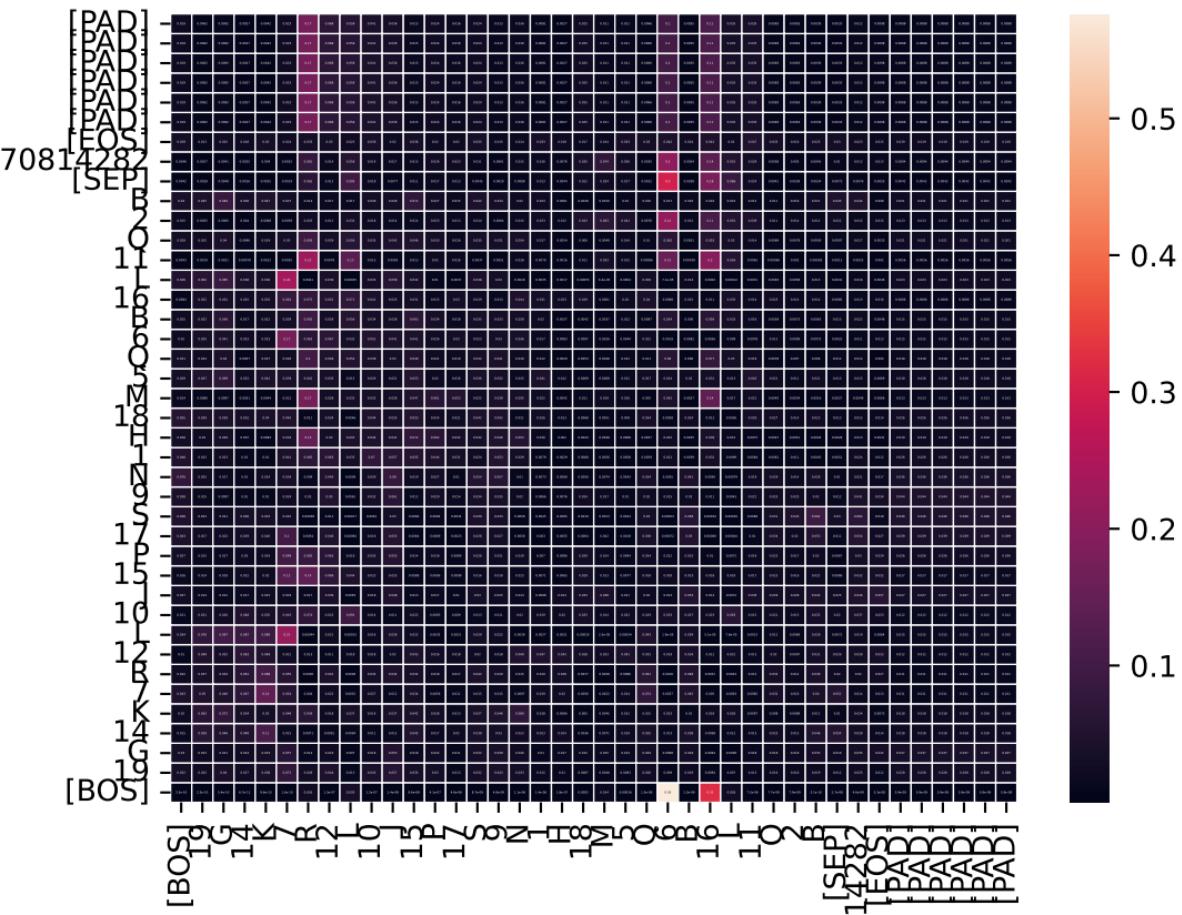
Step 29500 attention weights at layer 0 head 3, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



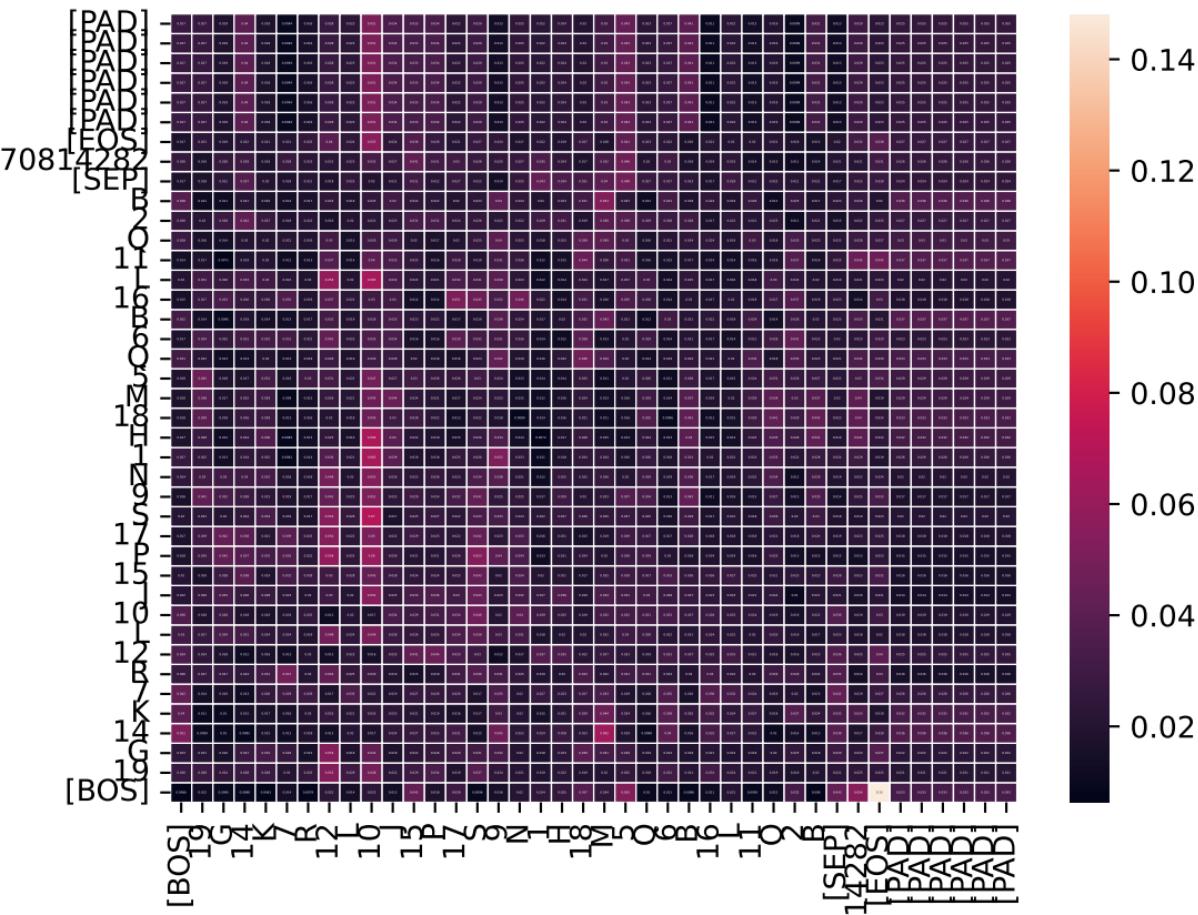
Step 29500 attention weights at layer 1 head 0, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



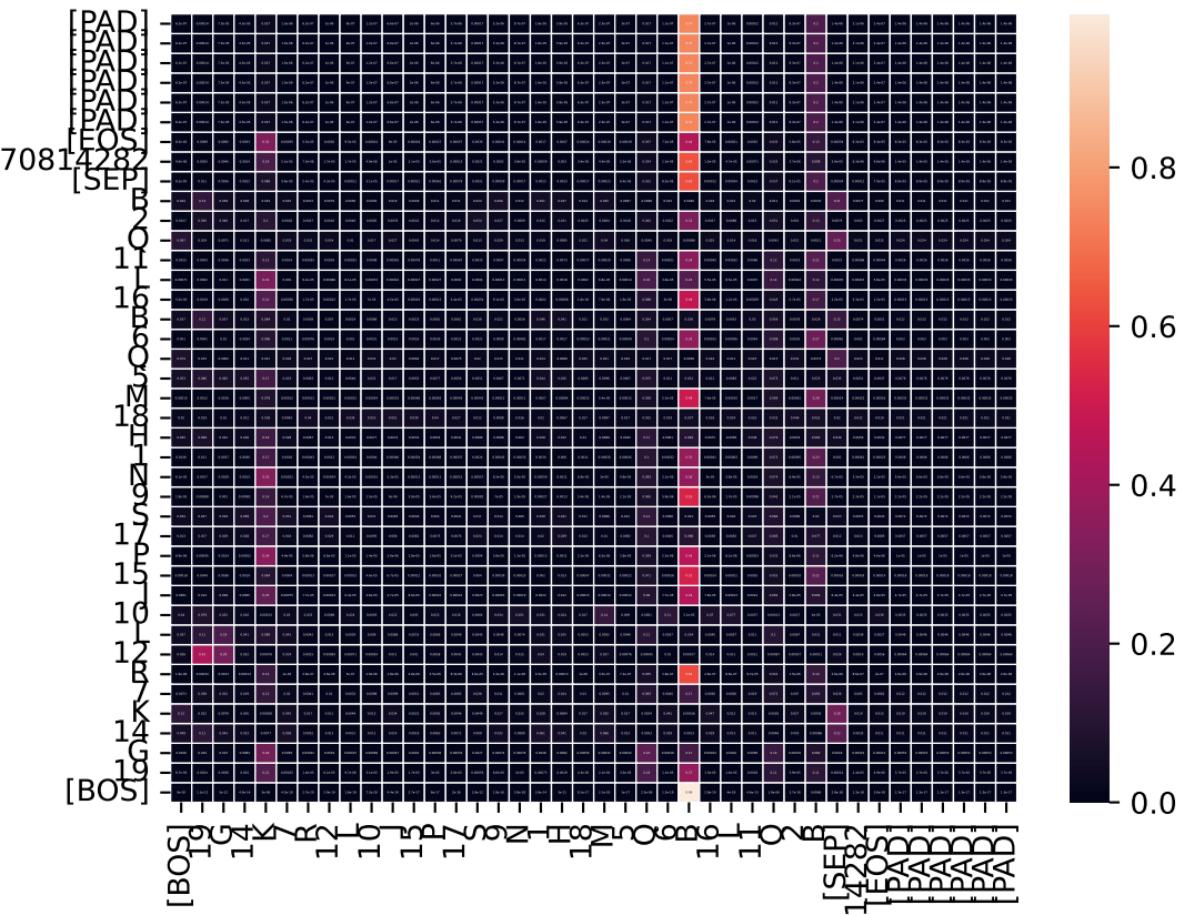
Step 29500 attention weights at layer 1 head 1, with accuracy 0.282

Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



Step 29500 attention weights at layer 1 head 2, with accuracy 0.282

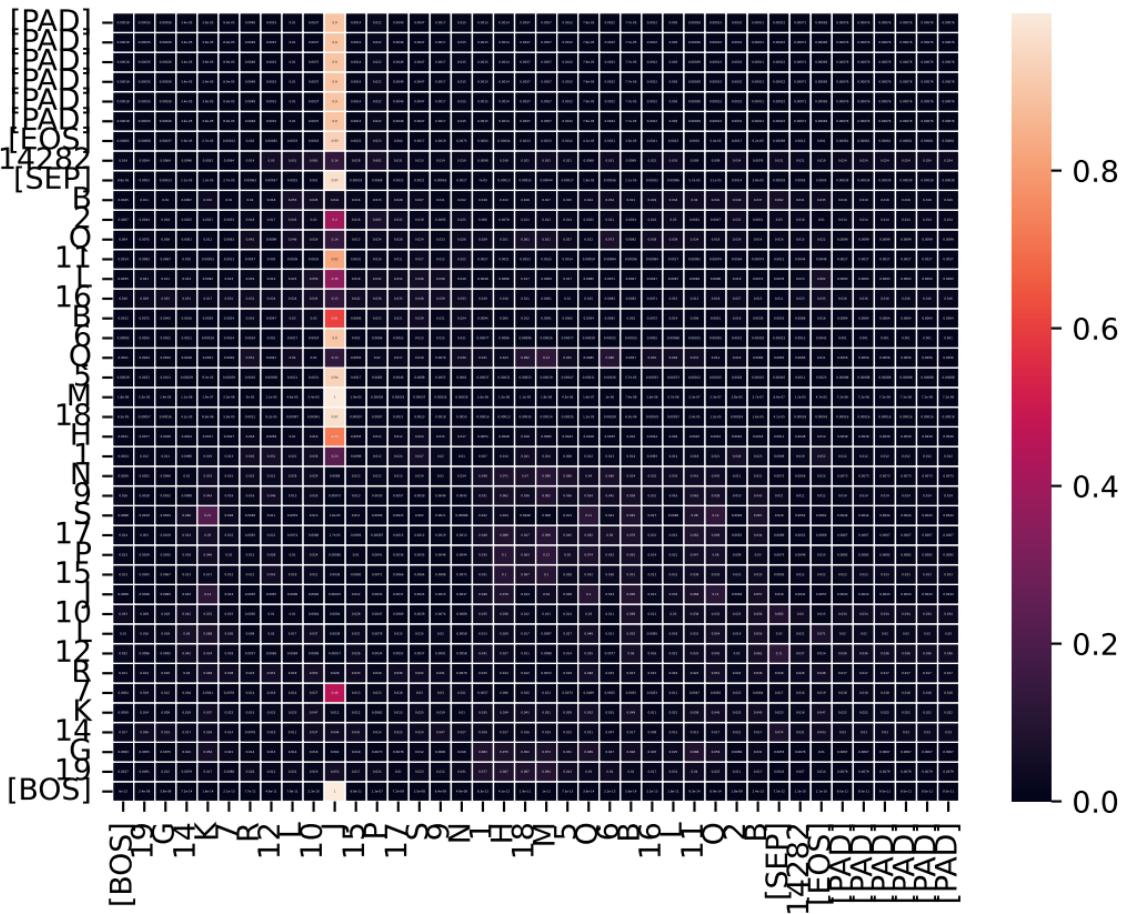
Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 5 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P



Step 29500 attention weights at layer 1 head 3, with accuracy 0.282

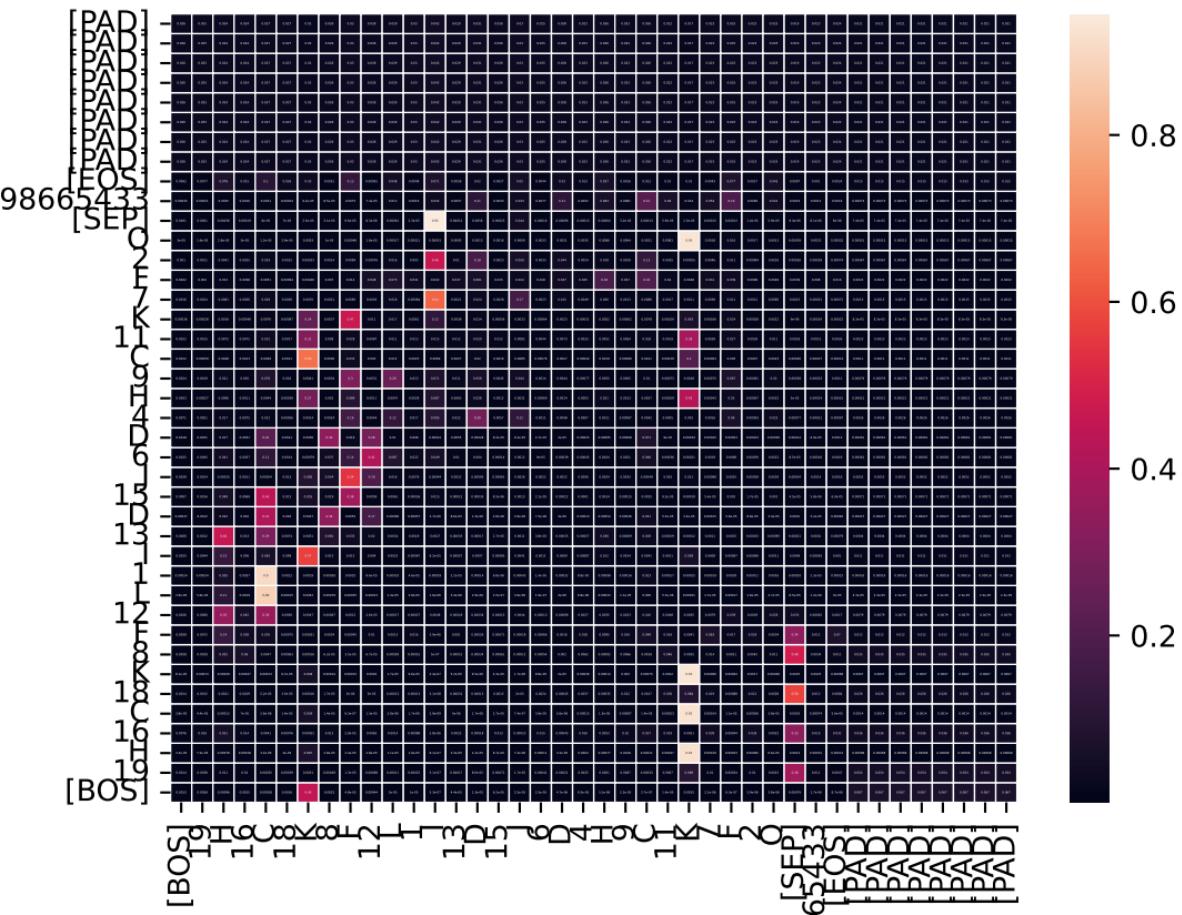
Sequence:[BOS] 19 G 14 K 7 R 12 L 10 J 15 P 17 S 9 N 1 H 18 M 5 Q 6 B 16 L 11 Q 2 B [SEP] 14.87 [EOS] SOL:P

70814282



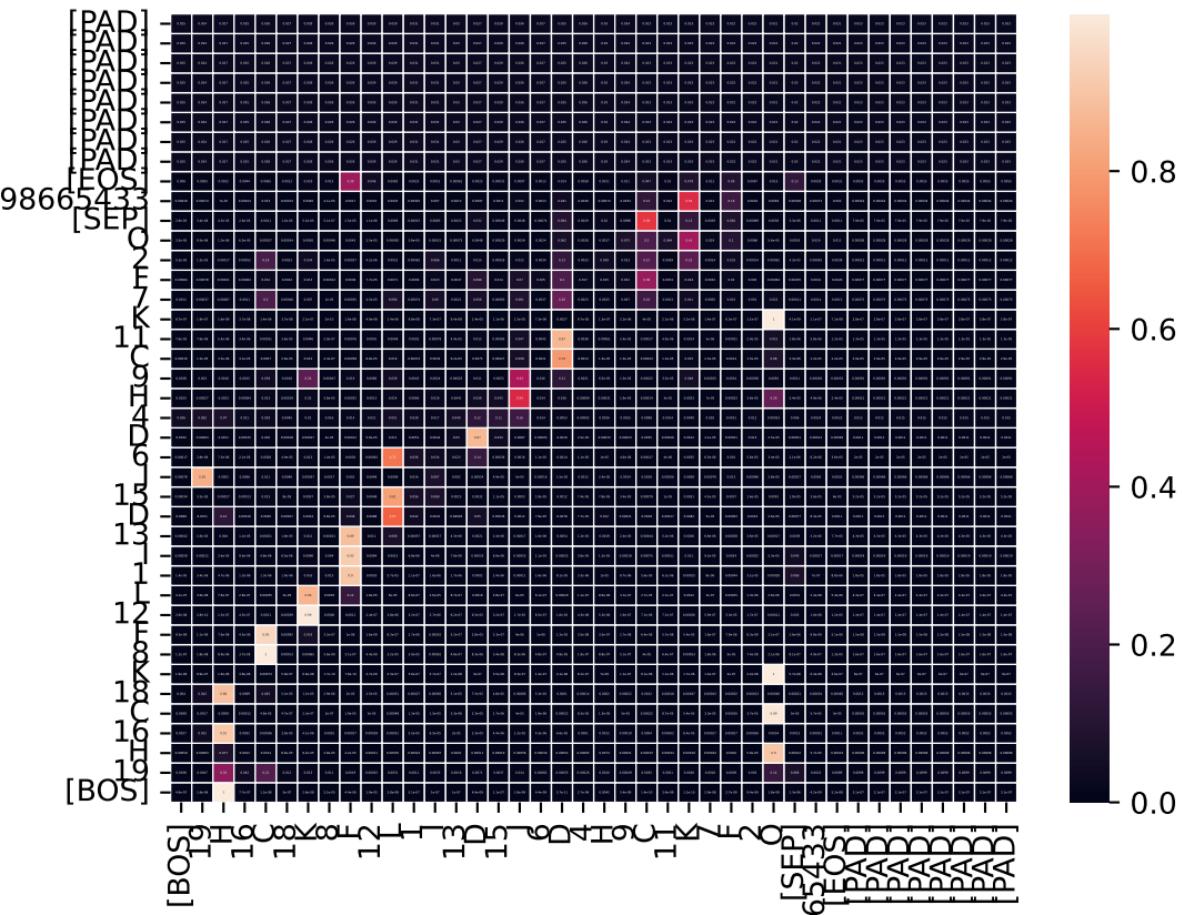
Step 30000 attention weights at layer 0 head 0, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



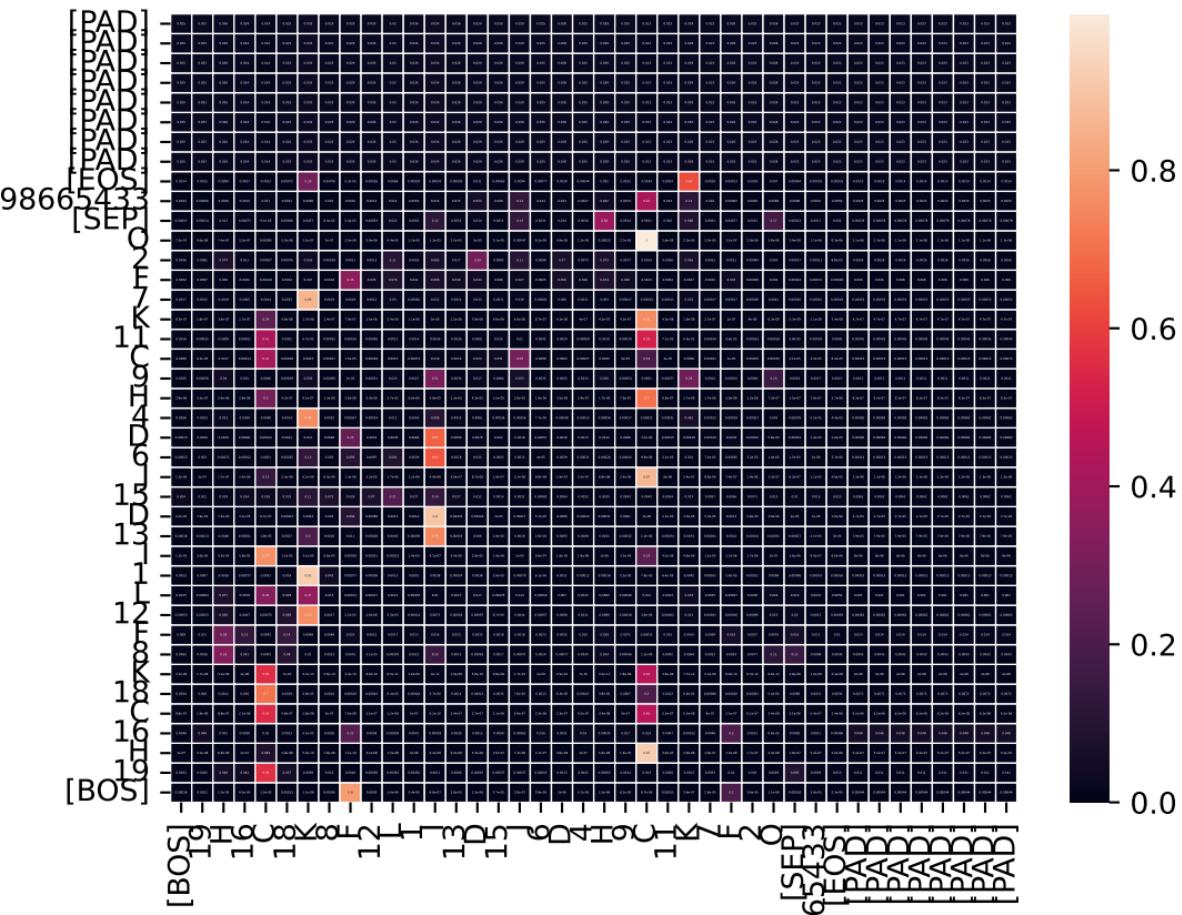
Step 30000 attention weights at layer 0 head 1, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



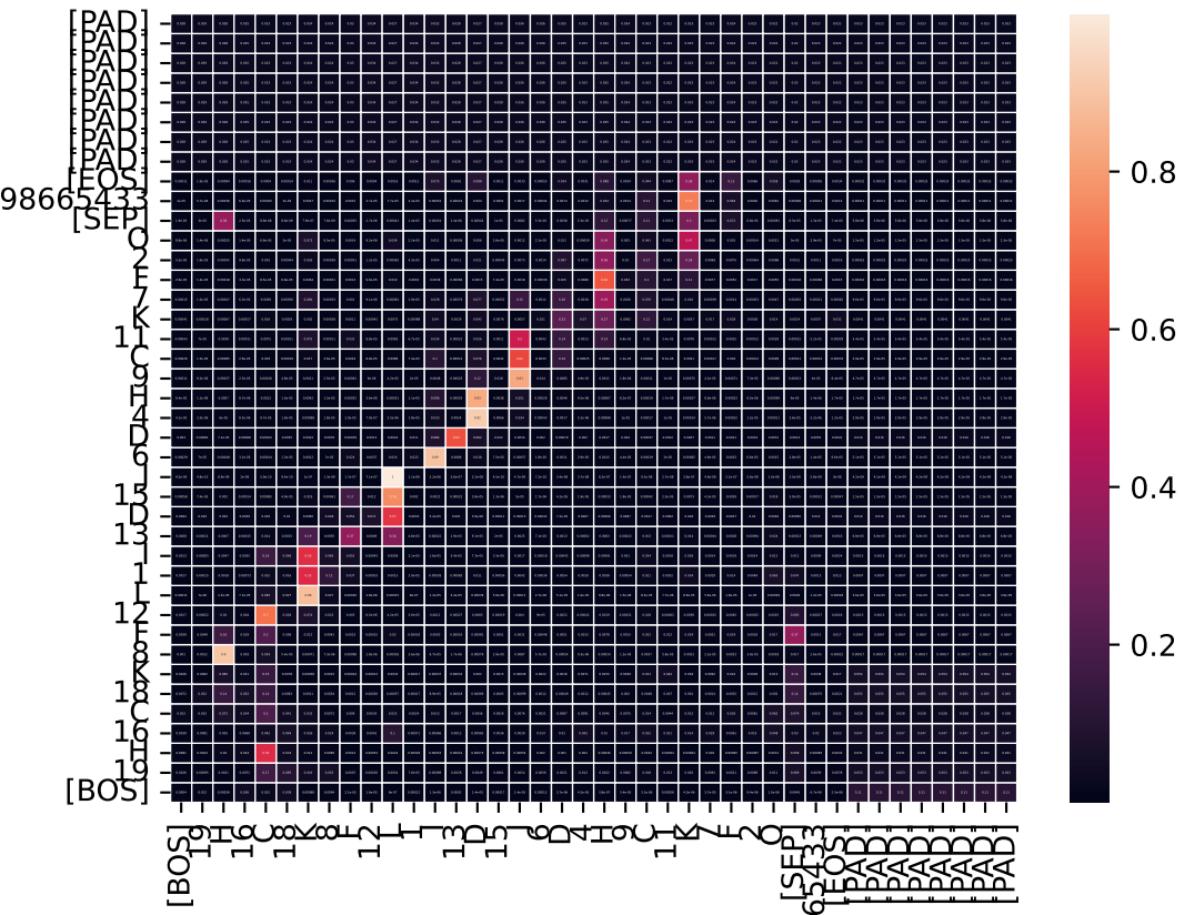
Step 30000 attention weights at layer 0 head 2, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



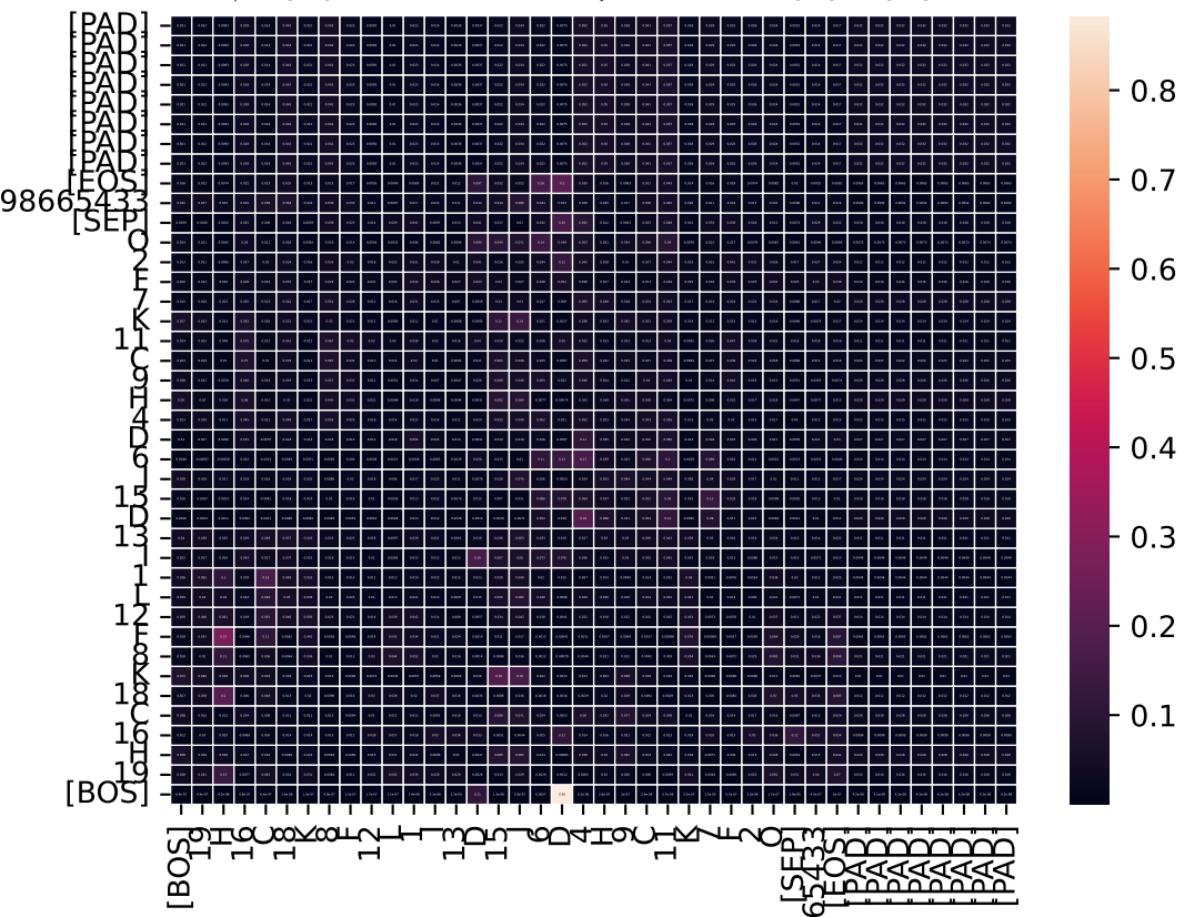
Step 30000 attention weights at layer 0 head 3, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



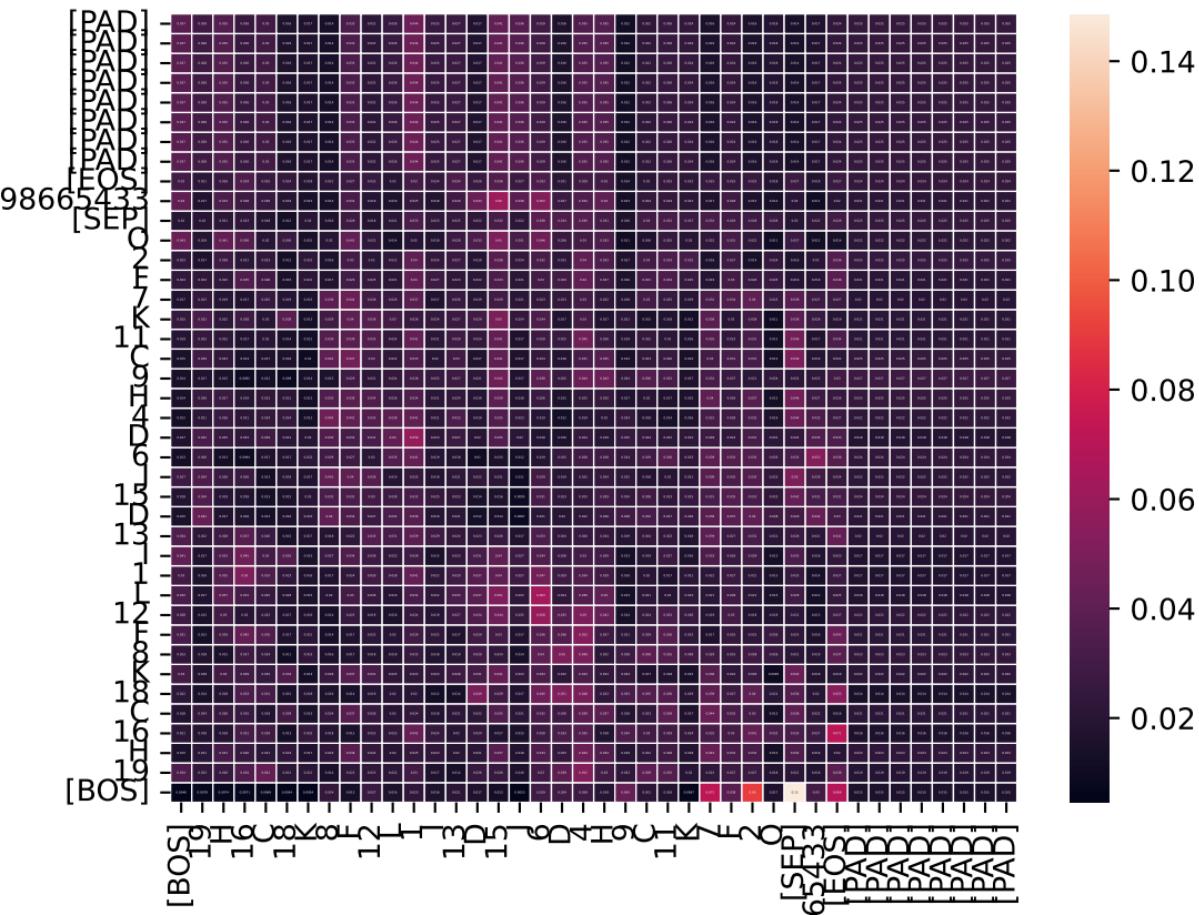
Step 30000 attention weights at layer 1 head 0, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



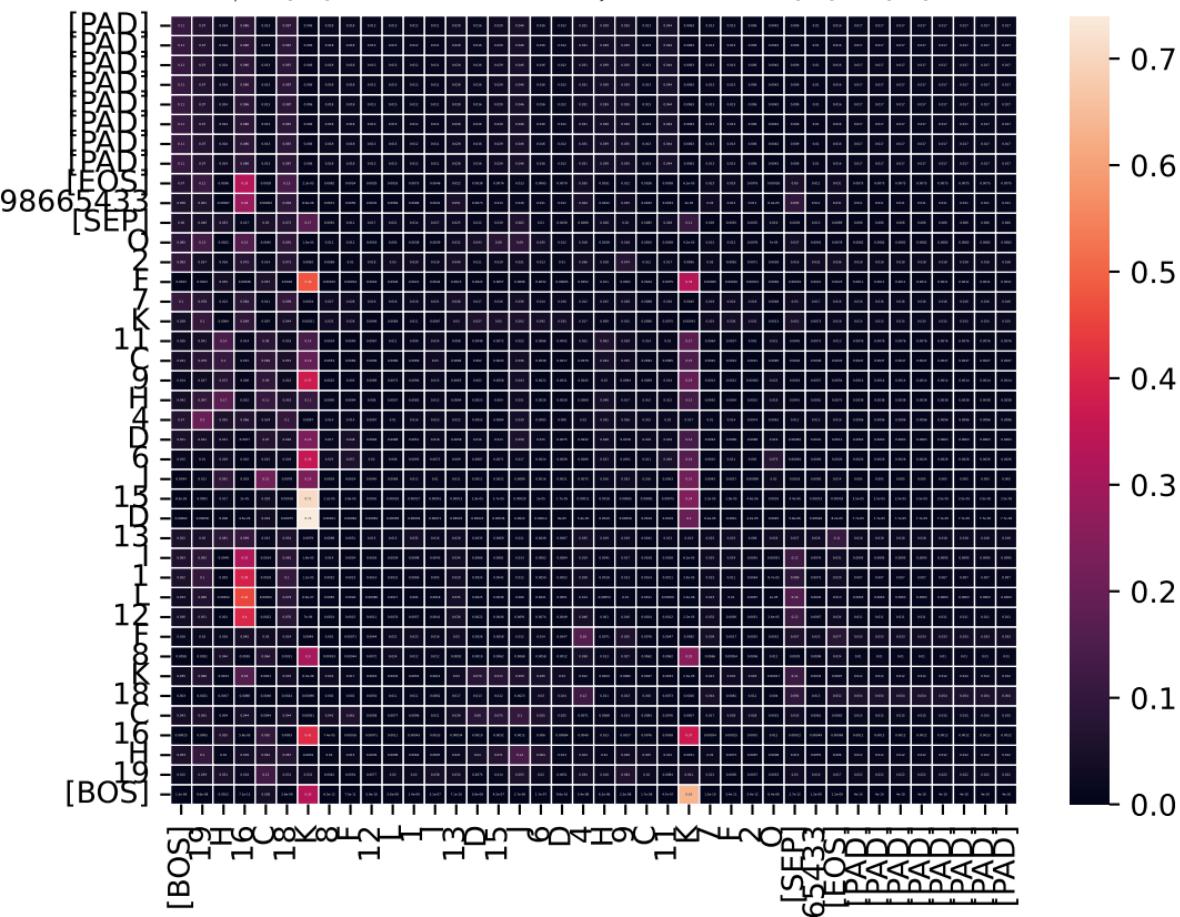
Step 30000 attention weights at layer 1 head 1, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



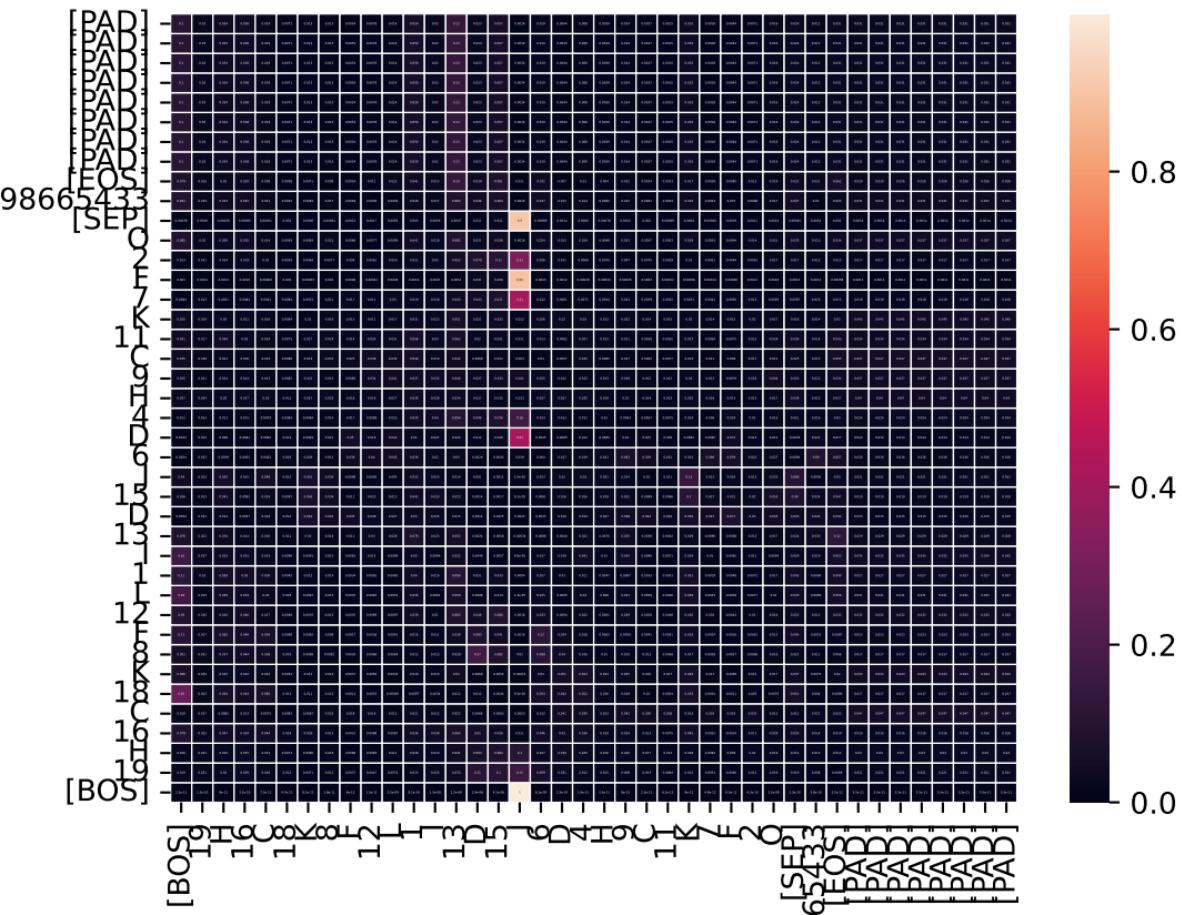
Step 30000 attention weights at layer 1 head 2, with accuracy 0.259

Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K

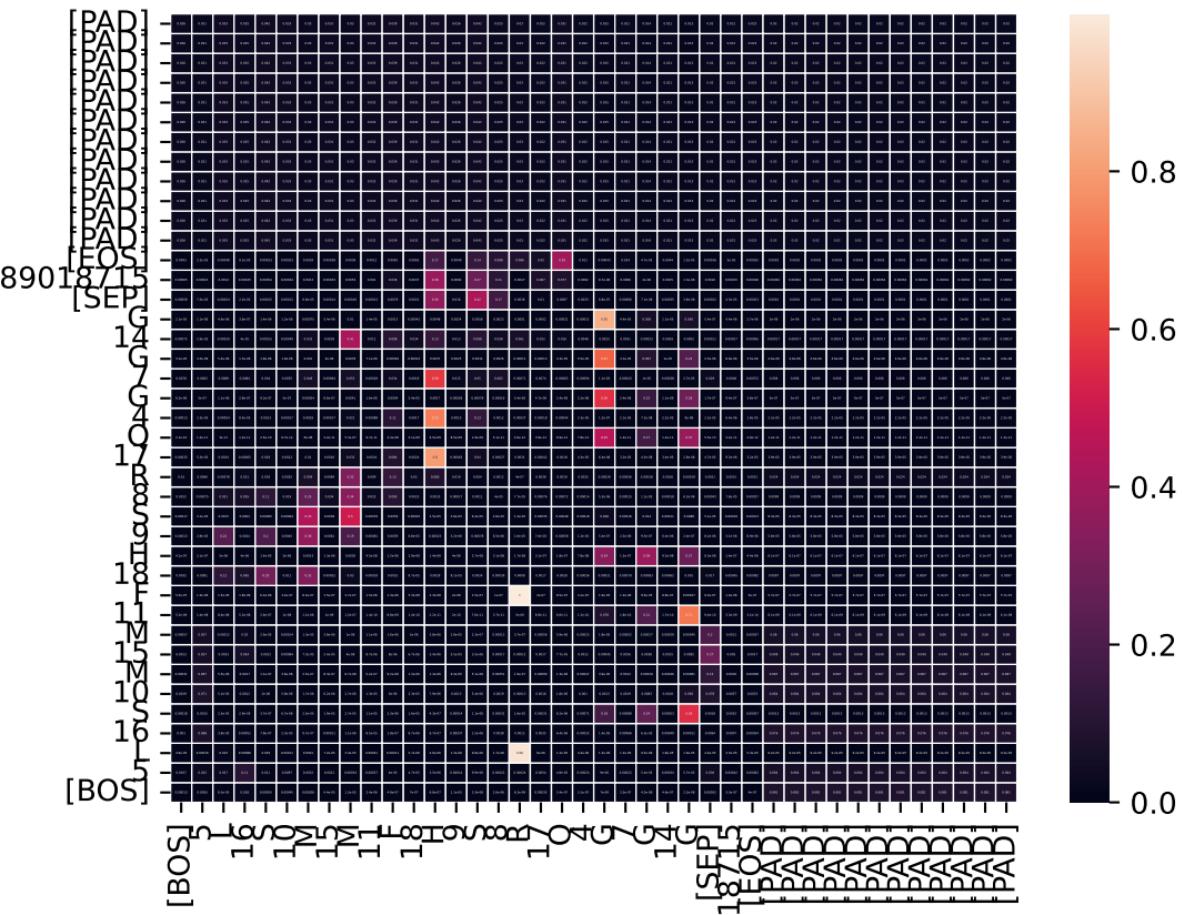


Step 30000 attention weights at layer 1 head 3, with accuracy 0.259

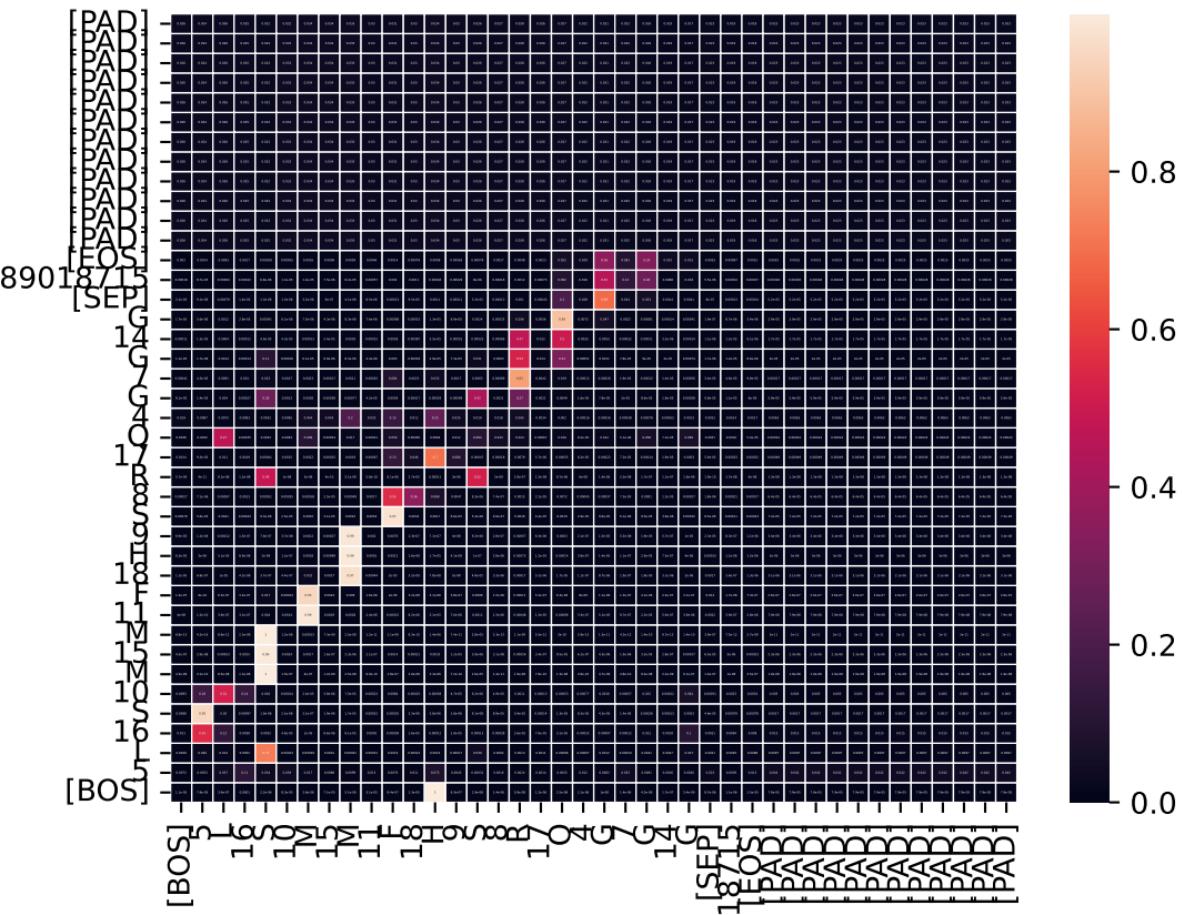
Sequence:[BOS] 19 H 16 C 18 K 8 F 12 L 11 13 D 15 J 6 D 4 H 9 C 11 K 7 F 2 O [SEP] 18.02 [EOS] SOL:K



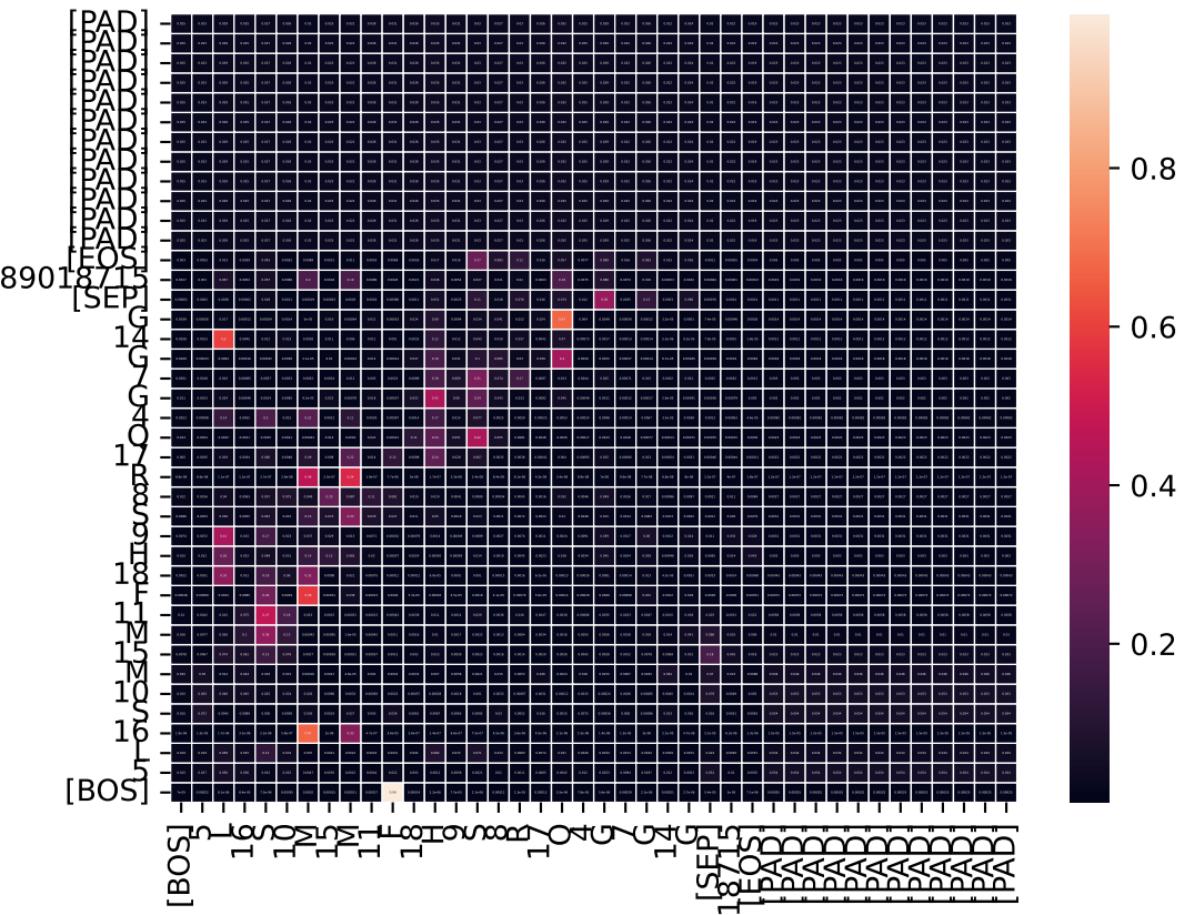
Step 39500 attention weights at layer 0 head 0, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G



Step 39500 attention weights at layer 0 head 1, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G

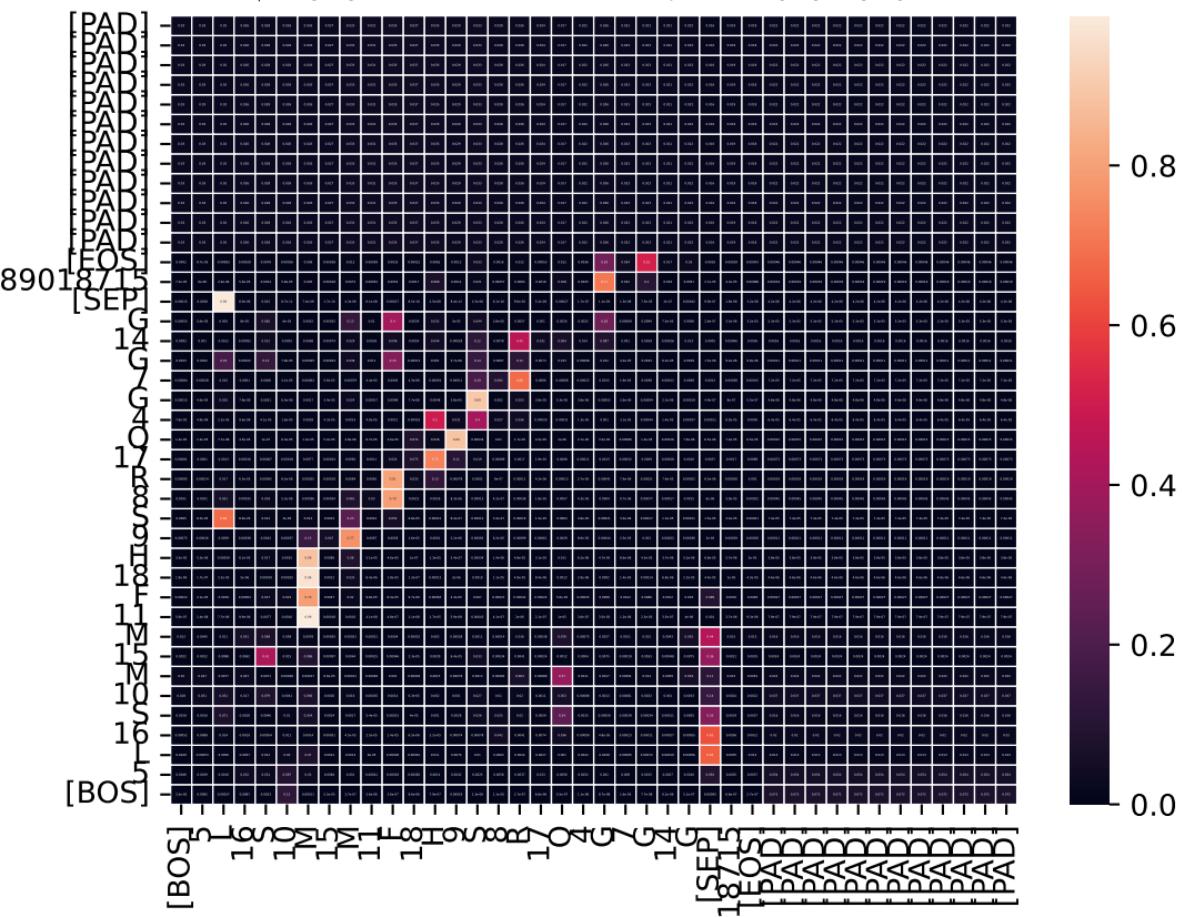


Step 39500 attention weights at layer 0 head 2, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G

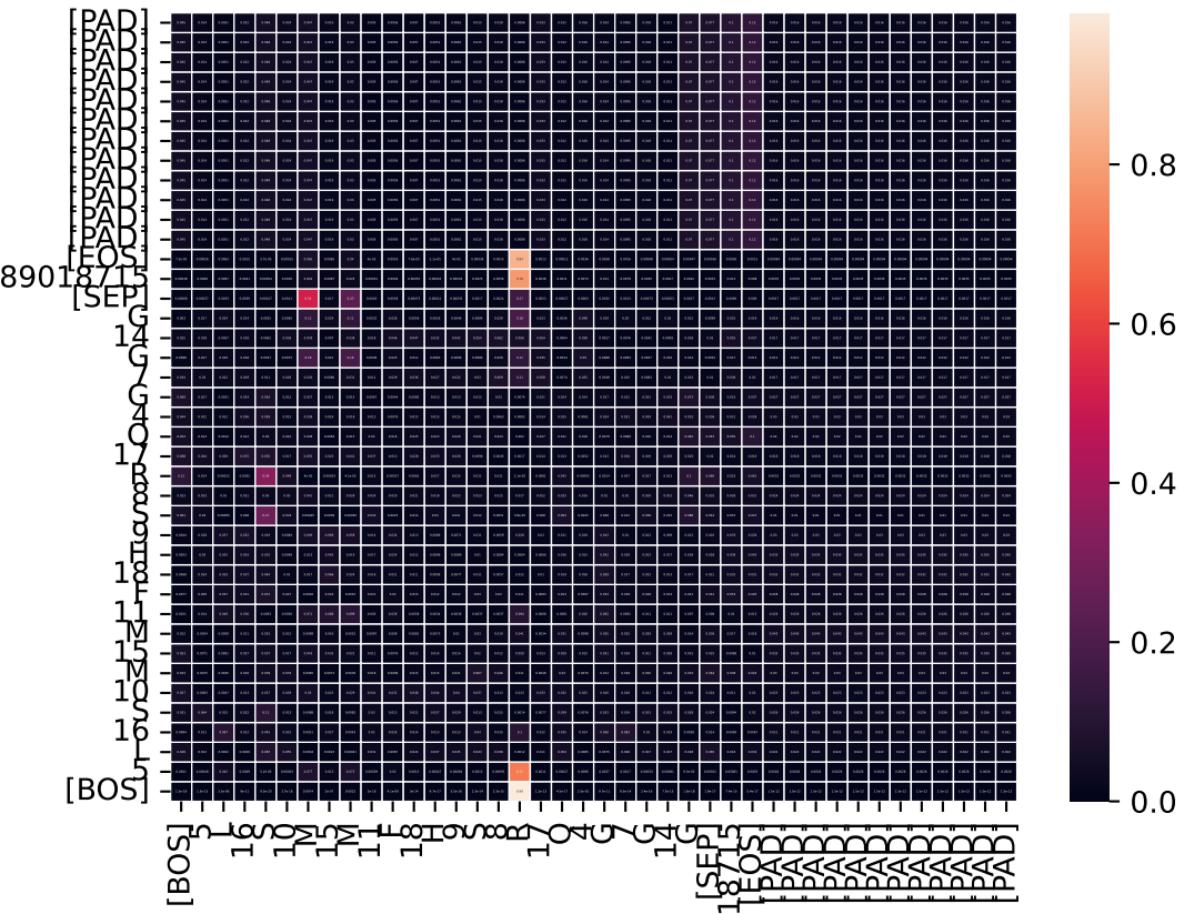


Step 39500 attention weights at layer 0 head 3, with accuracy 0.265

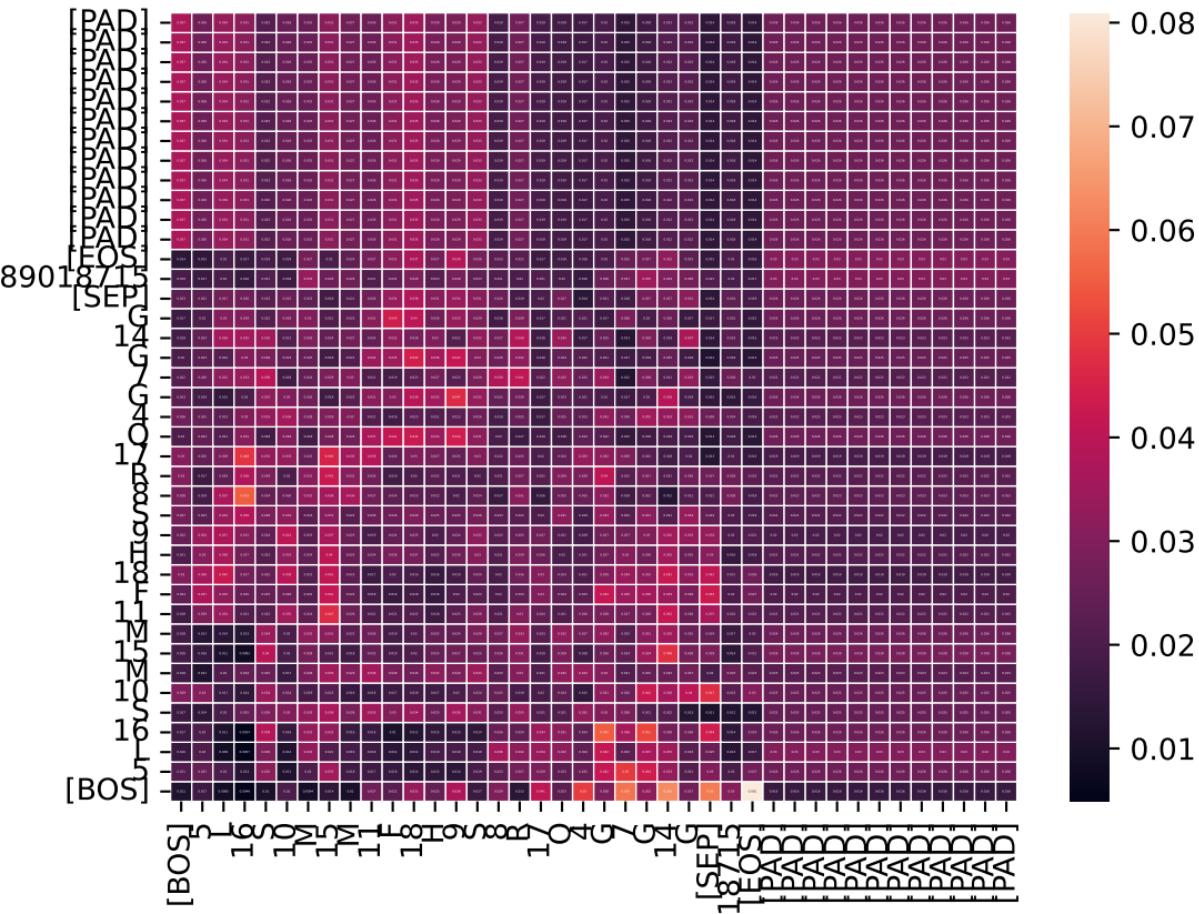
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 5 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G



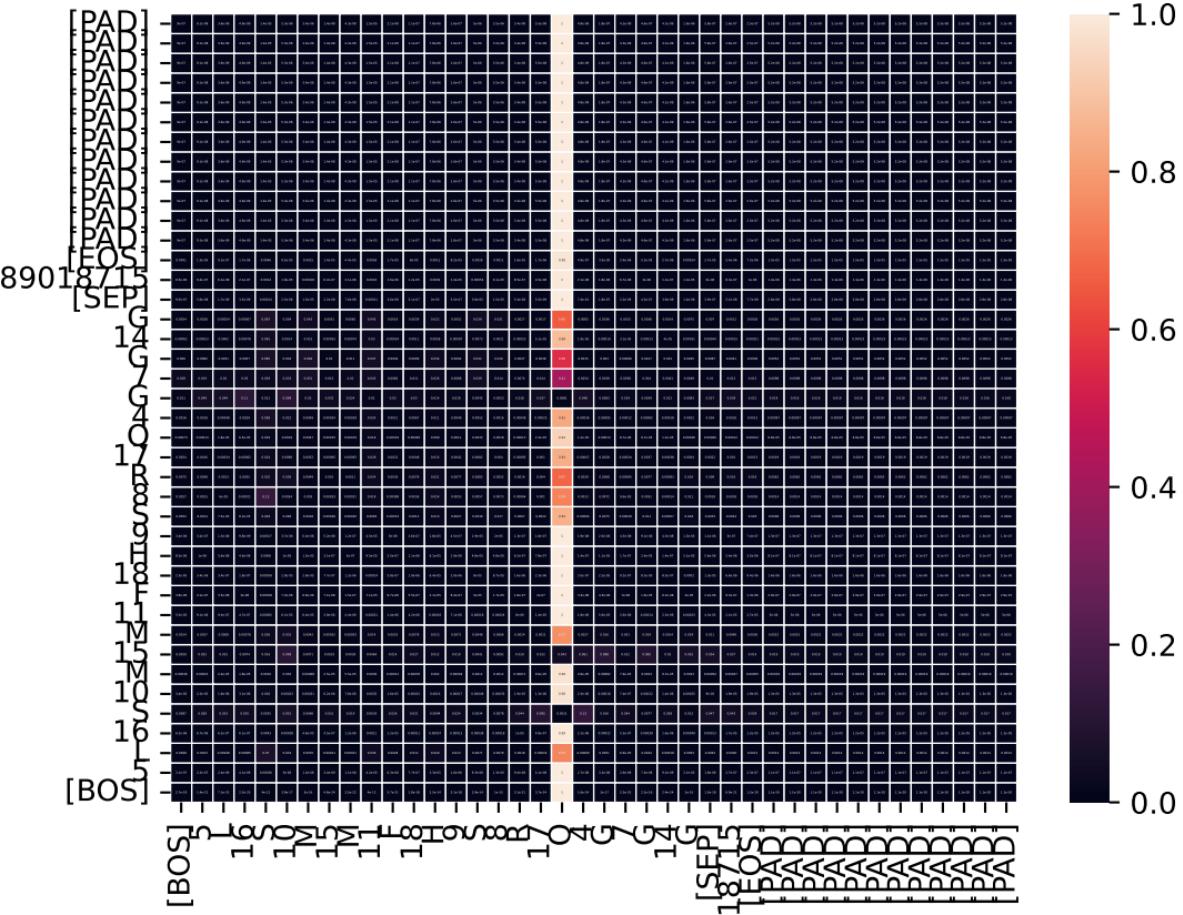
Step 39500 attention weights at layer 1 head 0, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G



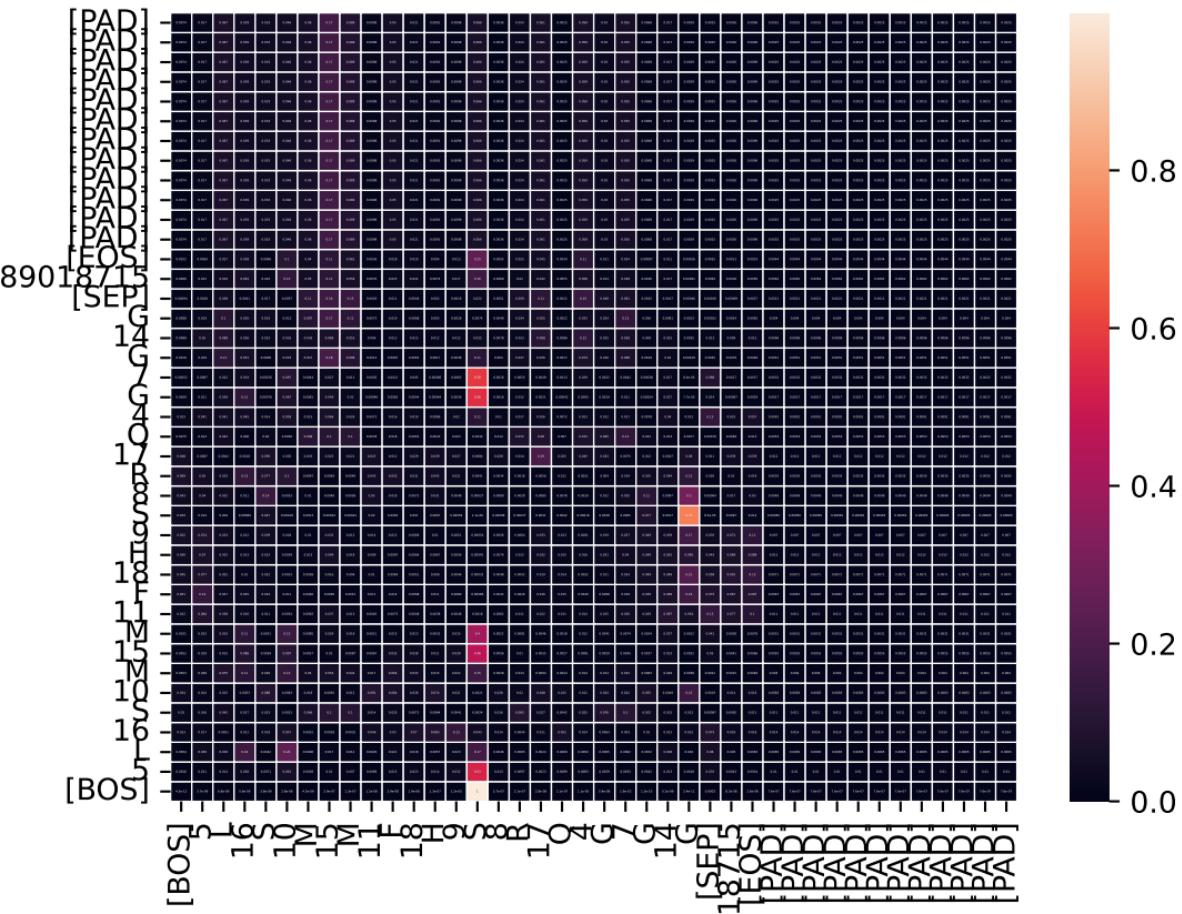
Step 39500 attention weights at layer 1 head 1, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G



Step 39500 attention weights at layer 1 head 2, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G

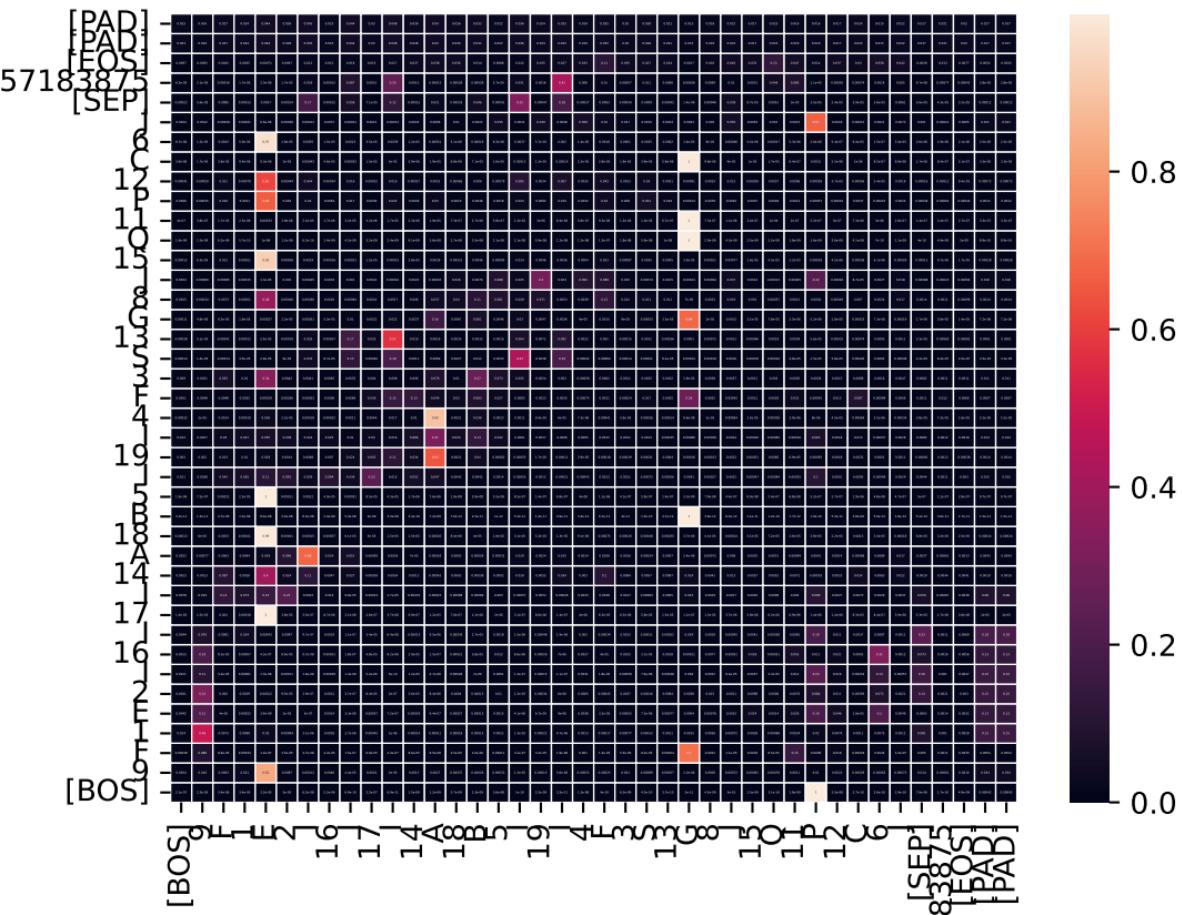


Step 39500 attention weights at layer 1 head 3, with accuracy 0.265
Sequence:[BOS] 5 L 16 S 10 M 15 M 11 F 18 H 9 S 8 R 17 Q 4 G 7 G 14 G [SEP] 3.99 [EOS] SOL:G



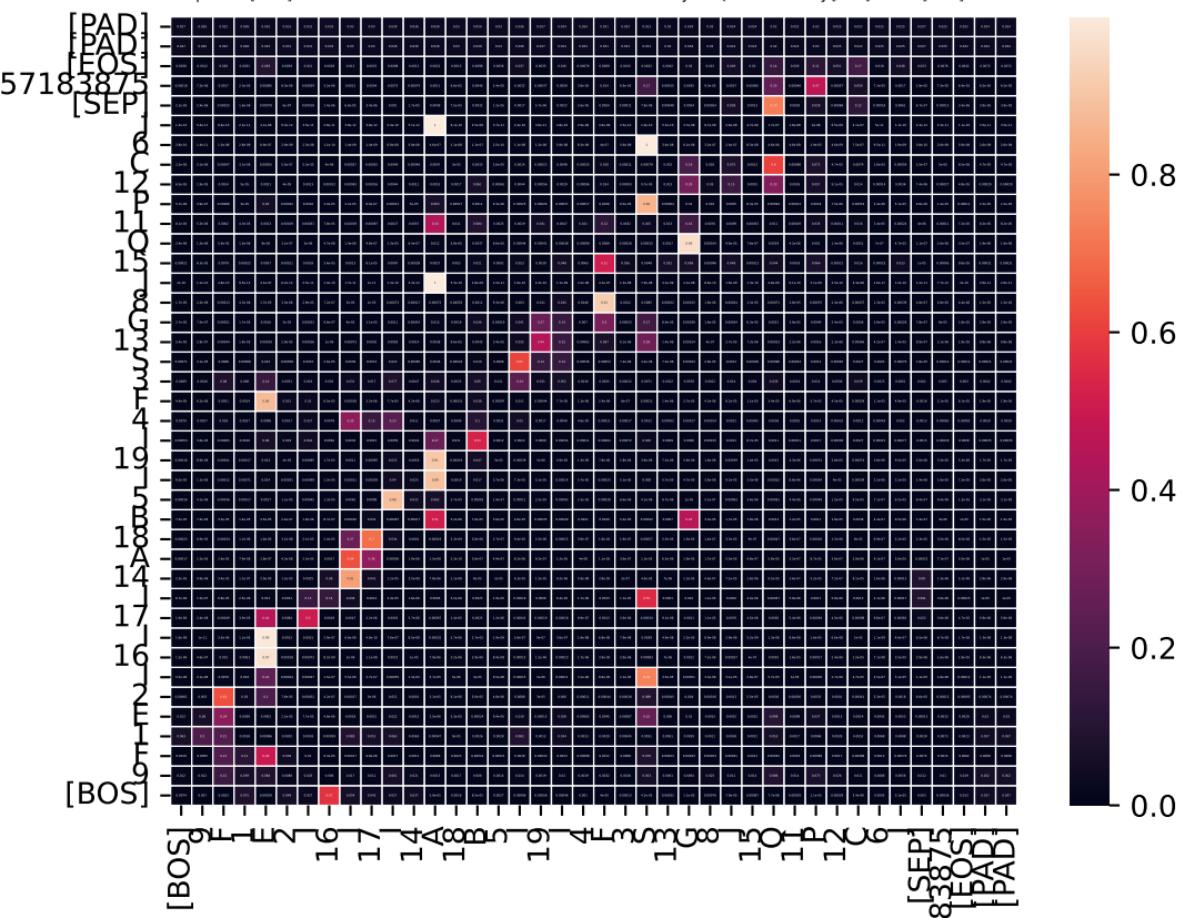
Step 40000 attention weights at layer 0 head 0, with accuracy 0.245

Sequence:[BOS] 9 F 1 E 2 I 16 | 17 | 14 A 18 B 5 | 19 14 F 3 S 13 G 8 J 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G



Step 40000 attention weights at layer 0 head 1, with accuracy 0.245

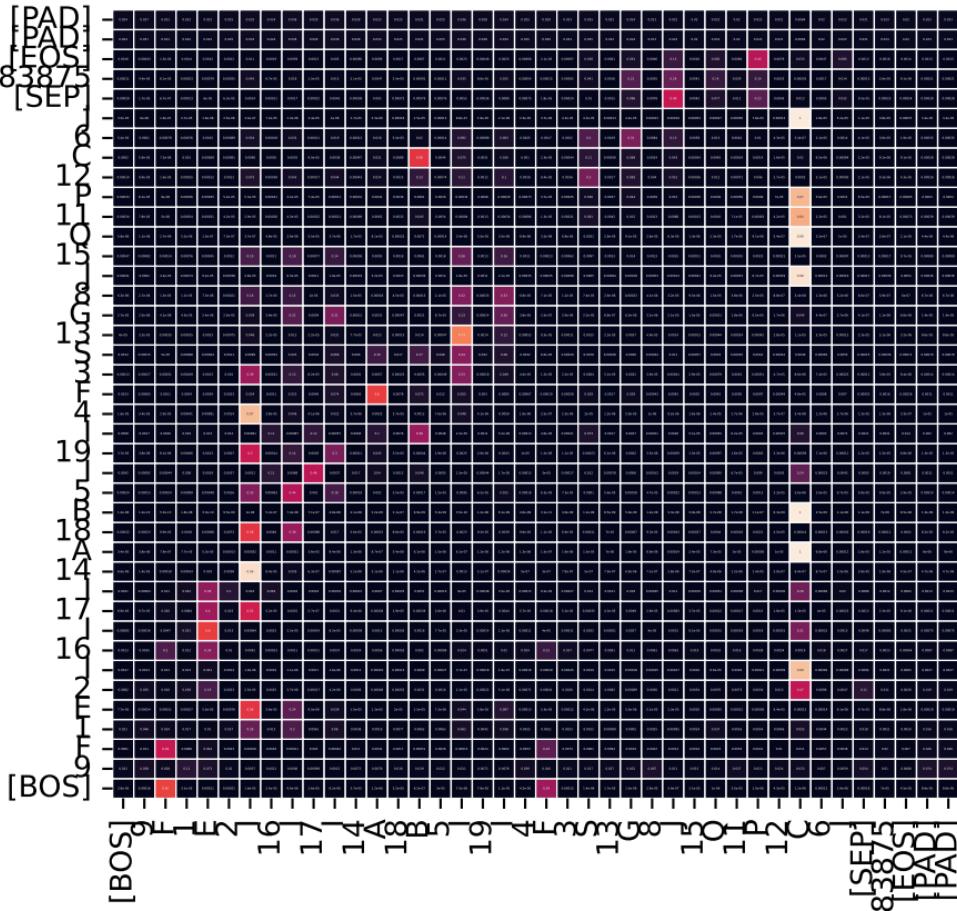
Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 14 F 3 S 13 G 8 | 15 Q 11 P 12 C 6 | [SEP] 13.14 [EOS] SOL:G



Step 40000 attention weights at layer 0 head 2, with accuracy 0.245

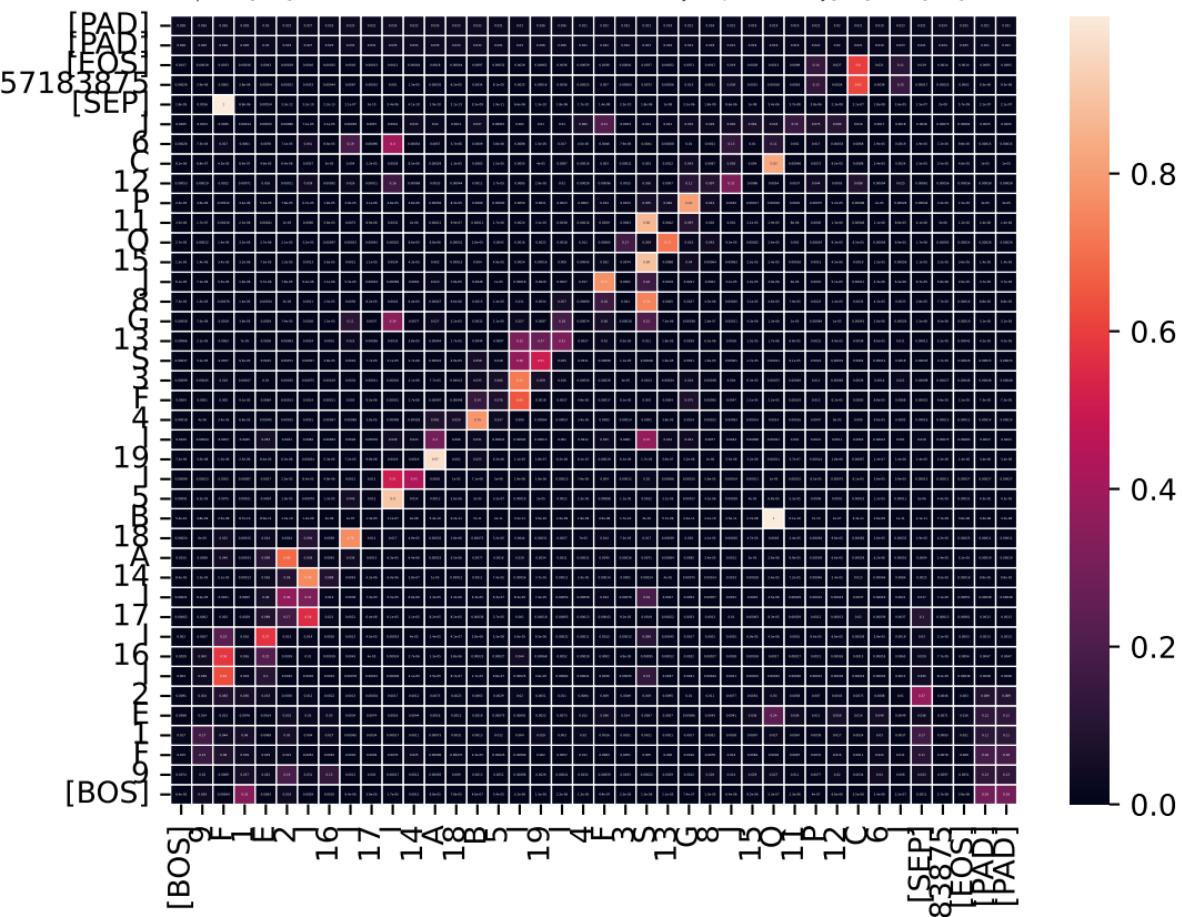
Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 | 4 F 3 S 13 G 8 | 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G

57183875



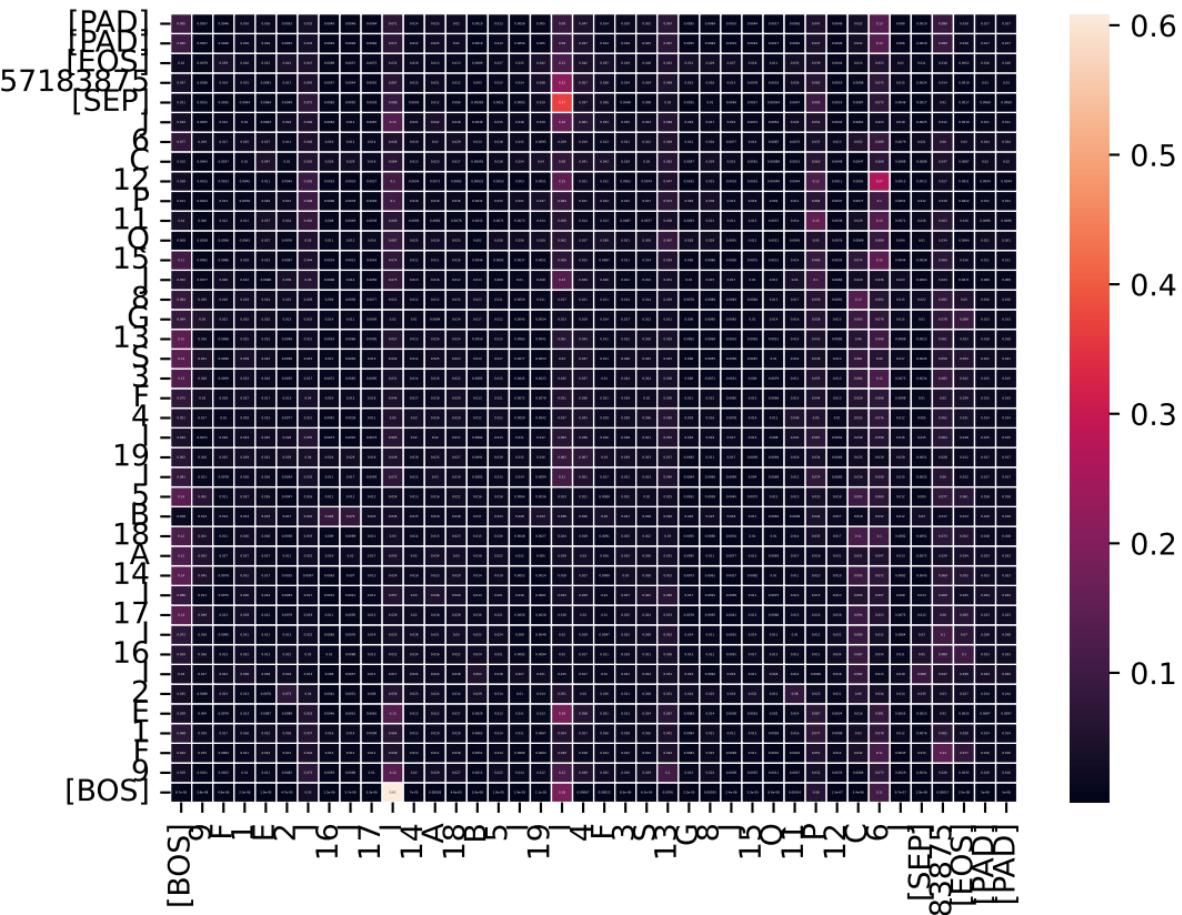
Step 40000 attention weights at layer 0 head 3, with accuracy 0.245

Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 14 F 3 S 13 G 8 J 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G



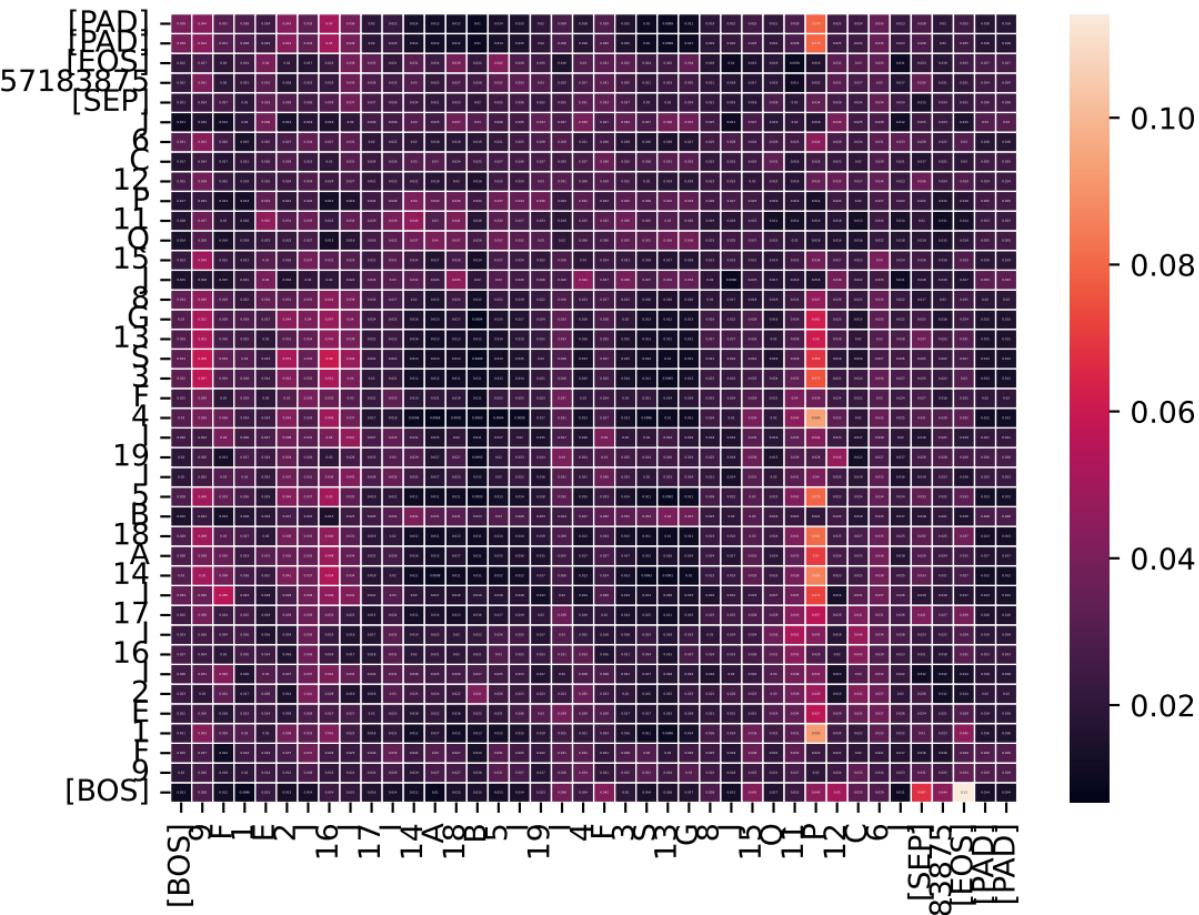
Step 40000 attention weights at layer 1 head 0, with accuracy 0.245

Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 14 F 3 S 13 G 8 | 15 Q 11 P 12 C 6 | [SEP] 13.14 [EOS] SOL:G



Step 40000 attention weights at layer 1 head 1, with accuracy 0.245

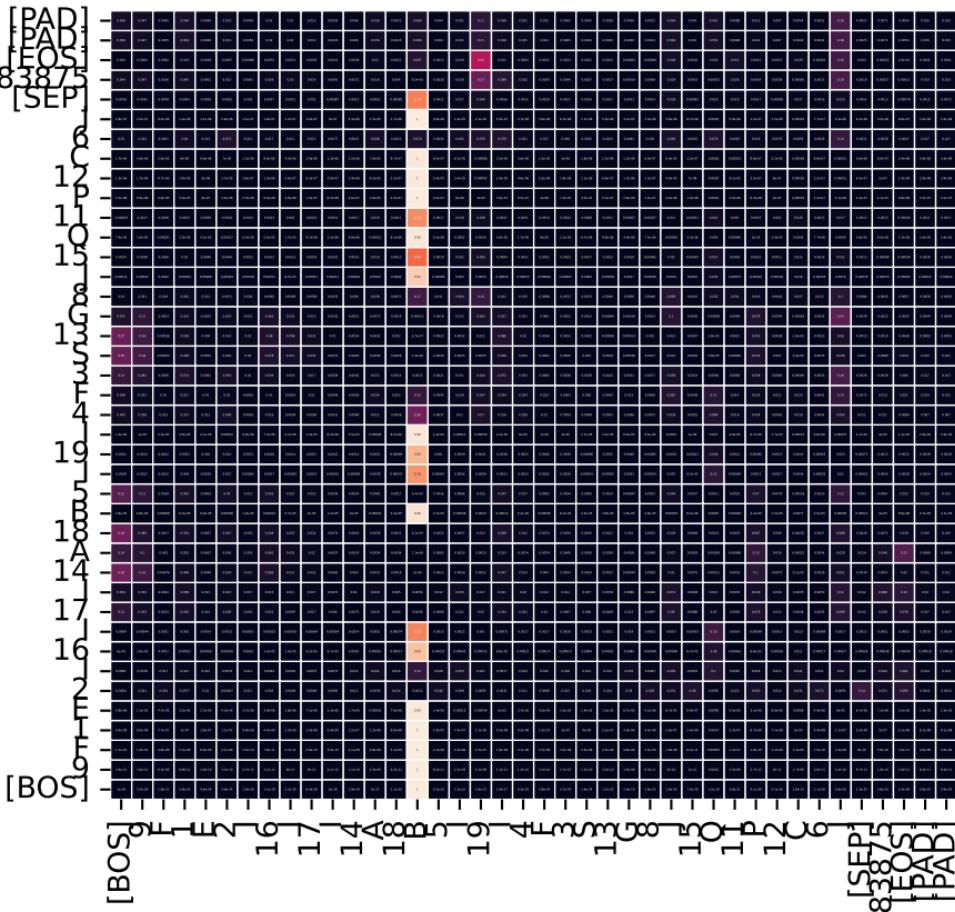
Sequence:[BOS] 9 F 1 E 2 I 16 | 17 | 14 A 18 B 5 | 19 14 F 3 S 13 G 8 J 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G



Step 40000 attention weights at layer 1 head 2, with accuracy 0.245

Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 | 4 F 3 S 13 G 8 | 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G

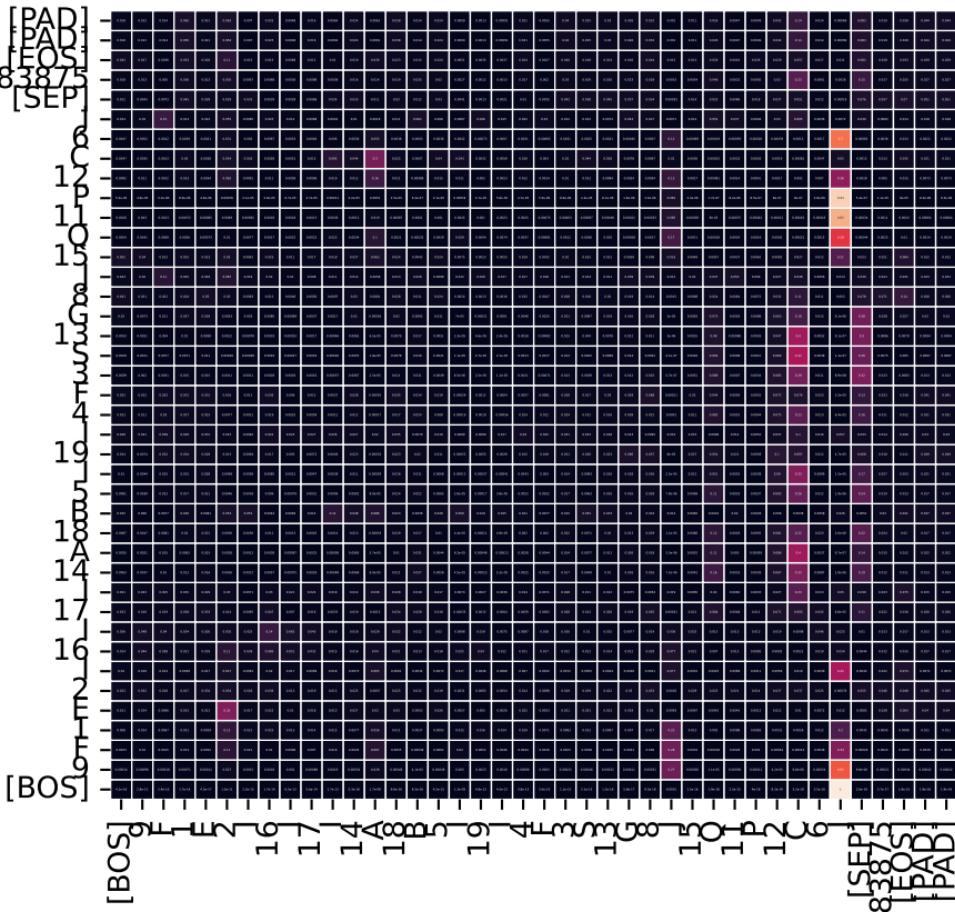
57183875



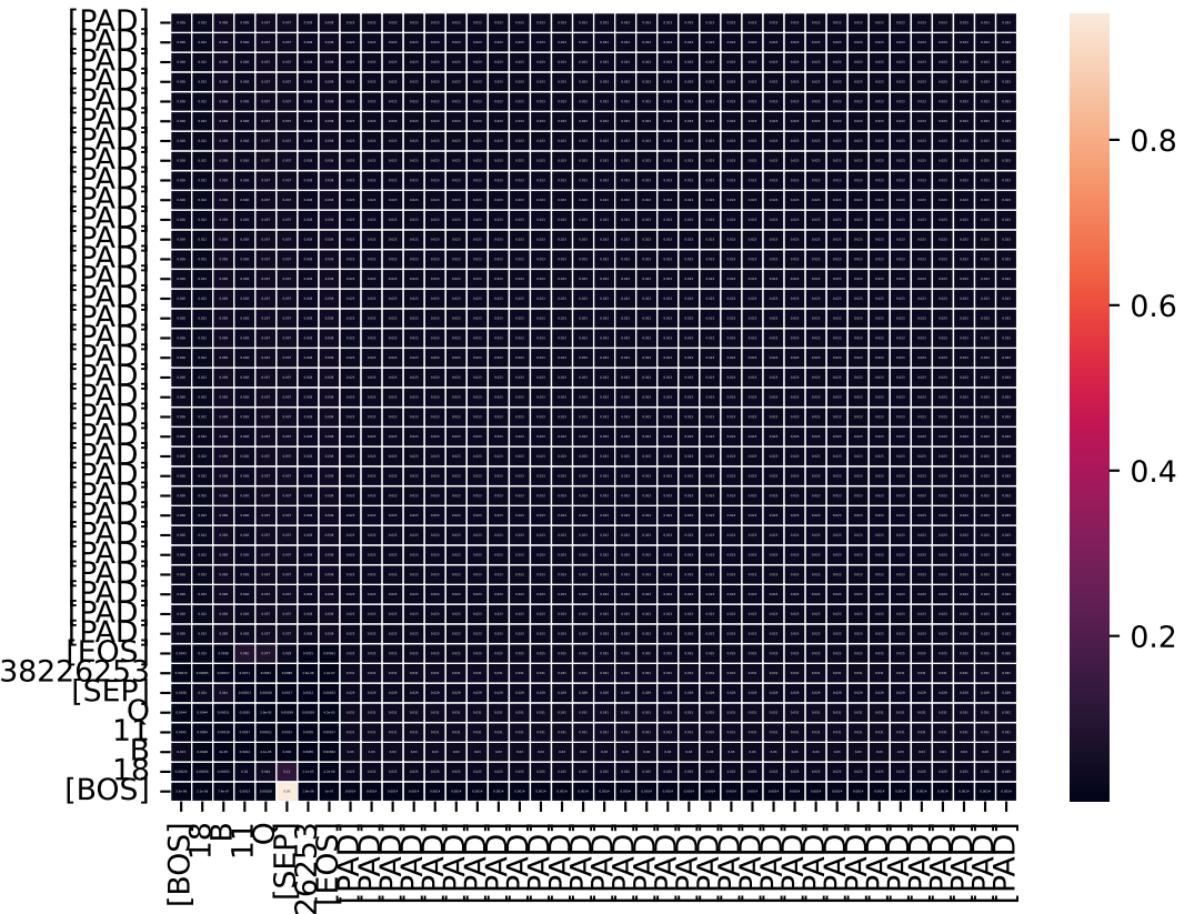
Step 40000 attention weights at layer 1 head 3, with accuracy 0.245

Sequence:[BOS] 9 F 1 E 2 | 16 | 17 | 14 A 18 B 5 | 19 | 4 F 3 S 13 G 8 | 15 Q 11 P 12 C 6 J [SEP] 13.14 [EOS] SOL:G

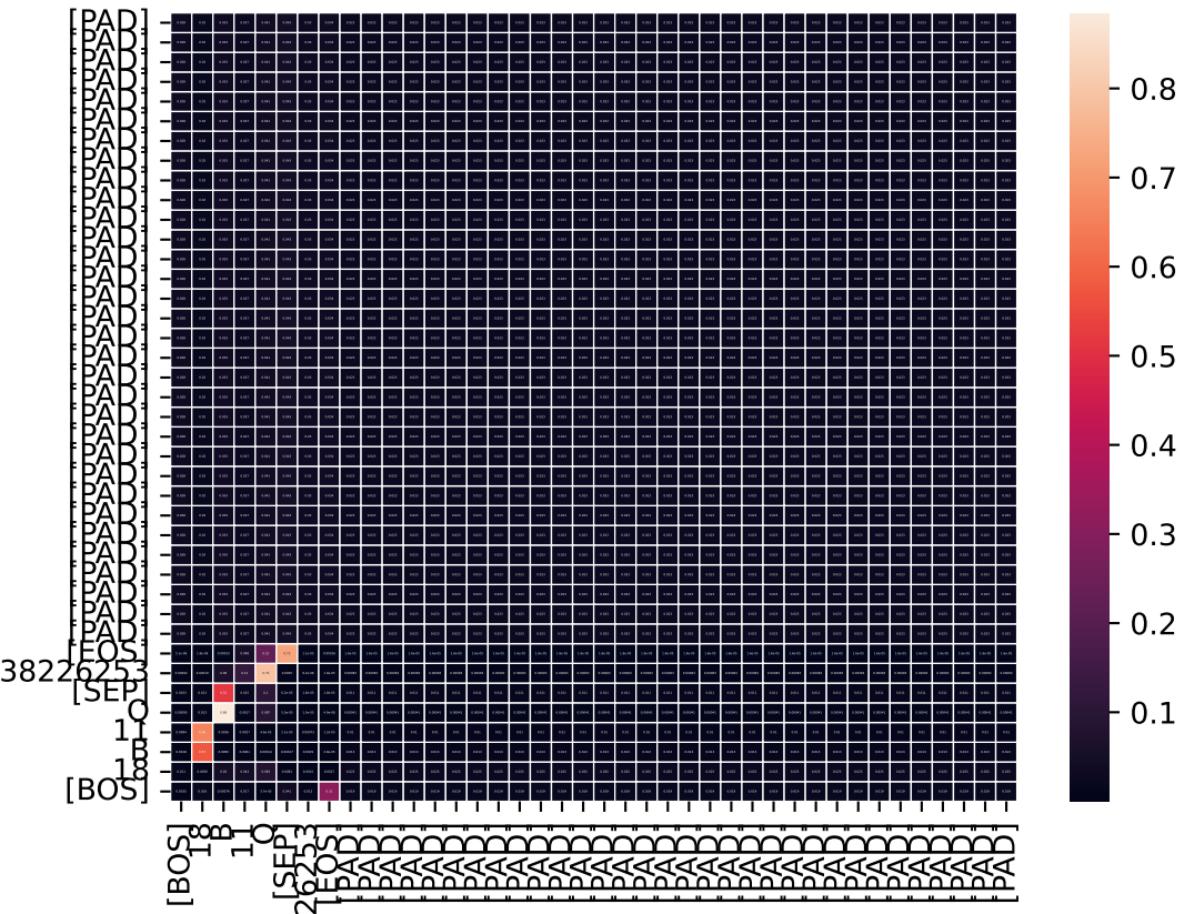
57183875



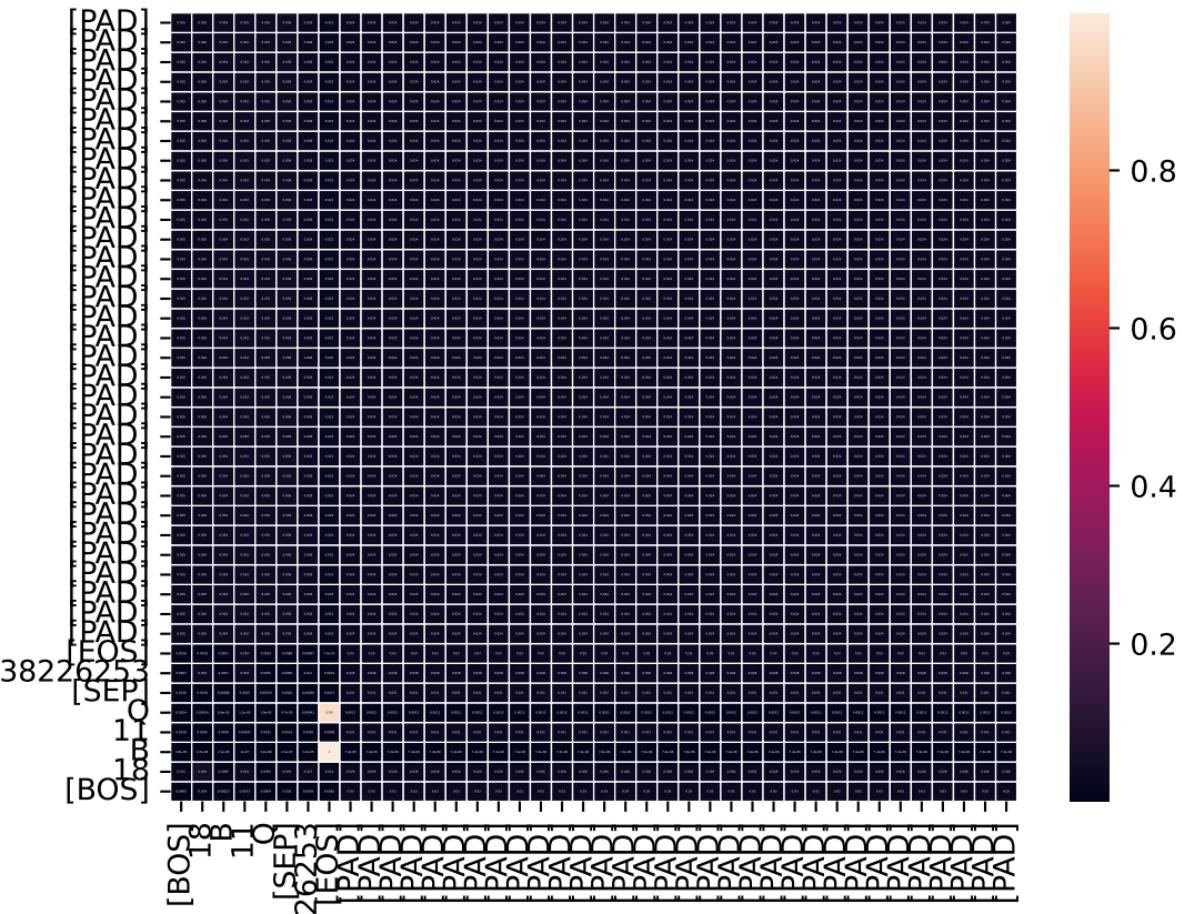
Step 49500 attention weights at layer 0 head 0, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



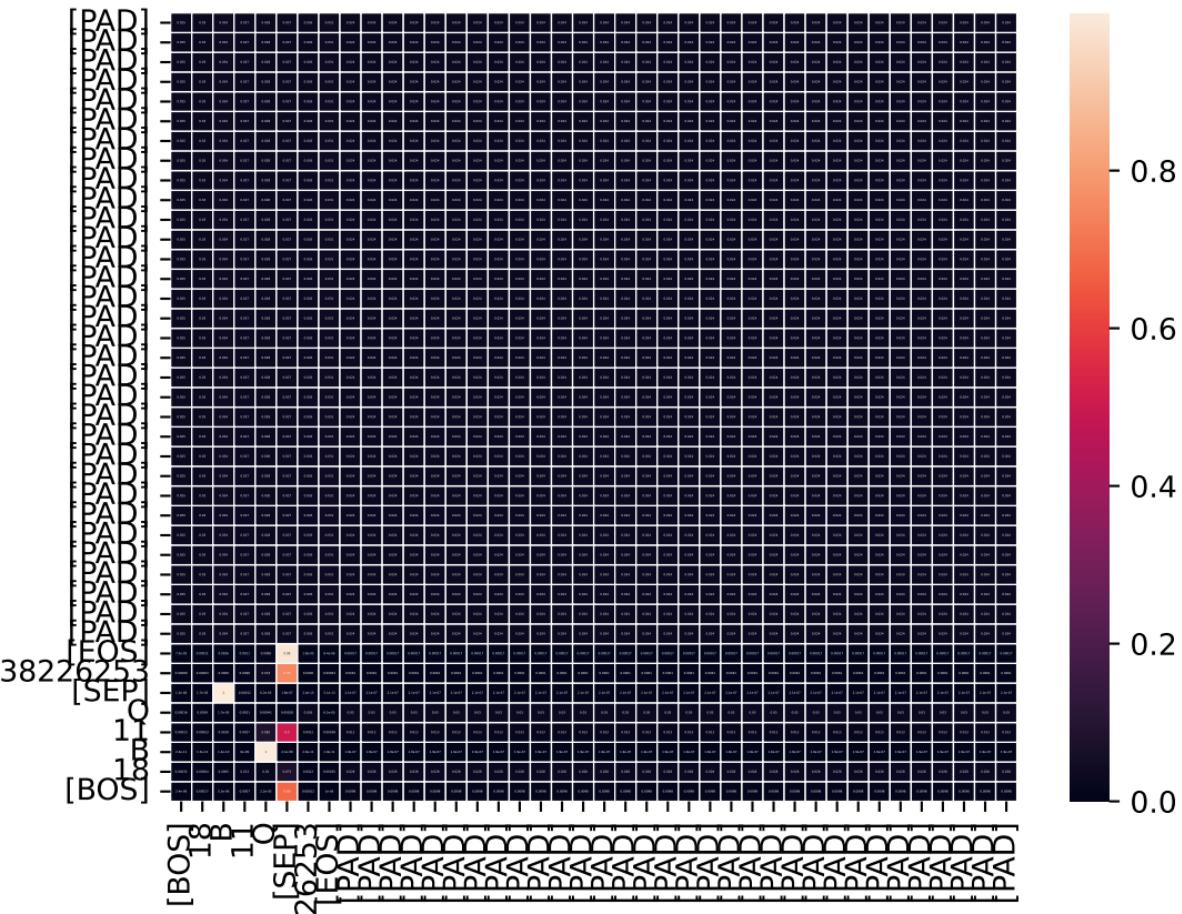
Step 49500 attention weights at layer 0 head 1, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



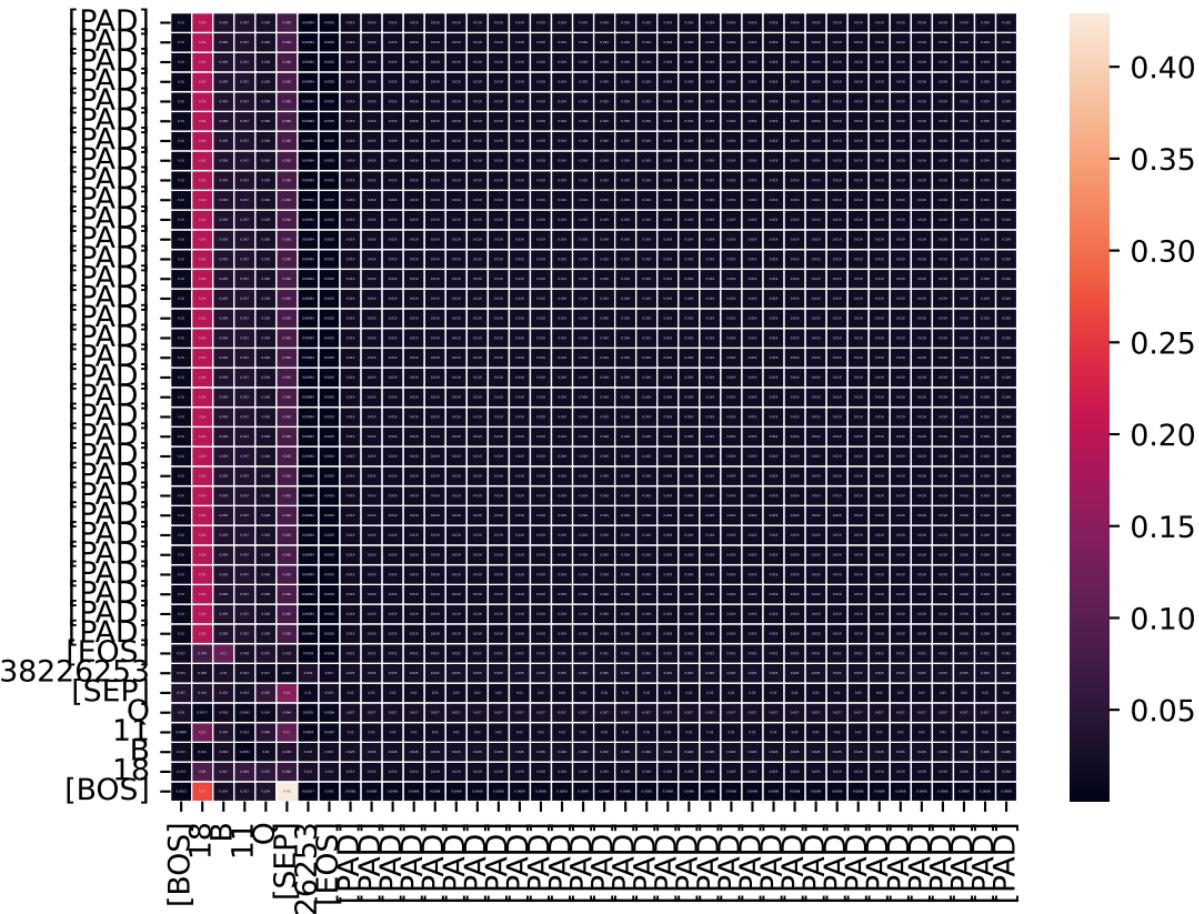
Step 49500 attention weights at layer 0 head 2, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



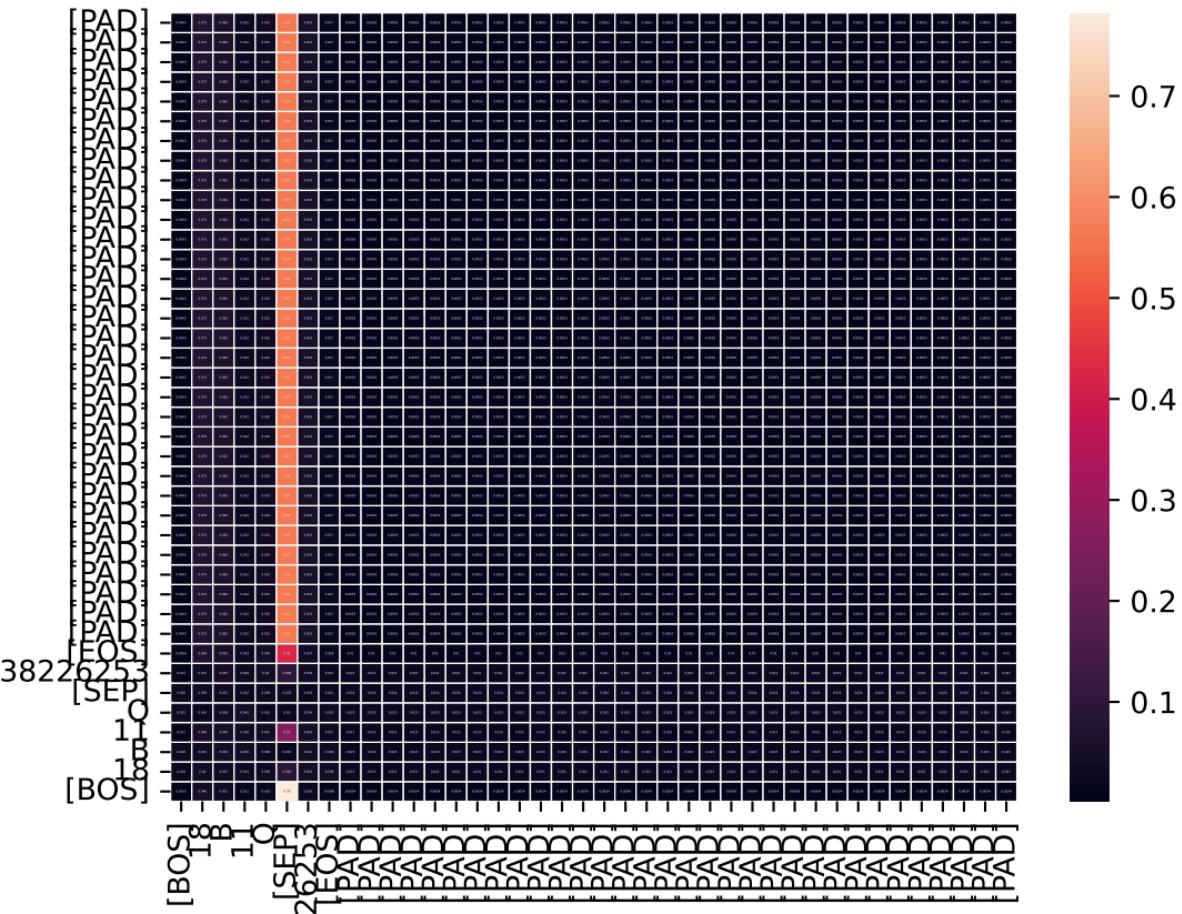
Step 49500 attention weights at layer 0 head 3, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



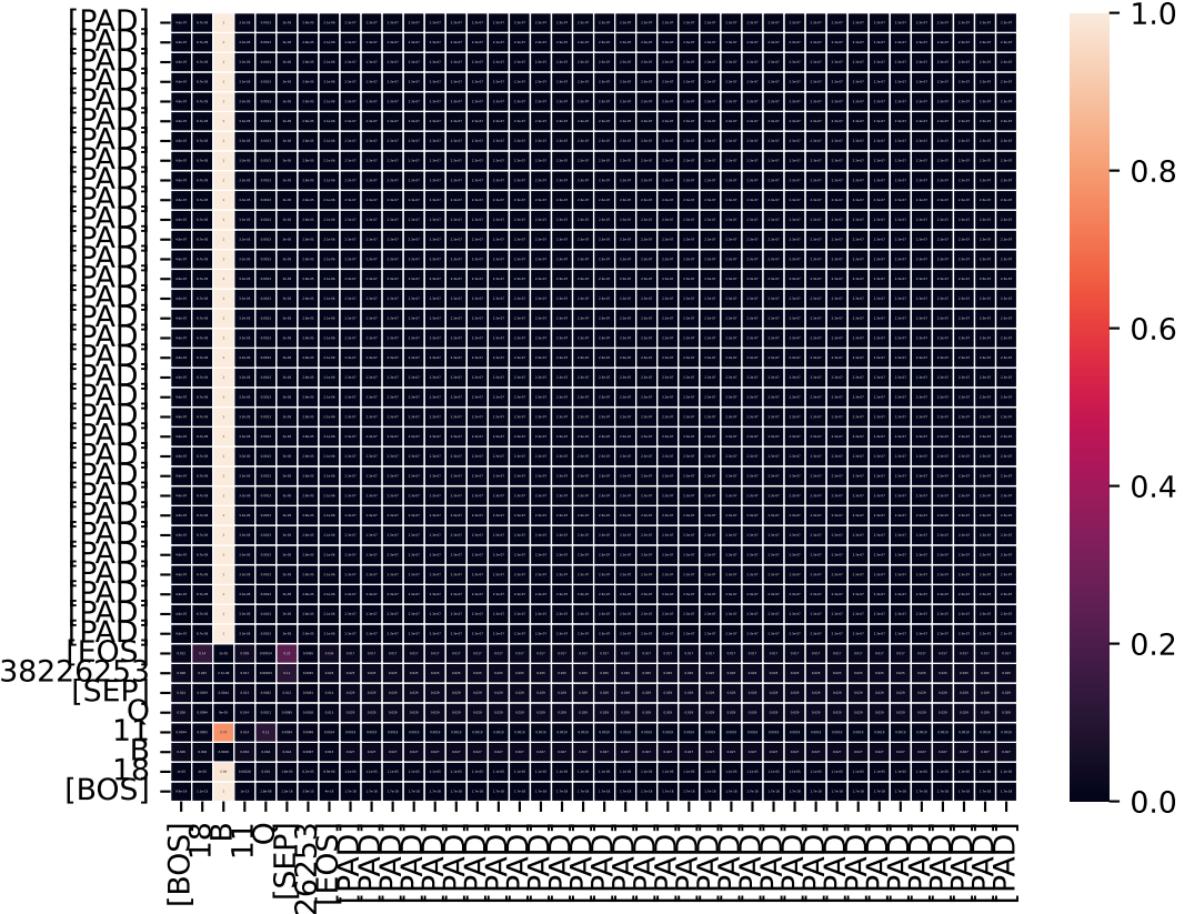
Step 49500 attention weights at layer 1 head 0, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



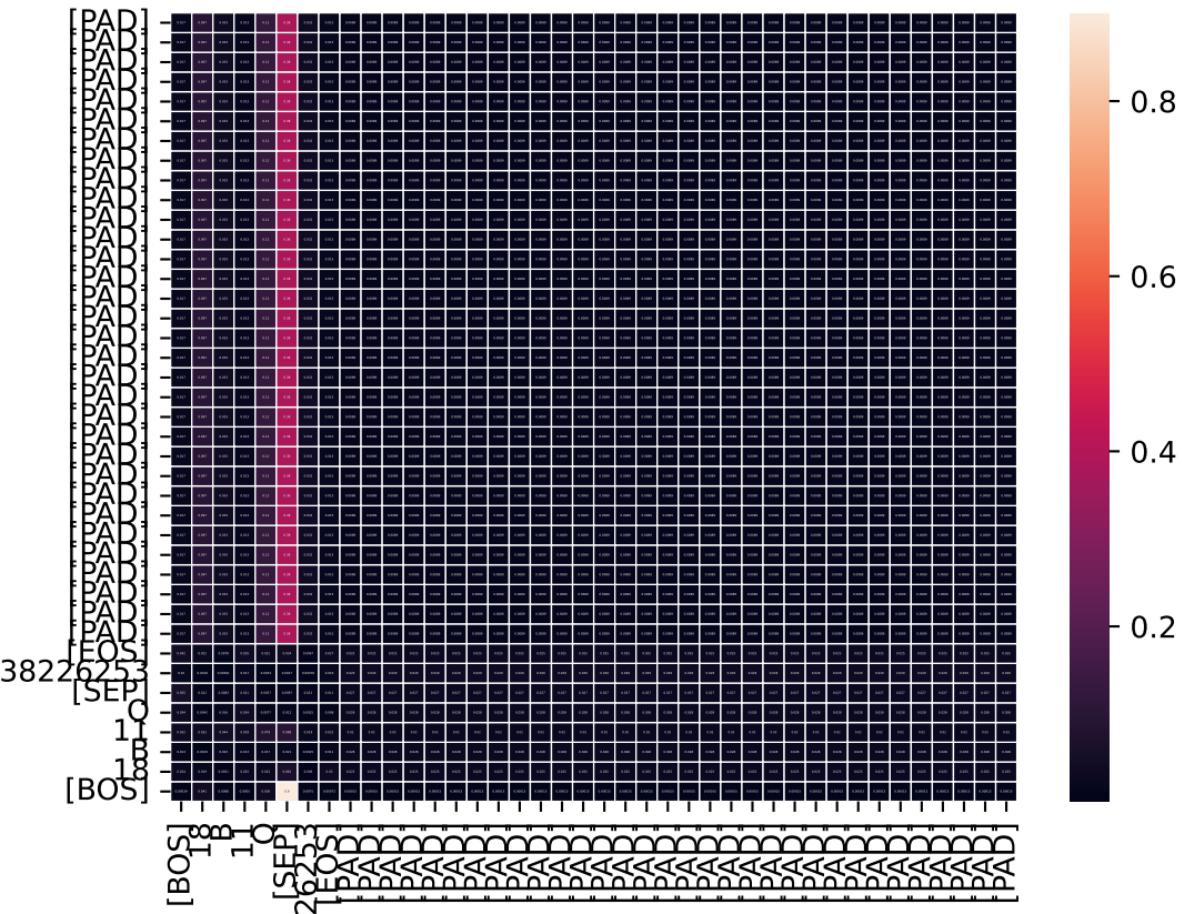
Step 49500 attention weights at layer 1 head 1, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



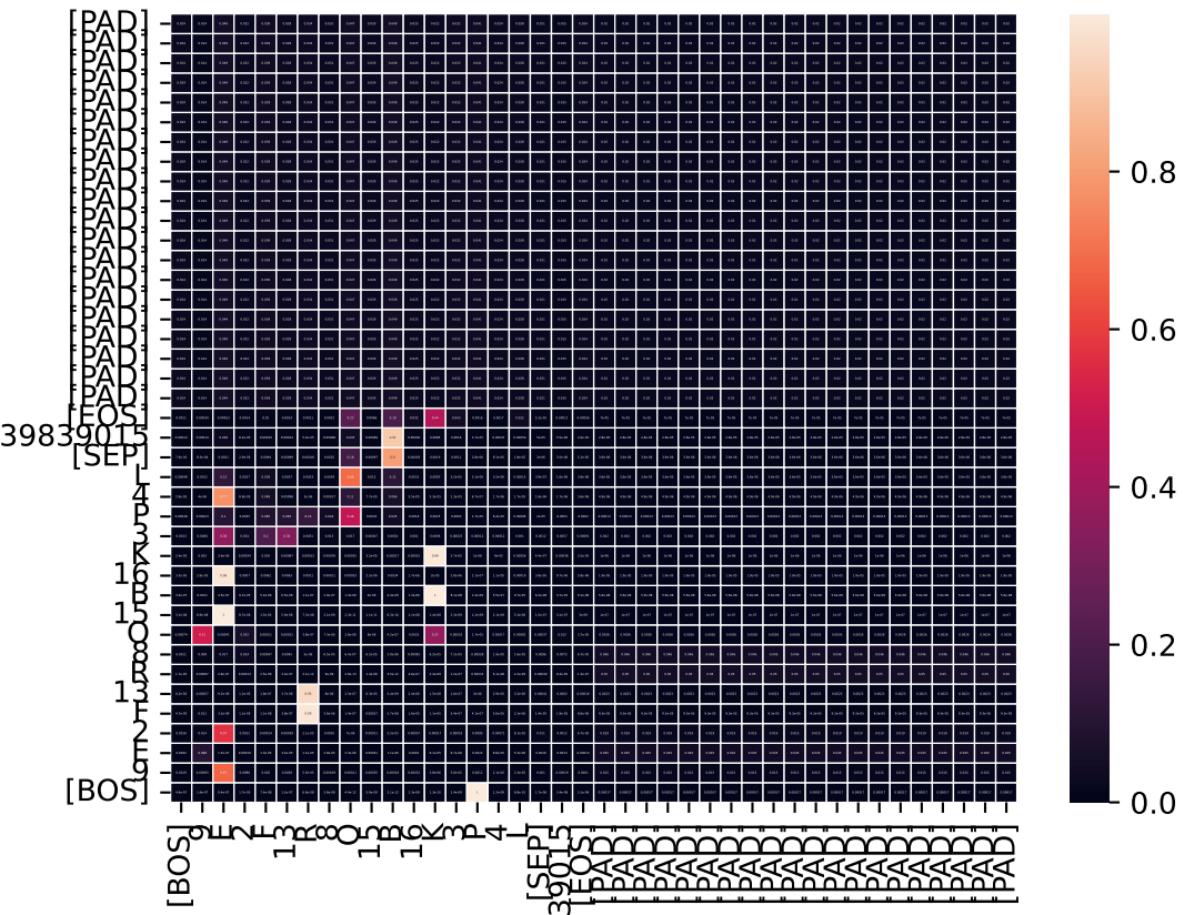
Step 49500 attention weights at layer 1 head 2, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:B



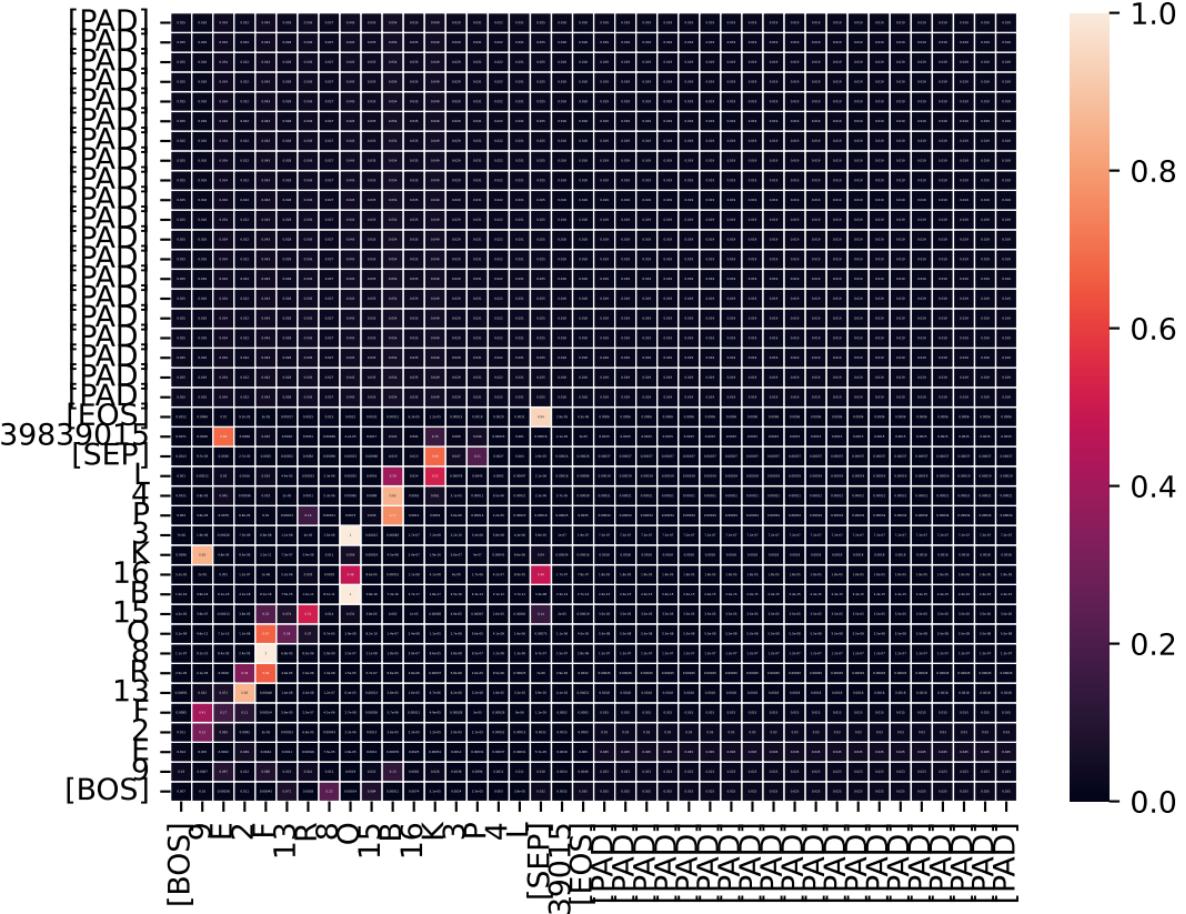
Step 49500 attention weights at layer 1 head 3, with accuracy 0.261
Sequence:[BOS] 18 B 11 Q [SEP] 17.93 [EOS] SOL:8



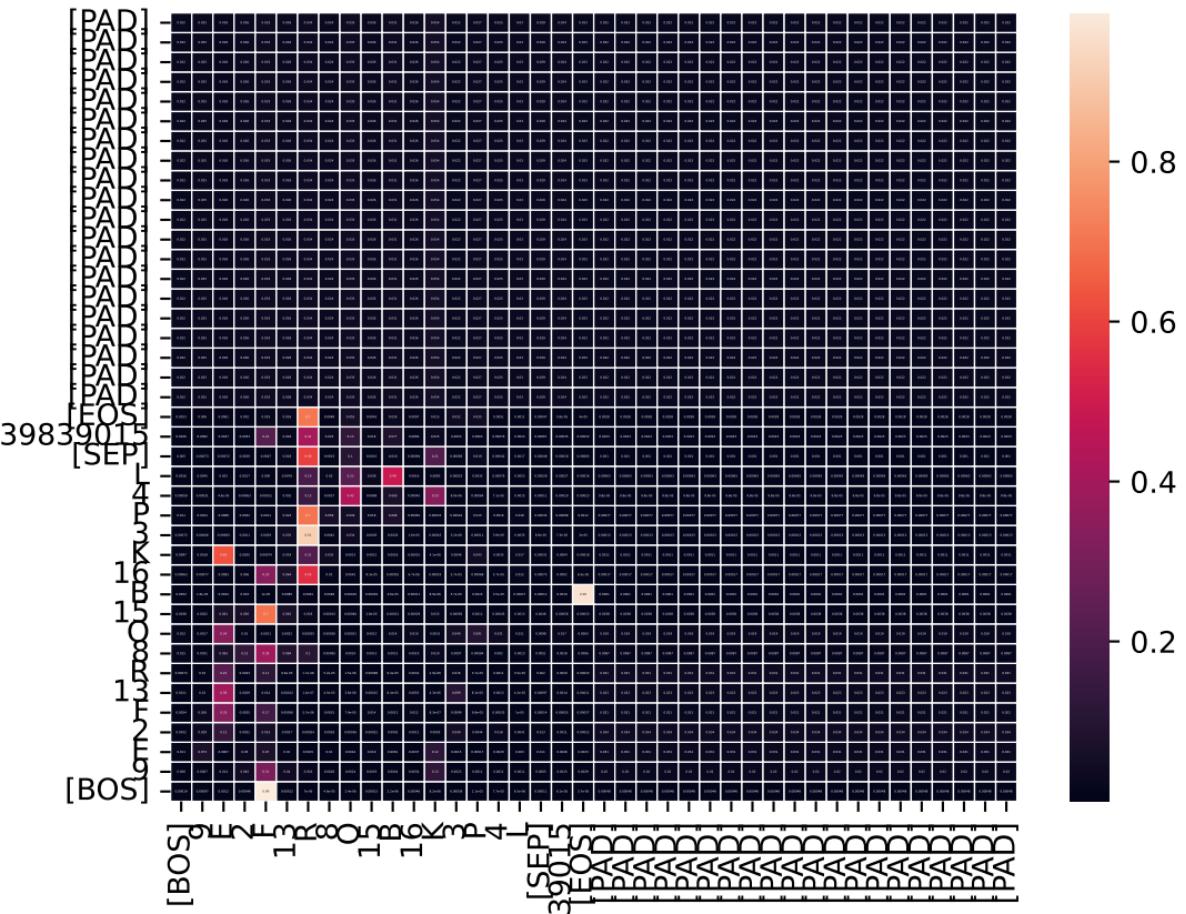
Step 50000 attention weights at layer 0 head 0, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



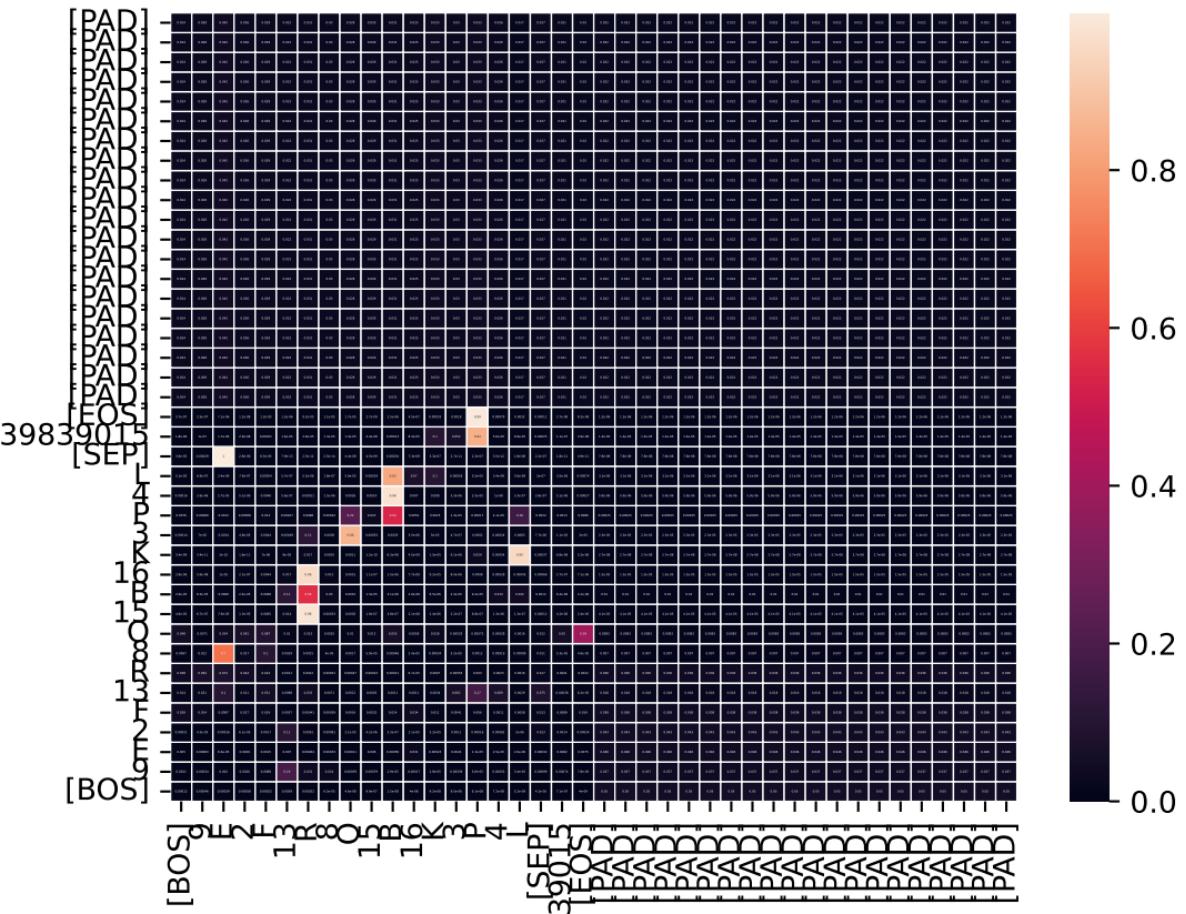
Step 50000 attention weights at layer 0 head 1, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



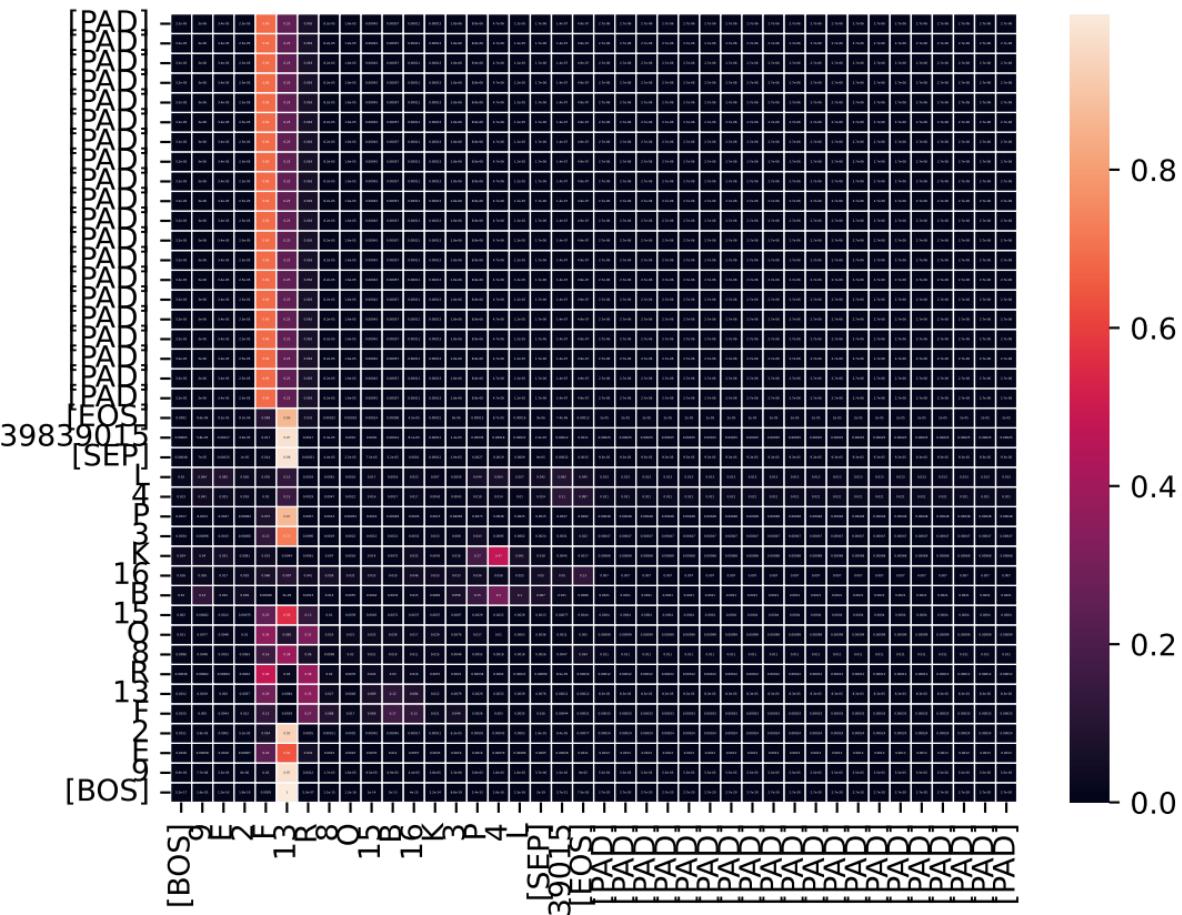
Step 50000 attention weights at layer 0 head 2, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



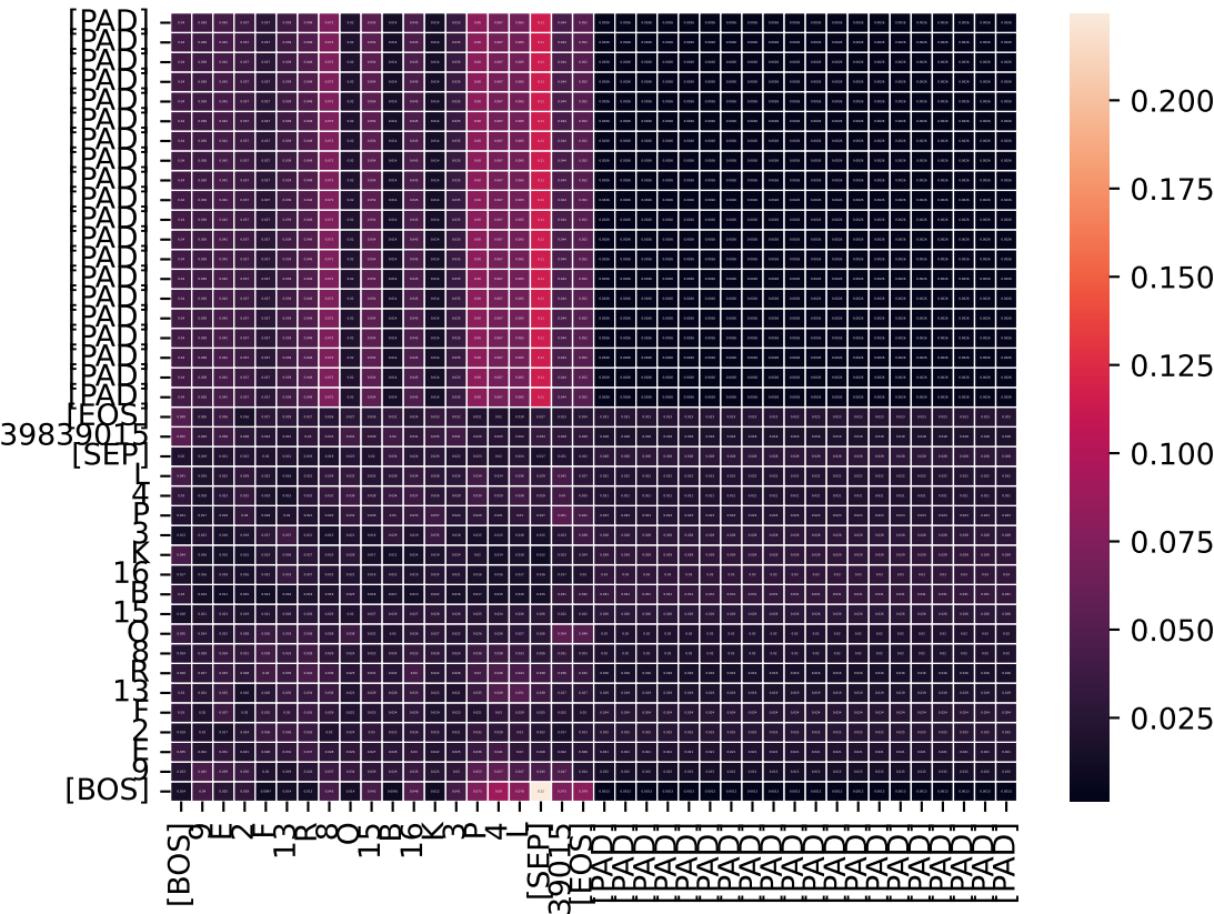
Step 50000 attention weights at layer 0 head 3, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



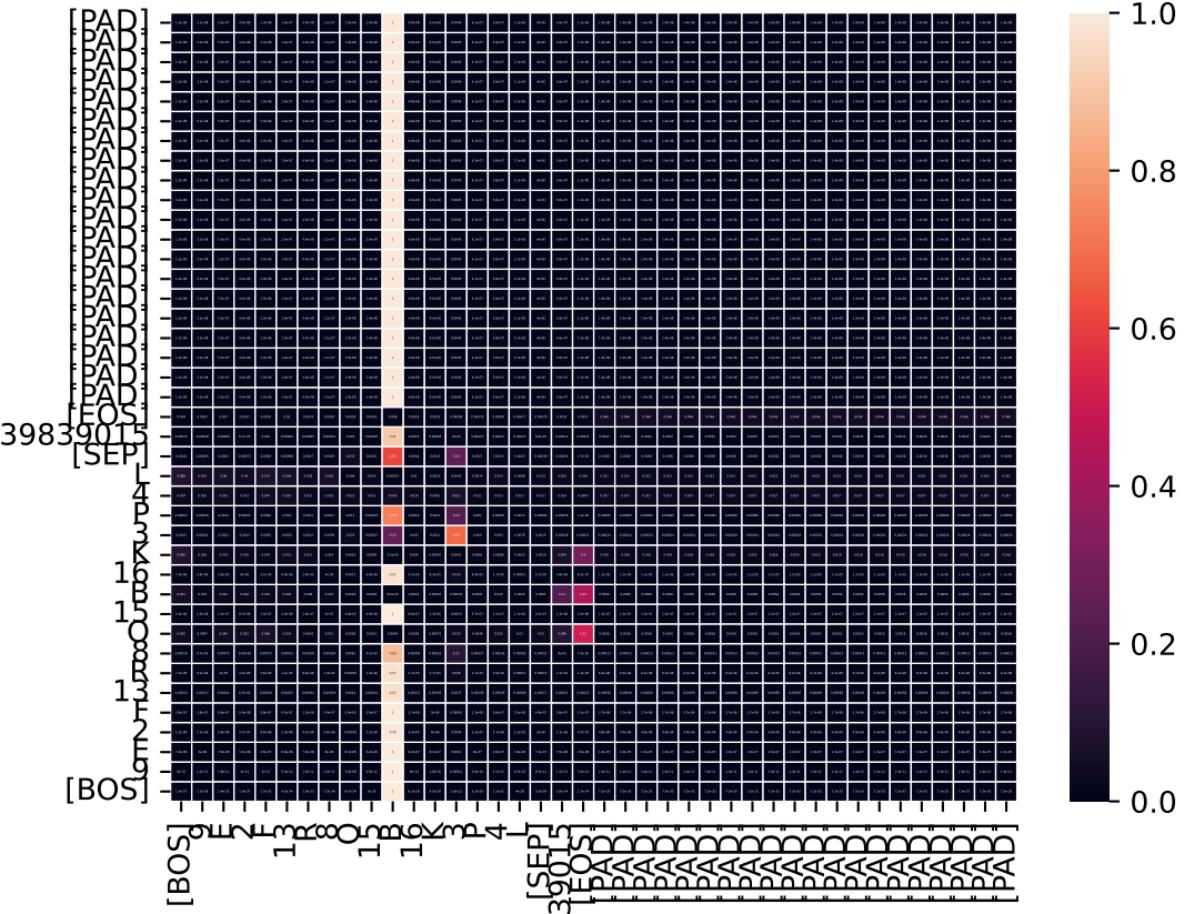
Step 50000 attention weights at layer 1 head 0, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



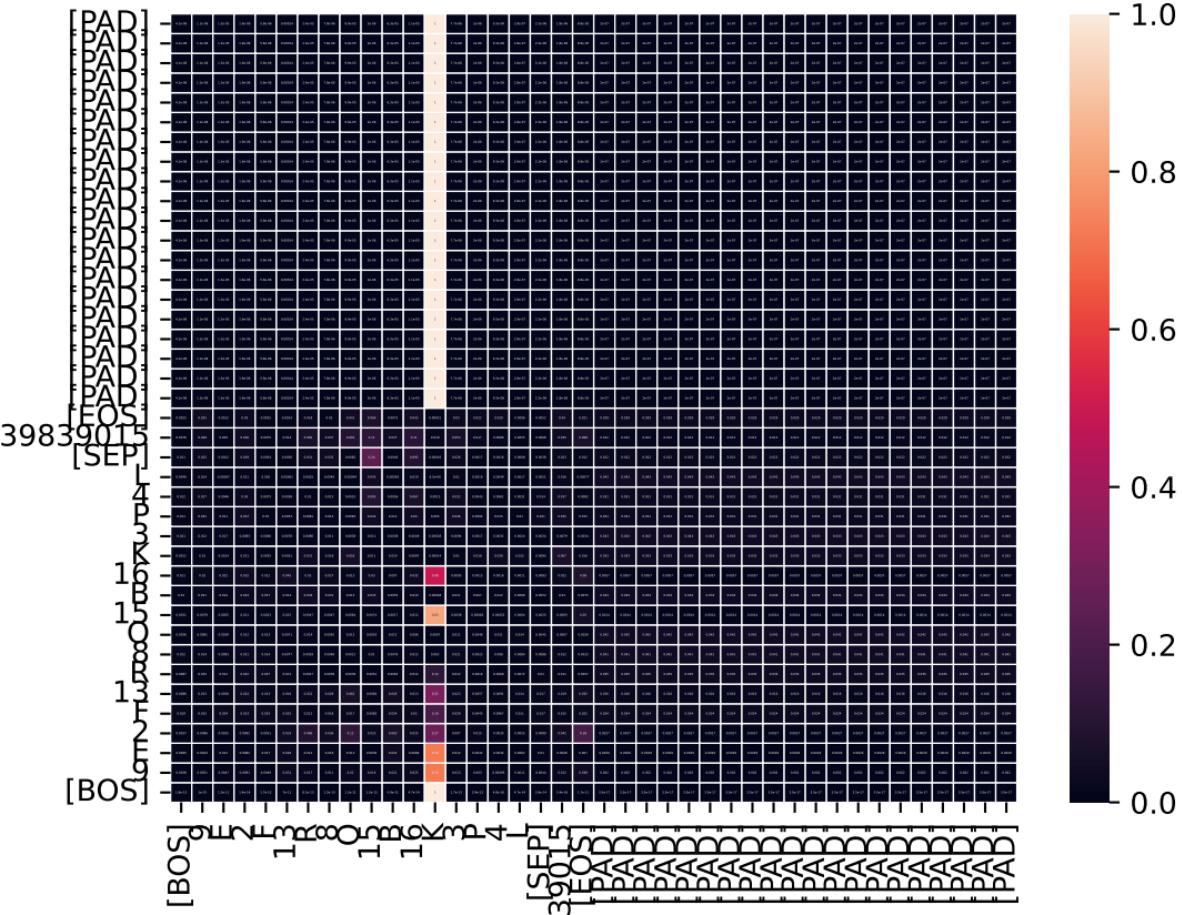
Step 50000 attention weights at layer 1 head 1, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



Step 50000 attention weights at layer 1 head 2, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K

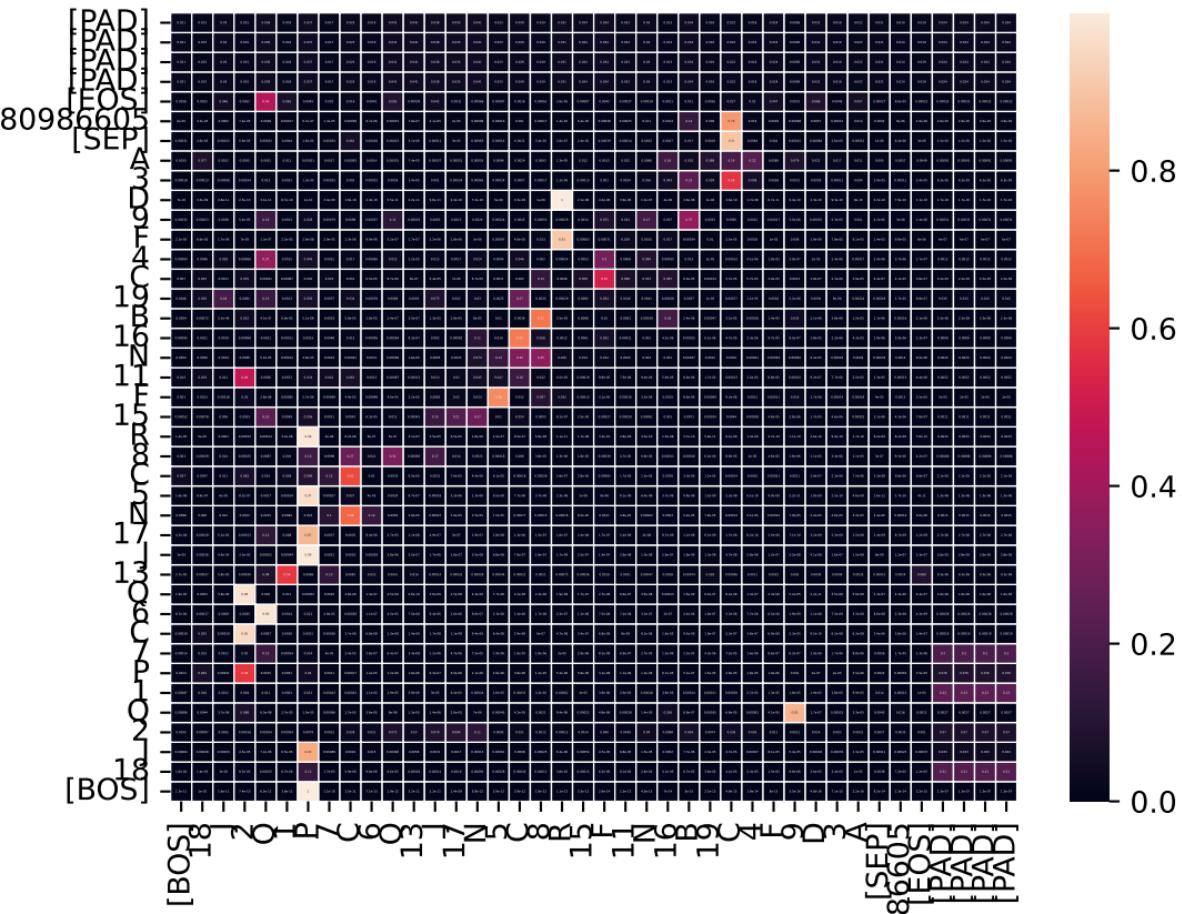


Step 50000 attention weights at layer 1 head 3, with accuracy 0.257
Sequence:[BOS] 9 E 2 F 13 R 8 O 15 B 16 K 3 P 4 L [SEP] 16.06 [EOS] SOL:K



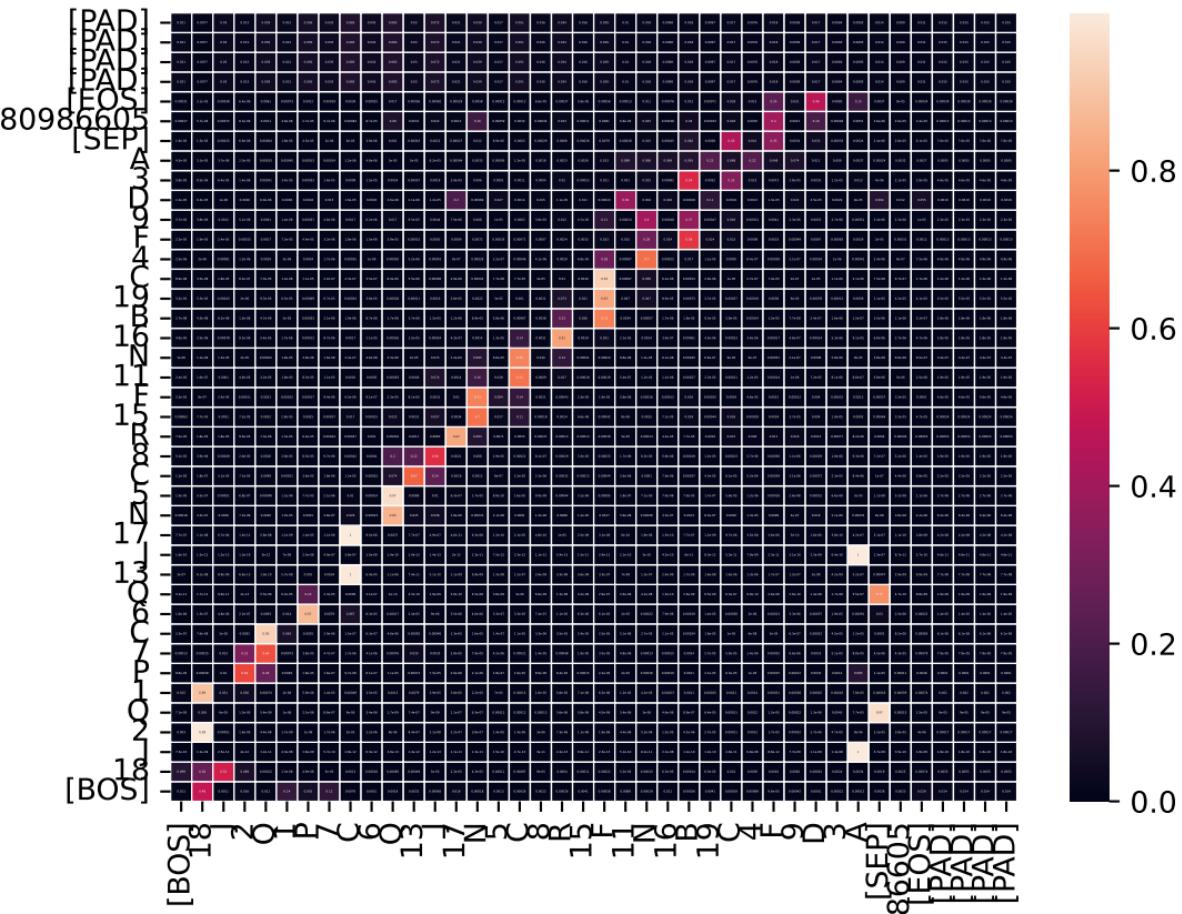
Step 59500 attention weights at layer 0 head 0, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:



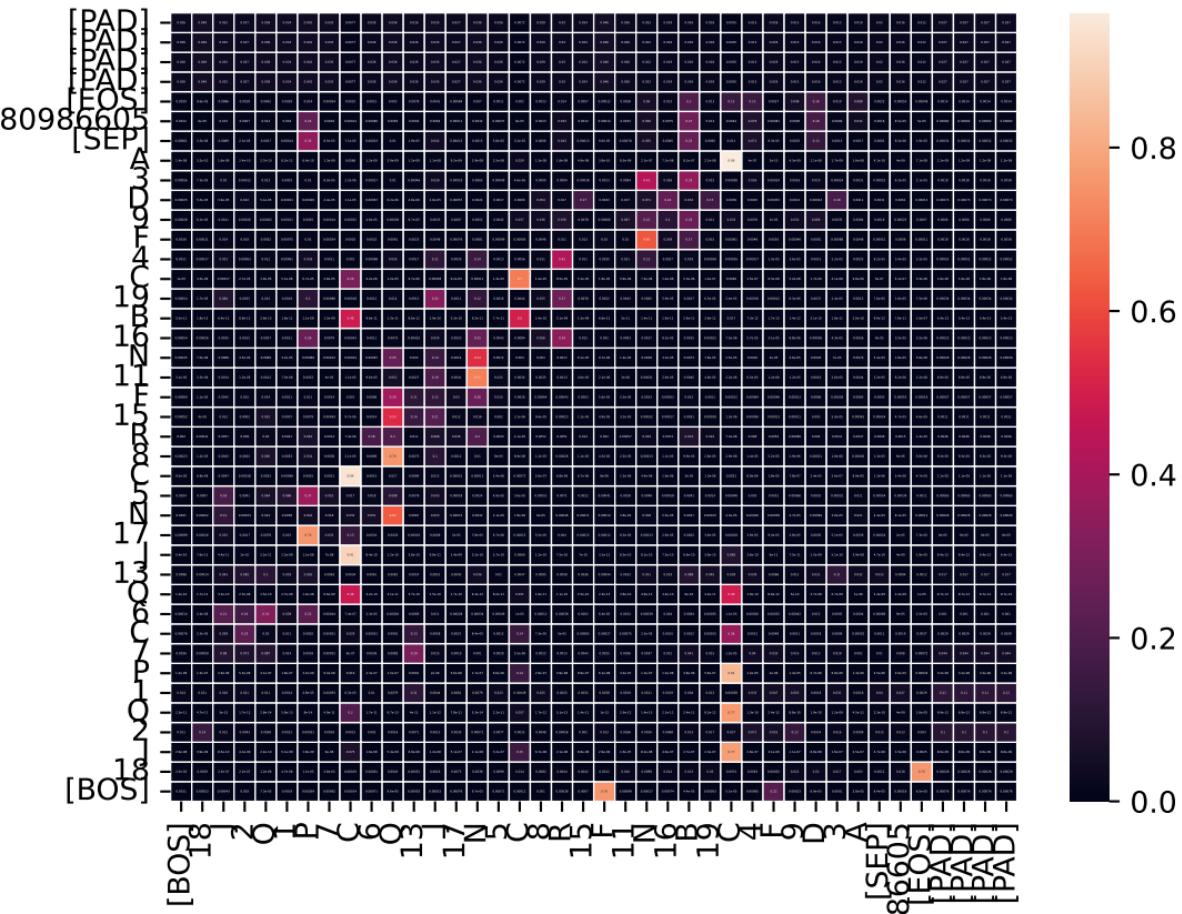
Step 59500 attention weights at layer 0 head 1, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:



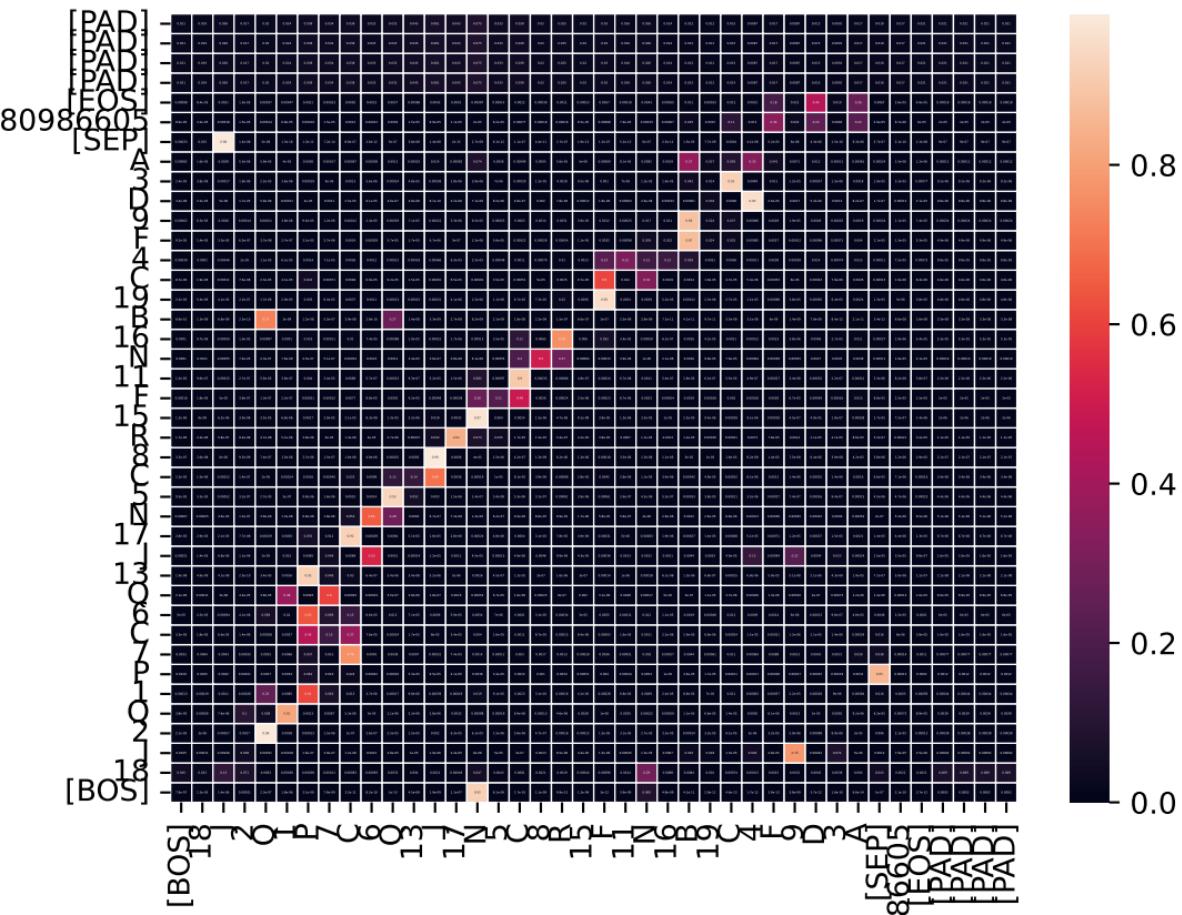
Step 59500 attention weights at layer 0 head 2, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:C



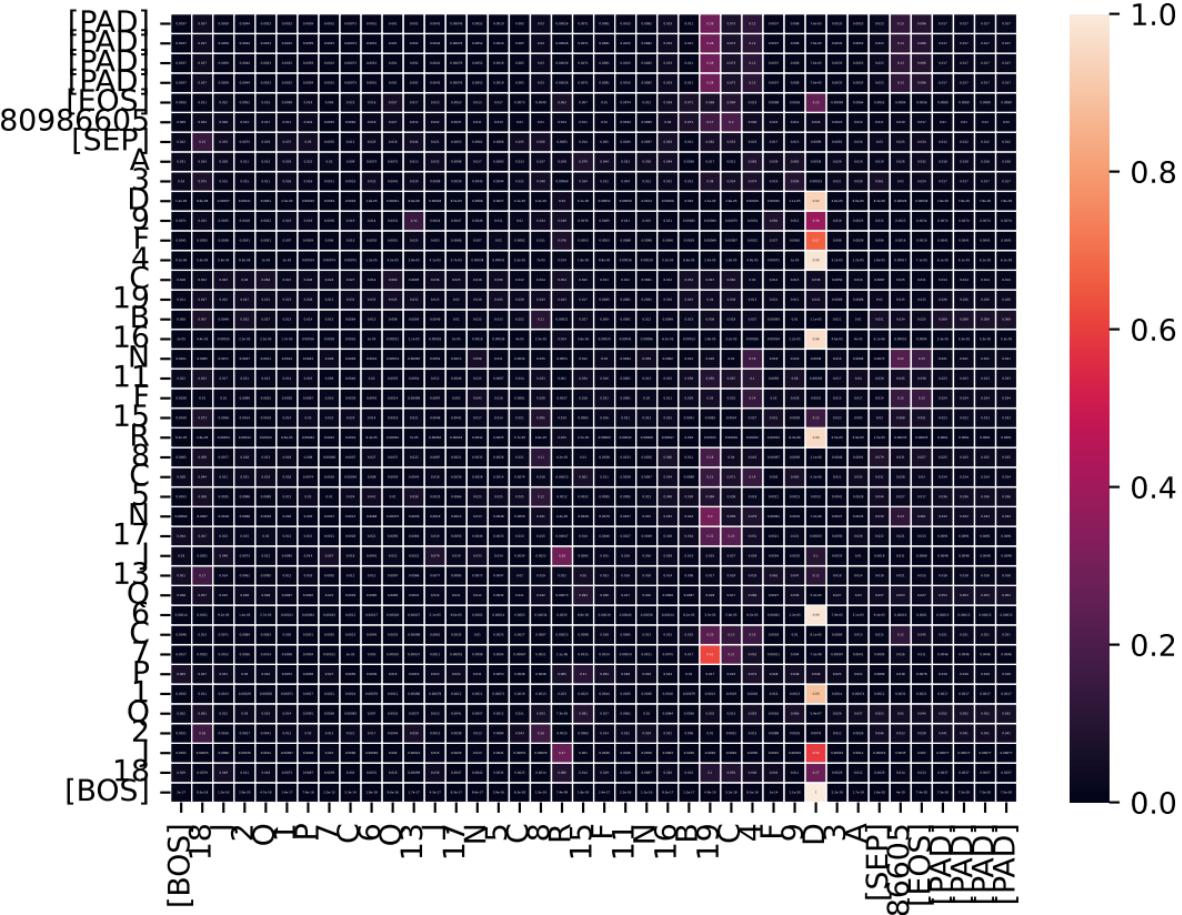
Step 59500 attention weights at layer 0 head 3, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL



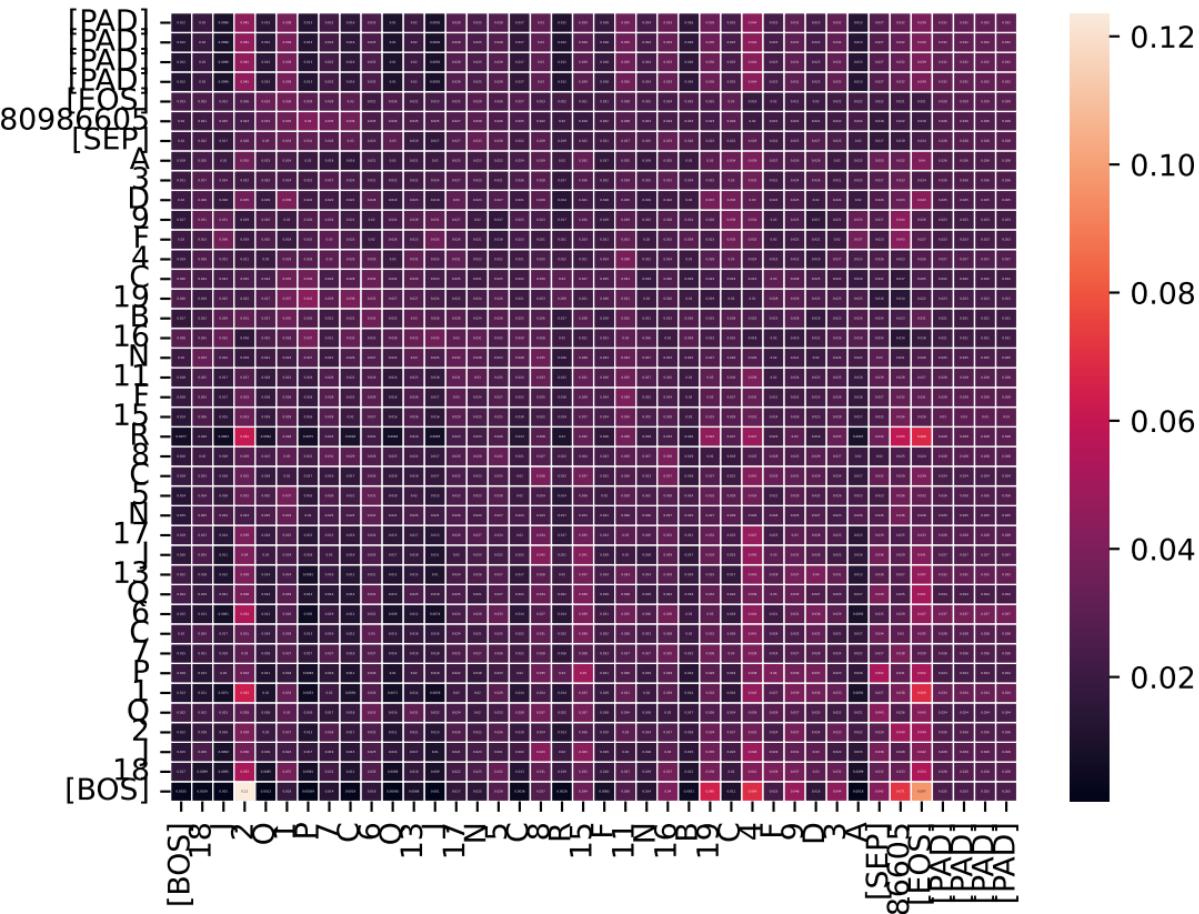
Step 59500 attention weights at layer 1 head 0, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:C



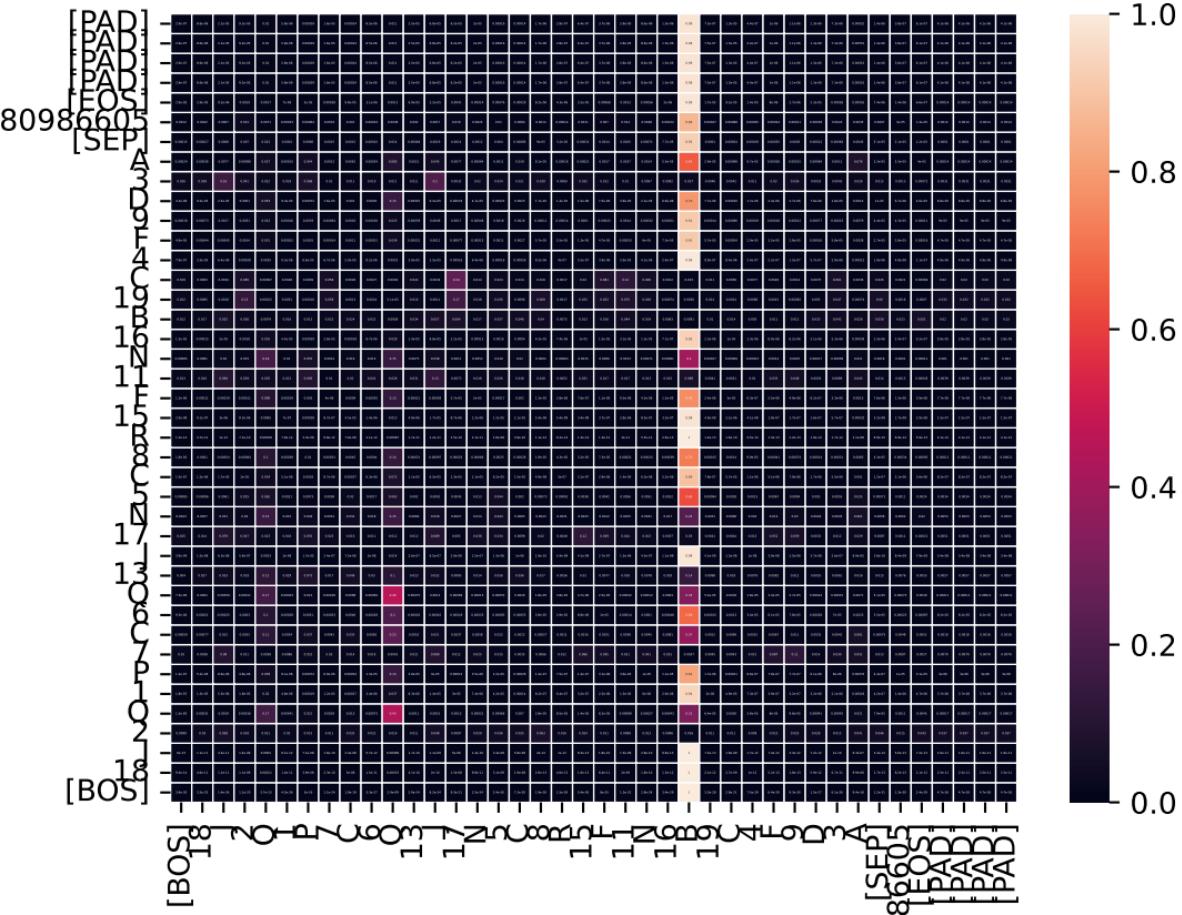
Step 59500 attention weights at layer 1 head 1, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:C



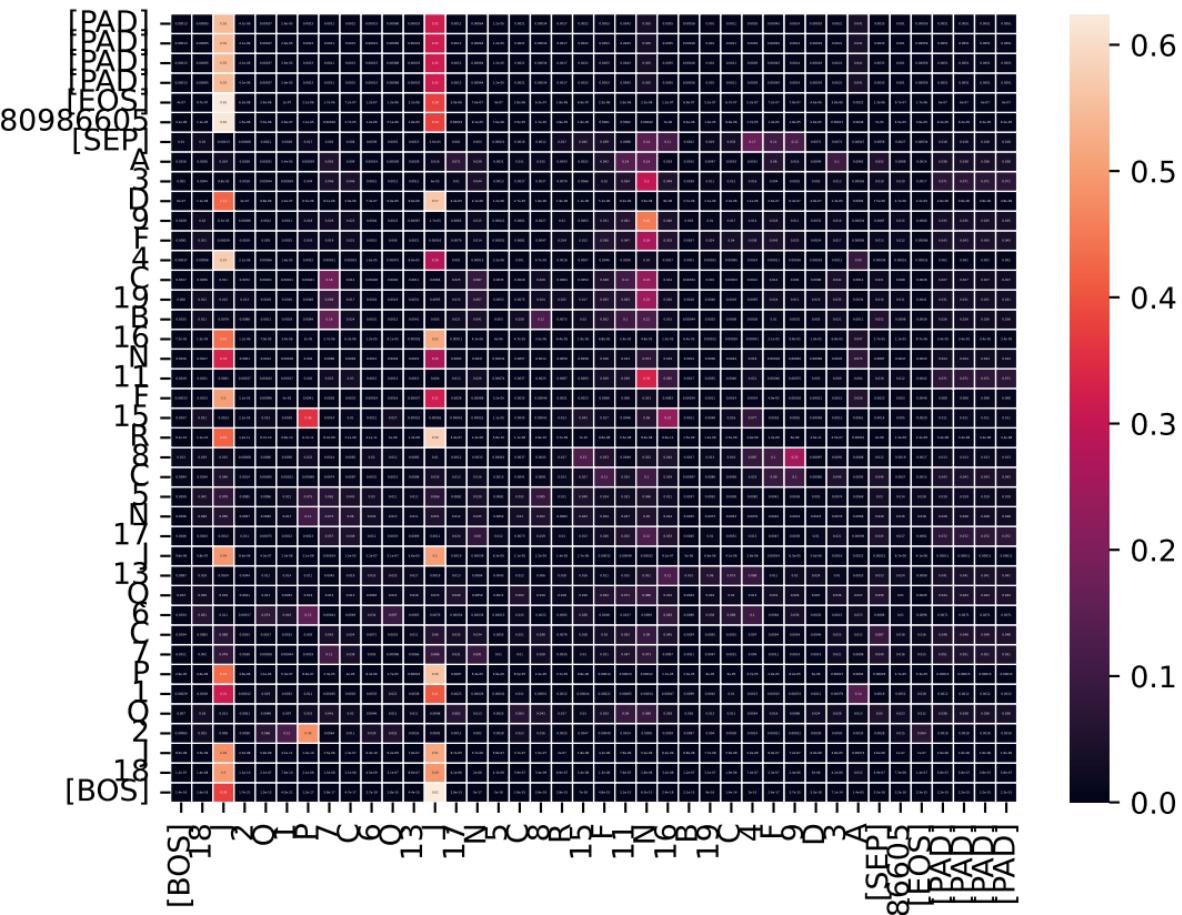
Step 59500 attention weights at layer 1 head 2, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:C



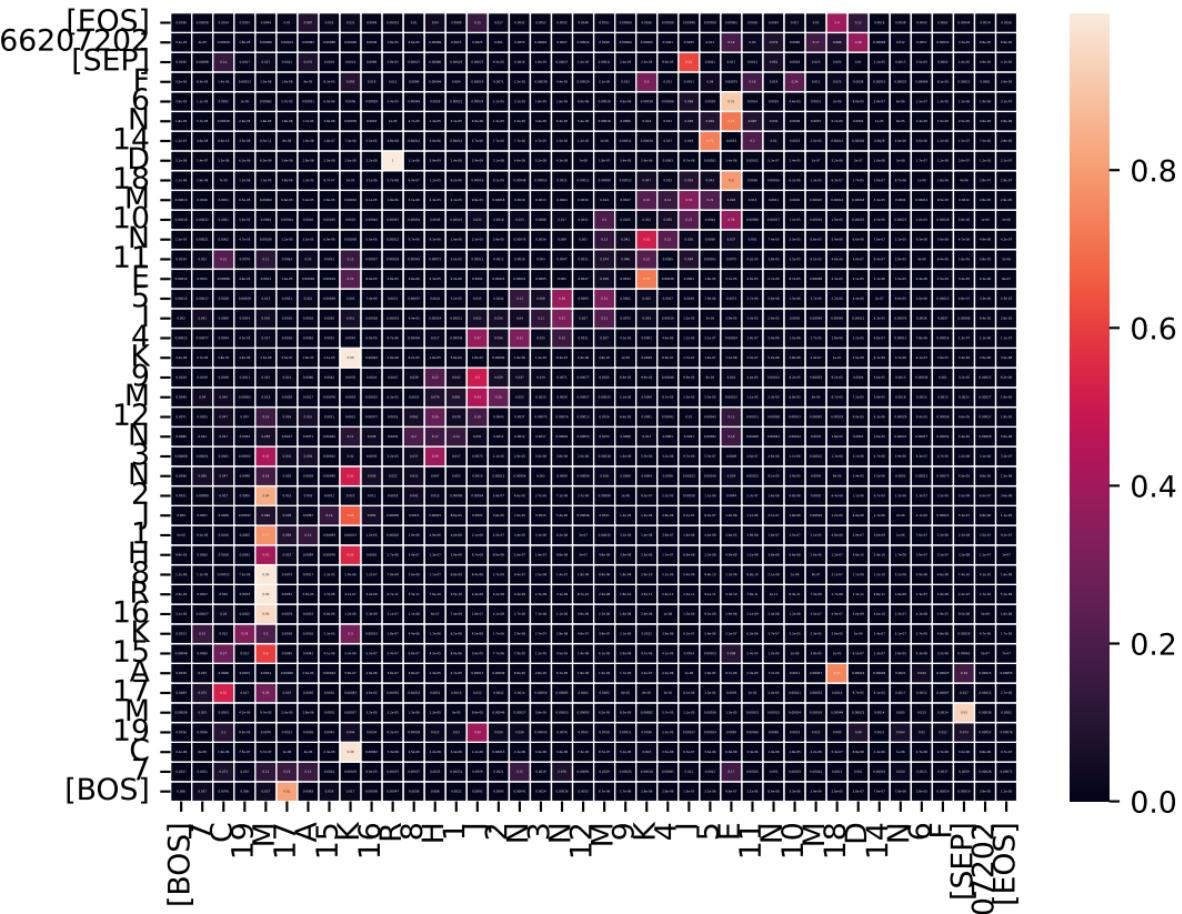
Step 59500 attention weights at layer 1 head 3, with accuracy 0.340

Sequence:[BOS] 18 J 2 Q 1 P 7 C 6 Q 13 J 17 N 5 C 8 R 15 F 11 N 16 B 19 C 4 F 9 D 3 A [SEP] 7.08 [EOS] SOL:C



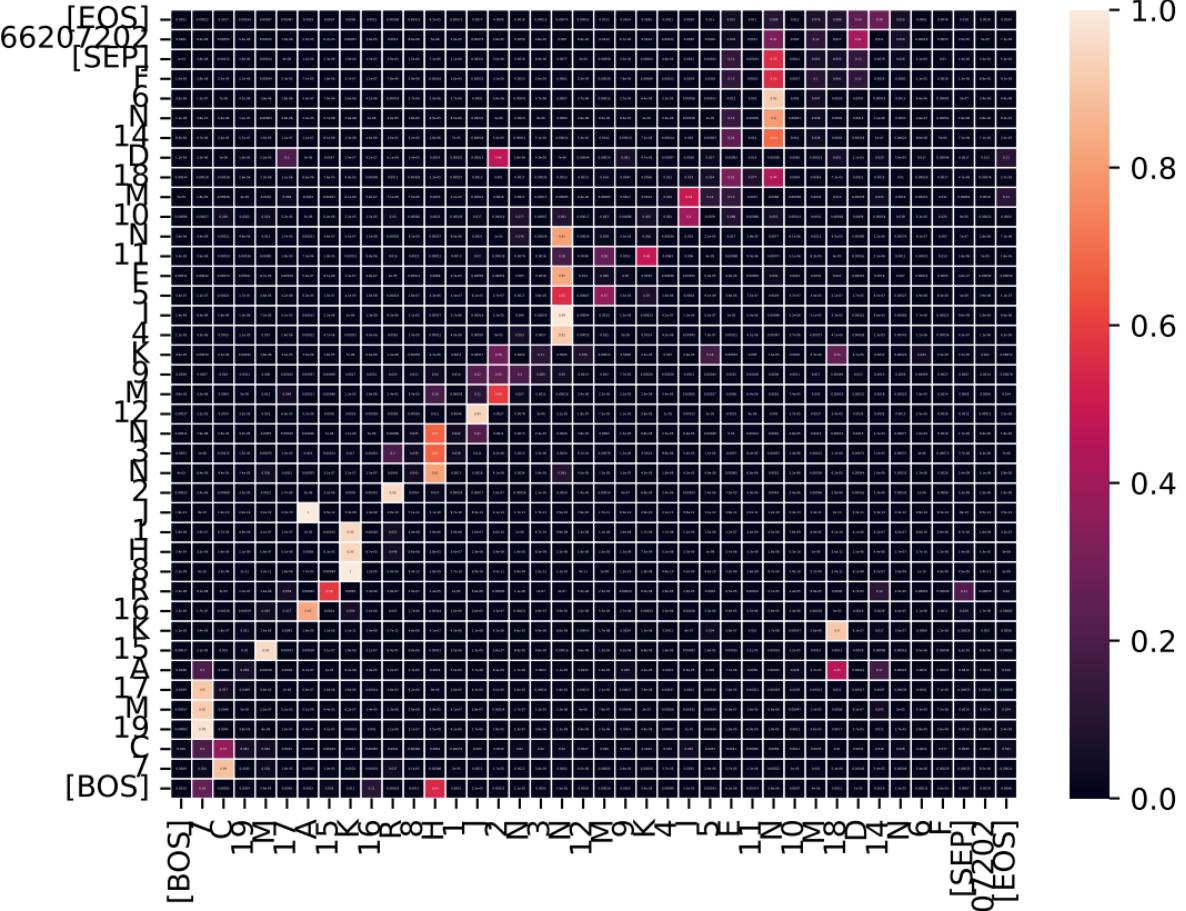
Step 60000 attention weights at layer 0 head 0, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



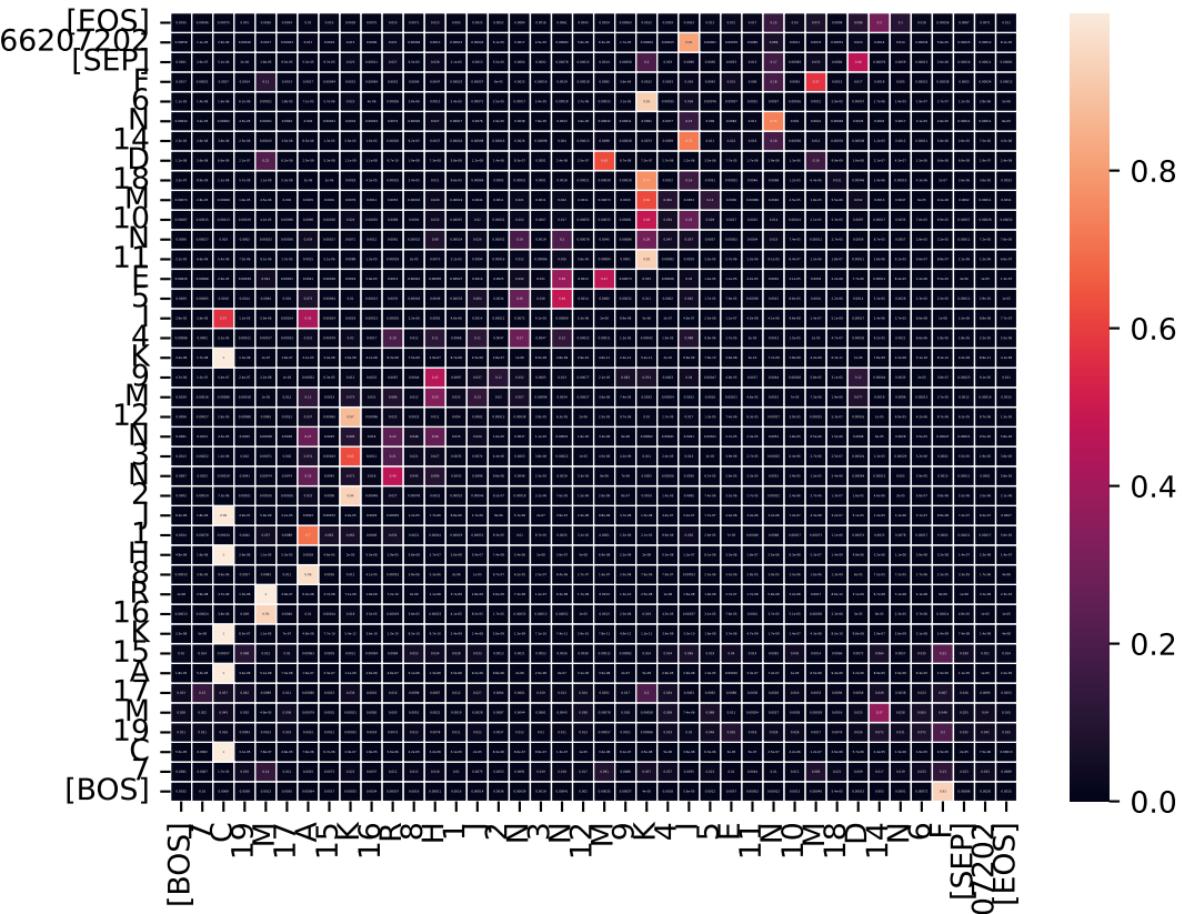
Step 60000 attention weights at layer 0 head 1, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



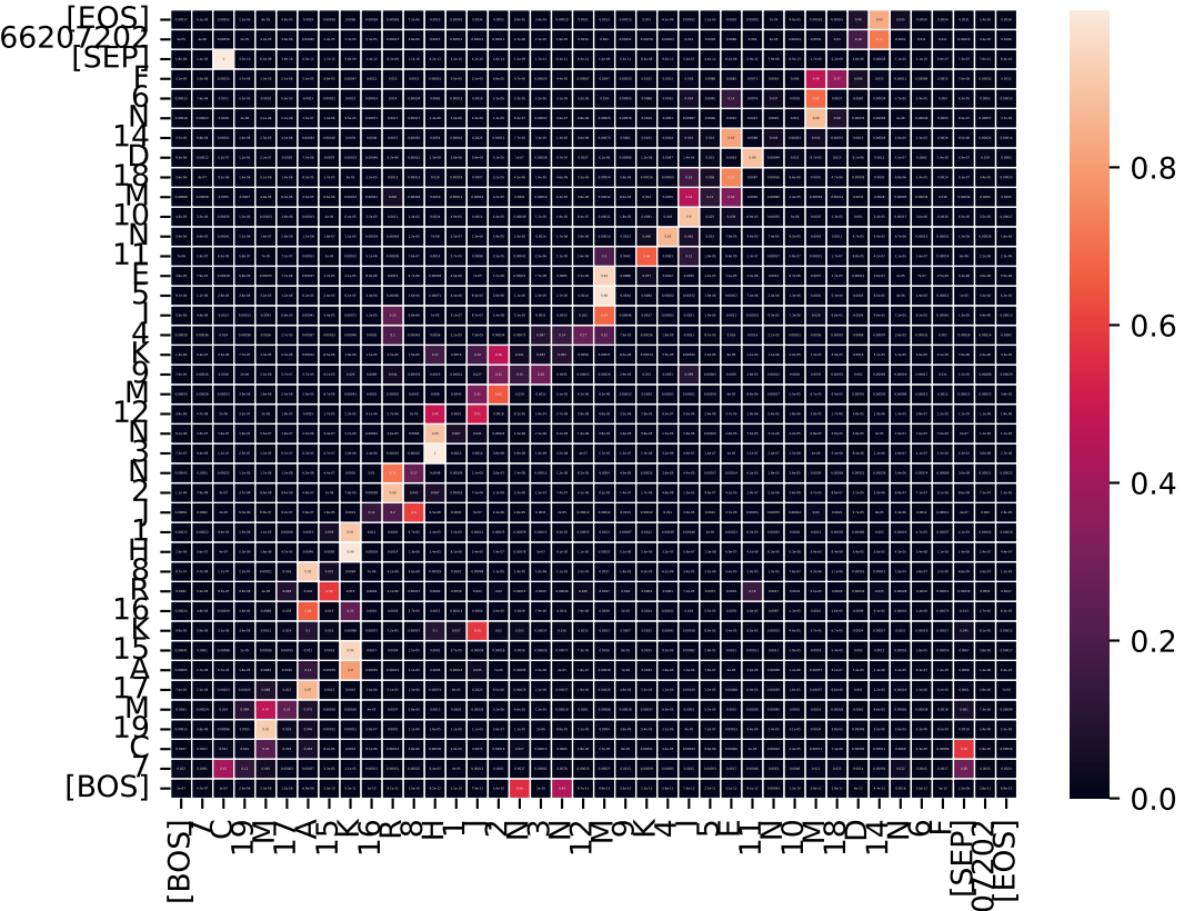
Step 60000 attention weights at layer 0 head 2, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



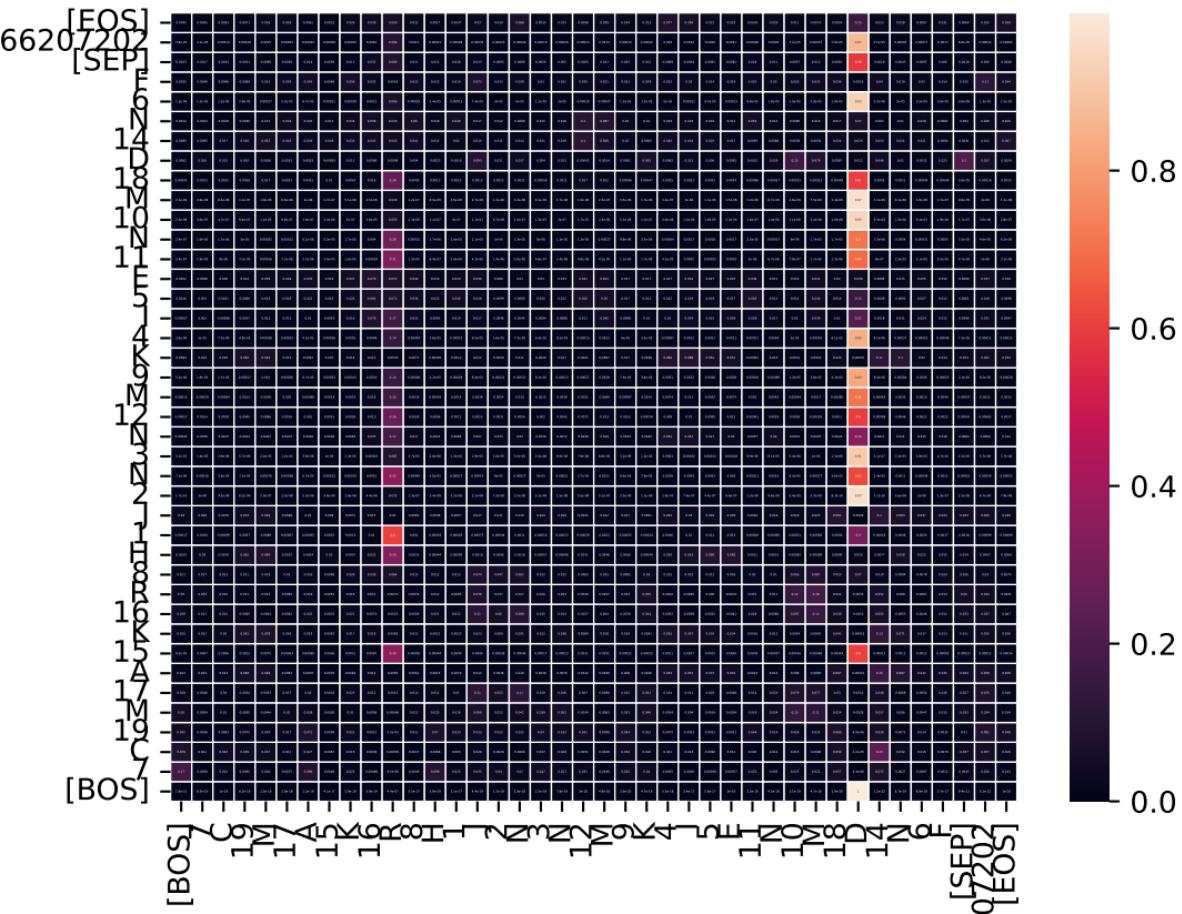
Step 60000 attention weights at layer 0 head 3, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



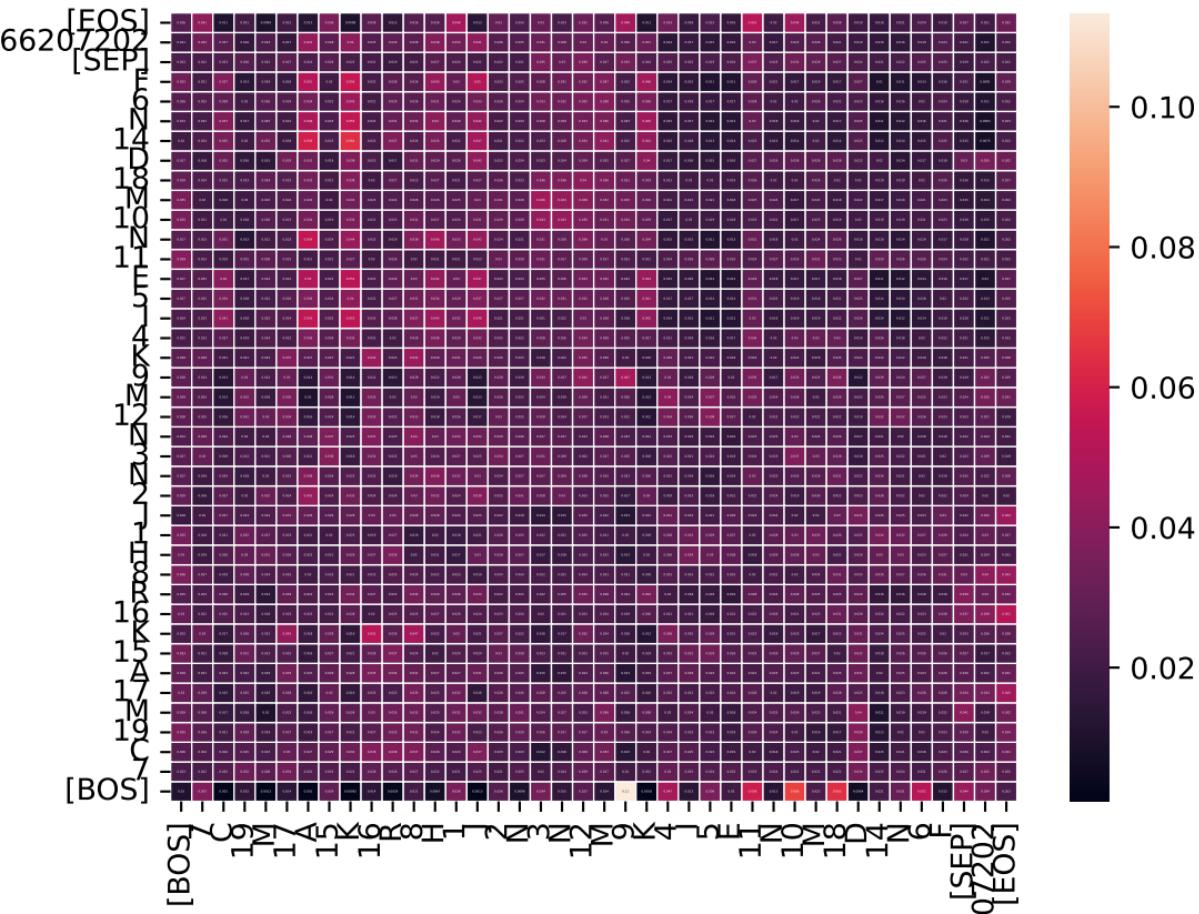
Step 60000 attention weights at layer 1 head 0, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



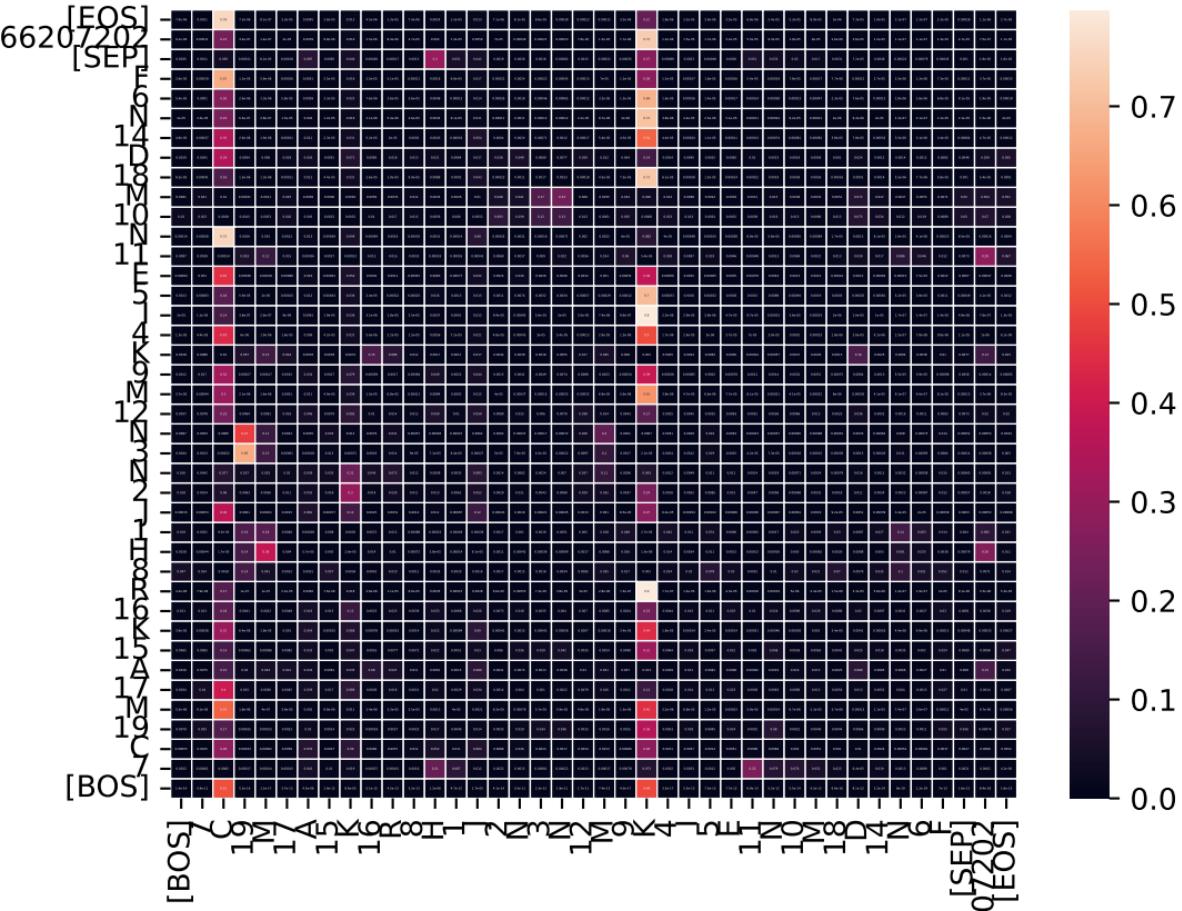
Step 60000 attention weights at layer 1 head 1, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



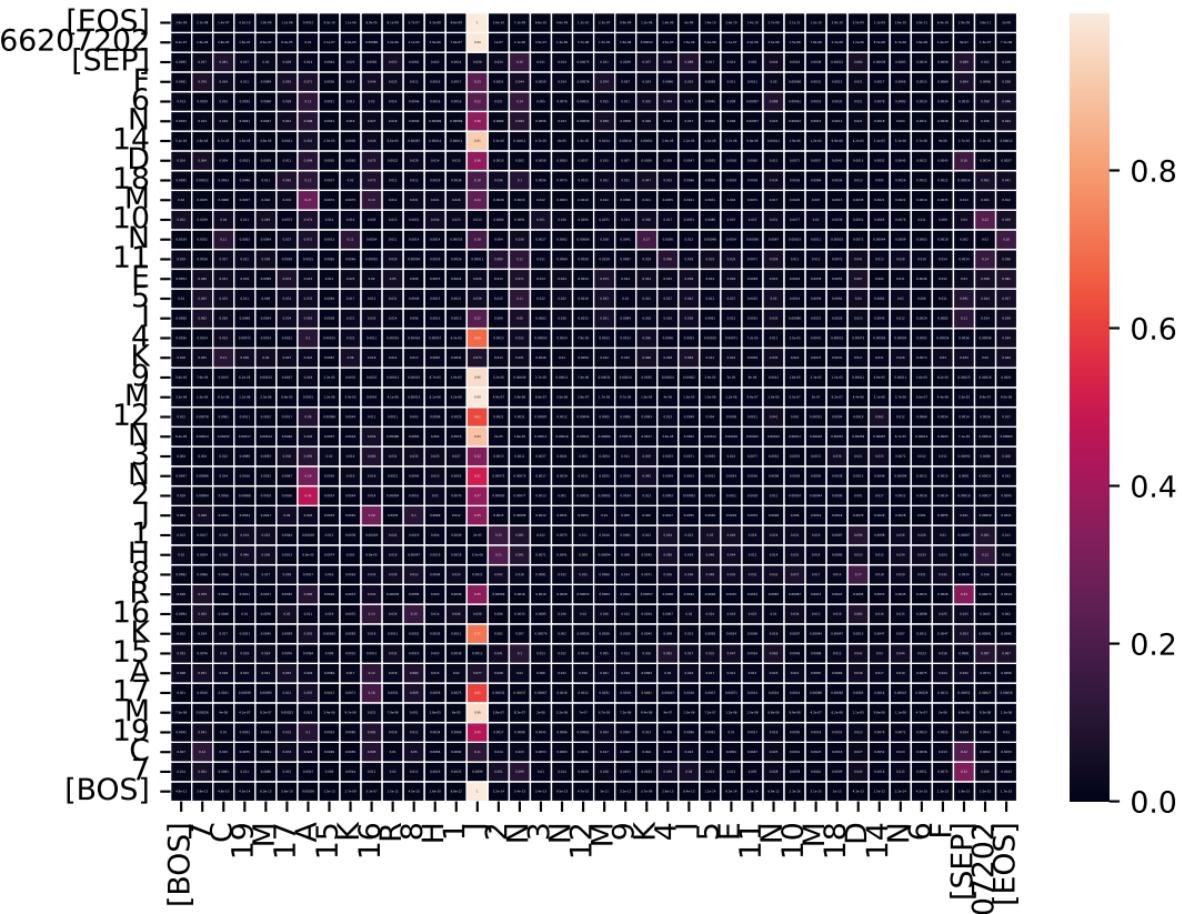
Step 60000 attention weights at layer 1 head 2, with accuracy 0.365

Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H

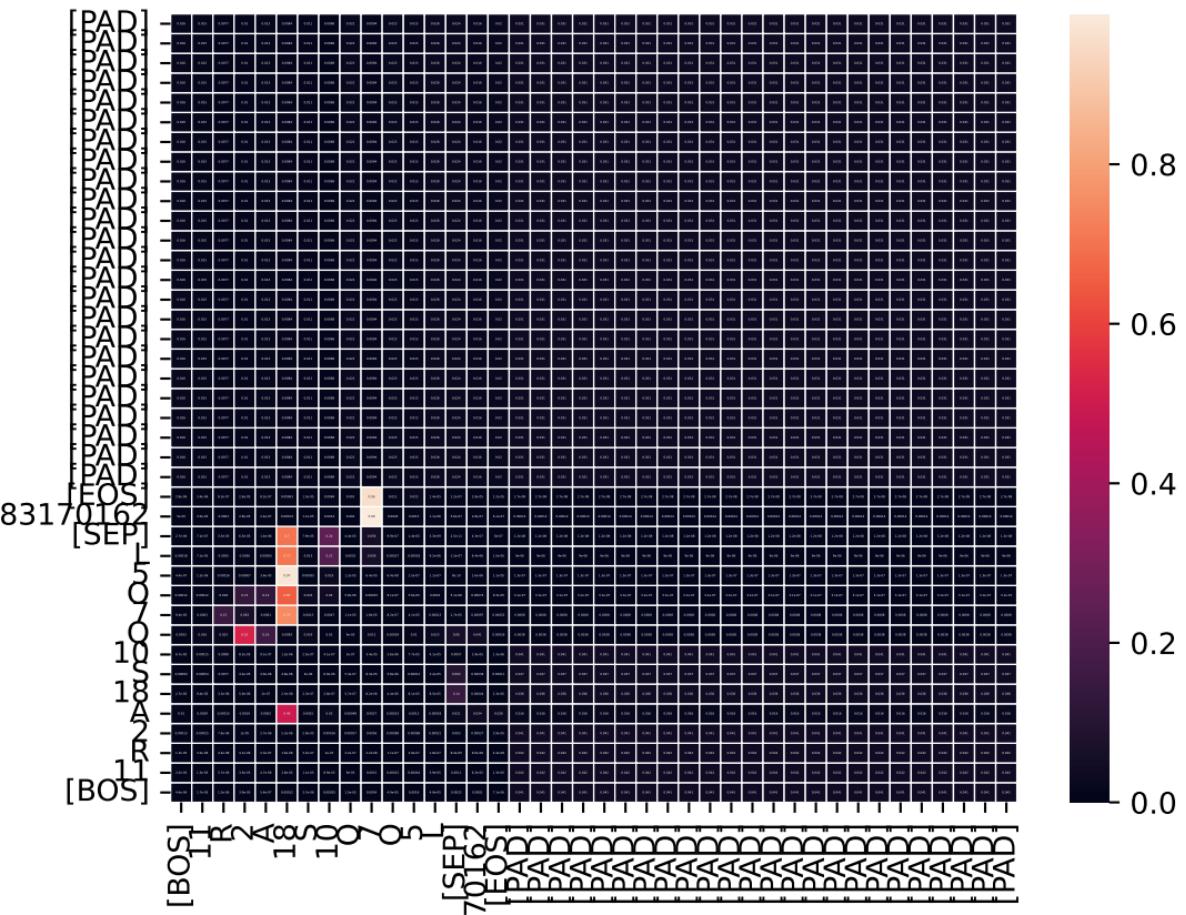


Step 60000 attention weights at layer 1 head 3, with accuracy 0.365

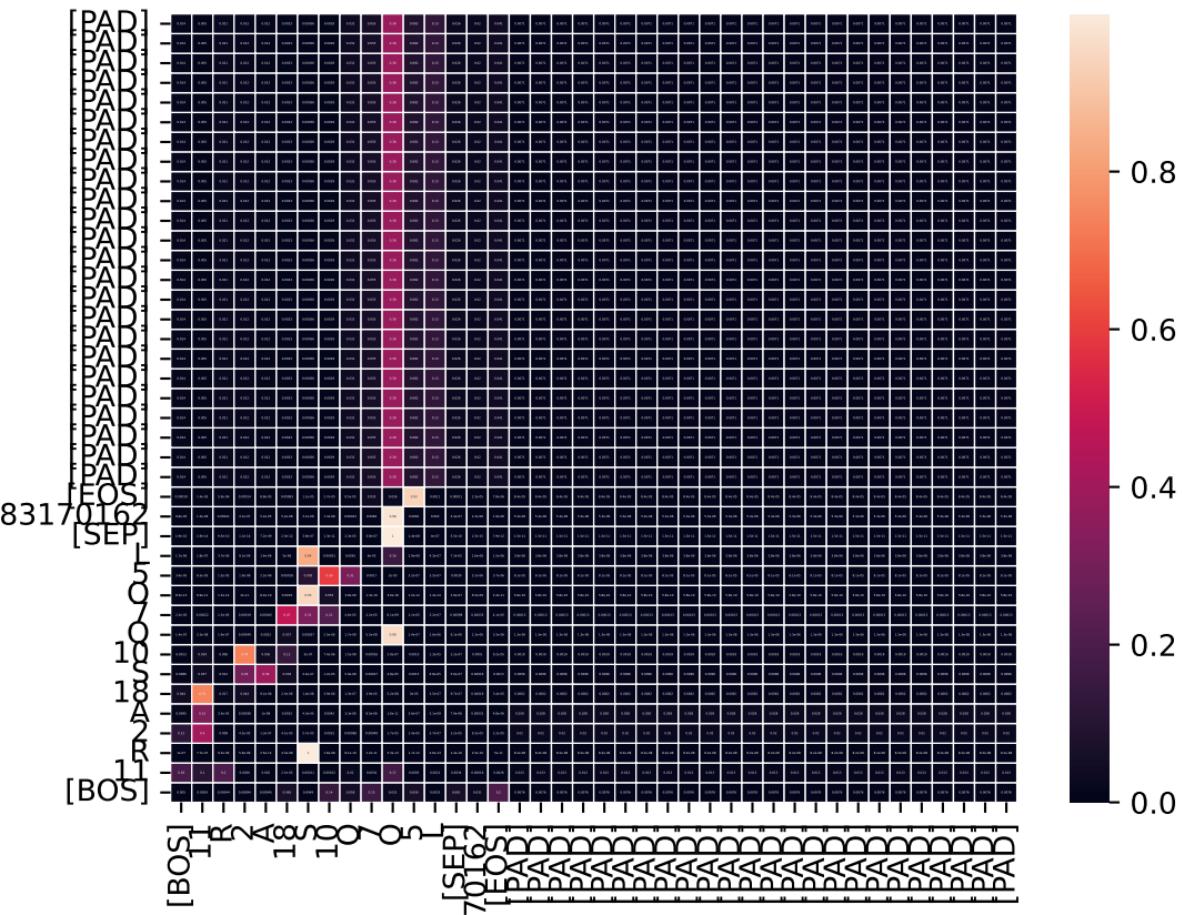
Sequence:[BOS] 7 C 19 M 17 A 15 K 16 R 8 H 1 J 2 N 3 N 12 M 9 K 4 I 5 E 11 N 10 M 18 D 14 N 6 F [SEP] 8.08 [EOS] SOL:H



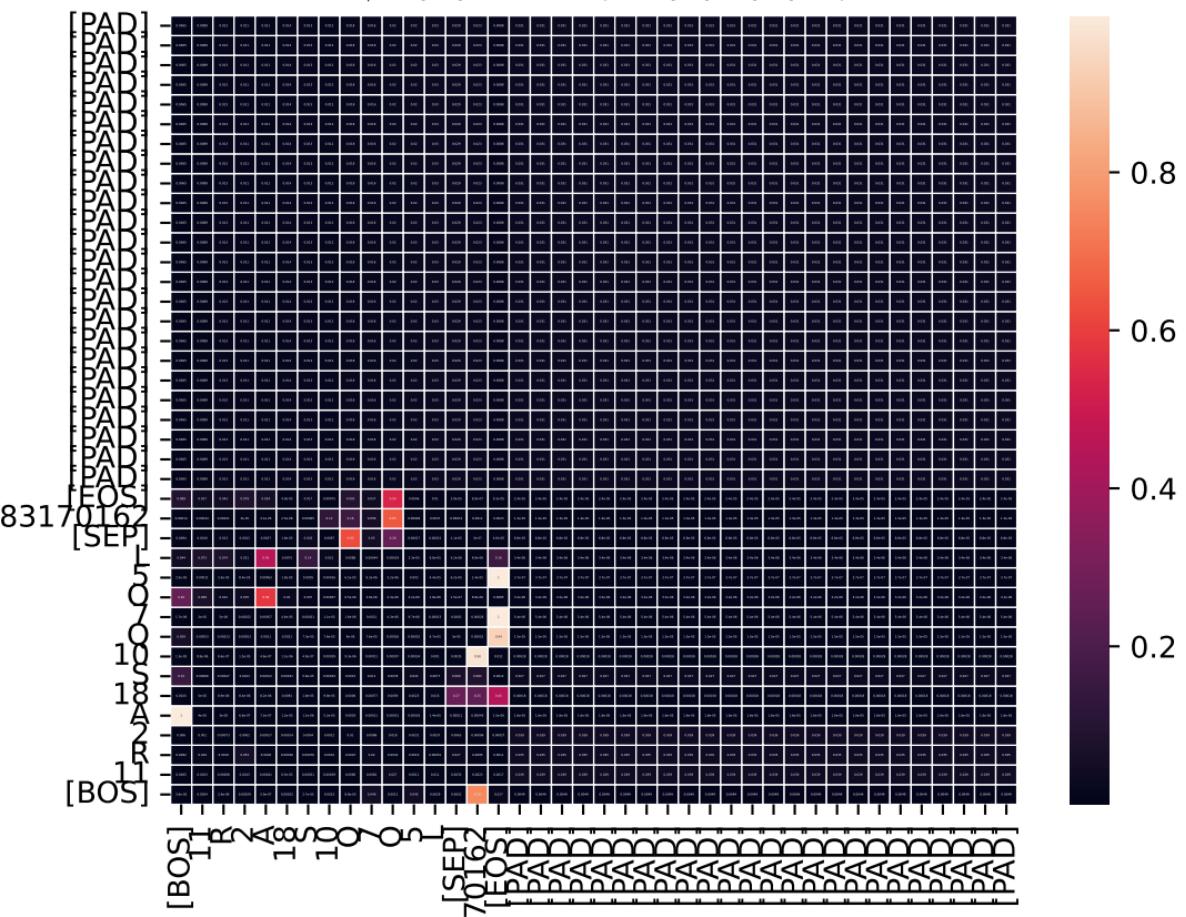
Step 68500 attention weights at layer 0 head 0, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



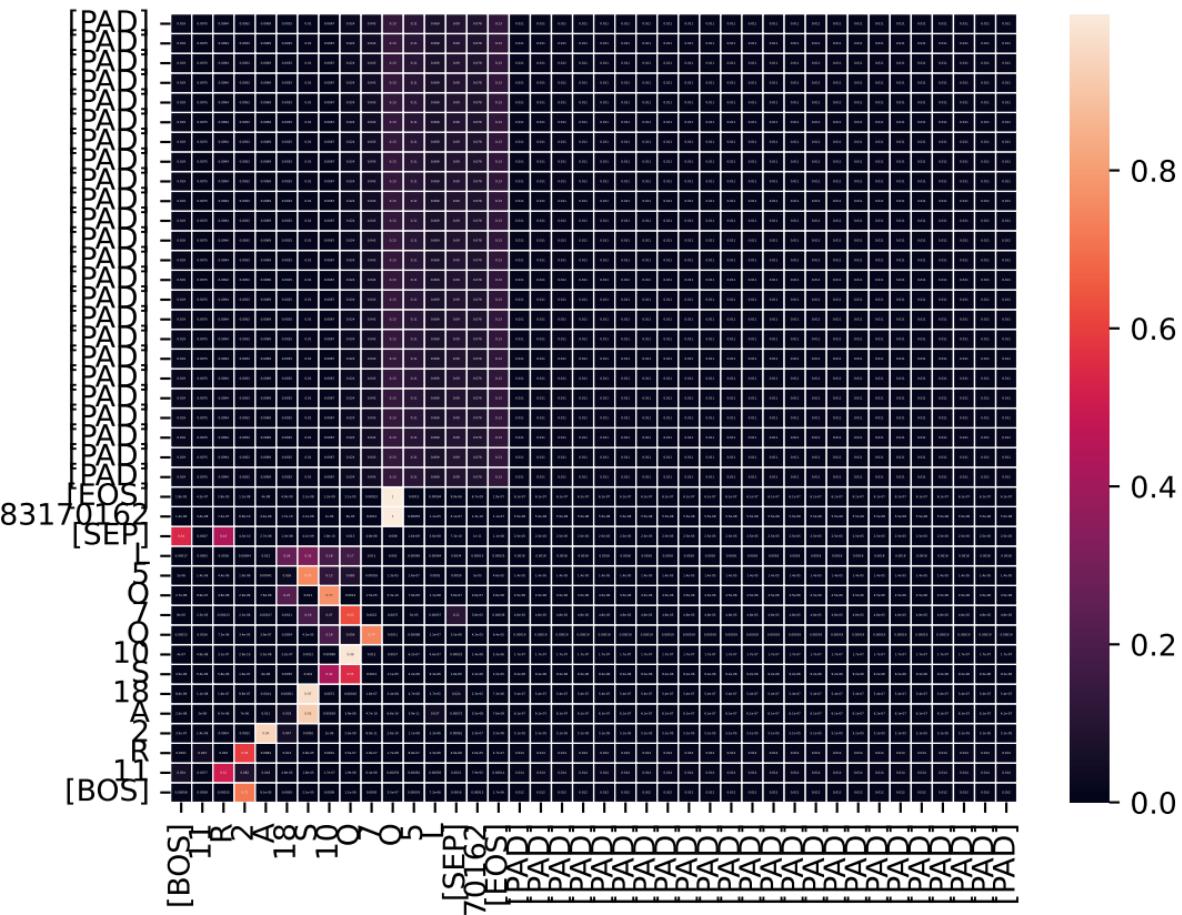
Step 68500 attention weights at layer 0 head 1, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



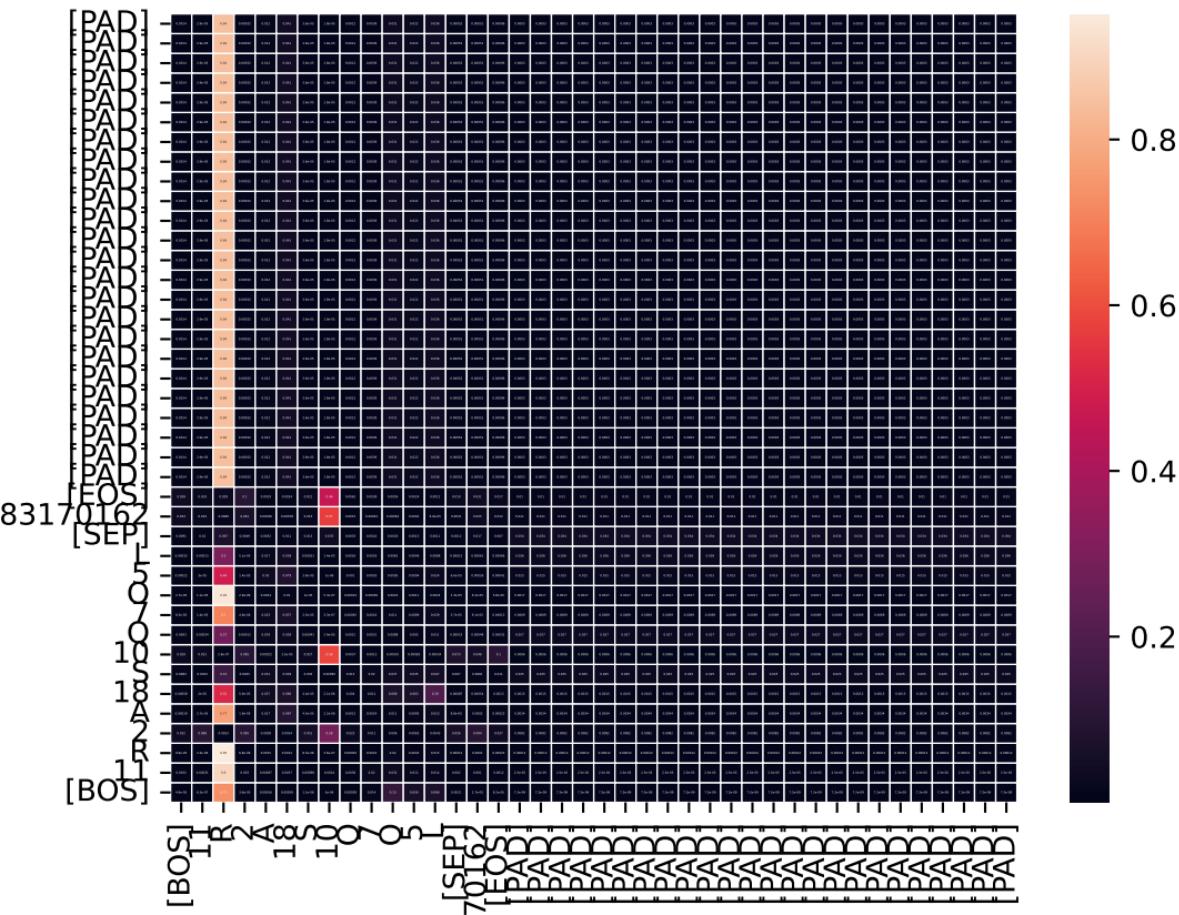
Step 68500 attention weights at layer 0 head 2, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



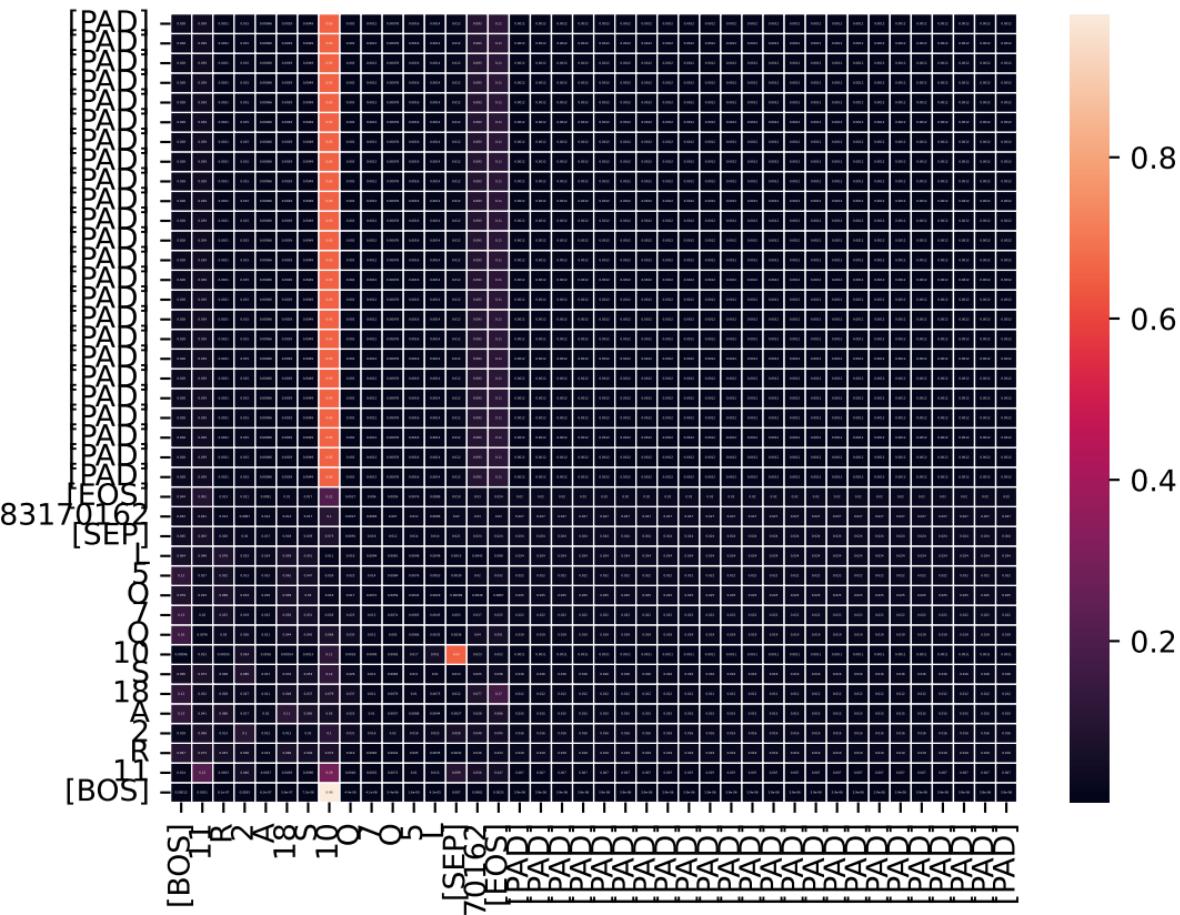
Step 68500 attention weights at layer 0 head 3, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



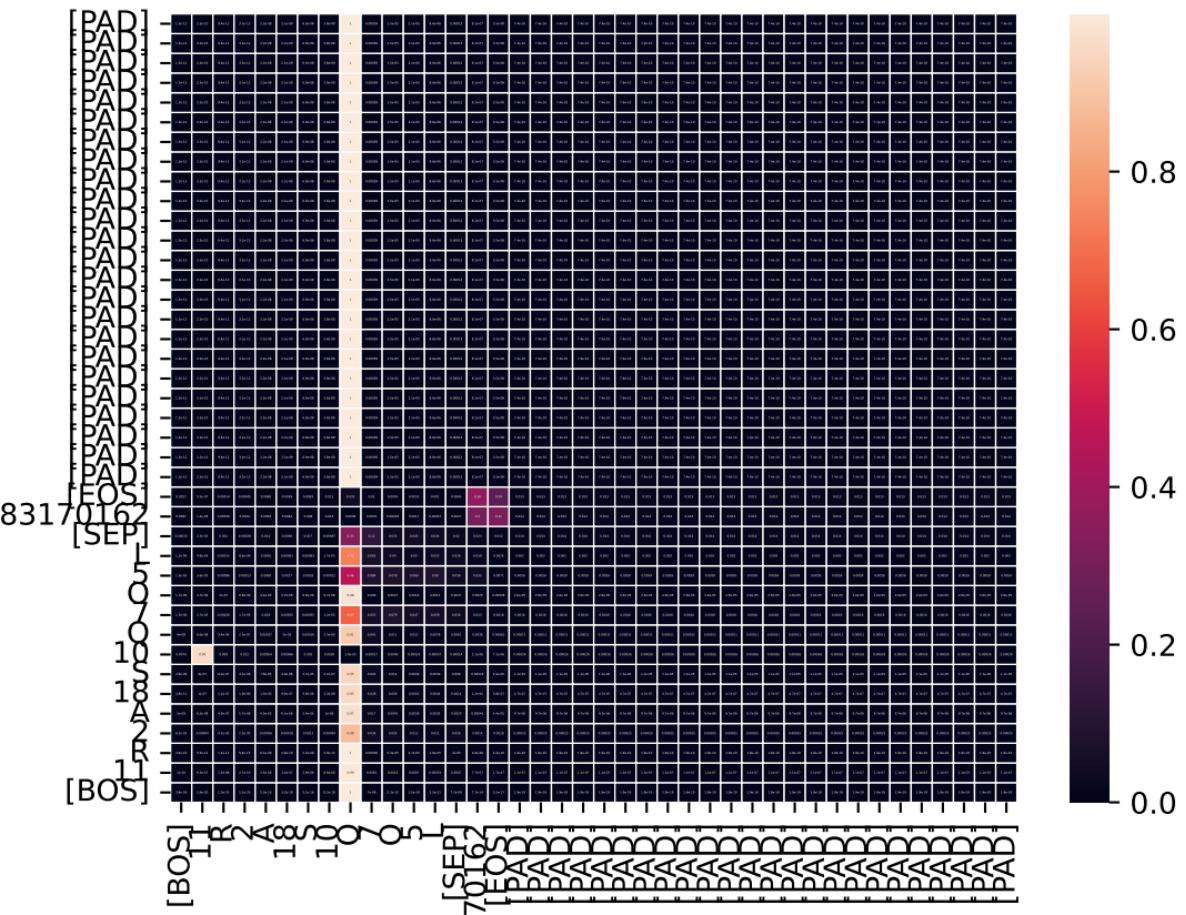
Step 68500 attention weights at layer 1 head 0, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



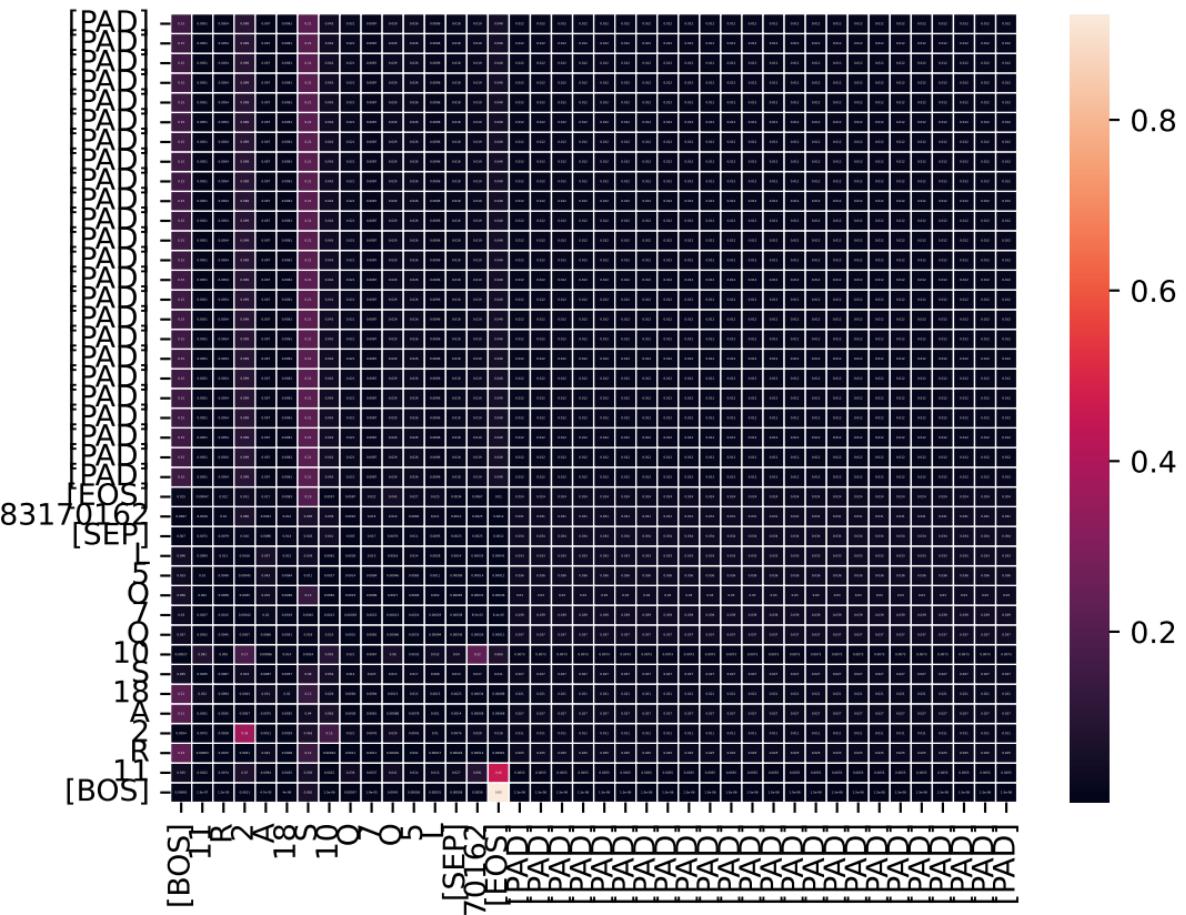
Step 68500 attention weights at layer 1 head 1, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



Step 68500 attention weights at layer 1 head 2, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q

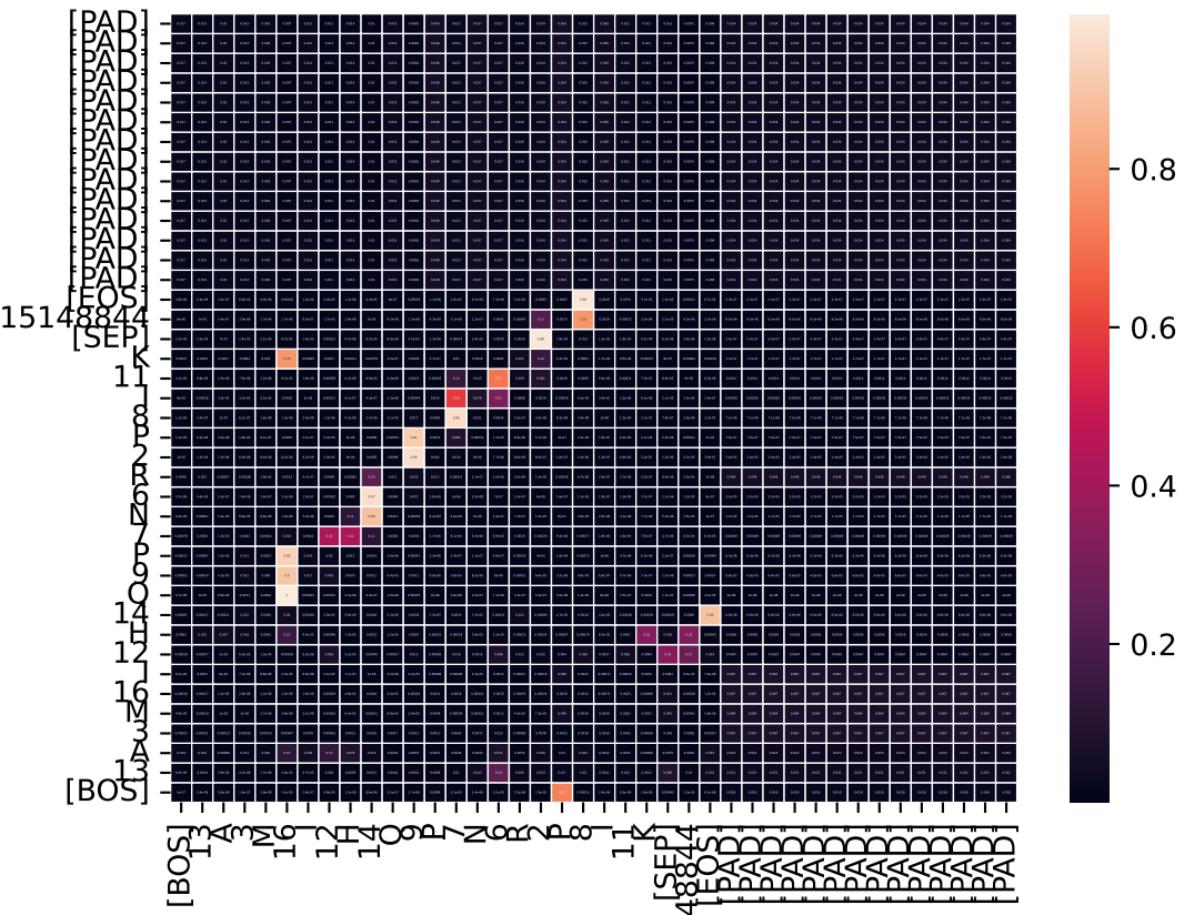


Step 68500 attention weights at layer 1 head 3, with accuracy 0.854
Sequence:[BOS] 11 R 2 A 18 S 10 Q 7 O 5 L [SEP] 9.91 [EOS] SOL:Q



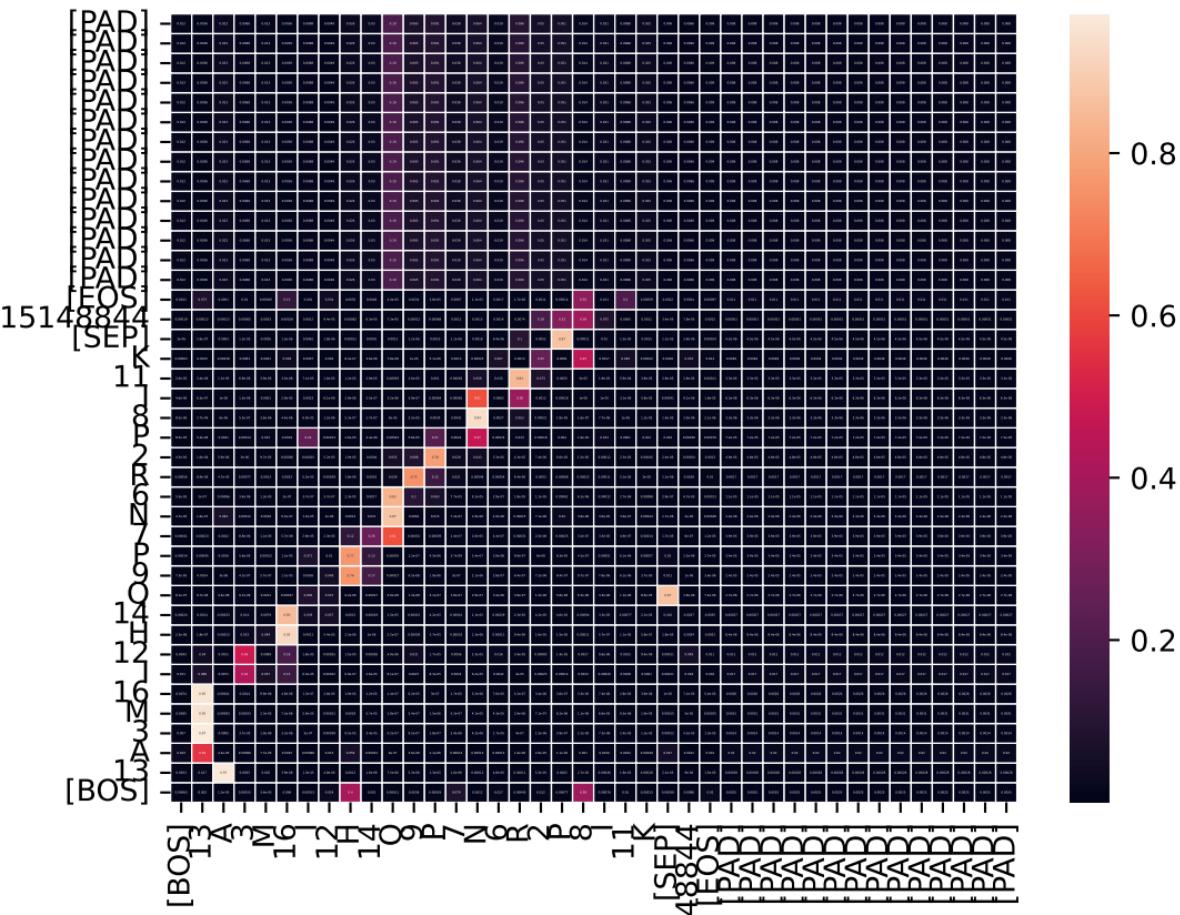
Step 69000 attention weights at layer 0 head 0, with accuracy 0.956

Sequence:[BOS] 13 A 3 M 16! 12 H 14 Q 9 P 7 N 6 R 2 P 8! 11 K [SEP] 16.06 [EOS] SOL:1



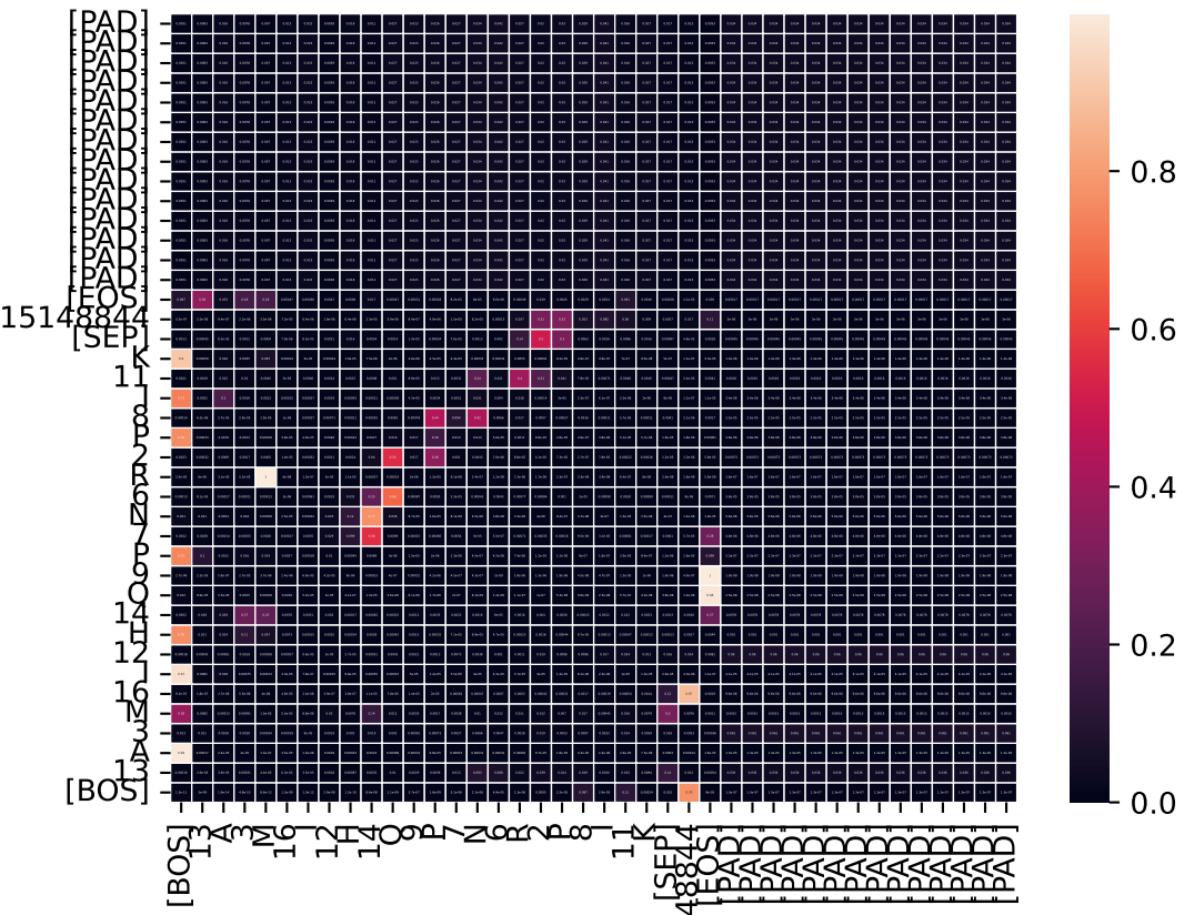
Step 69000 attention weights at layer 0 head 1, with accuracy 0.956

Sequence:[BOS] 13 A 3 M 16 | 12 H 14 Q 9 P 7 N 6 R 2 P 8 | 11 K [SEP] 16.06 [EOS] SOL:1



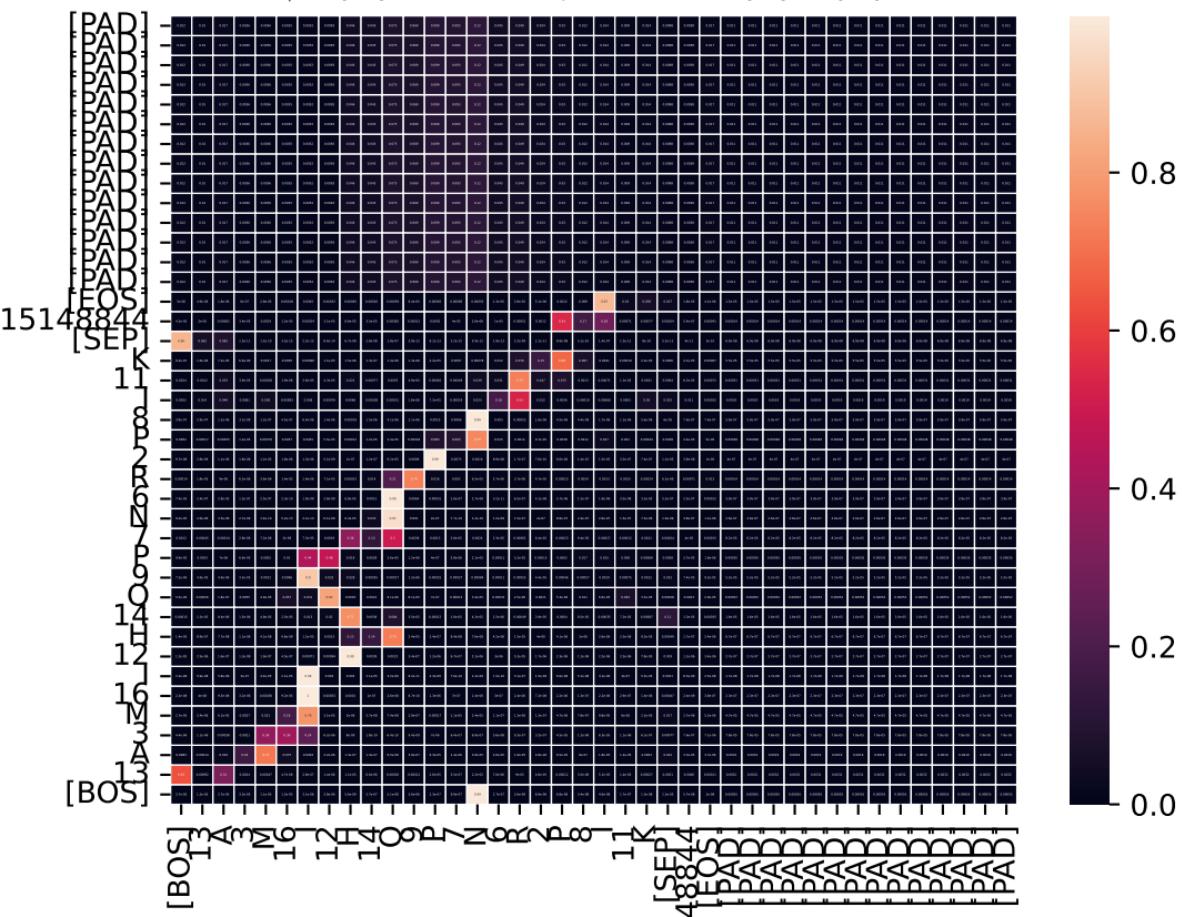
Step 69000 attention weights at layer 0 head 2, with accuracy 0.956
[PQSLAP, A, P, M, A, S, L, P, H, A, C, S, P, T, N, G, P, D, D, S, L, I, K, [SEP]] 16/26

Sequence:[BOS] 13 A 3 M 16! 12 H 14 Q 9 P 7 N 6 R 2 P 8! 11 K [SEP] 16.06 [EOS] SOL:1

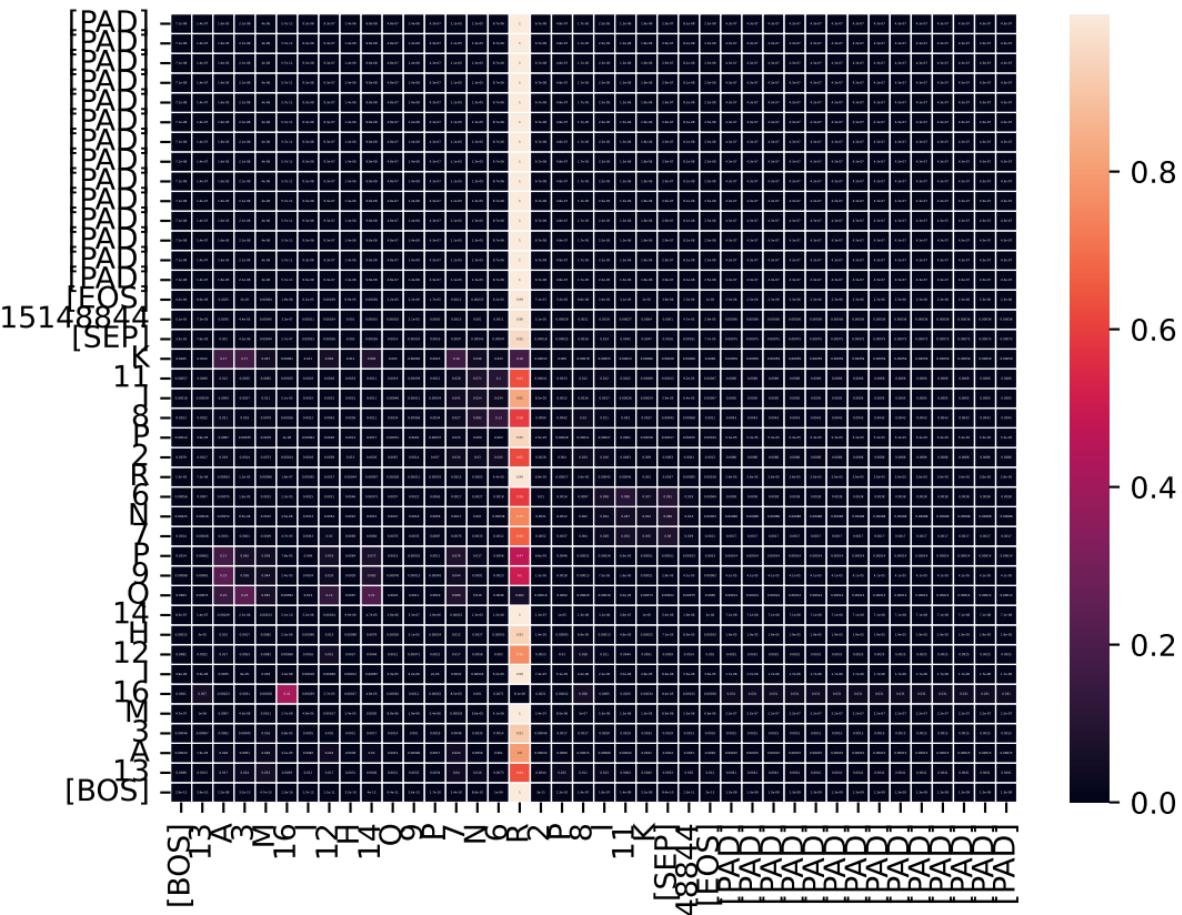


Step 69000 attention weights at layer 0 head 3, with accuracy 0.956
[P,O,S,I,L,A,B,M,A,C,I,D,H,A,G,C,P,T,N,C,B,D,P,G,L,I,I,K] [SEP] 16 26 16

Sequence:[BOS] 13 A 3 M 16! 12 H 14 Q 9 P 7 N 6 R 2 P 8! 11 K [SEP] 16.06 [EOS] SOL:1

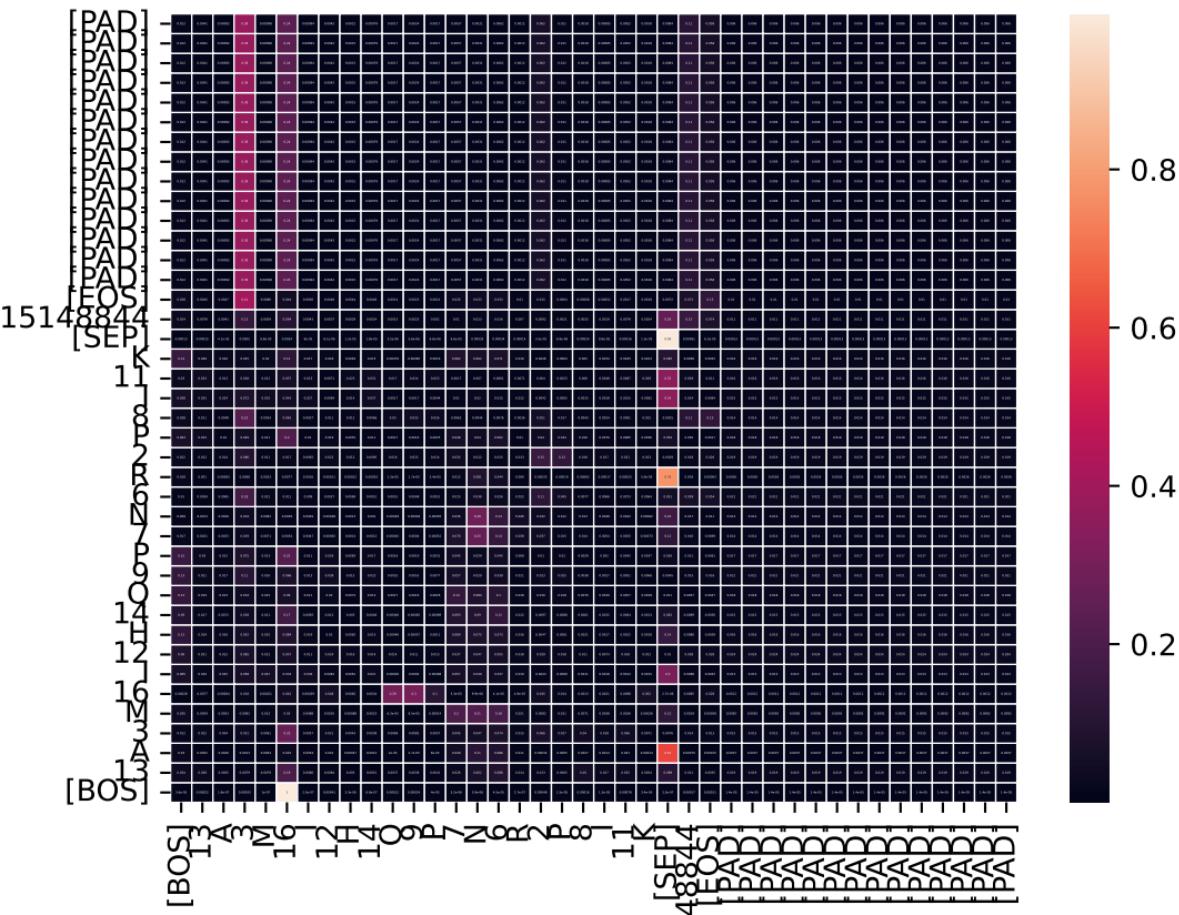


Step 69000 attention weights at layer 1 head 0, with accuracy 0.956
Sequence:[BOS] 13 A 3 M 16 I 12 H 14 Q 9 P 7 N 6 R 2 P 8 I 11 K [SEP] 16.06 [EOS] SOL:I

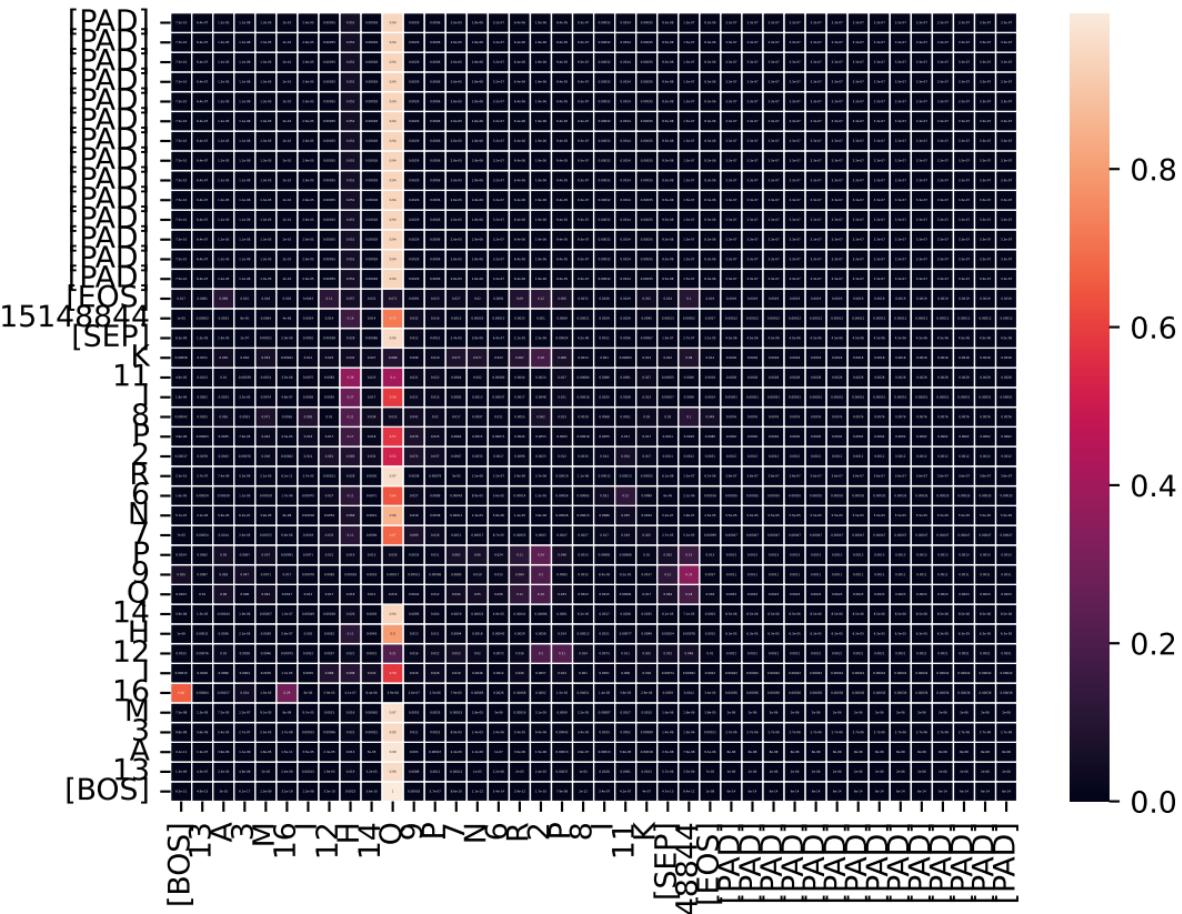


Step 69000 attention weights at layer 1 head 1, with accuracy 0.956

Sequence:[BOS] 13 A 3 M 16 | 12 H 14 Q 9 P 7 N 6 R 2 P 8 | 11 K [SEP] 16.06 [EOS] SOL:I

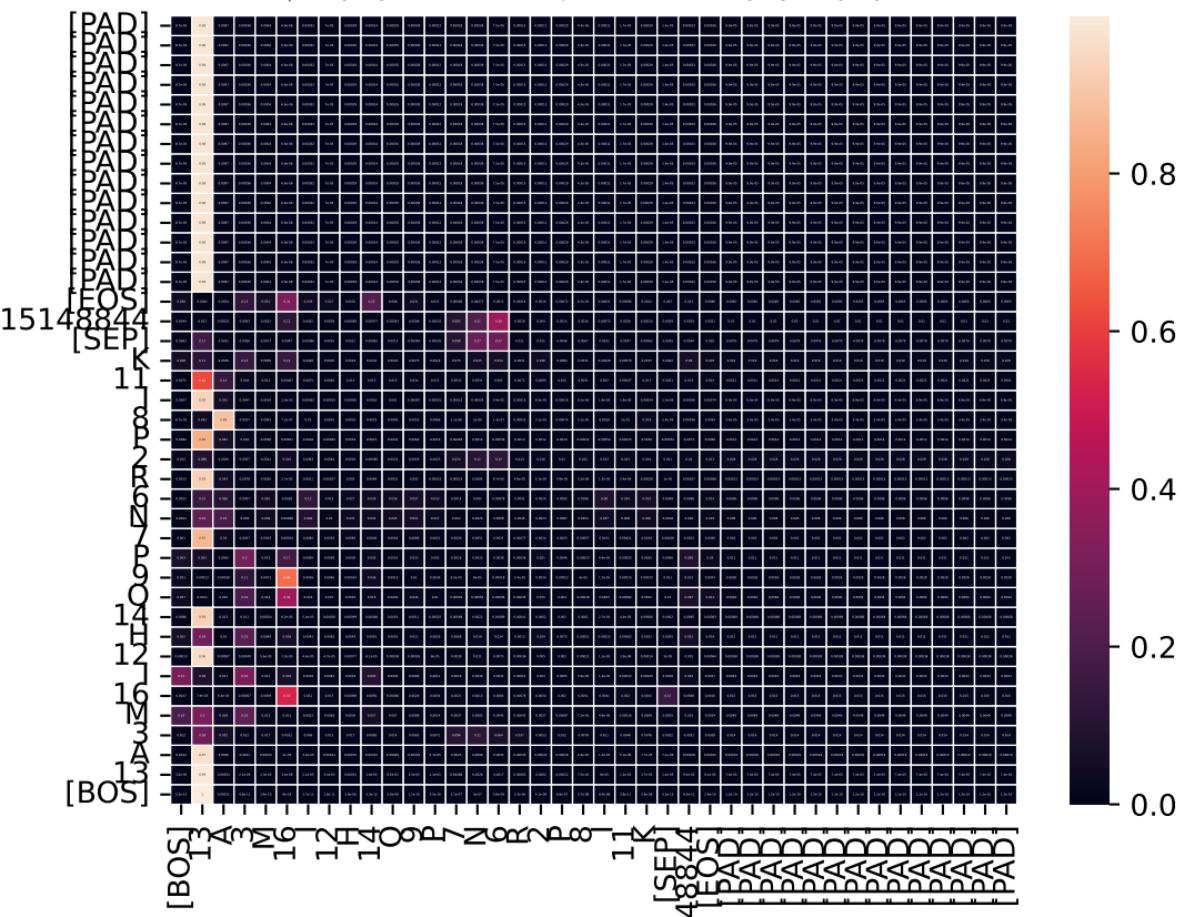


Step 69000 attention weights at layer 1 head 2, with accuracy 0.956
Sequence:[BOS] 13 A 3 M 16 I 12 H 14 Q 9 P 7 N 6 R 2 P 8 I 11 K [SEP] 16.06 [EOS] SOL:I



Step 69000 attention weights at layer 1 head 3, with accuracy 0.956

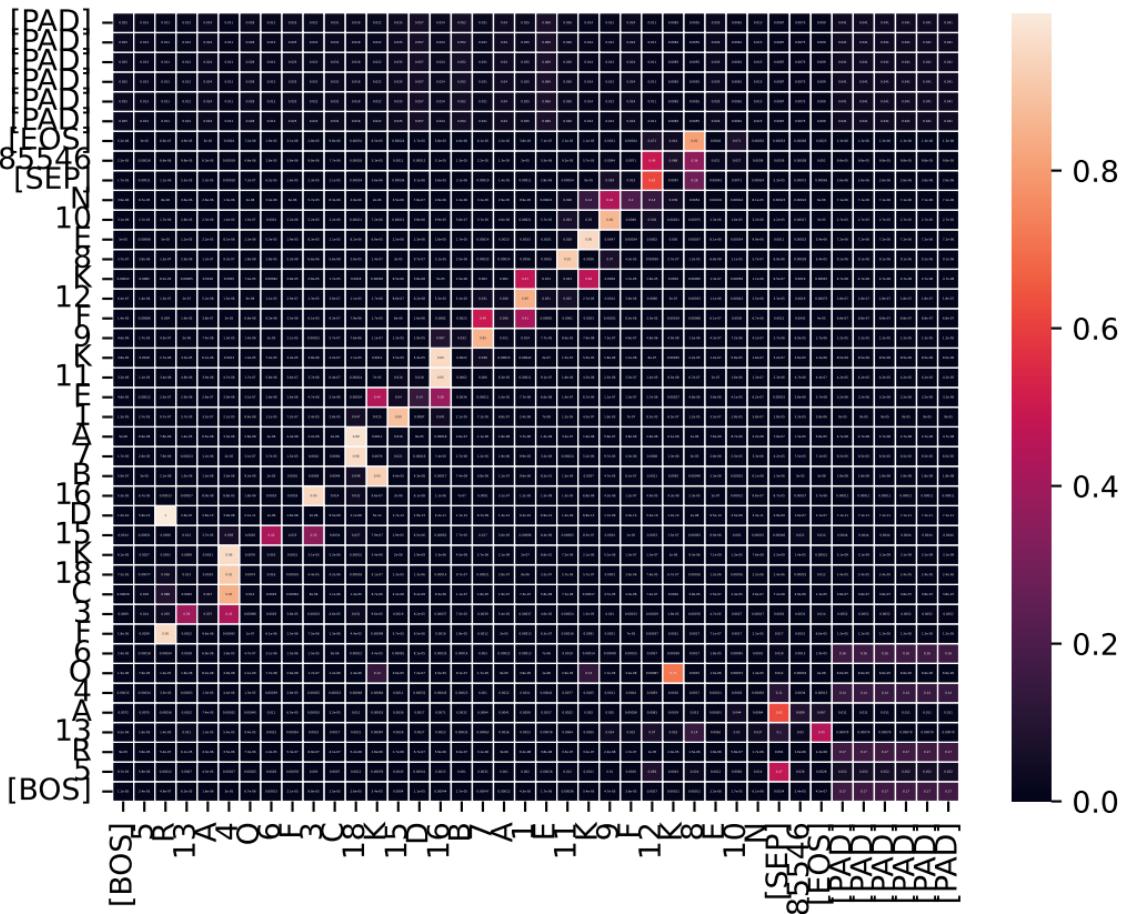
Sequence:[BOS] 13 A 3 M 16! 12 H 14 Q 9 P 7 N 6 R 2 P 8! 11 K [SEP] 16.06 [EOS] SOL:1



Step 69500 attention weights at layer 0 head 0, with accuracy 0.984

Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A

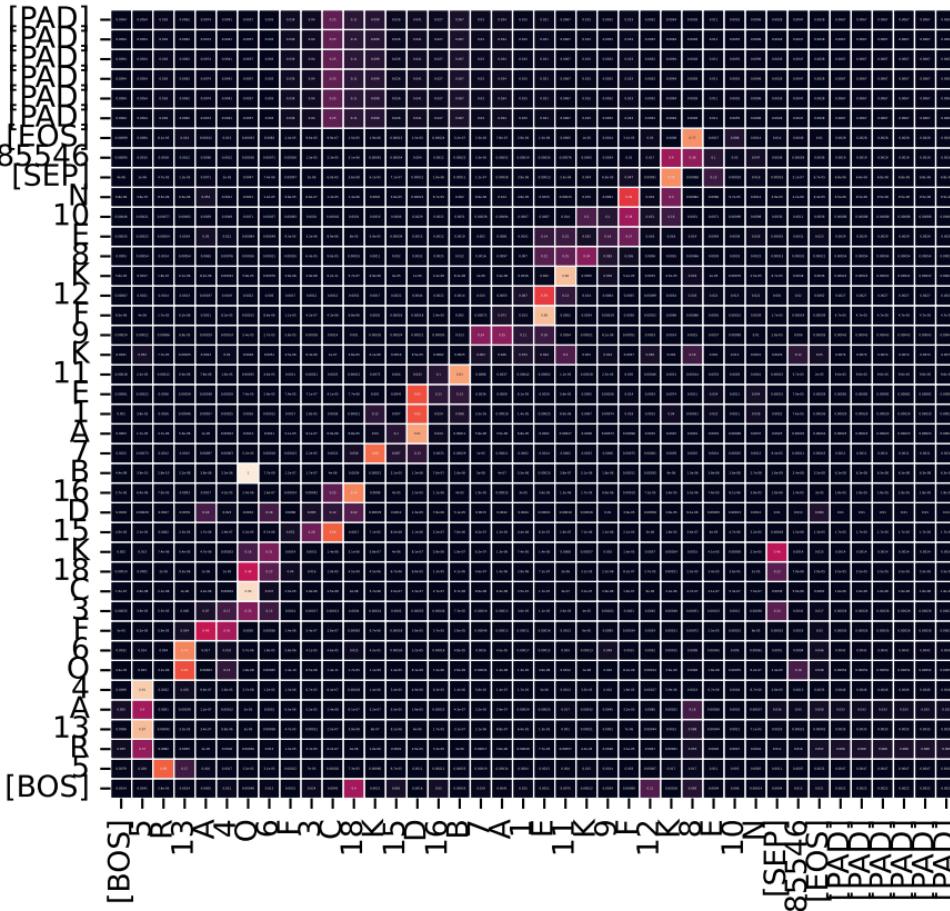
47285546



Step 69500 attention weights at layer 0 head 1, with accuracy 0.984

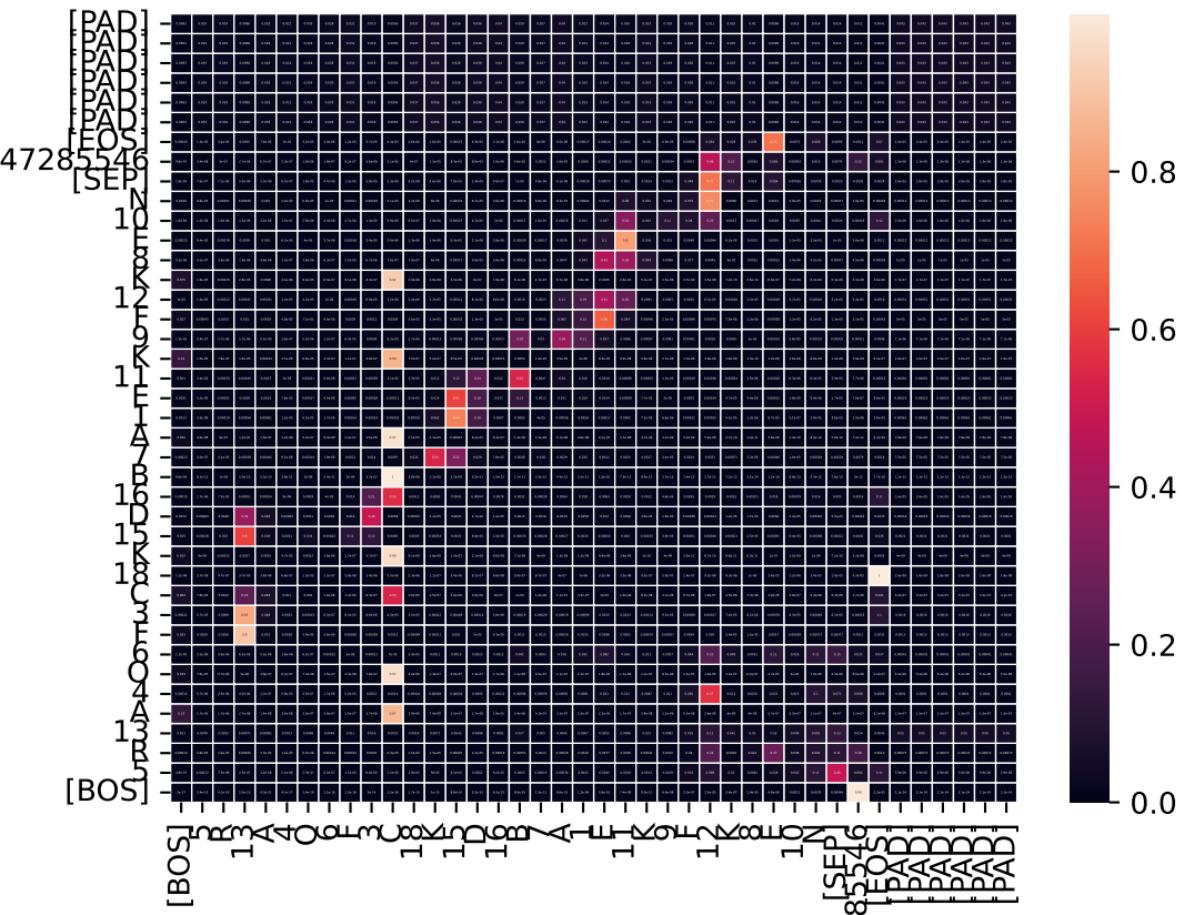
Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A

4728



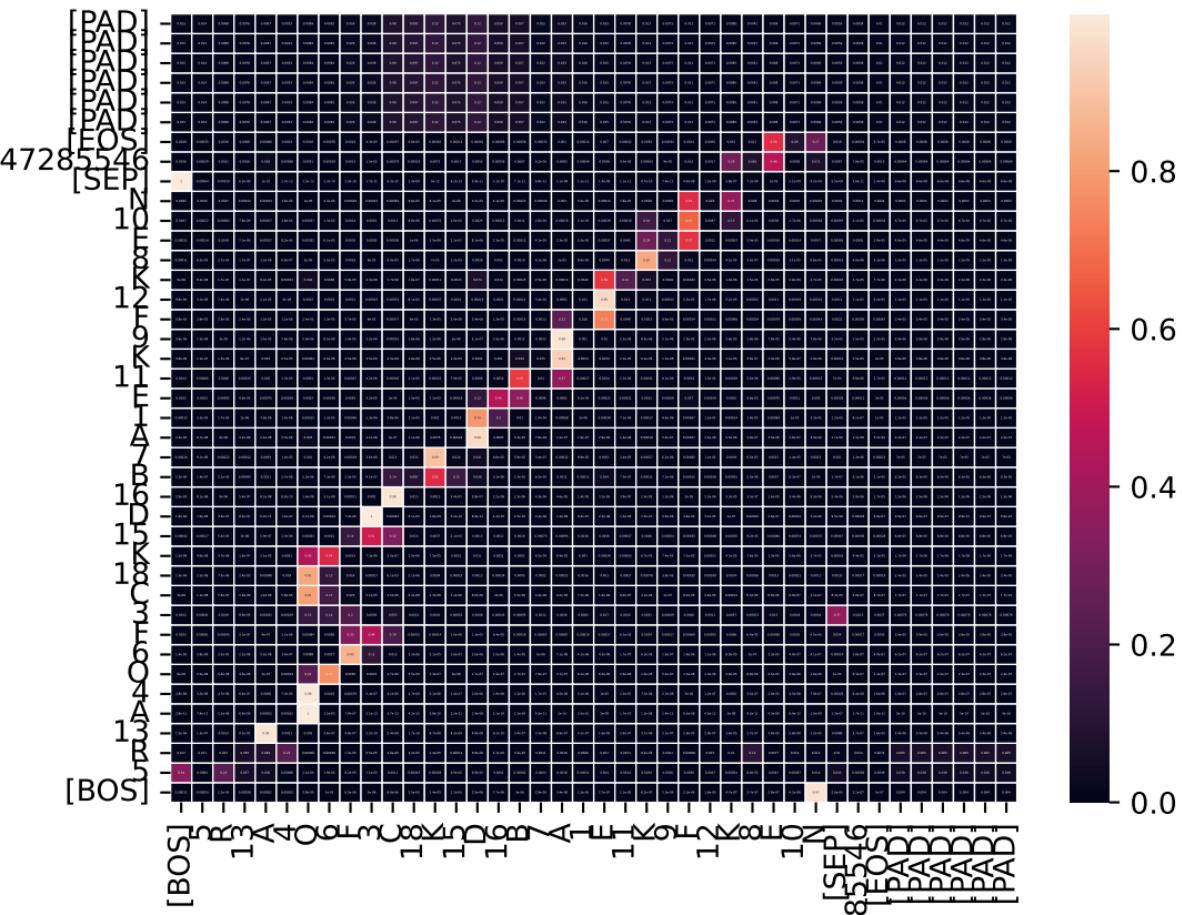
Step 69500 attention weights at layer 0 head 2, with accuracy 0.984

Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A



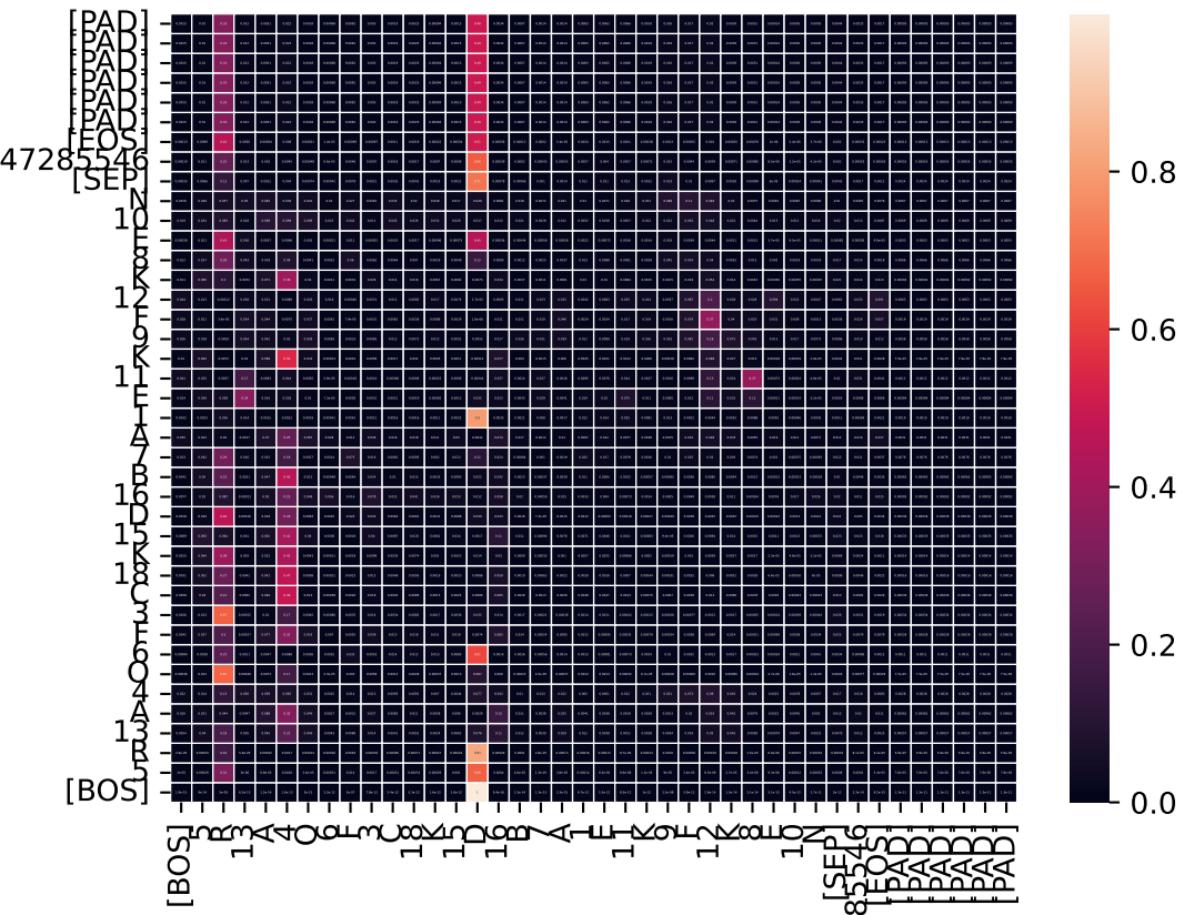
Step 69500 attention weights at layer 0 head 3, with accuracy 0.984

Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A



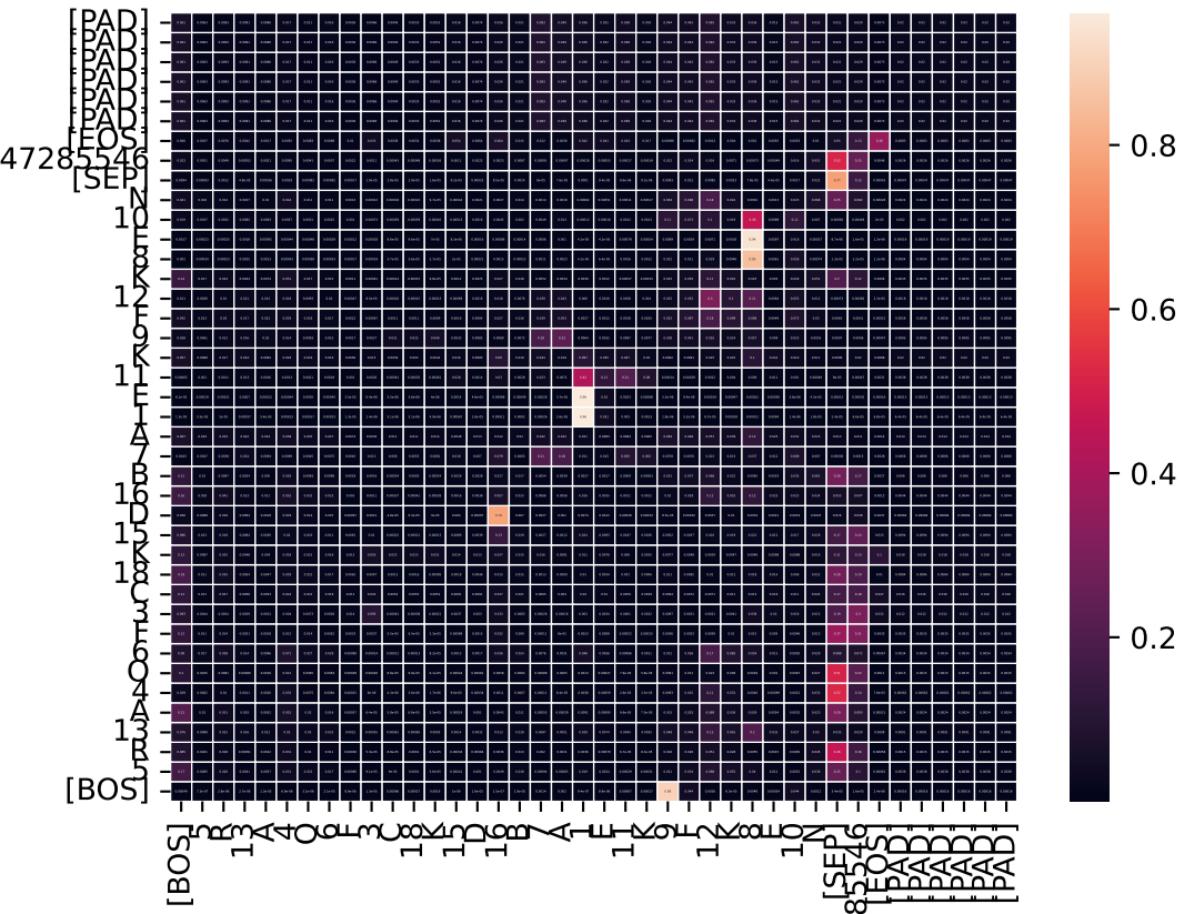
Step 69500 attention weights at layer 1 head 0, with accuracy 0.984

Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A



Step 69500 attention weights at layer 1 head 1, with accuracy 0.984

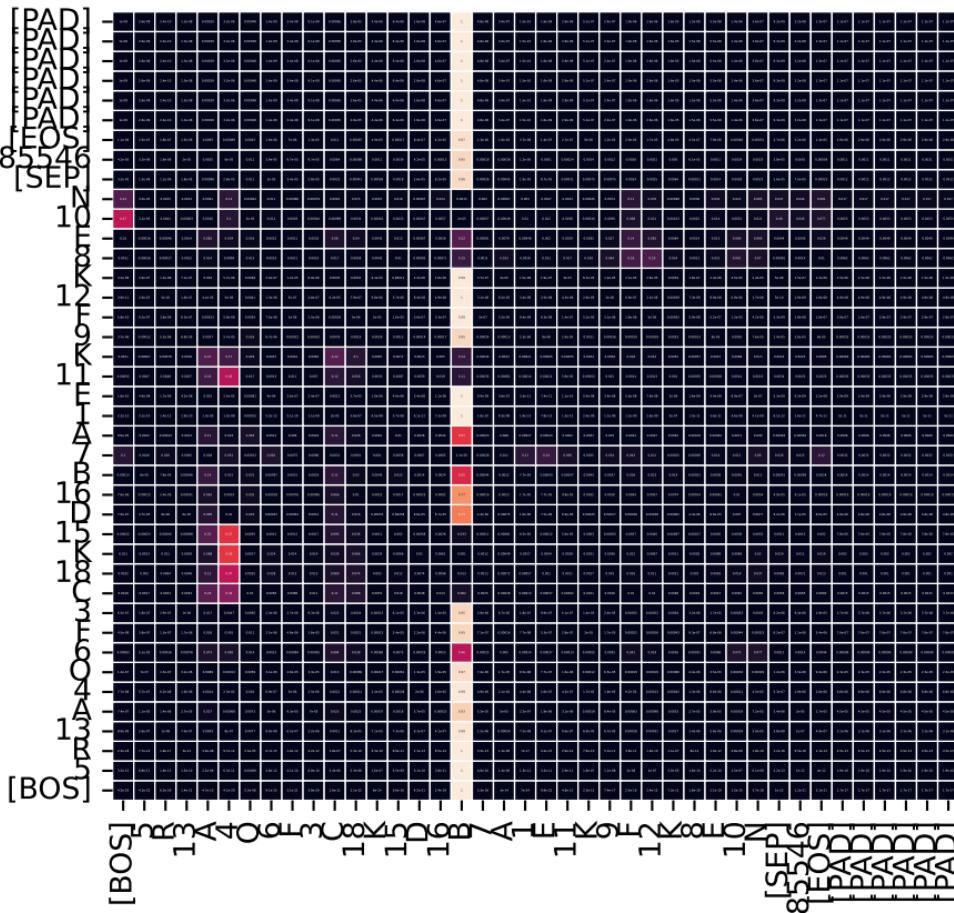
Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A



Step 69500 attention weights at layer 1 head 2, with accuracy 0.984

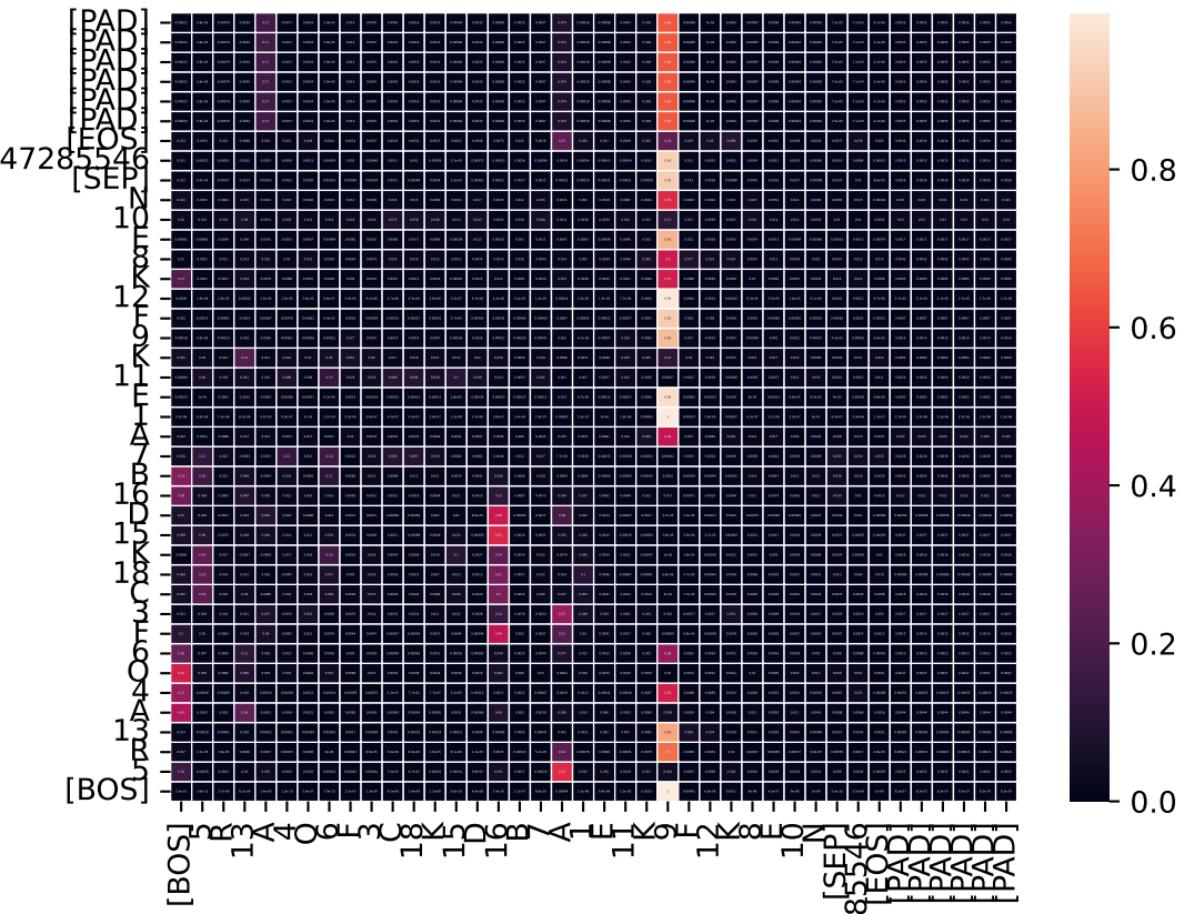
Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:A

47285546

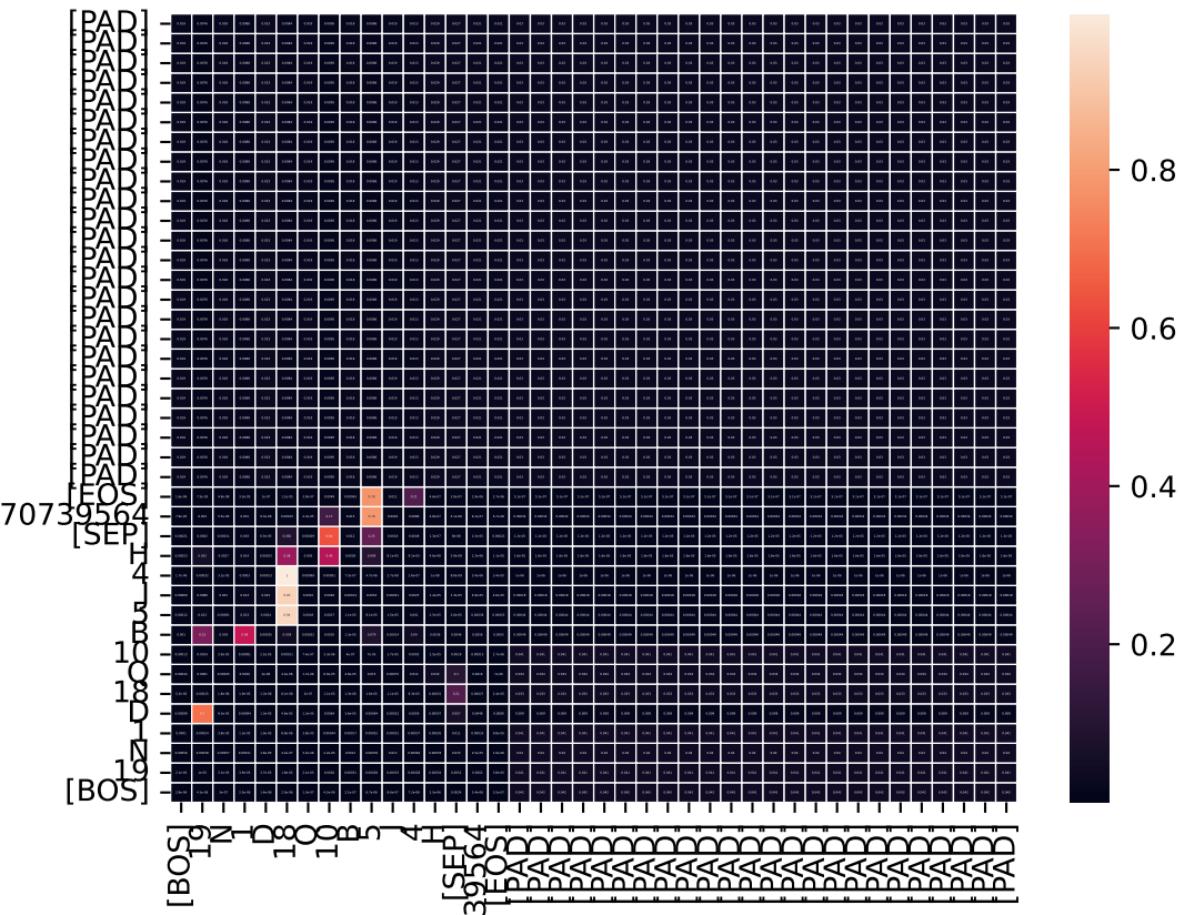


Step 69500 attention weights at layer 1 head 3, with accuracy 0.984

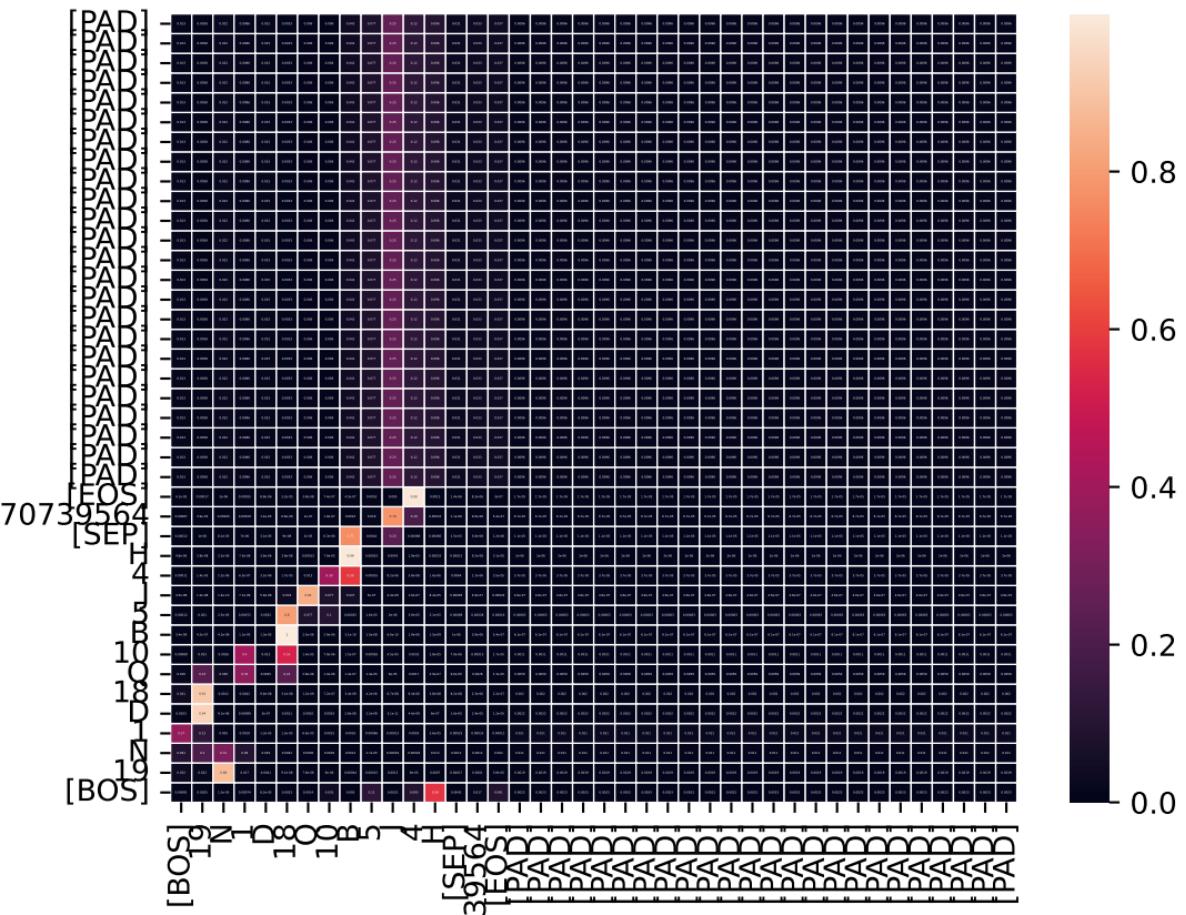
Sequence:[BOS] 5 R 13 A 4 O 6 F 3 C 18 K 15 D 16 B 7 A 1 E 11 K 9 F 12 K 8 E 10 N [SEP] 6.97 [EOS] SOL:AI



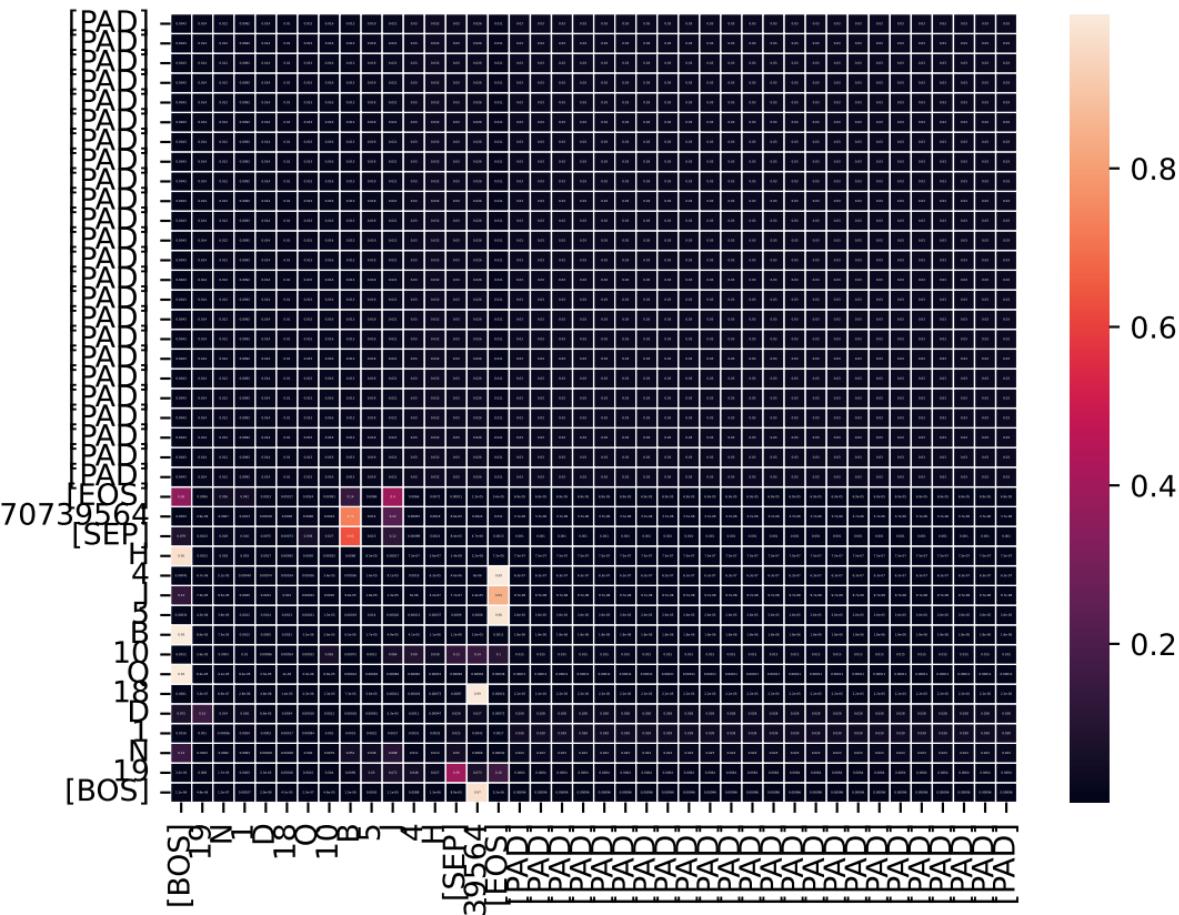
Step 70000 attention weights at layer 0 head 0, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



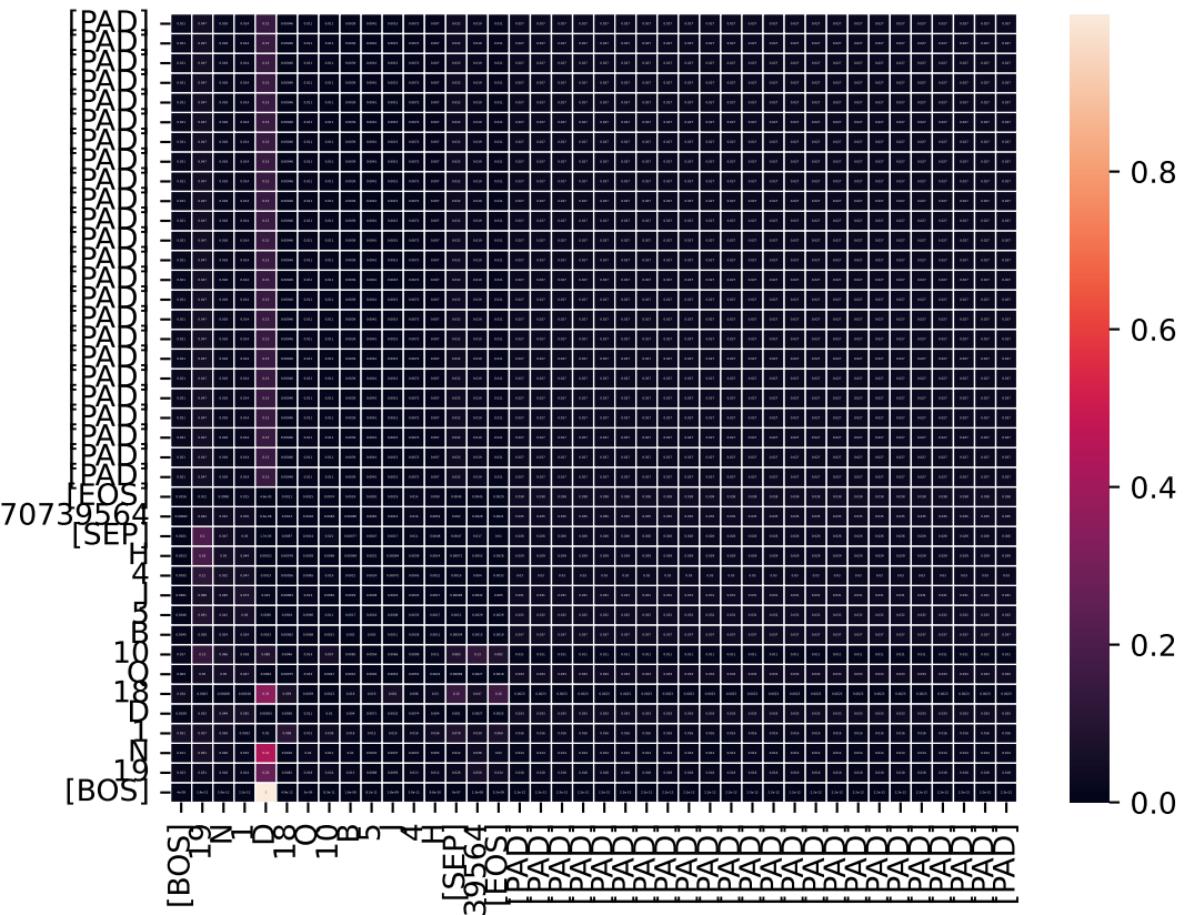
Step 70000 attention weights at layer 0 head 1, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



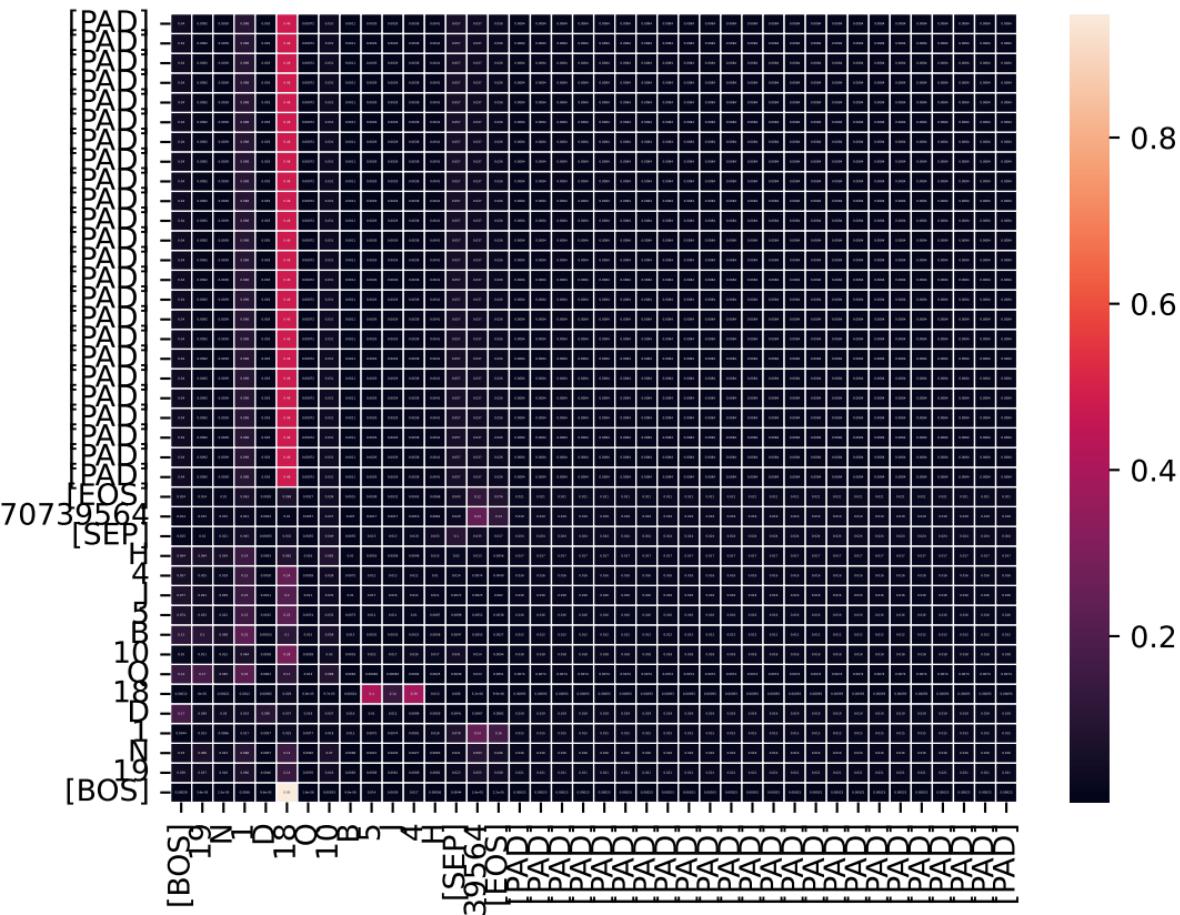
Step 70000 attention weights at layer 0 head 2, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



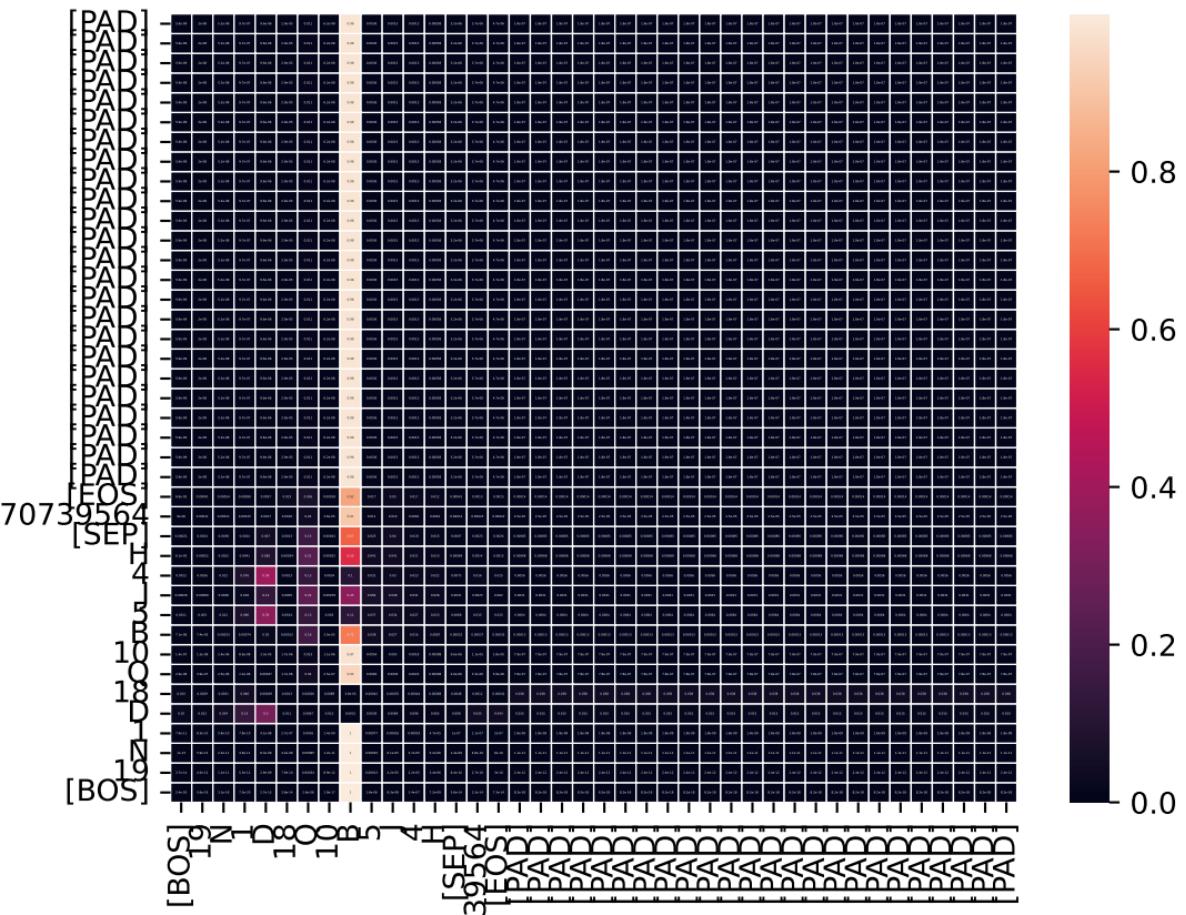
Step 70000 attention weights at layer 1 head 0, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



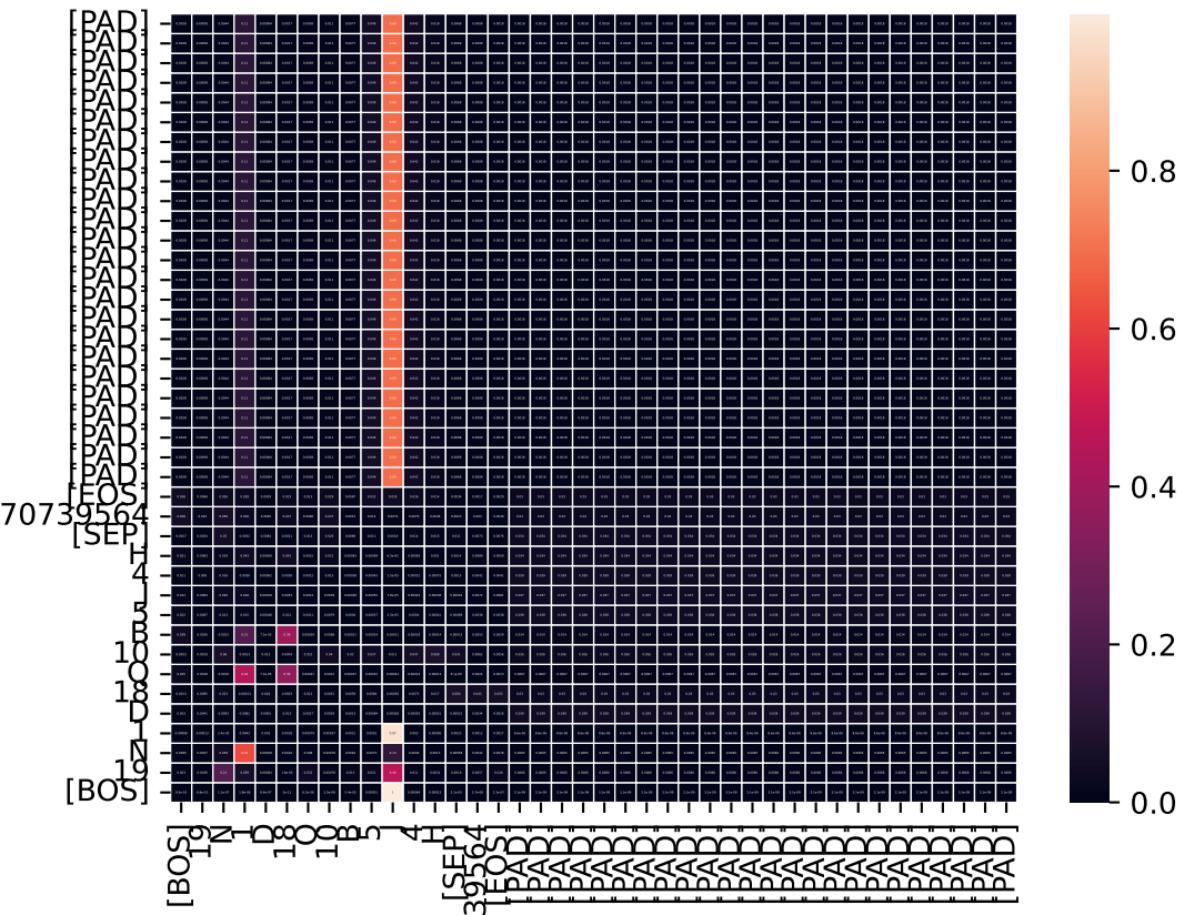
Step 70000 attention weights at layer 1 head 1, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



Step 70000 attention weights at layer 1 head 2, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q

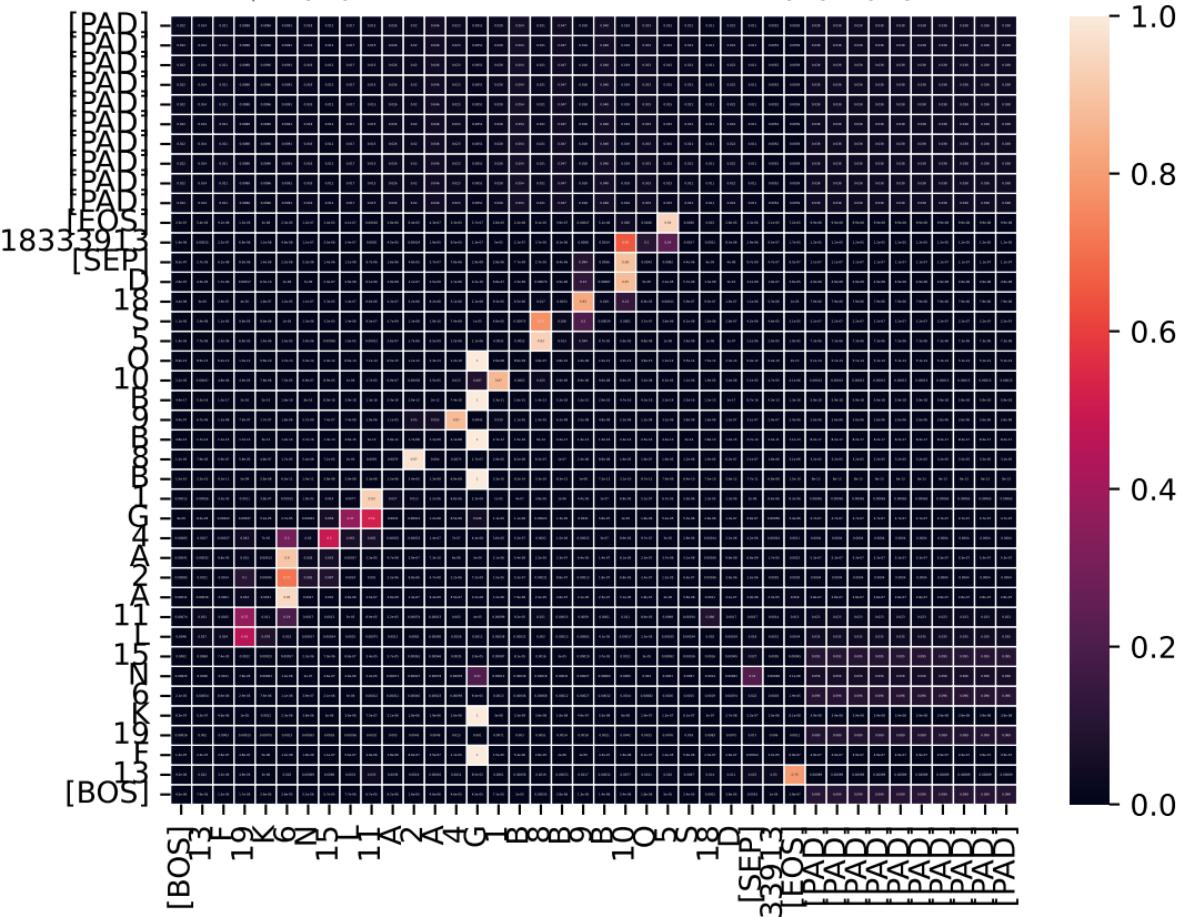


Step 70000 attention weights at layer 1 head 3, with accuracy 0.987
Sequence:[BOS] 19 N 1 D 18 Q 10 B 5 J 4 H [SEP] 18.04 [EOS] SOL:Q



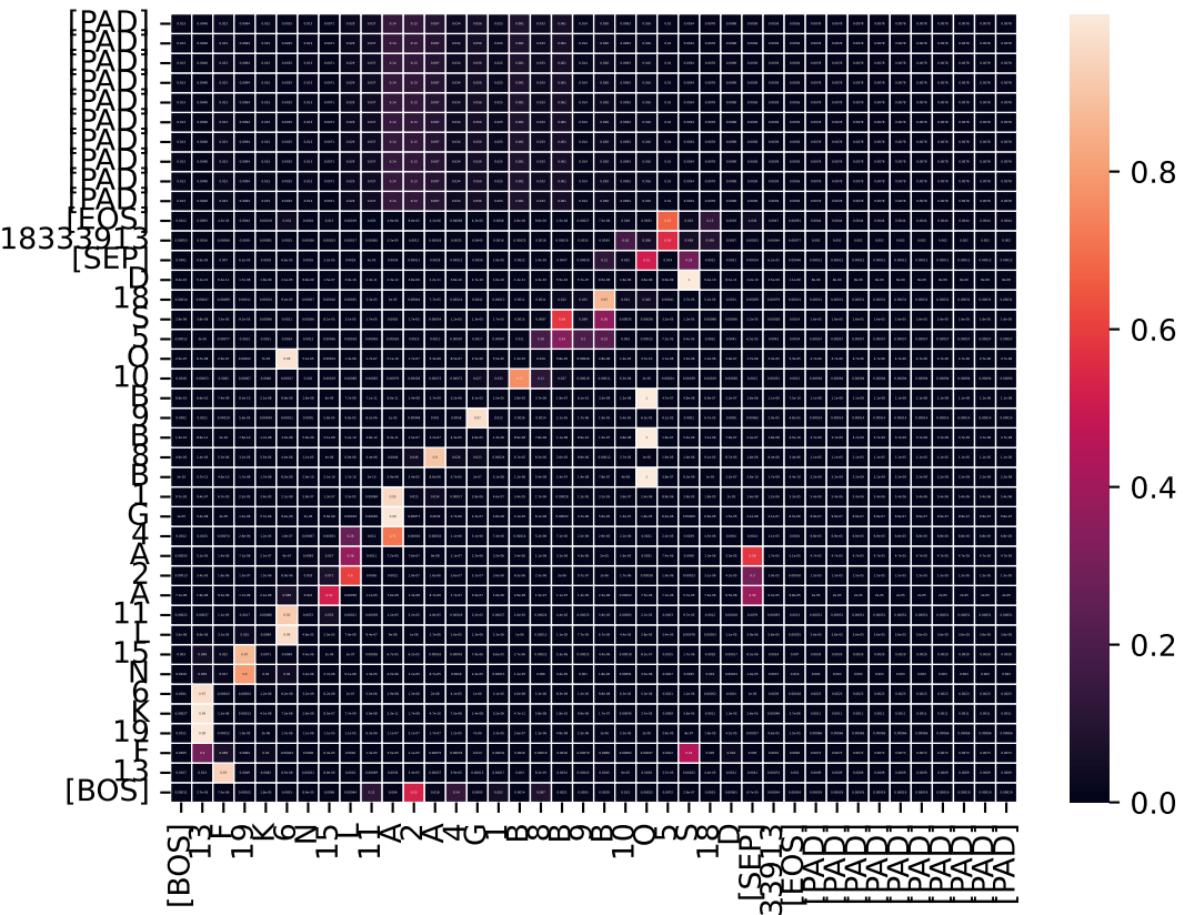
Step 70500 attention weights at layer 0 head 0, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



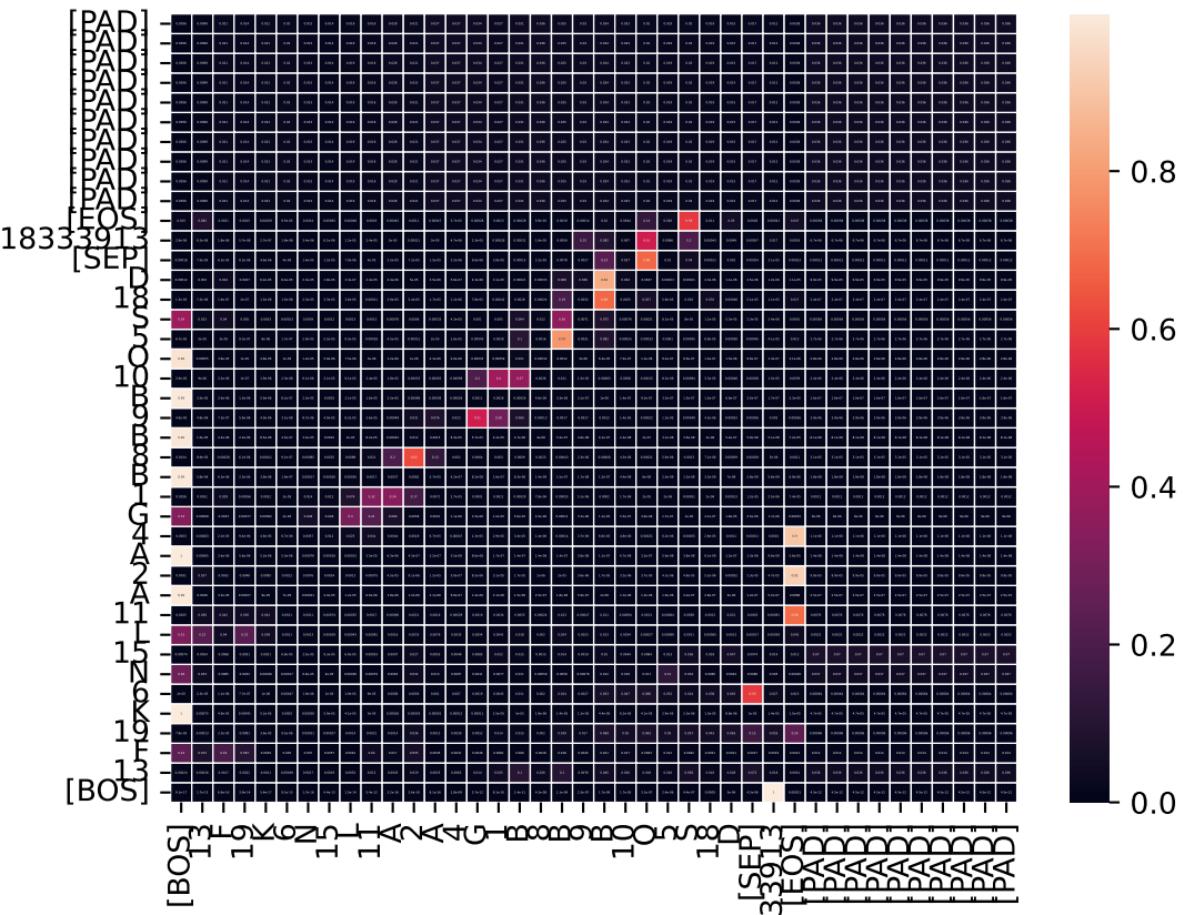
Step 70500 attention weights at layer 0 head 1, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



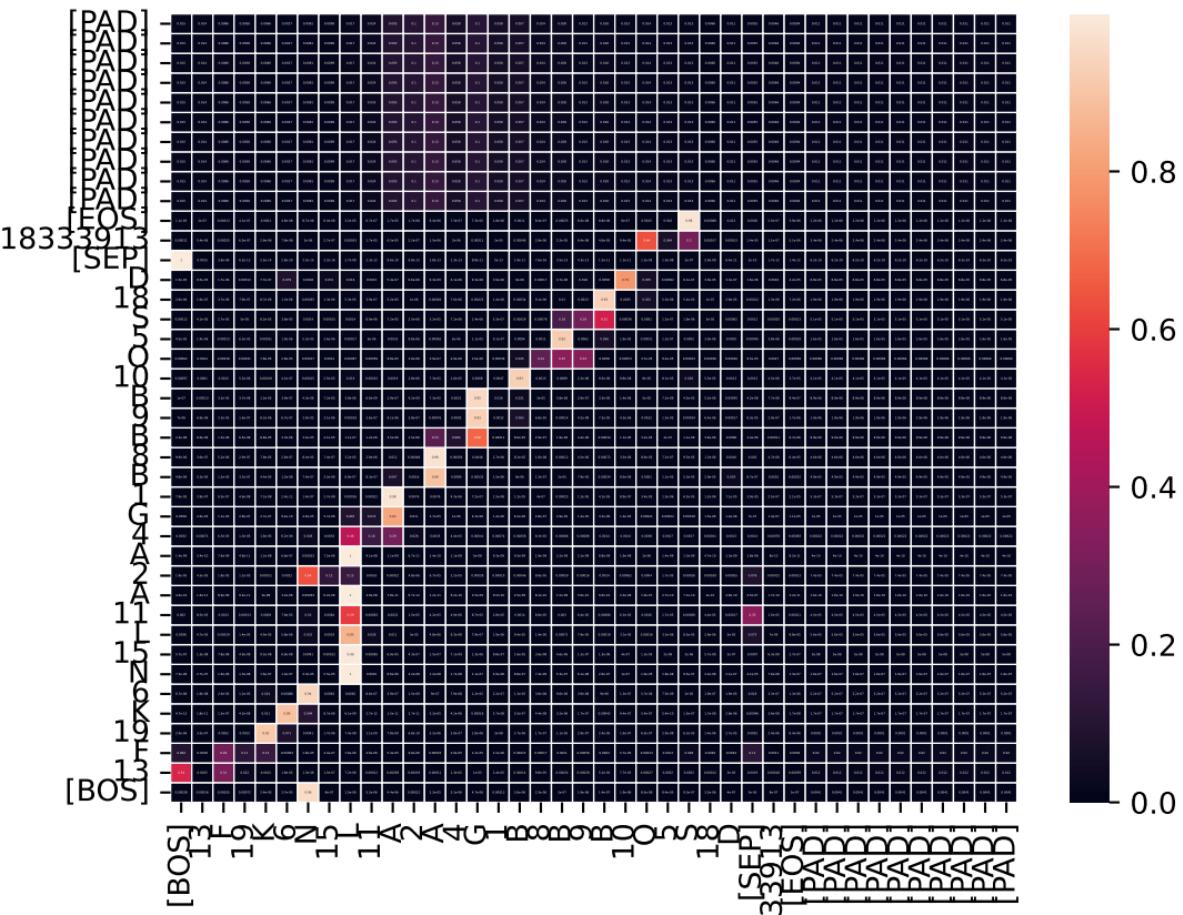
Step 70500 attention weights at layer 0 head 2, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



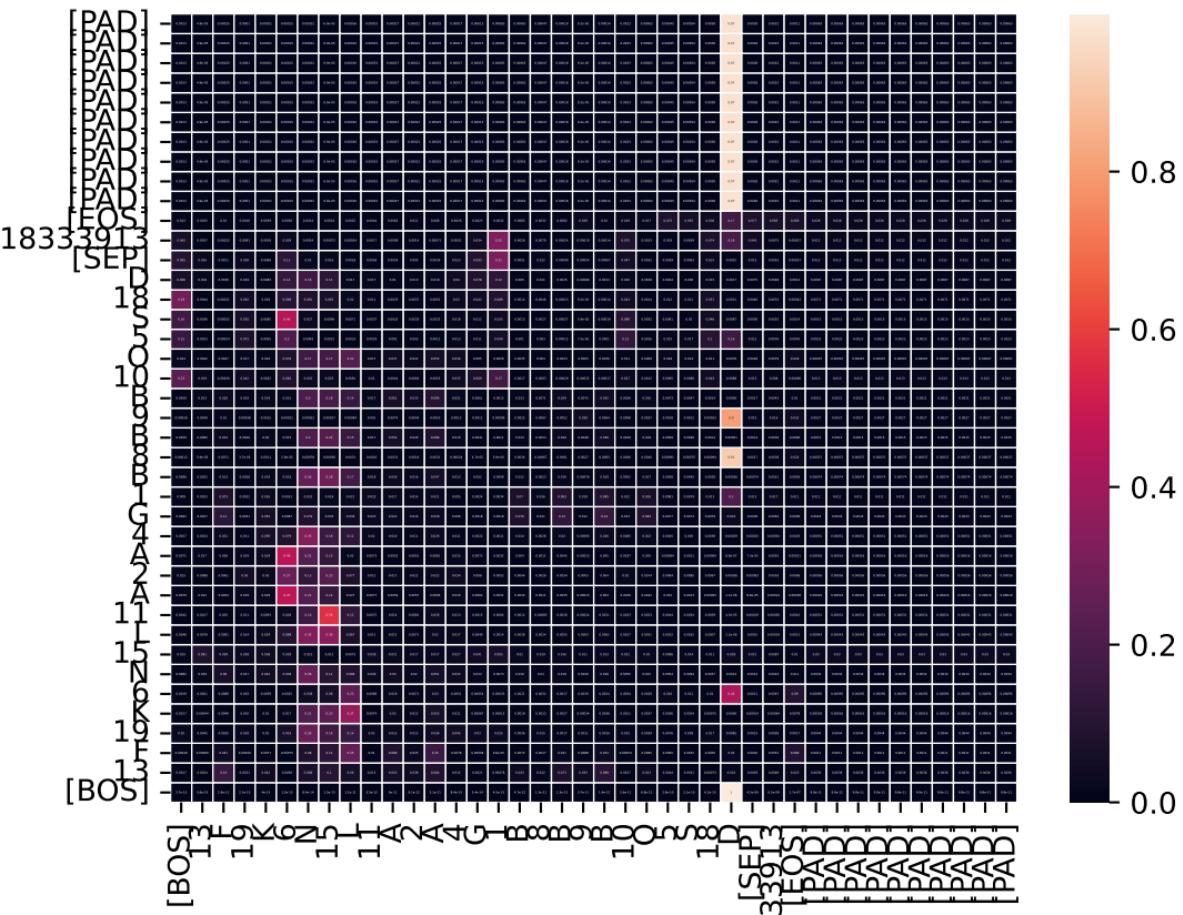
Step 70500 attention weights at layer 0 head 3, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



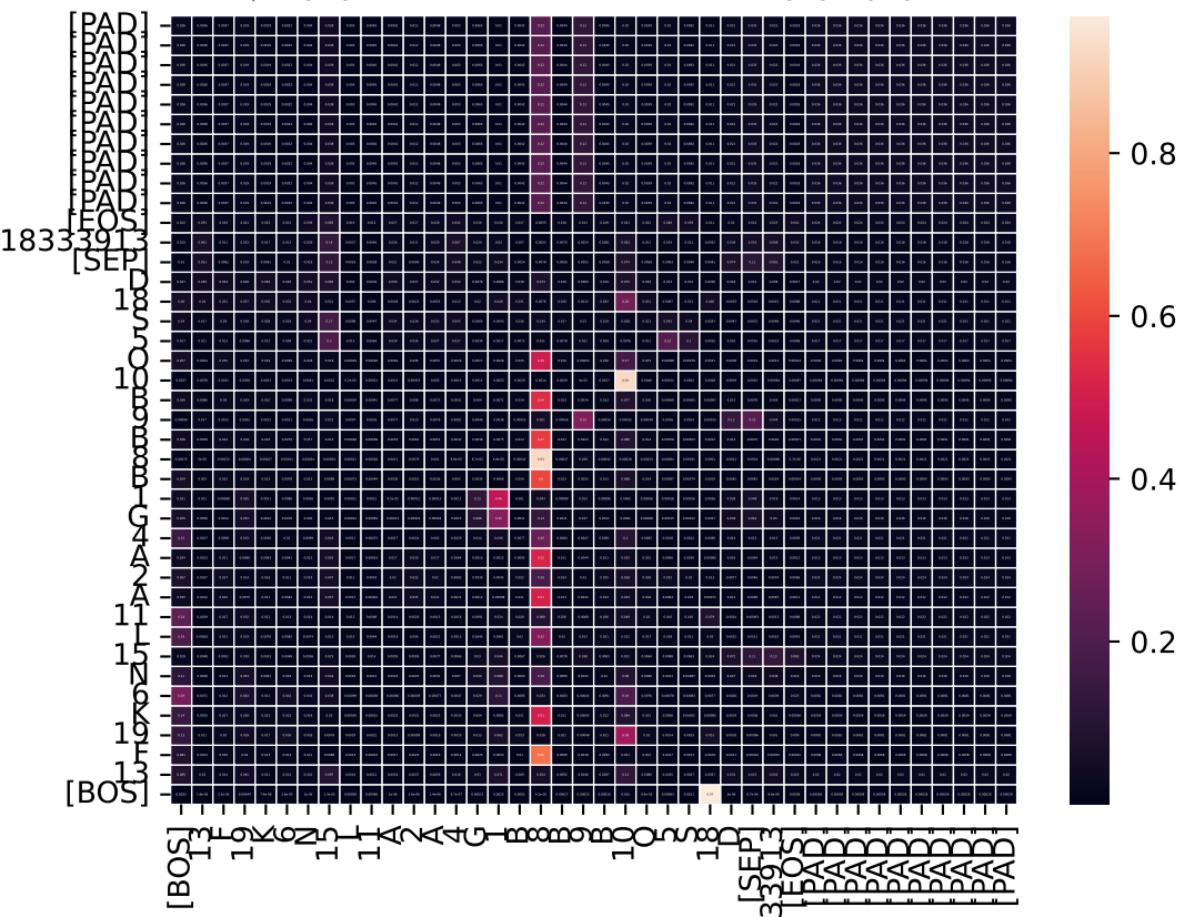
Step 70500 attention weights at layer 1 head 0, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



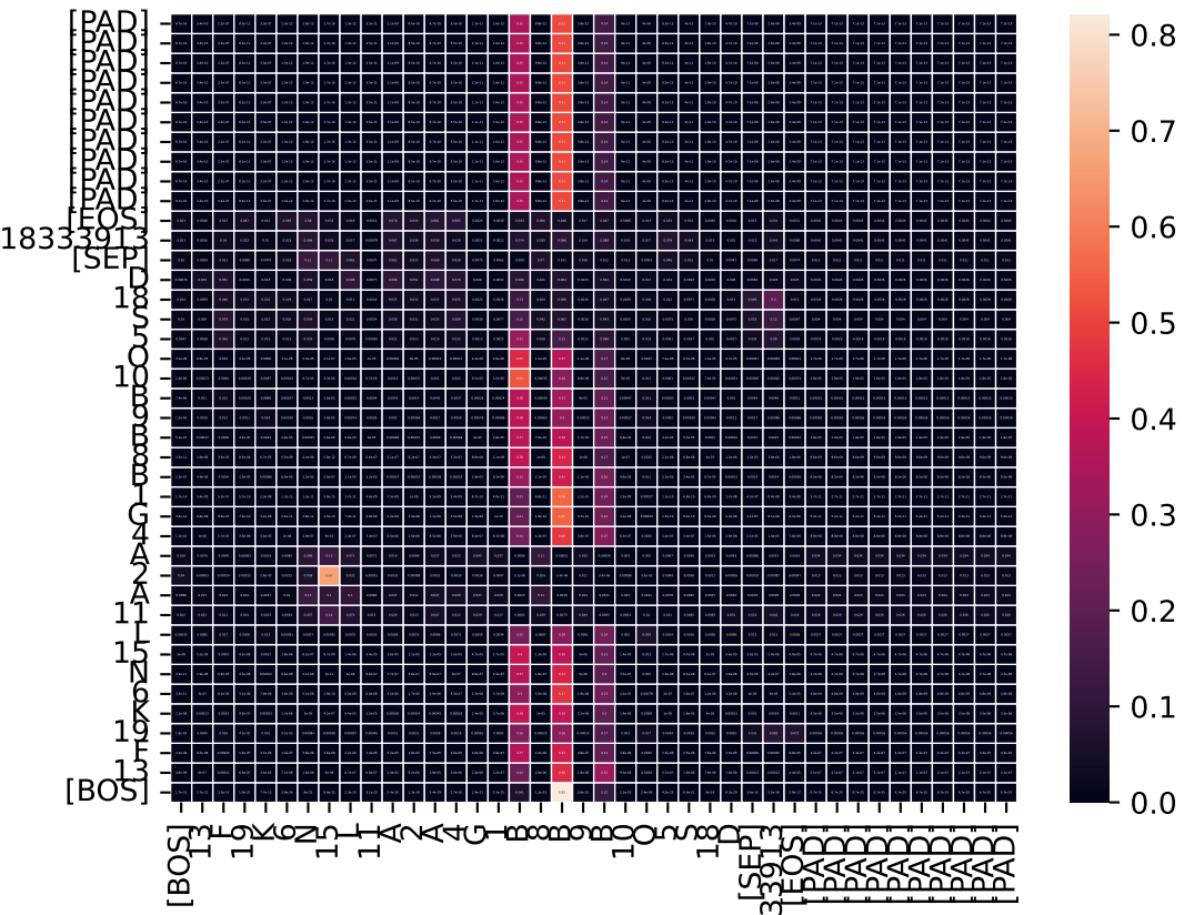
Step 70500 attention weights at layer 1 head 1, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



Step 70500 attention weights at layer 1 head 2, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B



Step 70500 attention weights at layer 1 head 3, with accuracy 0.992

Sequence:[BOS] 13 F 19 K 6 N 15 L 11 A 2 A 4 G 1 B 8 B 9 B 10 O 5 S 18 D [SEP] 8.86 [EOS] SOL:B

