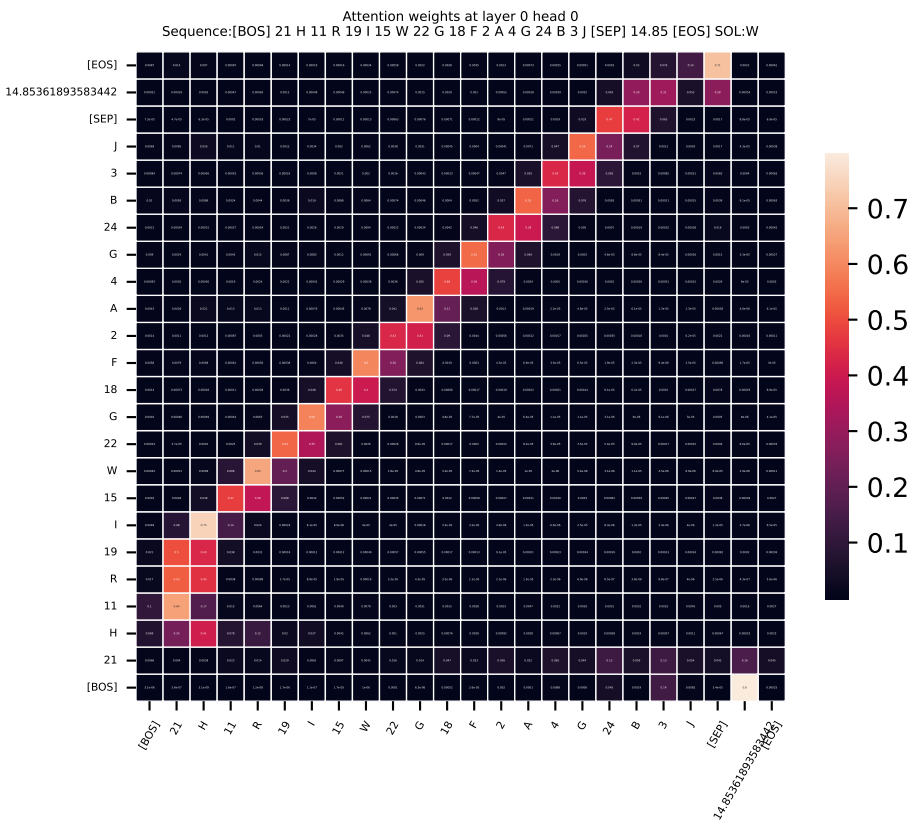
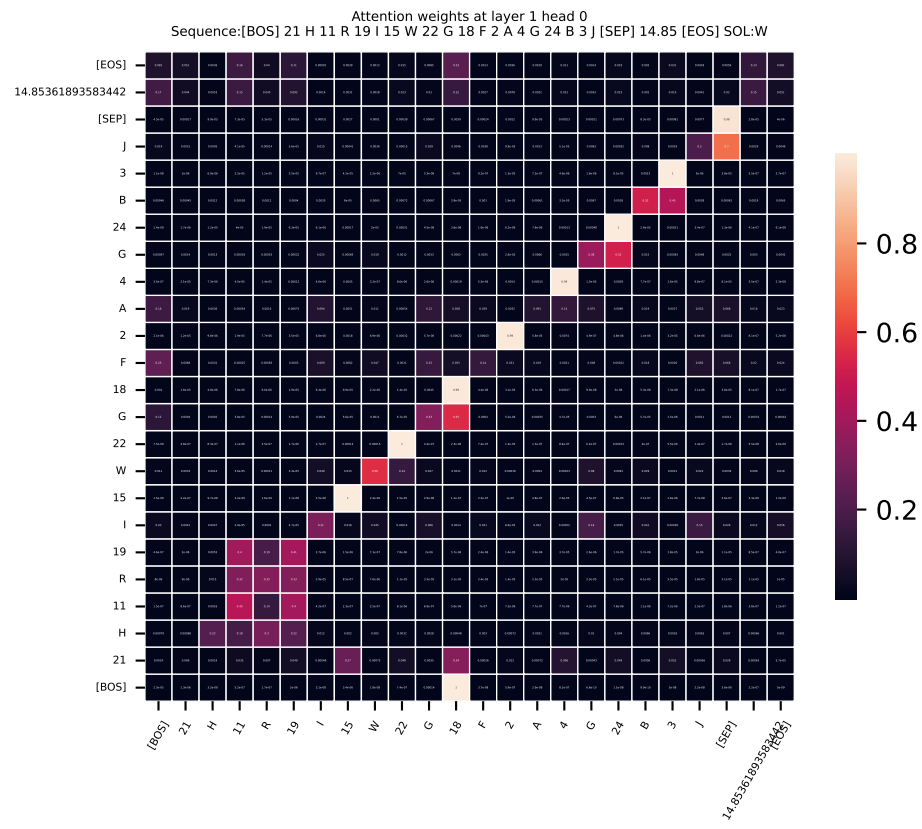
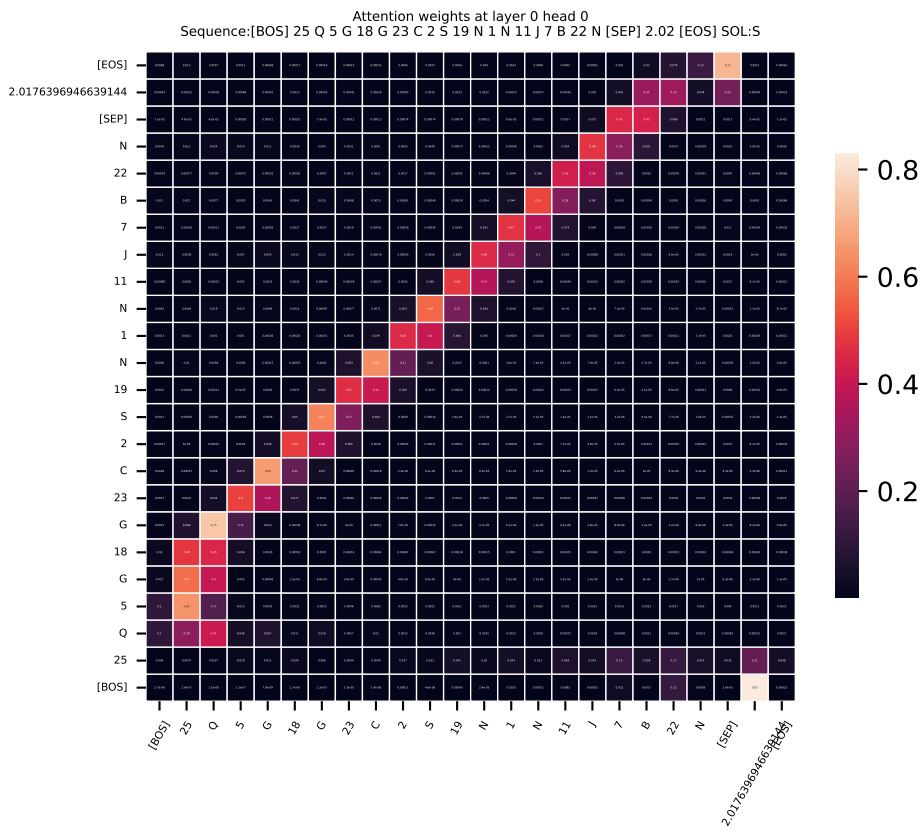
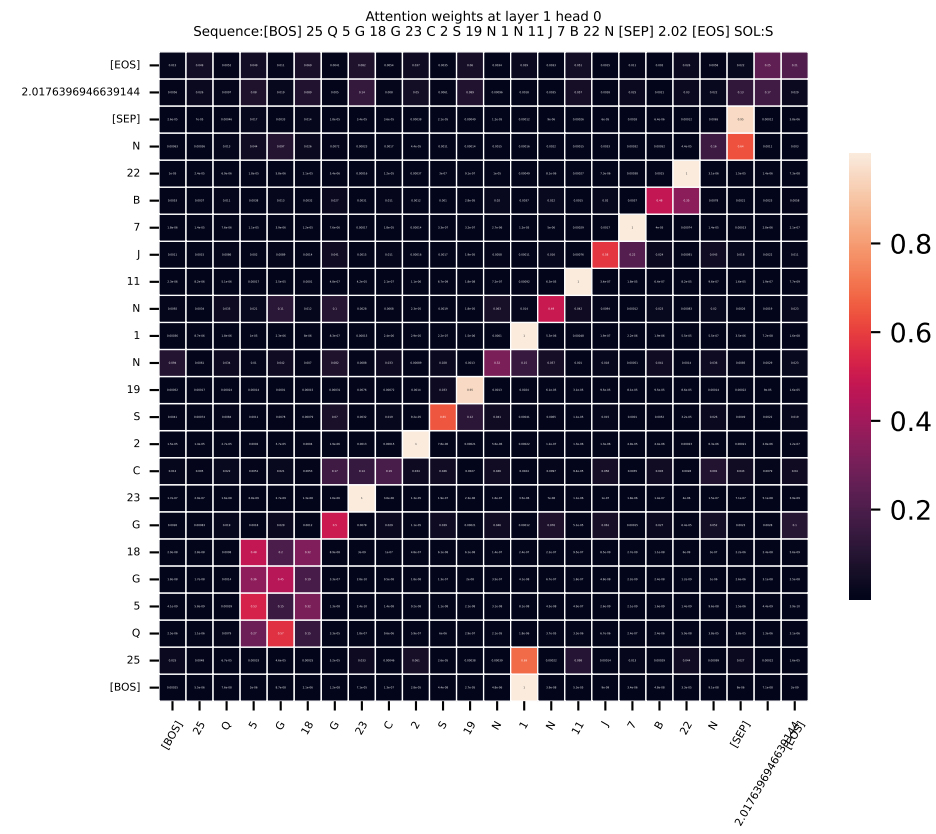
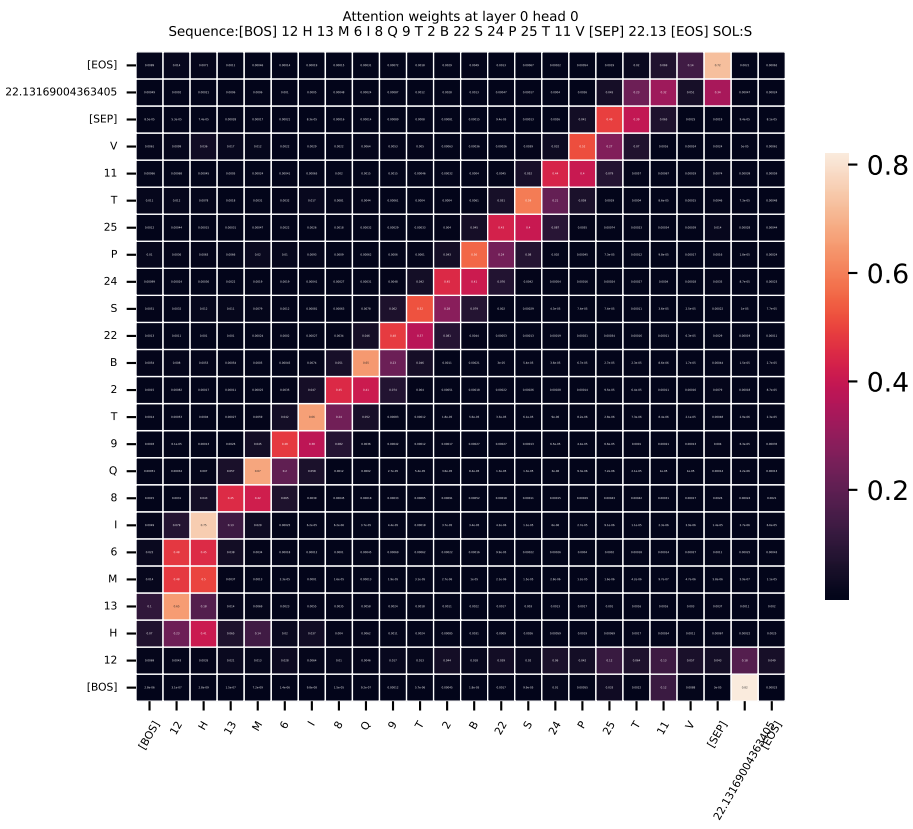
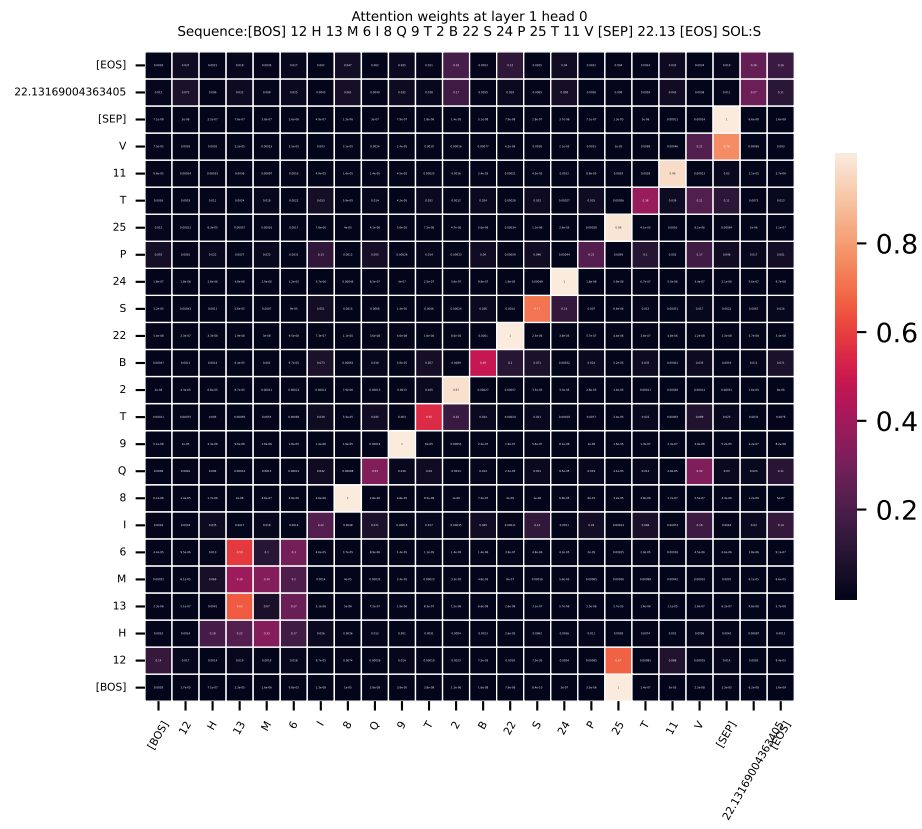

Attention weights at layer 0 head 0
Sequence:[BOS] 21 H 11 R 19 I 15 W 22 G 18 F 2 A 4 G 24 B 3 J [SEP] 14.85 [EOS] SOL:W
[EOS]
14.85361893583442
[SEP]
J
3
B
24
G
4
A
2
F
18
G
22
W
15
I
19
R
11
H
21
[BOS]
[BOS] 21 H 11 R 19 I 15 W 22 G 18 F 2 A 4 G 24 B 3 J [SEP] 14.85 [EOS]
0.7
0.6
0.5
0.4
0.3
0.2
0.1

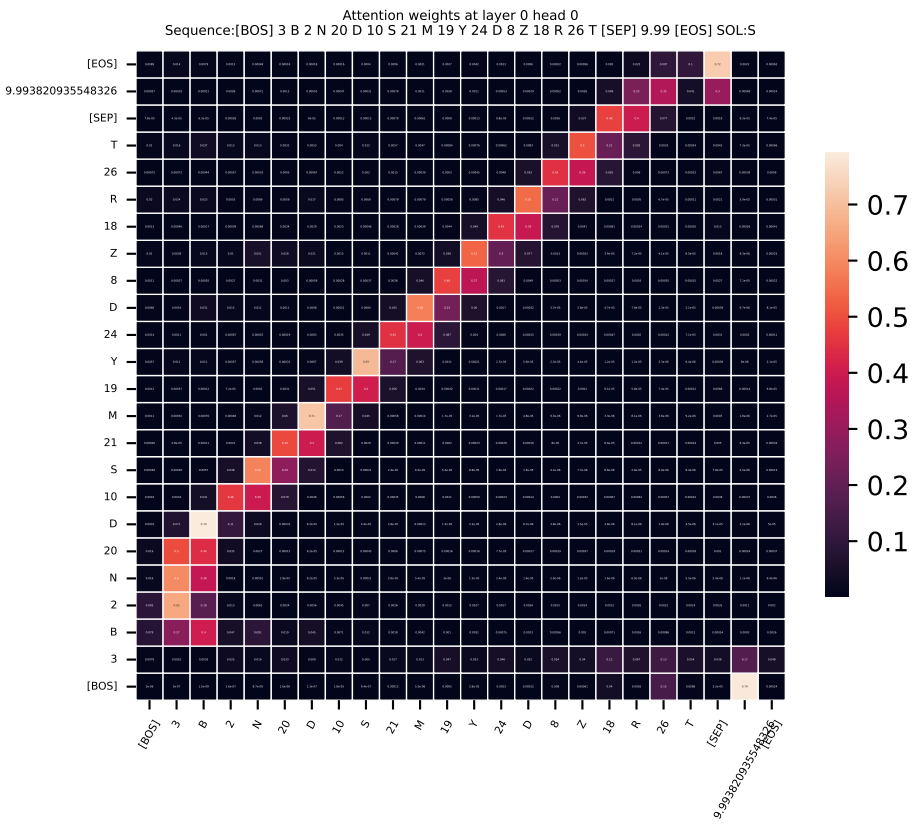

Attention weights at layer 1 head 0
Sequence:[BOS] 21 H 11 R 19 I 15 W 22 G 18 F 2 A 4 G 24 B 3 J [SEP] 14.85 [EOS] SOL:W
[EOS]
14.85361893583442
[SEP]
J
3
B
24
G
4
A
2
F
18
G
22
W
15
I
19
R
11
H
21
[BOS]
[BOS] 21 H 11 R 19 I 15 W 22 G 18 F 2 A 4 G 24 B 3 J [SEP] 14.85 [EOS]
0.8
0.6
0.4
0.2

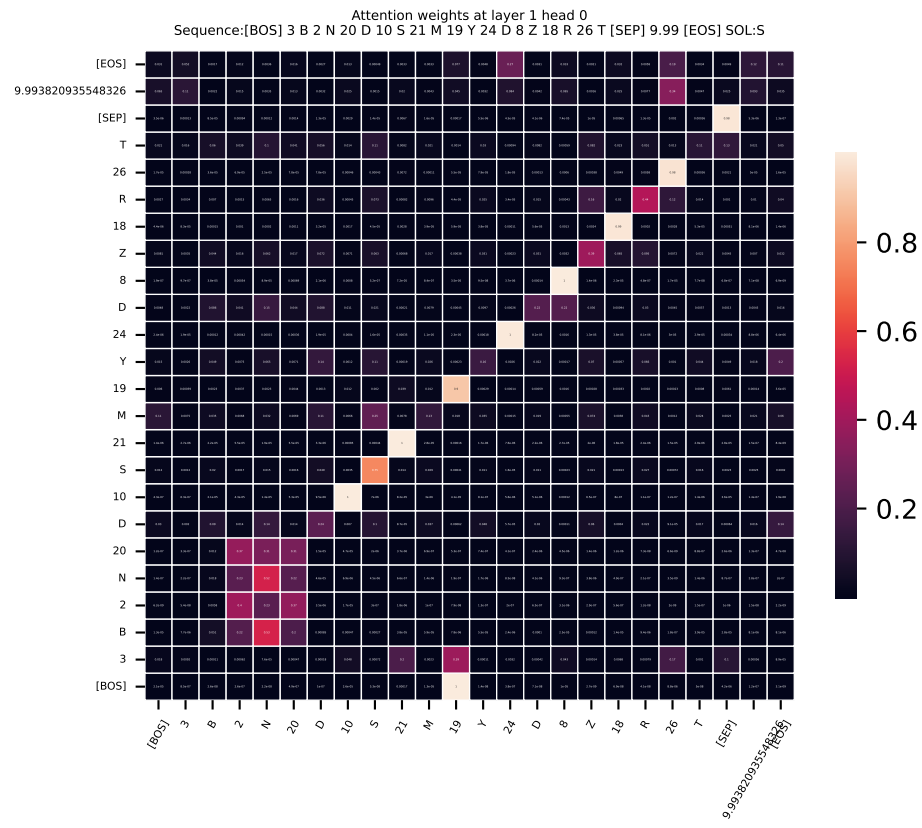

Attention weights at layer 0 head 0
Sequence:[BOS] 25 Q 5 G 18 G 23 C 2 S 19 N 1 N 11 J 7 B 22 N [SEP] 2.02 [EOS] SOL:S
[EOS]
2.0176396946639144
[SEP]
N
22
B
7
J
11
N
1
N
19
S
2
C
23
G
18
G
5
Q
25
[BOS]
[BOS]
25
Q
5
G
18
G
23
C
2
S
19
N
1
N
11
J
7
B
22
N
[SEP]
0.8
0.6
0.4
0.2

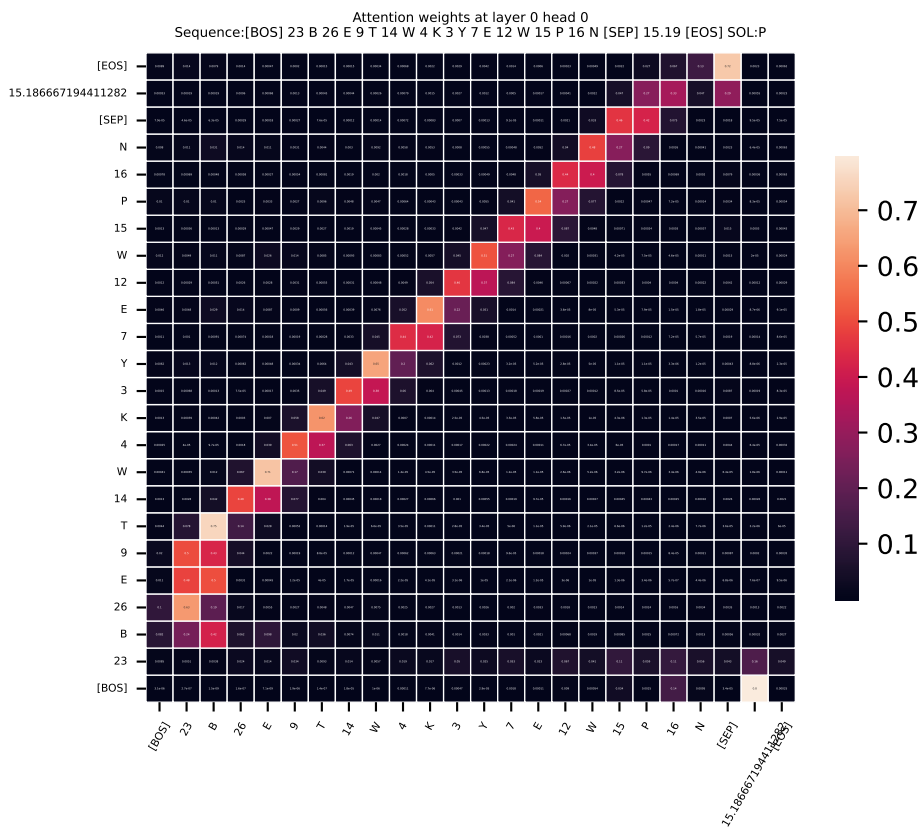

Attention weights at layer 1 head 0
Sequence:[BOS] 25 Q 5 G 18 G 23 C 2 S 19 N 1 N 11 J 7 B 22 N [SEP] 2.02 [EOS] SOL:S
[EOS]
2.0176396946639144
[SEP]
N
22
B
7
J
11
N
1
N
19
S
2
C
23
G
18
G
5
Q
25
[BOS]
[BOS]
25
Q
5
G
18
G
23
C
2
S
19
N
1
N
11
J
7
B
22
N
[SEP]
0.8
0.6
0.4
0.2

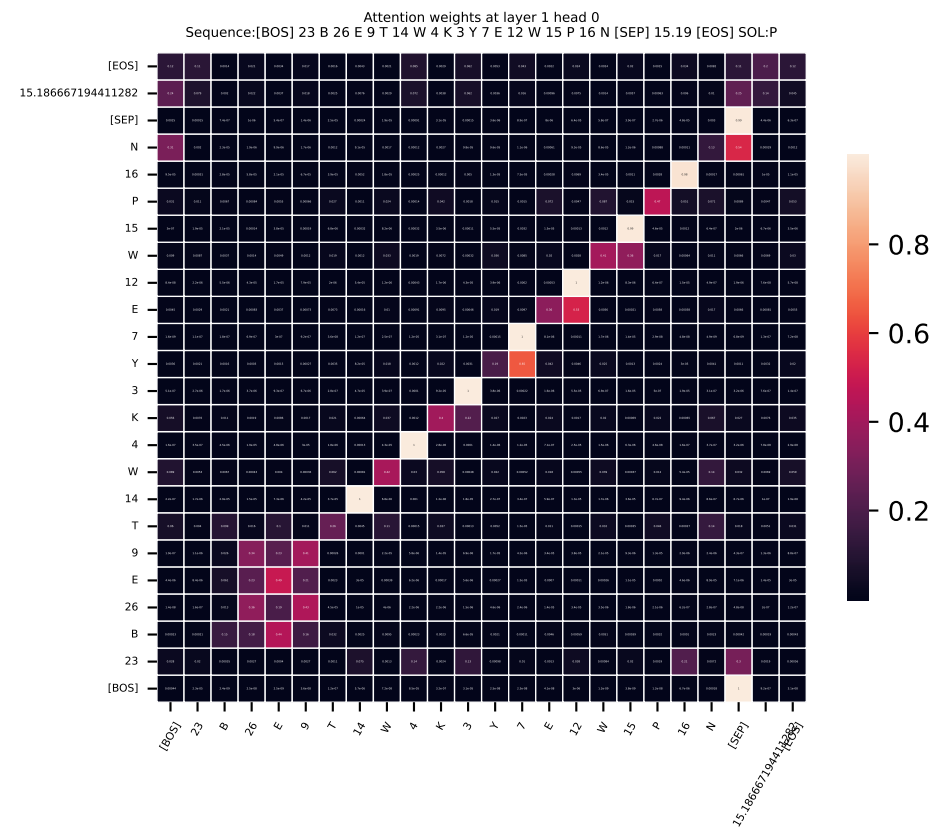

Attention weights at layer 0 head 0
Sequence:[BOS] 12 H 13 M 6 I 8 Q 9 T 2 B 22 S 24 P 25 T 11 V [SEP] 22.13 [EOS] SOL:S
[EOS]
22.13169004363405
[SEP]
V
11
T
25
P
24
S
22
B
2
T
9
Q
8
I
6
M
13
H
12
[BOS]
[BOS]
12
H
13
M
6
I
8
Q
9
T
2
B
22
S
24
P
25
T
11
V
[SEP]
22.13169004363405
[EOS]
0.8
0.6
0.4
0.2

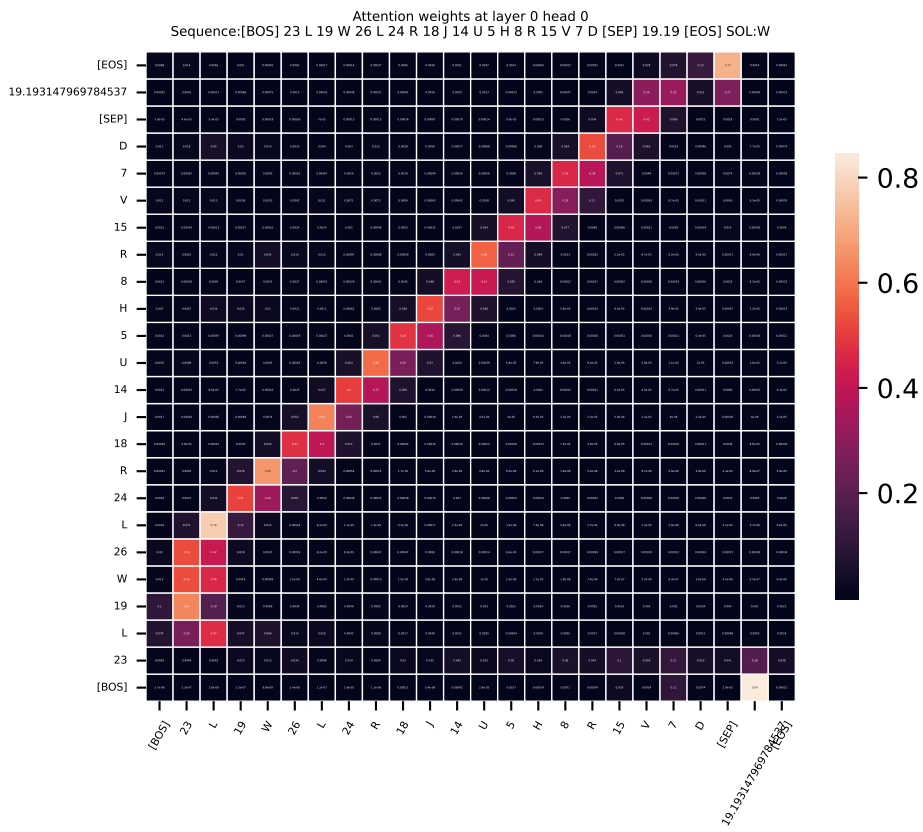

Attention weights at layer 1 head 0
Sequence:[BOS] 12 H 13 M 6 I 8 Q 9 T 2 B 22 S 24 P 25 T 11 V [SEP] 22.13 [EOS] SOL:S
[EOS]
22.13169004363405
[SEP]
V
11
T
25
P
24
S
22
B
2
T
9
Q
8
I
6
M
13
H
12
[BOS]
[BOS]
12
H
13
M
6
I
8
Q
9
T
2
B
22
S
24
P
25
T
11
V
[SEP]
22.13169004363405
[EOS]
0.8
0.6
0.4
0.2

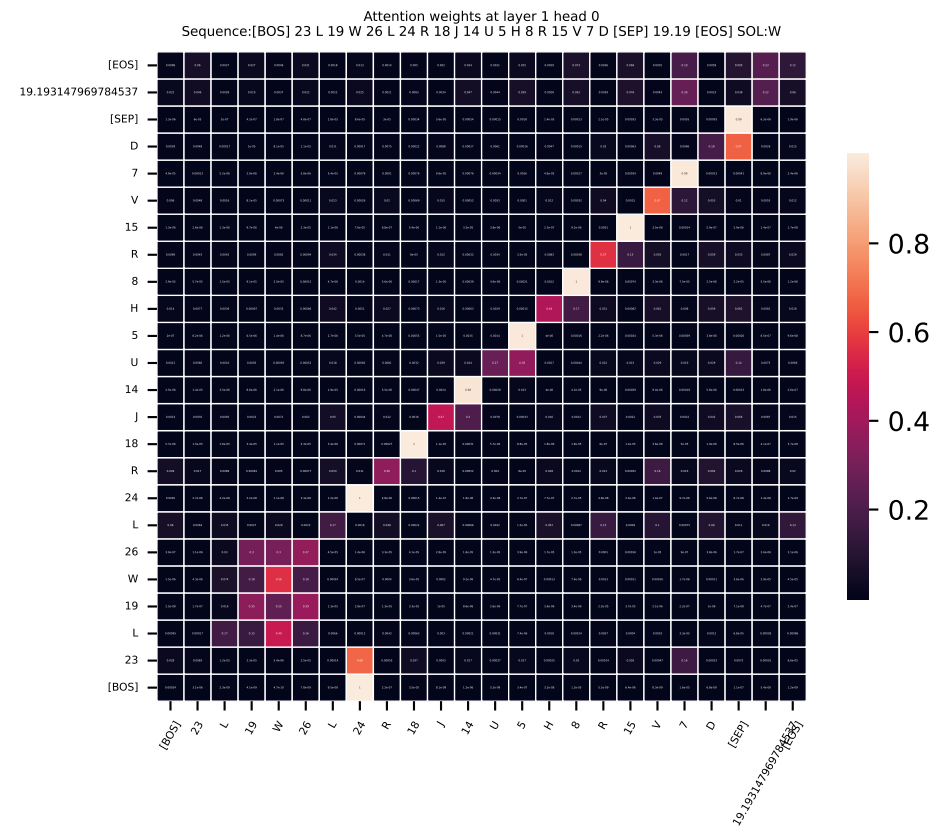

Attention weights at layer 0 head 0
Sequence:[BOS] 3 B 2 N 20 D 10 S 21 M 19 Y 24 D 8 Z 18 R 26 T [SEP] 9.99 [EOS] SOL:S
[EOS]
9.993820935548326
[SEP]
T
26
R
18
Z
8
D
24
Y
19
M
21
S
10
D
20
N
2
B
3
[BOS]
[BOS] 3 B 2 N 20 D 10 S 21 M 19 Y 24 D 8 Z 18 R 26 T [SEP] 9.99 [EOS]
0.7
0.6
0.5
0.4
0.3
0.2
0.1

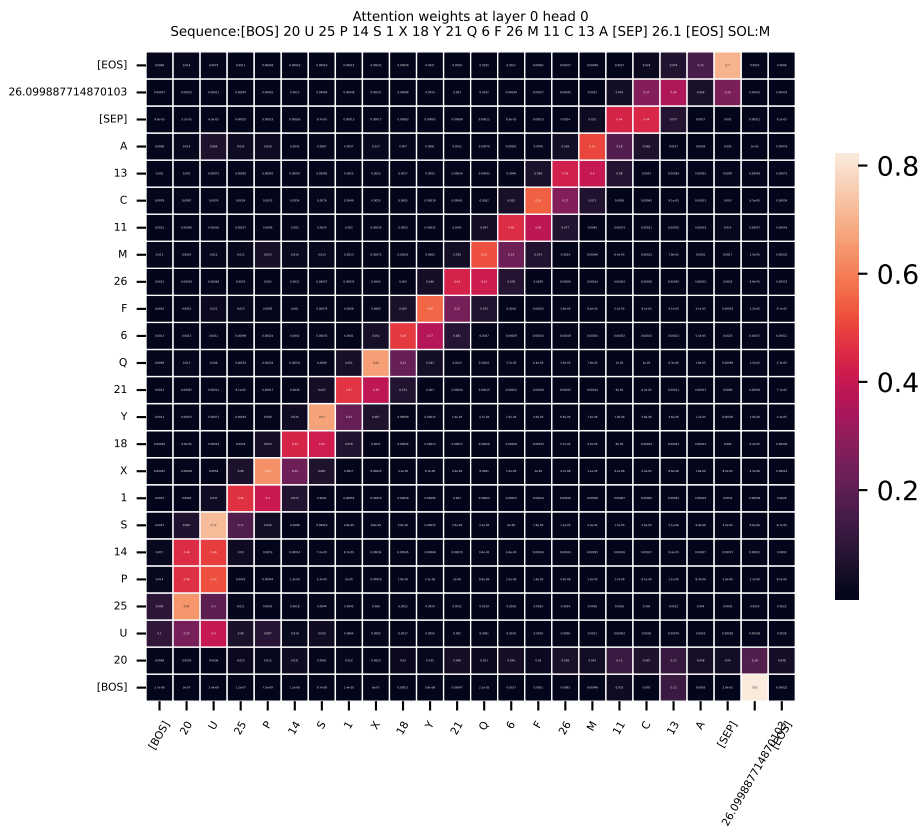

Attention weights at layer 1 head 0
Sequence:[BOS] 3 B 2 N 20 D 10 S 21 M 19 Y 24 D 8 Z 18 R 26 T [SEP] 9.99 [EOS] SOL:S
[EOS]
9.993820935548326
[SEP]
T
26
R
18
Z
8
D
24
Y
19
M
21
S
10
D
20
N
2
B
3
[BOS]
[BOS] 3 B 2 N 20 D 10 S 21 M 19 Y 24 D 8 Z 18 R 26 T [SEP] 9.99 [EOS]
0.8
0.6
0.4
0.2

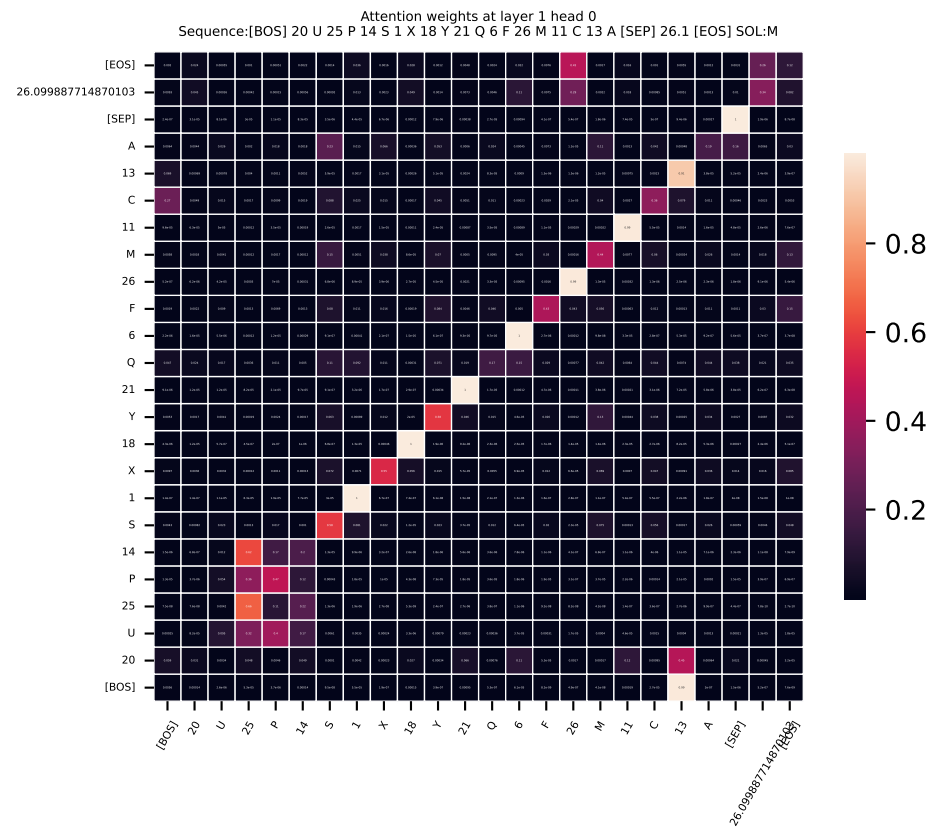

Attention weights at layer 0 head 0
Sequence:[BOS] 23 B 26 E 9 T 14 W 4 K 3 Y 7 E 12 W 15 P 16 N [SEP] 15.19 [EOS] SOL:P
[EOS]
15.186667194411282
[SEP]
N
16
P
15
W
12
E
7
Y
3
K
4
W
14
T
9
E
26
B
23
[BOS]
[BOS]
23
B
26
E
9
T
14
W
4
K
3
Y
7
E
12
W
15
P
16
N
[SEP]
0.7
0.6
0.5
0.4
0.3
0.2
0.1

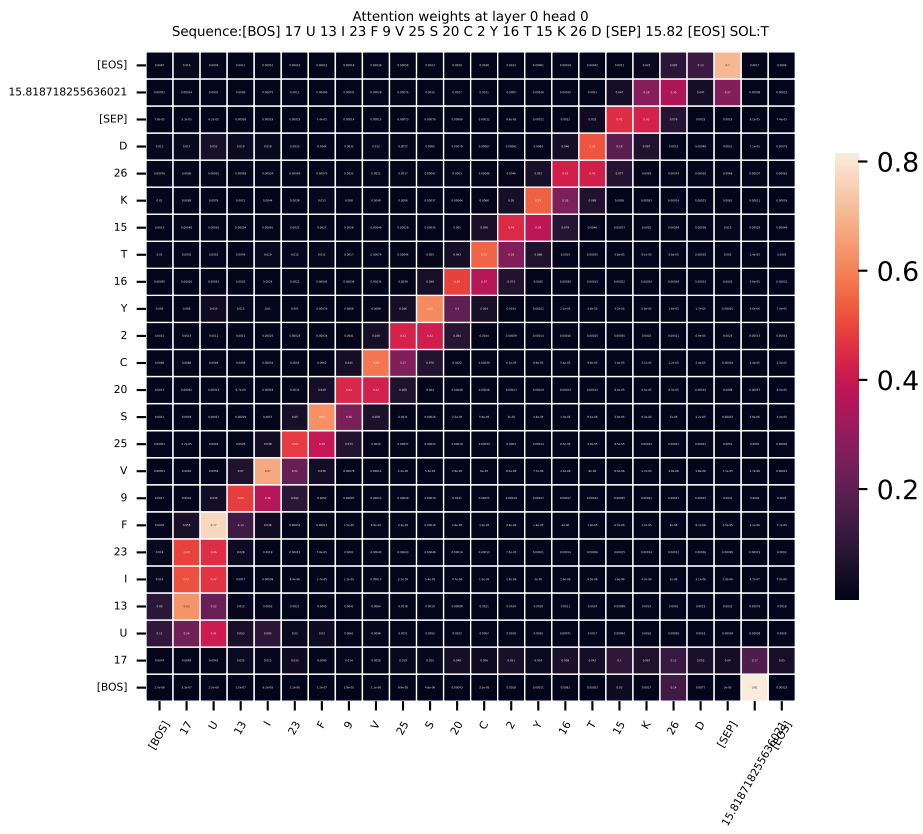

Attention weights at layer 1 head 0
Sequence:[BOS] 23 B 26 E 9 T 14 W 4 K 3 Y 7 E 12 W 15 P 16 N [SEP] 15.19 [EOS] SOL:P
[EOS]
15.186667194411282
[SEP]
N
16
P
15
W
12
E
7
Y
3
K
4
W
14
T
9
E
26
B
23
[BOS]
[BOS]
23
B
26
E
9
T
14
W
4
K
3
Y
7
E
12
W
15
P
16
N
[SEP]
0.8
0.6
0.4
0.2

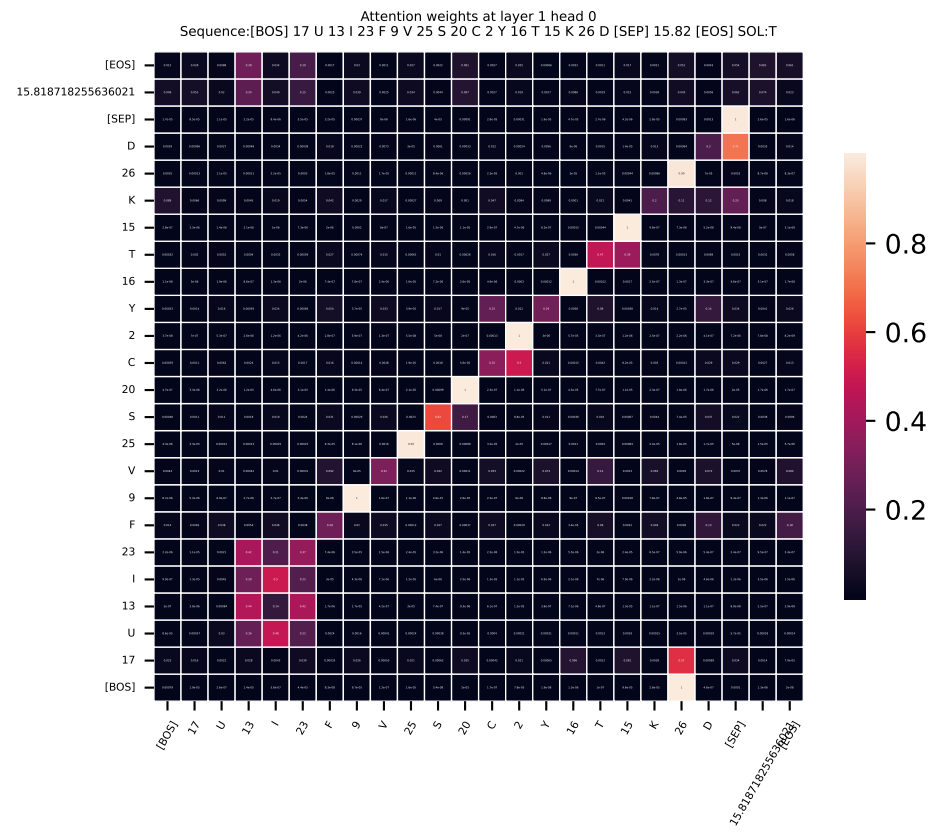

Attention weights at layer 0 head 0
Sequence:[BOS] 23 L 19 W 26 L 24 R 18 J 14 U 5 H 8 R 15 V 7 D [SEP] 19.19 [EOS] SOL:W
[EOS]
19.193147969784537
[SEP]
D
7
V
15
R
8
H
5
U
14
J
18
R
24
L
26
W
19
L
23
[BOS]
[BOS]
23
L
19
W
26
L
24
R
18
J
14
U
5
H
8
R
15
V
7
D
[SEP]
19.193147969784537
[EOS]
0.8
0.6
0.4
0.2

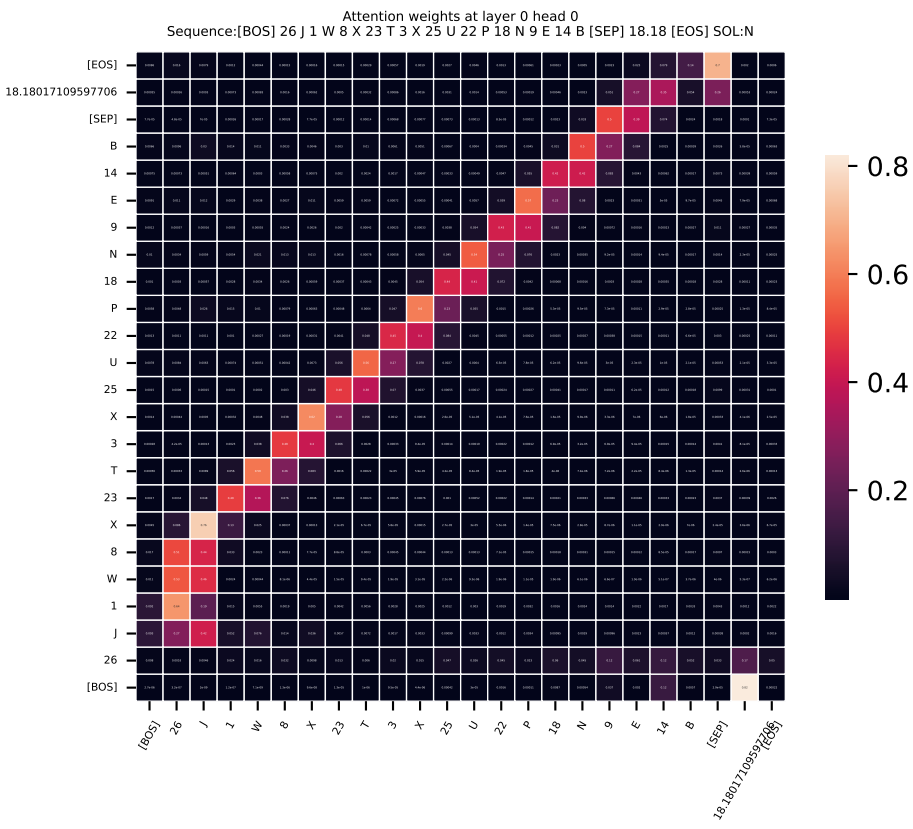

Attention weights at layer 1 head 0
Sequence:[BOS] 23 L 19 W 26 L 24 R 18 J 14 U 5 H 8 R 15 V 7 D [SEP] 19.19 [EOS] SOL:W
[EOS]
19.193147969784537
[SEP]
D
7
V
15
R
8
H
5
U
14
J
18
R
24
L
26
W
19
L
23
[BOS]
[BOS]
23
L
19
W
26
L
24
R
18
J
14
U
5
H
8
R
15
V
7
D
[SEP]
19.193147969784537
[EOS]
0.8
0.6
0.4
0.2

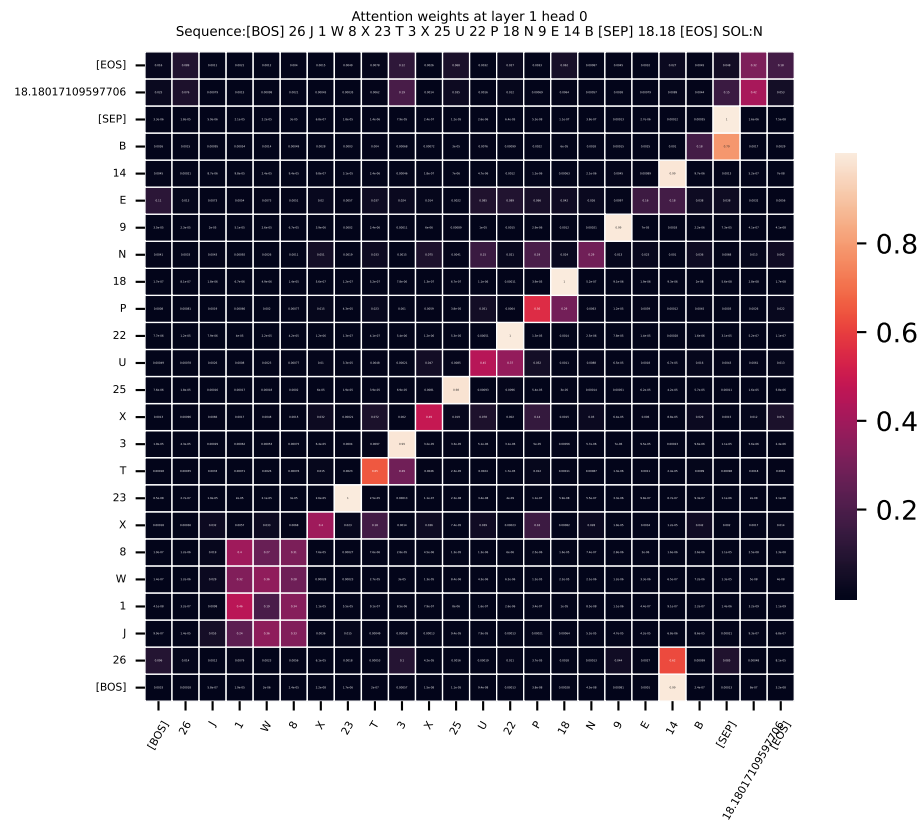

Attention weights at layer 0 head 0
Sequence:[BOS] 20 U 25 P 14 S 1 X 18 Y 21 Q 6 F 26 M 11 C 13 A [SEP] 26.1 [EOS] SOL:M
[EOS]
26.099887714870103
[SEP]
A
13
C
11
M
26
F
6
Q
21
Y
18
X
1
S
14
P
25
U
20
[BOS]
[BOS] 20 U 25 P 14 S 1 X 18 Y 21 Q 6 F 26 M 11 C 13 A [SEP] 26.1 [EOS] [BOS]
0.8
0.6
0.4
0.2

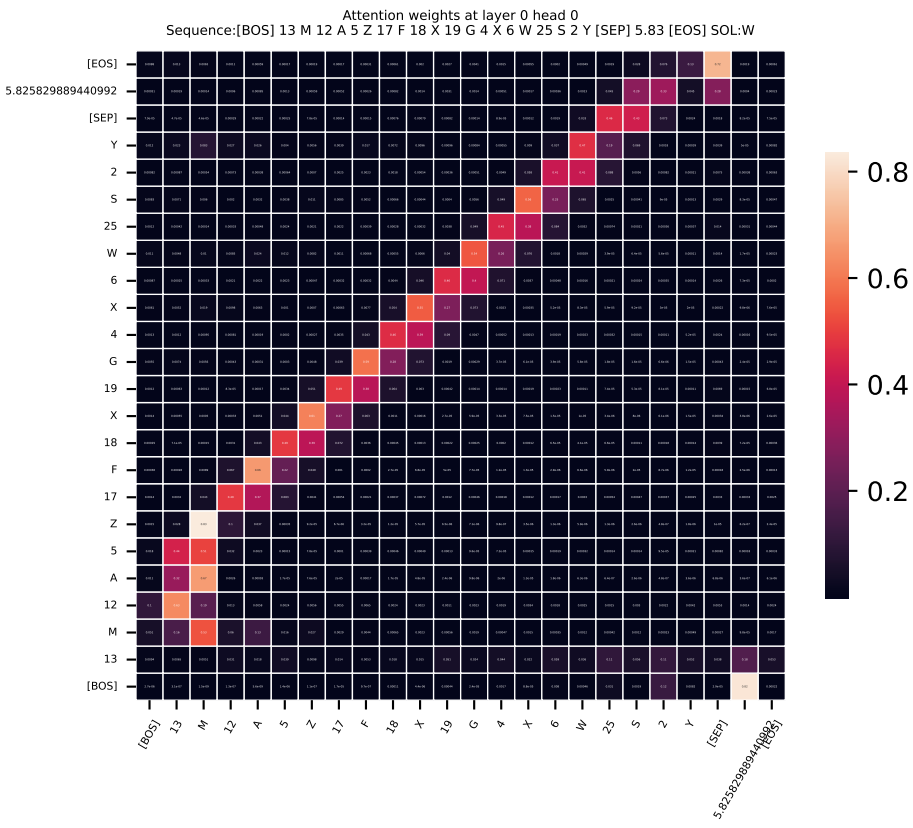

Attention weights at layer 1 head 0
Sequence:[BOS] 20 U 25 P 14 S 1 X 18 Y 21 Q 6 F 26 M 11 C 13 A [SEP] 26.1 [EOS] SOL:M
[EOS]
26.099887714870103
[SEP]
A
13
C
11
M
26
F
6
Q
21
Y
18
X
1
S
14
P
25
U
20
[BOS]
[BOS] 20 U 25 P 14 S 1 X 18 Y 21 Q 6 F 26 M 11 C 13 A [SEP] 26.1 [EOS] [BOS]
0.8
0.6
0.4
0.2

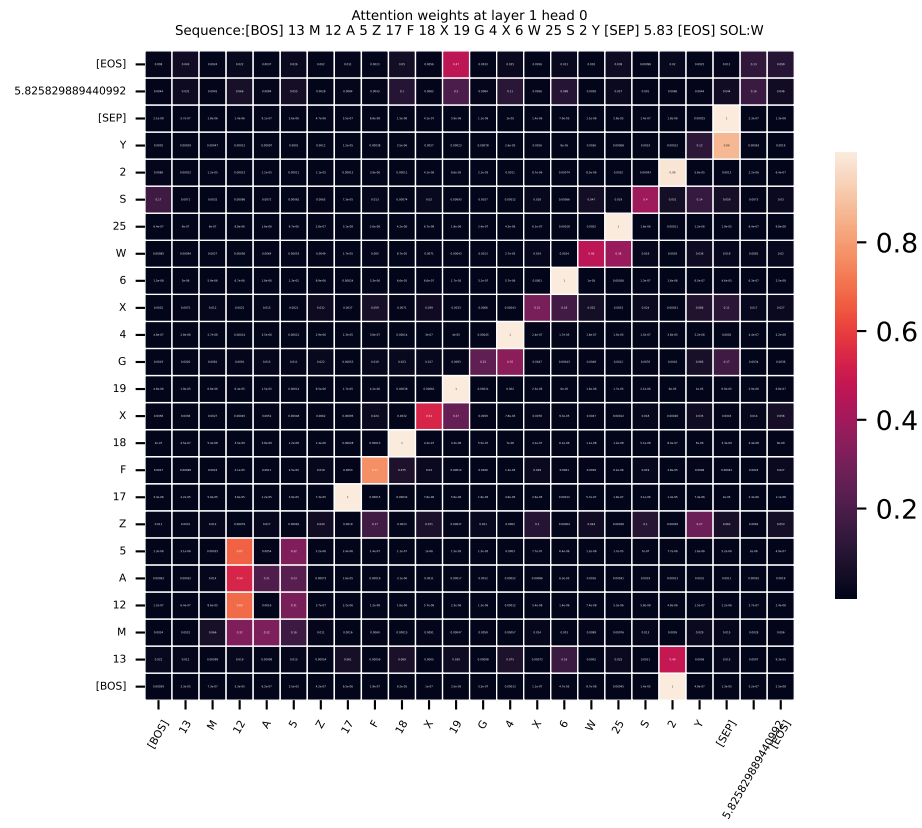

Attention weights at layer 0 head 0
Sequence:[BOS] 17 U 13 I 23 F 9 V 25 S 20 C 2 Y 16 T 15 K 26 D [SEP] 15.82 [EOS] SOL:T
[EOS]
15.818718255636021
[SEP]
D
26
K
15
T
16
Y
2
C
20
S
25
V
9
F
23
I
13
U
17
[BOS]
0.8
0.6
0.4
0.2

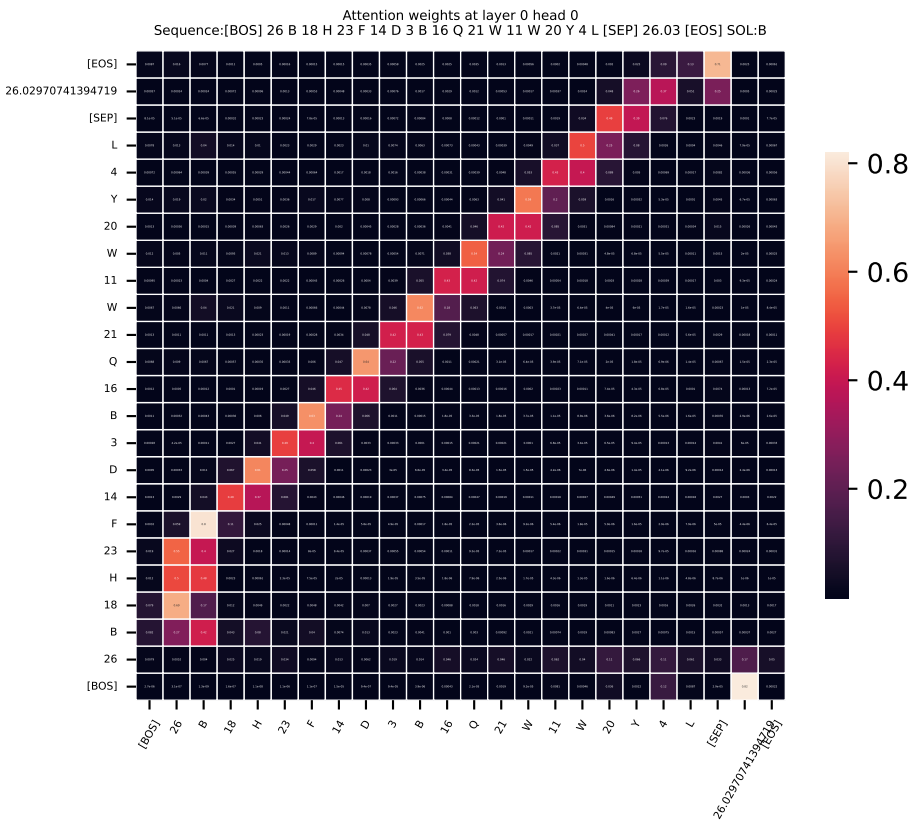

Attention weights at layer 1 head 0
Sequence:[BOS] 17 U 13 I 23 F 9 V 25 S 20 C 2 Y 16 T 15 K 26 D [SEP] 15.82 [EOS] SOL:T
[EOS]
15.818718255636021
[SEP]
D
26
K
15
T
16
Y
2
C
20
S
25
V
9
F
23
I
13
U
17
[BOS]
0.8
0.6
0.4
0.2

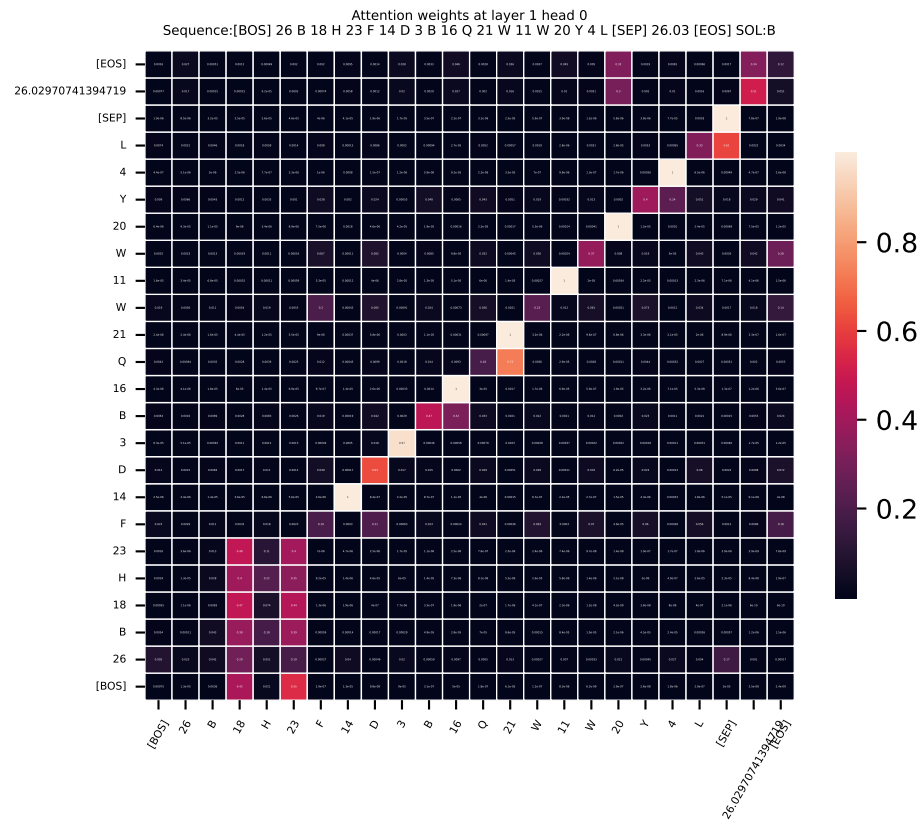

Attention weights at layer 0 head 0
Sequence:[BOS] 26 J 1 W 8 X 23 T 3 X 25 U 22 P 18 N 9 E 14 B [SEP] 18.18 [EOS] SOL:N
[EOS]
18.18017109597706
[SEP]
B
14
E
9
N
18
P
22
U
25
X
3
T
23
X
8
W
1
J
26
[BOS]
[BOS] 26 J 1 W 8 X 23 T 3 X 25 U 22 P 18 N 9 E 14 B [SEP] 18.18 [EOS]
0.8
0.6
0.4
0.2

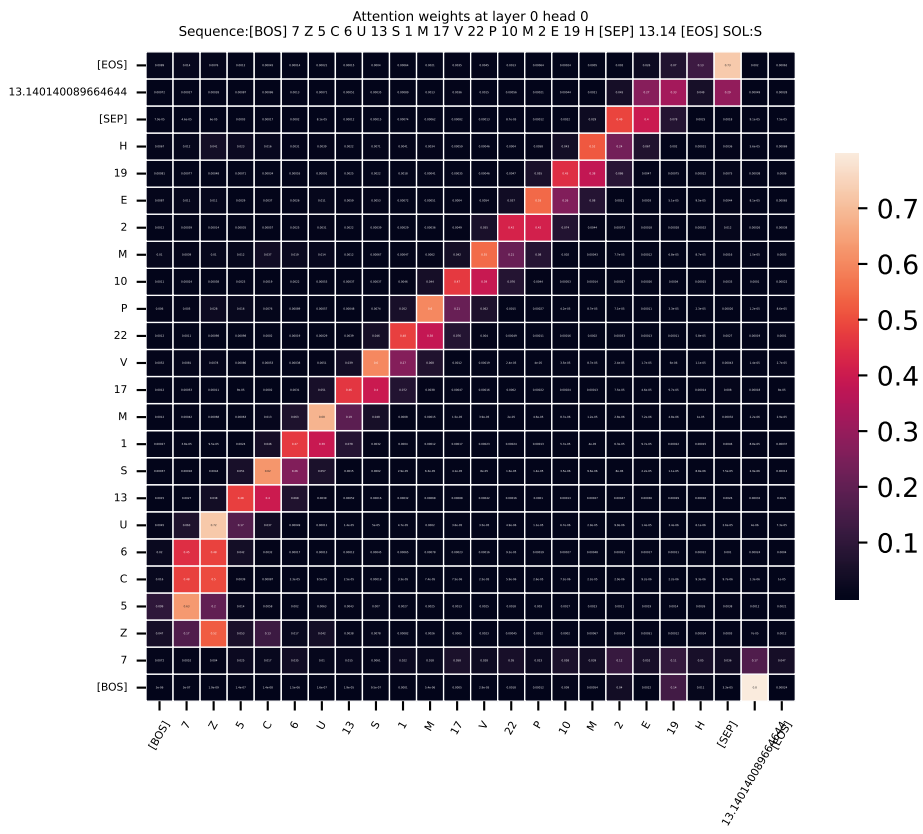

Attention weights at layer 1 head 0
Sequence:[BOS] 26 J 1 W 8 X 23 T 3 X 25 U 22 P 18 N 9 E 14 B [SEP] 18.18 [EOS] SOL:N
[EOS]
18.18017109597706
[SEP]
B
14
E
9
N
18
P
22
U
25
X
3
T
23
X
8
W
1
J
26
[BOS]
[BOS] 26 J 1 W 8 X 23 T 3 X 25 U 22 P 18 N 9 E 14 B [SEP] 18.18 [EOS]
0.8
0.6
0.4
0.2

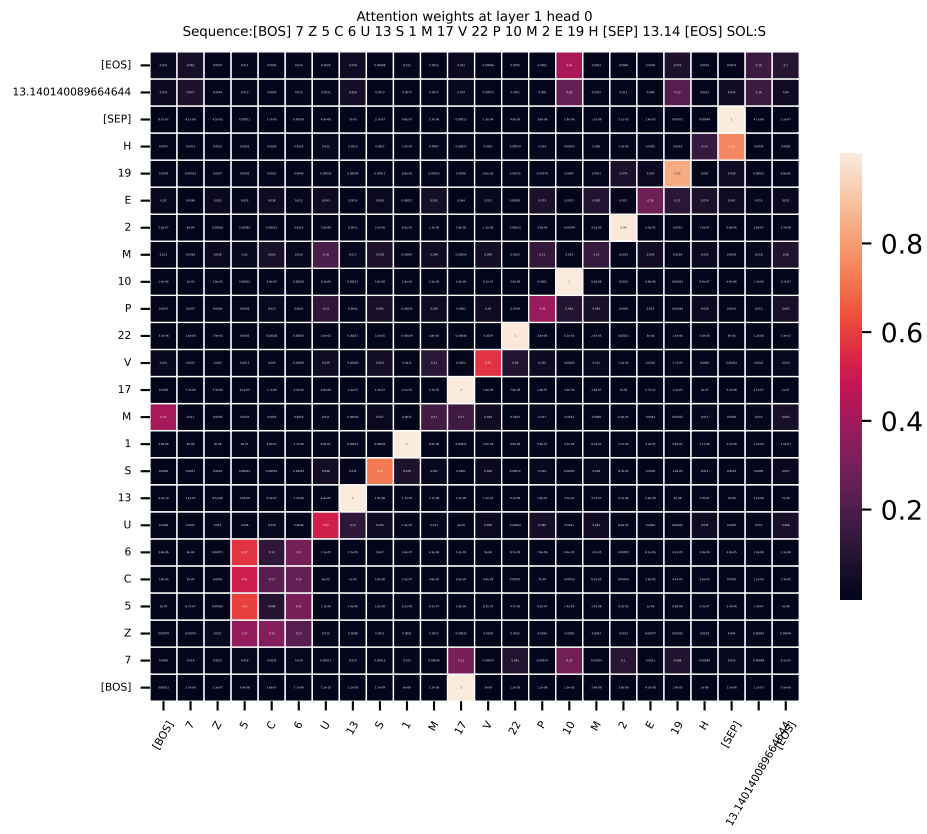

Attention weights at layer 0 head 0
Sequence:[BOS] 13 M 12 A 5 Z 17 F 18 X 19 G 4 X 6 W 25 S 2 Y [SEP] 5.83 [EOS] SOL:W
[EOS]
5.825829889440992
[SEP]
Y
2
S
25
W
6
X
4
G
19
X
18
F
17
Z
5
A
12
M
13
[BOS]
[BOS] 13 M 12 A 5 Z 17 F 18 X 19 G 4 X 6 W 25 S 2 Y [SEP] 5.83 [EOS] SOL:W
0.8
0.6
0.4
0.2

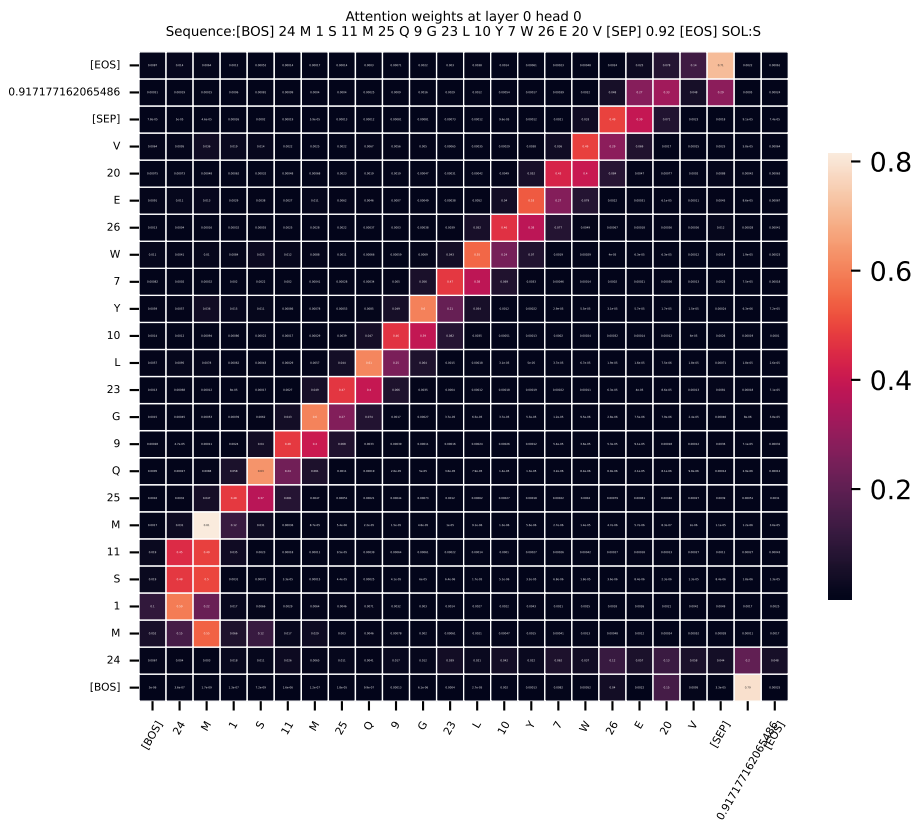

Attention weights at layer 1 head 0
Sequence:[BOS] 13 M 12 A 5 Z 17 F 18 X 19 G 4 X 6 W 25 S 2 Y [SEP] 5.83 [EOS] SOL:W
[EOS]
5.825829889440992
[SEP]
Y
2
S
25
W
6
X
4
G
19
X
18
F
17
Z
5
A
12
M
13
[BOS]
[BOS] 13 M 12 A 5 Z 17 F 18 X 19 G 4 X 6 W 25 S 2 Y [SEP] 5.83 [EOS] SOL:W
0.8
0.6
0.4
0.2

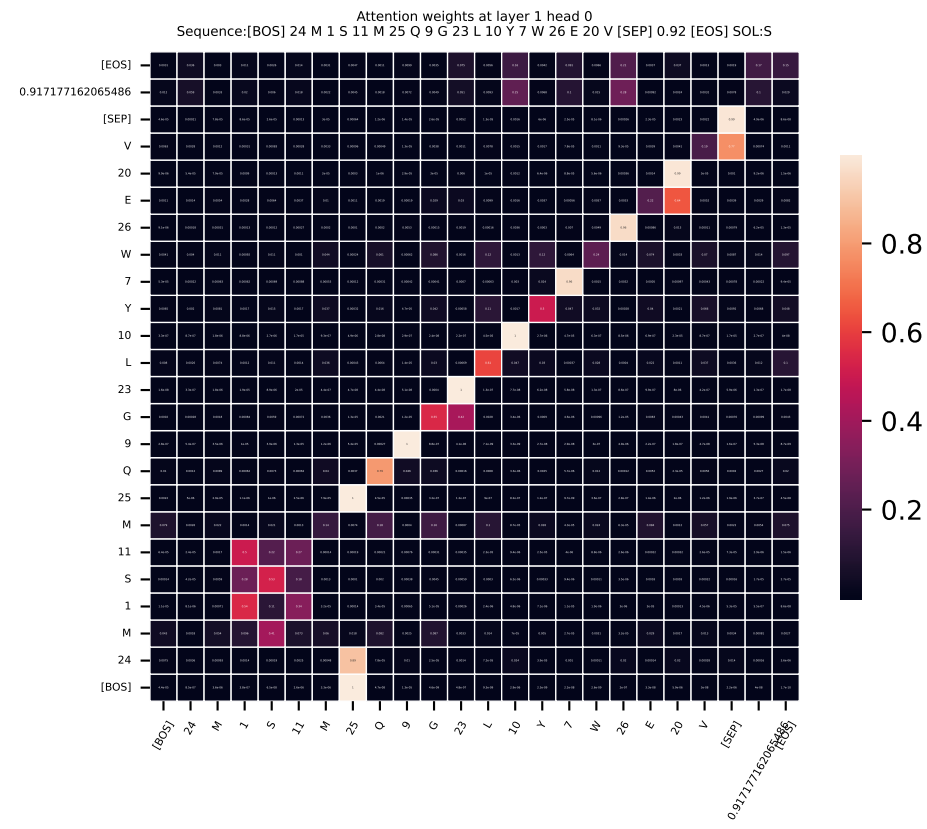

Attention weights at layer 0 head 0
Sequence:[BOS] 26 B 18 H 23 F 14 D 3 B 16 Q 21 W 11 W 20 Y 4 L [SEP] 26.03 [EOS] SOL:B
[EOS]
26.02970741394719
[SEP]
L
4
Y
20
W
11
W
21
Q
16
B
3
D
14
F
23
H
18
B
26
[BOS]
[BOS]
26
B
18
H
23
F
14
D
3
B
16
Q
21
W
11
W
20
Y
4
L
[SEP]
26.02970741394719
[EOS]
0.8
0.6
0.4
0.2

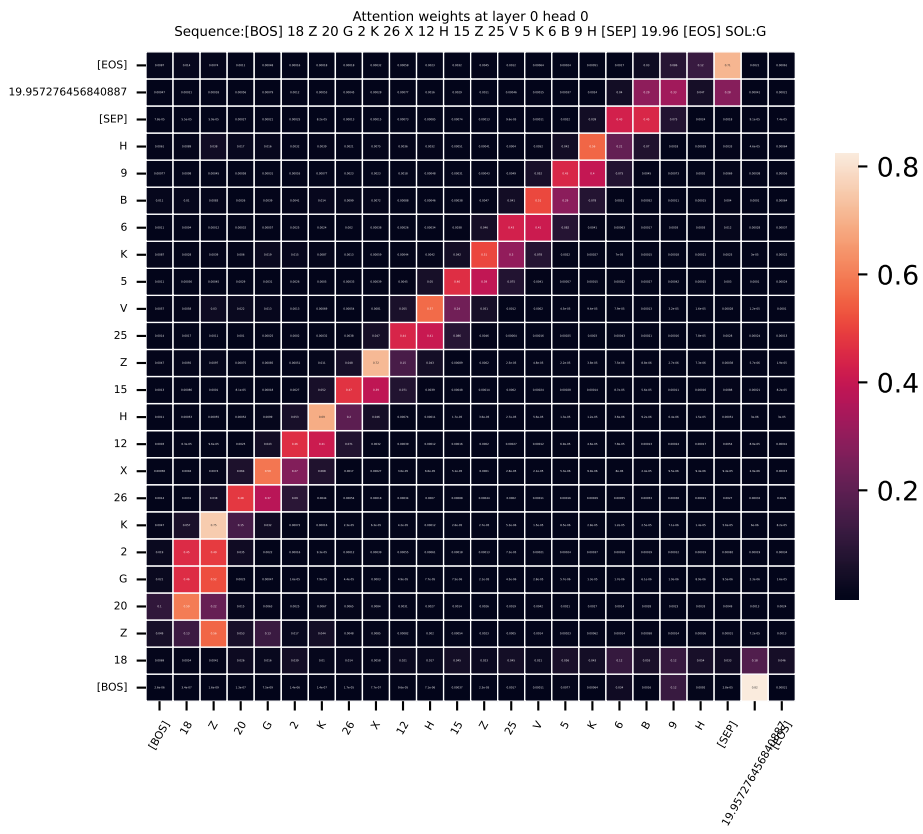

Attention weights at layer 1 head 0
Sequence:[BOS] 26 B 18 H 23 F 14 D 3 B 16 Q 21 W 11 W 20 Y 4 L [SEP] 26.03 [EOS] SOL:B
[EOS]
26.02970741394719
[SEP]
L
4
Y
20
W
11
W
21
Q
16
B
3
D
14
F
23
H
18
B
26
[BOS]
[BOS]
26
B
18
H
23
F
14
D
3
B
16
Q
21
W
11
W
20
Y
4
L
[SEP]
26.02970741394719
[EOS]
0.8
0.6
0.4
0.2

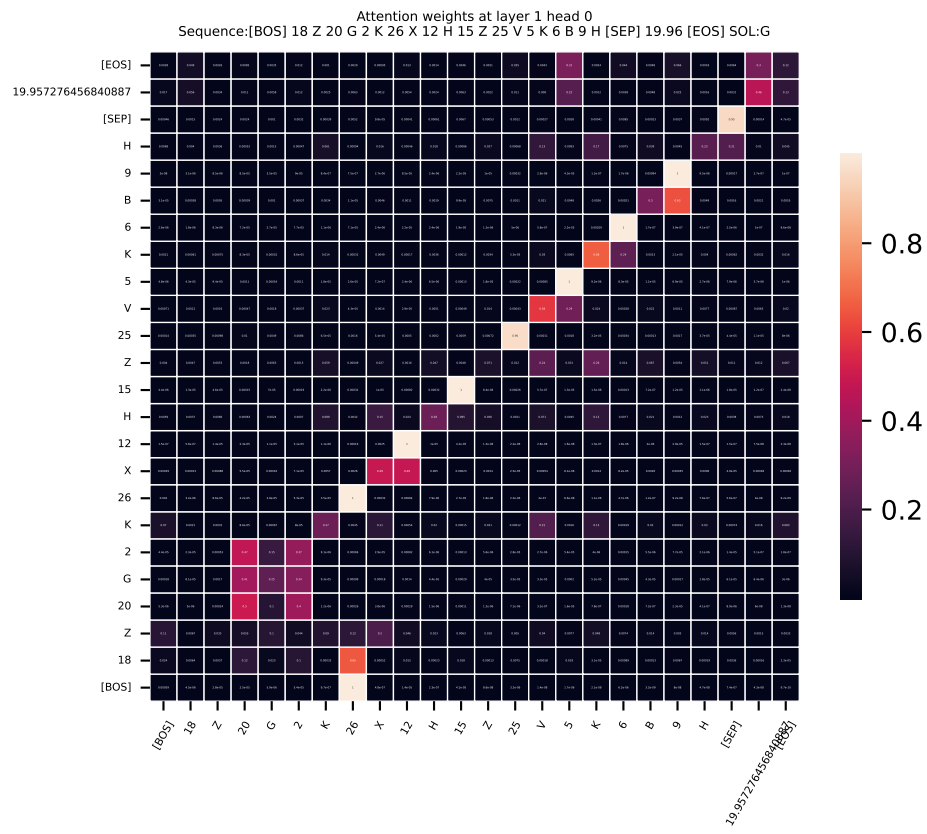

Attention weights at layer 0 head 0
Sequence:[BOS] 7 Z 5 C 6 U 13 S 1 M 17 V 22 P 10 M 2 E 19 H [SEP] 13.14 [EOS] SOL:S
[EOS]
13.140140089664644
[SEP]
H
19
E
2
M
10
P
22
V
17
M
1
S
13
U
6
C
5
Z
7
[BOS]
[BOS] 7 Z 5 C 6 U 13 S 1 M 17 V 22 P 10 M 2 E 19 H [SEP] 13.14 [EOS] SOL:S
0.7
0.6
0.5
0.4
0.3
0.2
0.1

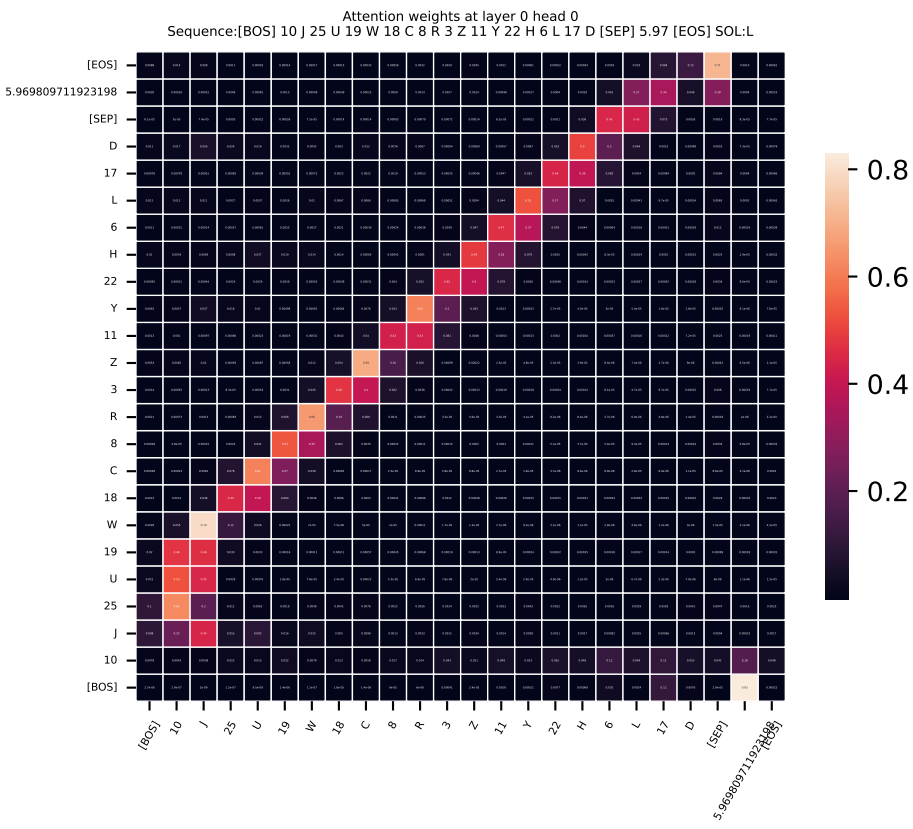

Attention weights at layer 1 head 0
Sequence:[BOS] 7 Z 5 C 6 U 13 S 1 M 17 V 22 P 10 M 2 E 19 H [SEP] 13.14 [EOS] SOL:S
[EOS]
13.140140089664644
[SEP]
H
19
E
2
M
10
P
22
V
17
M
1
S
13
U
6
C
5
Z
7
[BOS]
[BOS] 7 Z 5 C 6 U 13 S 1 M 17 V 22 P 10 M 2 E 19 H [SEP] 13.14 [EOS] SOL:S
0.8
0.6
0.4
0.2

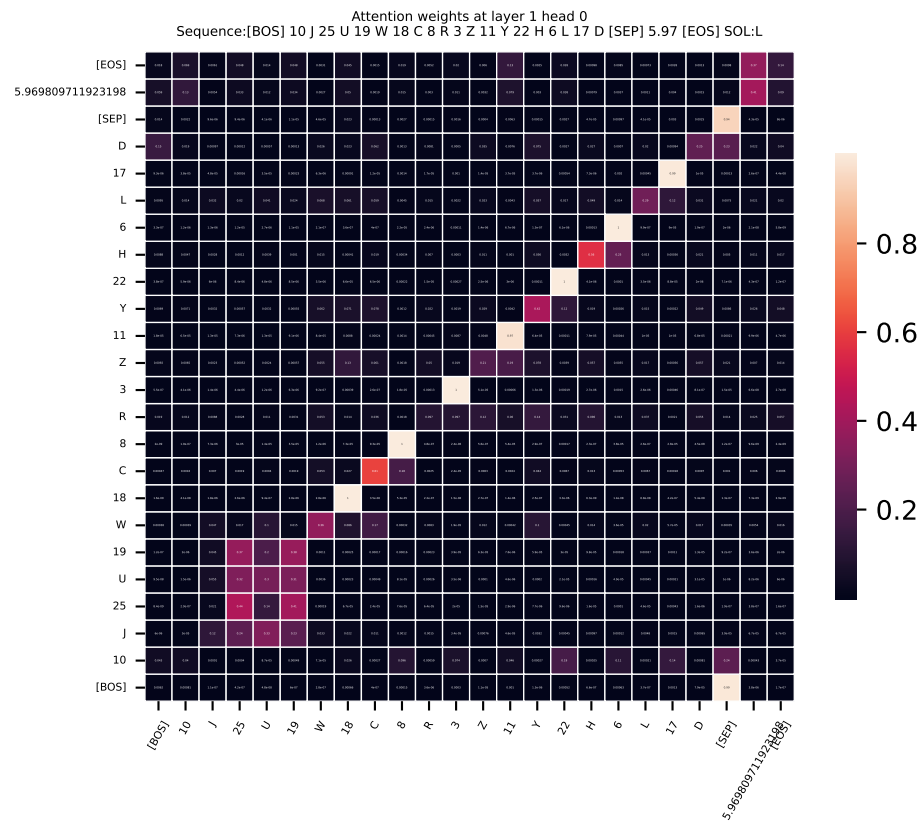

Attention weights at layer 0 head 0
Sequence:[BOS] 24 M 1 S 11 M 25 Q 9 G 23 L 10 Y 7 W 26 E 20 V [SEP] 0.92 [EOS] SOL:S
[EOS]
0.917177162065486
[SEP]
V
20
E
26
W
7
Y
10
L
23
G
9
Q
25
M
11
S
1
M
24
[BOS]
0.8
0.6
0.4
0.2

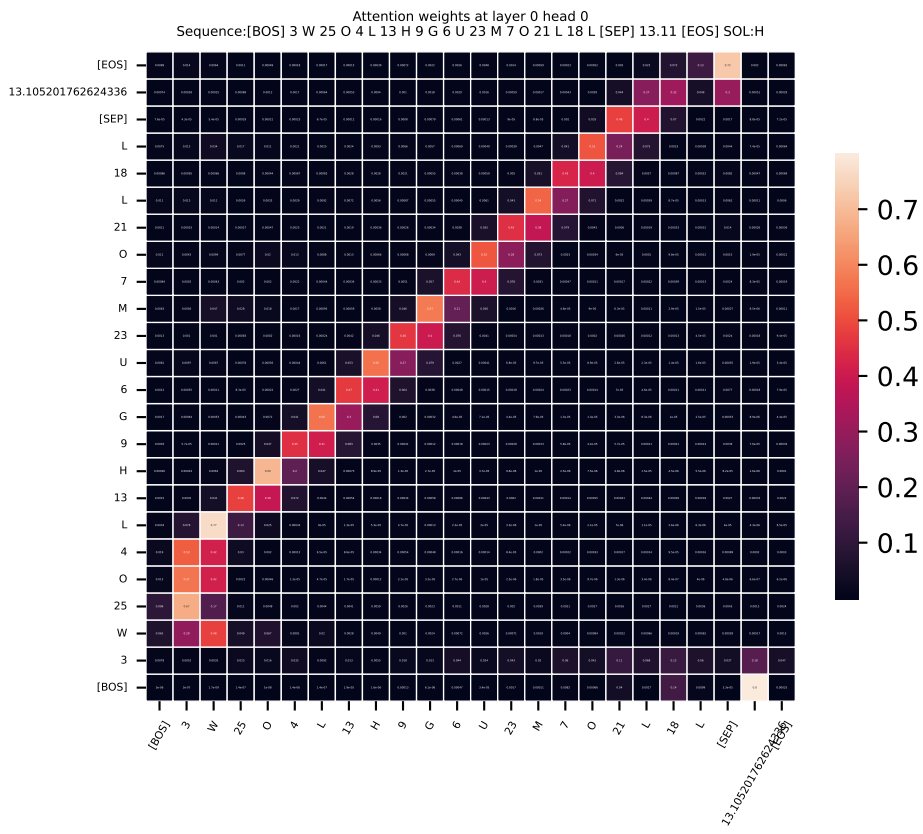

Attention weights at layer 1 head 0
Sequence:[BOS] 24 M 1 S 11 M 25 Q 9 G 23 L 10 Y 7 W 26 E 20 V [SEP] 0.92 [EOS] SOL:S
[EOS]
0.917177162065486
[SEP]
V
20
E
26
W
7
Y
10
L
23
G
9
Q
25
M
11
S
1
M
24
[BOS]
0.8
0.6
0.4
0.2

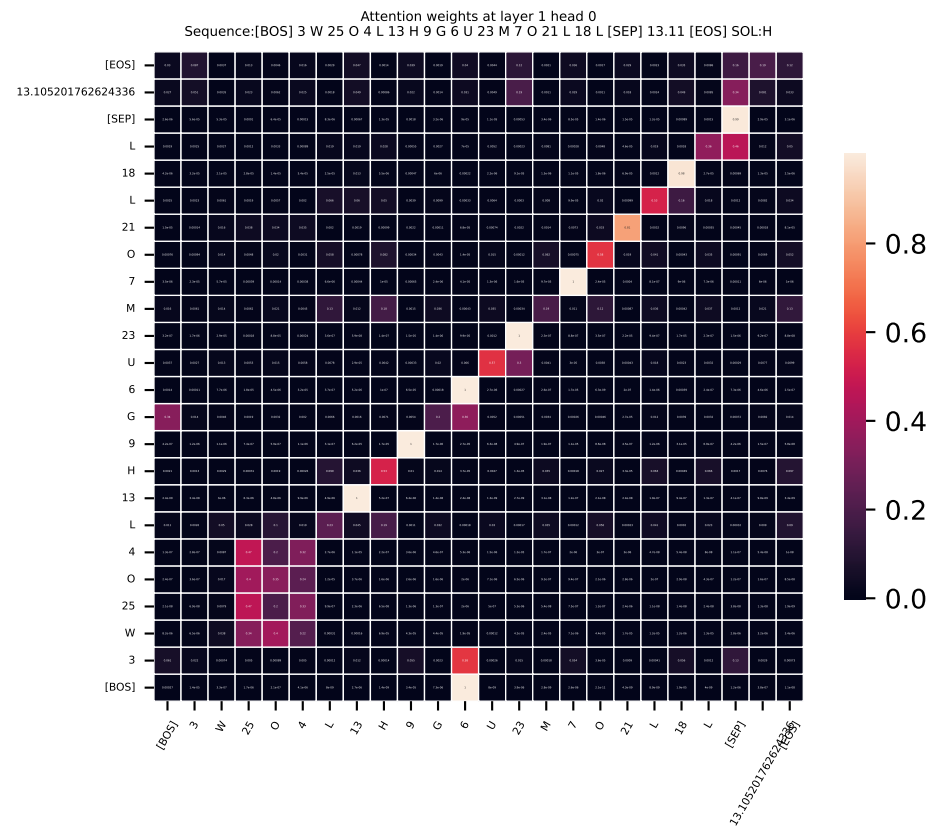

Attention weights at layer 0 head 0
Sequence:[BOS] 18 Z 20 G 2 K 26 X 12 H 15 Z 25 V 5 K 6 B 9 H [SEP] 19.96 [EOS] SOL:G
[EOS]
19.957276456840887
[SEP]
H
9
B
6
K
5
V
25
Z
15
H
12
X
26
K
2
G
20
Z
18
[BOS]
[BOS] 18 Z 20 G 2 K 26 X 12 H 15 Z 25 V 5 K 6 B 9 H [SEP] 19.96 [EOS] SOL:G
0.8
0.6
0.4
0.2

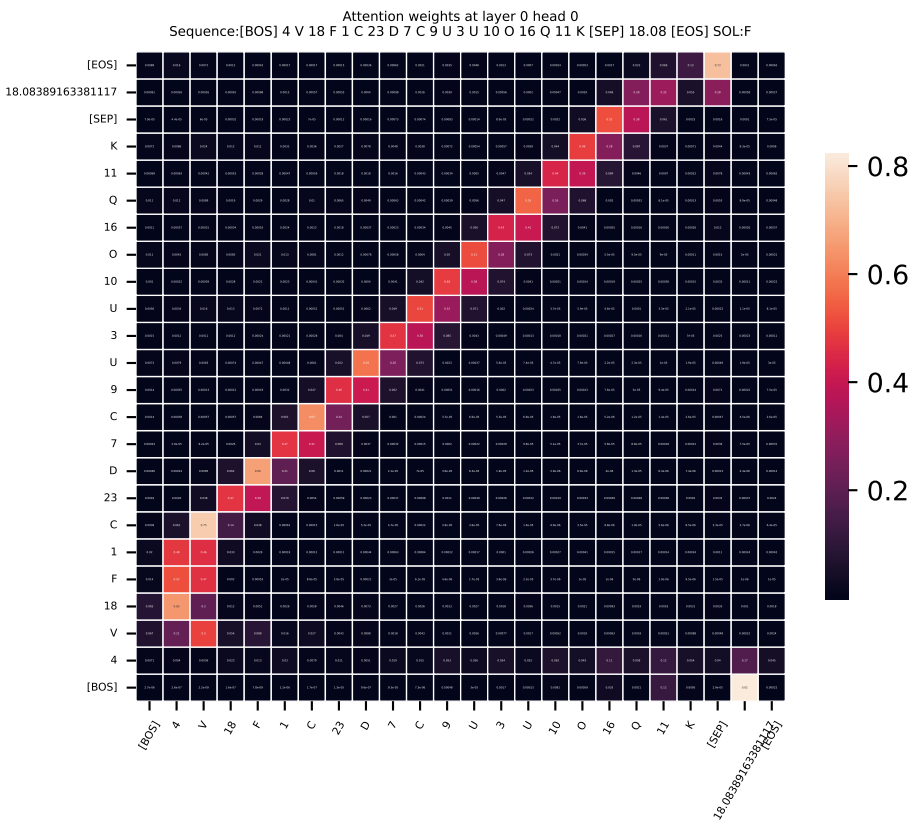

Attention weights at layer 1 head 0
Sequence:[BOS] 18 Z 20 G 2 K 26 X 12 H 15 Z 25 V 5 K 6 B 9 H [SEP] 19.96 [EOS] SOL:G
[EOS]
19.957276456840887
[SEP]
H
9
B
6
K
5
V
25
Z
15
H
12
X
26
K
2
G
20
Z
18
[BOS]
[BOS] 18 Z 20 G 2 K 26 X 12 H 15 Z 25 V 5 K 6 B 9 H [SEP] 19.96 [EOS] SOL:G
0.8
0.6
0.4
0.2

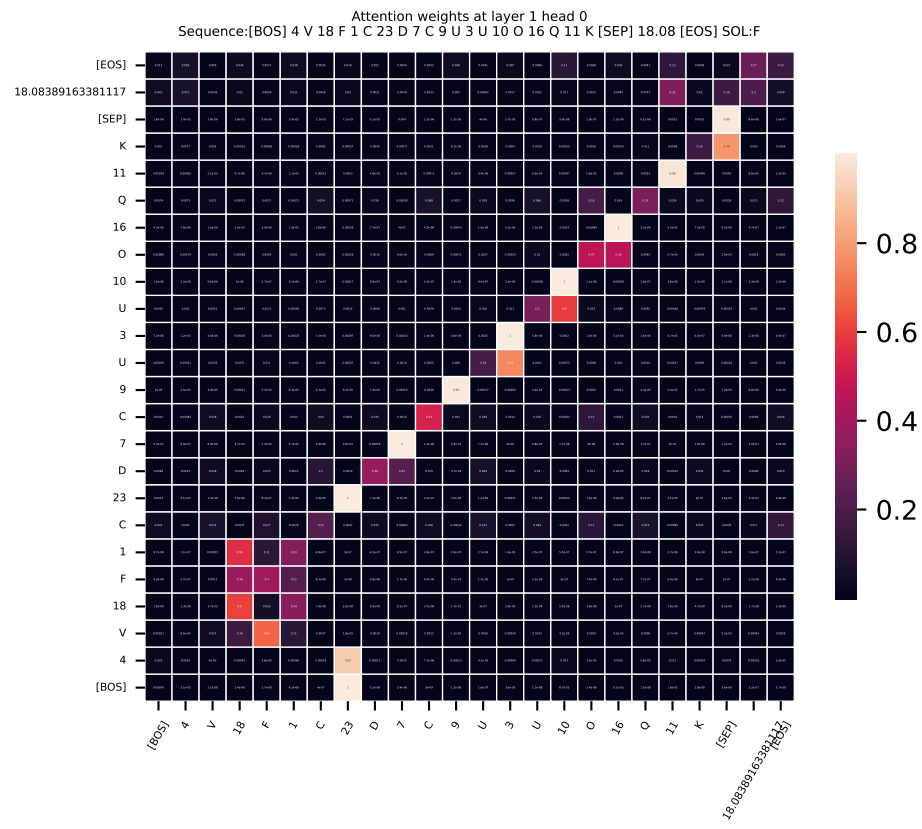

Attention weights at layer 0 head 0
Sequence:[BOS] 10 J 25 U 19 W 18 C 8 R 3 Z 11 Y 22 H 6 L 17 D [SEP] 5.97 [EOS] SOL:L
[EOS]
5.969809711923198
[SEP]
D
17
L
6
H
22
Y
11
Z
3
R
8
C
18
W
19
U
25
J
10
[BOS]
0.8
0.6
0.4
0.2

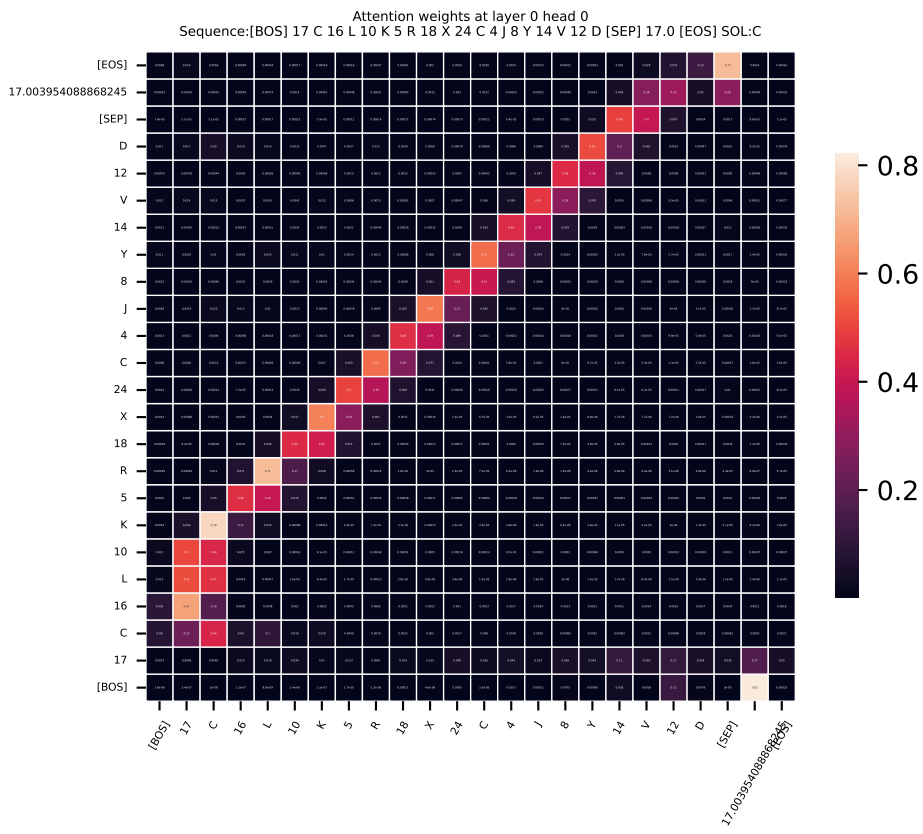

Attention weights at layer 1 head 0
Sequence:[BOS] 10 J 25 U 19 W 18 C 8 R 3 Z 11 Y 22 H 6 L 17 D [SEP] 5.97 [EOS] SOL:L
[EOS]
5.969809711923198
[SEP]
D
17
L
6
H
22
Y
11
Z
3
R
8
C
18
W
19
U
25
J
10
[BOS]
0.8
0.6
0.4
0.2

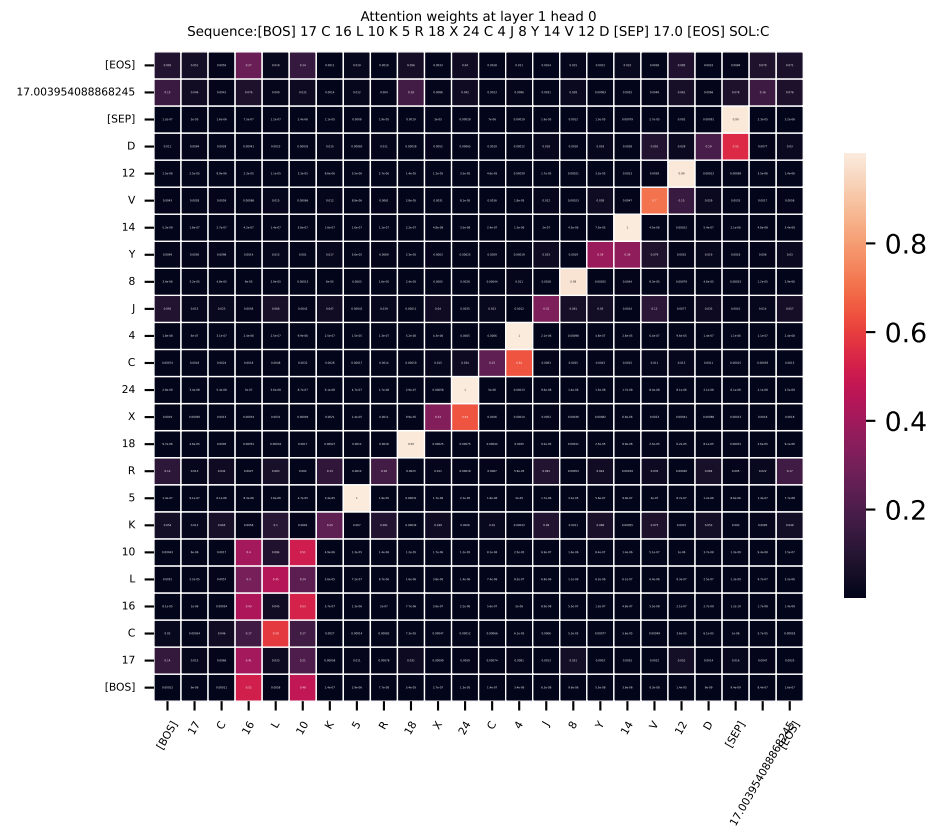

Attention weights at layer 0 head 0
Sequence:[BOS] 3 W 25 O 4 L 13 H 9 G 6 U 23 M 7 O 21 L 18 L [SEP] 13.11 [EOS] SOL:H
[EOS]
13.105201762624336
[SEP]
L
18
L
21
O
7
M
23
U
6
G
9
H
13
L
4
O
25
W
3
[BOS]
[BOS] 3 W 25 O 4 L 13 H 9 G 6 U 23 M 7 O 21 L 18 L [SEP]
0.7
0.6
0.5
0.4
0.3
0.2
0.1

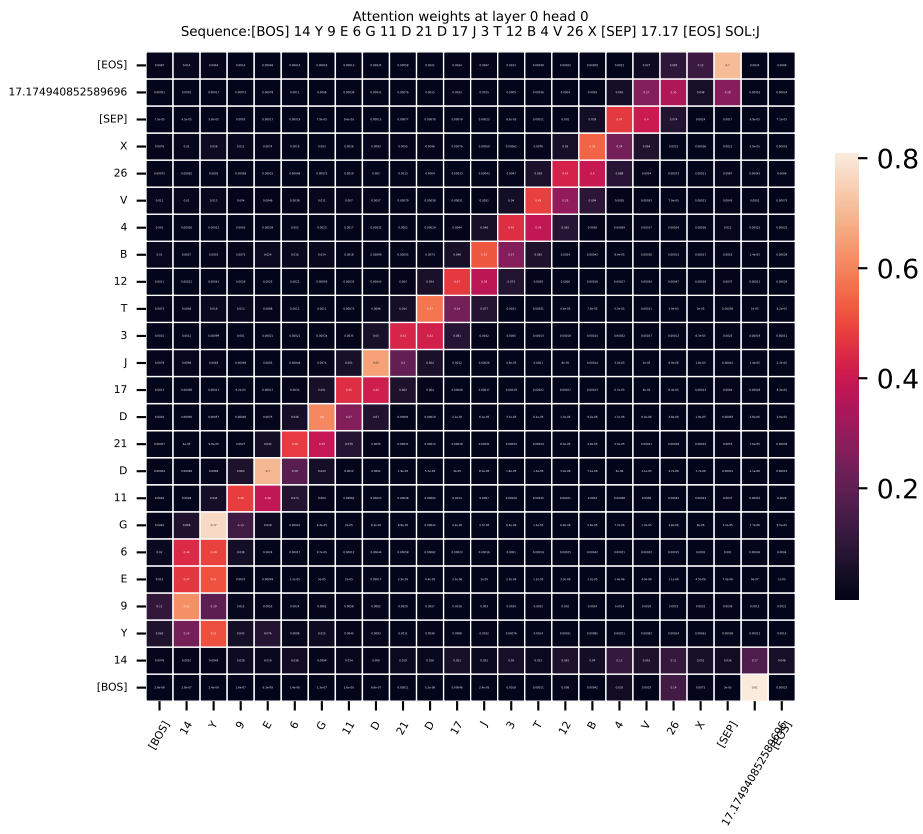

Attention weights at layer 1 head 0
Sequence:[BOS] 3 W 25 O 4 L 13 H 9 G 6 U 23 M 7 O 21 L 18 L [SEP] 13.11 [EOS] SOL:H
[EOS]
13.105201762624336
[SEP]
L
18
L
21
O
7
M
23
U
6
G
9
H
13
L
4
O
25
W
3
[BOS]
[BOS] 3 W 25 O 4 L 13 H 9 G 6 U 23 M 7 O 21 L 18 L [SEP]
0.8
0.6
0.4
0.2
0.0

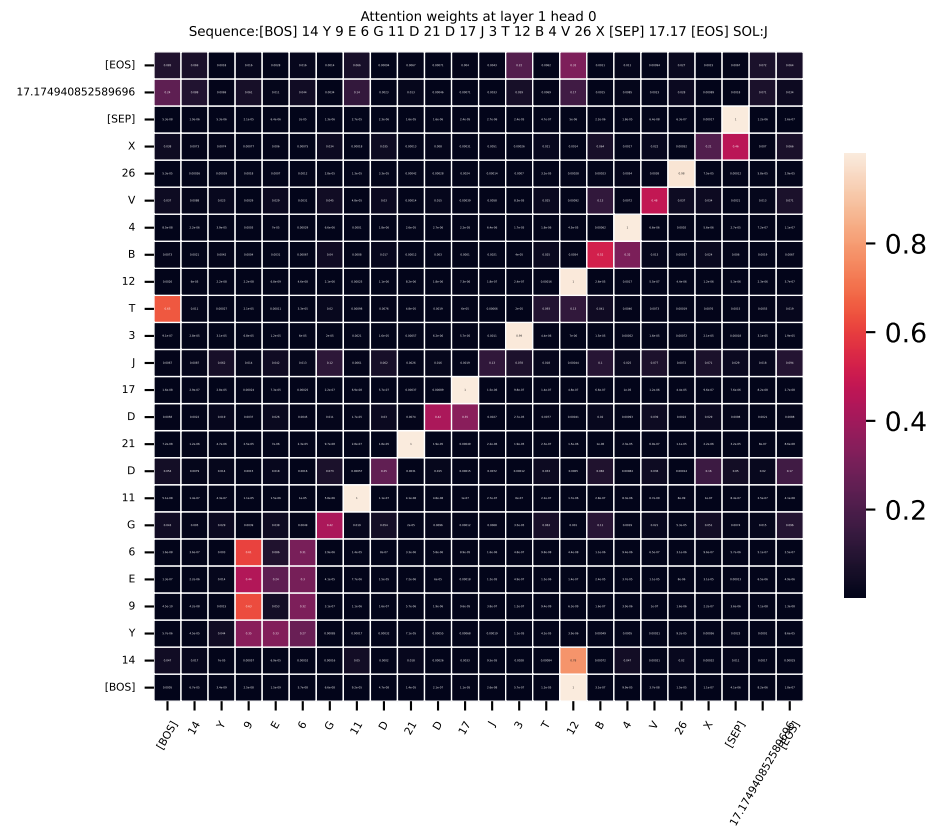

Attention weights at layer 0 head 0
Sequence:[BOS] 4 V 18 F 1 C 23 D 7 C 9 U 3 U 10 O 16 Q 11 K [SEP] 18.08 [EOS] SOL:F
[EOS]
18.08389163381117
[SEP]
K
11
Q
16
O
10
U
3
U
9
C
7
D
23
C
1
F
18
V
4
[BOS]
[BOS] 4 V 18 F 1 C 23 D 7 C 9 U 3 U 10 O 16 Q 11 K [SEP] 18.08389163381117 [EOS]
0.8
0.6
0.4
0.2

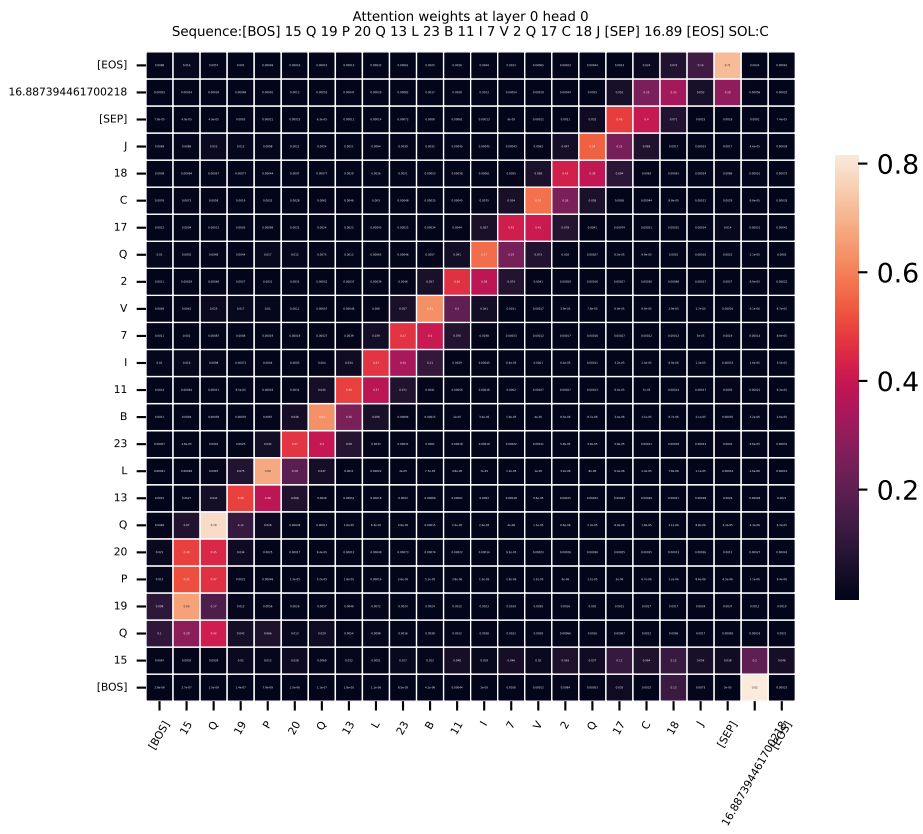

Attention weights at layer 1 head 0
Sequence:[BOS] 4 V 18 F 1 C 23 D 7 C 9 U 3 U 10 O 16 Q 11 K [SEP] 18.08 [EOS] SOL:F
[EOS]
18.08389163381117
[SEP]
K
11
Q
16
O
10
U
3
U
9
C
7
D
23
C
1
F
18
V
4
[BOS]
[BOS] 4 V 18 F 1 C 23 D 7 C 9 U 3 U 10 O 16 Q 11 K [SEP] 18.08389163381117 [EOS]
0.8
0.6
0.4
0.2

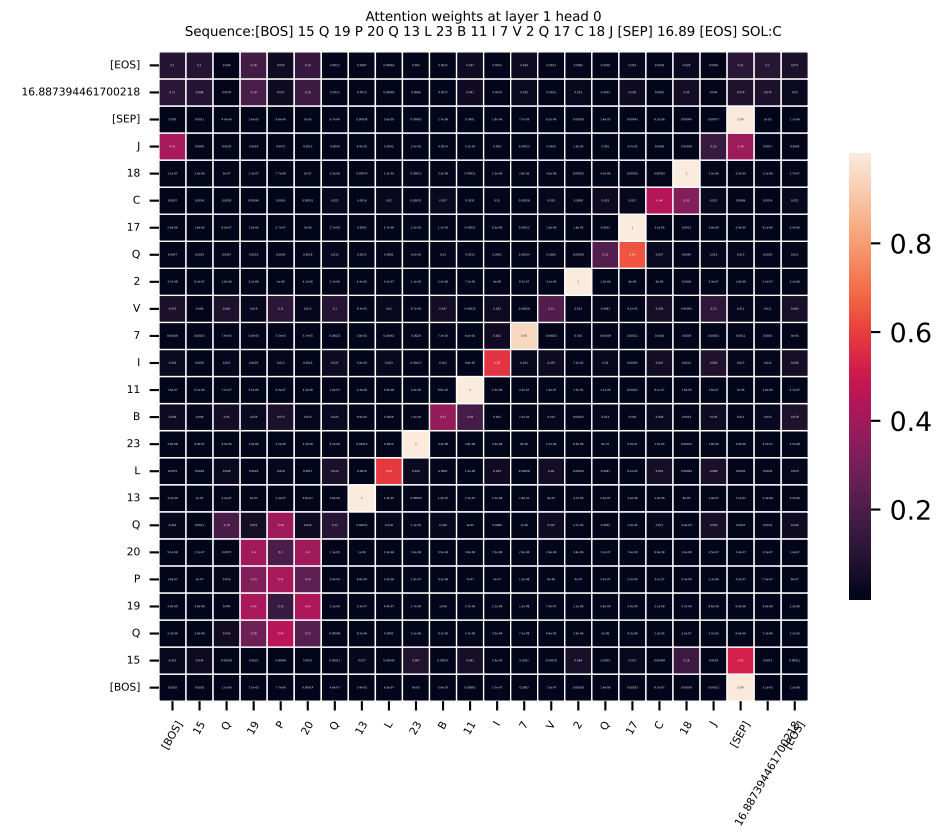

Attention weights at layer 0 head 0
Sequence:[BOS] 17 C 16 L 10 K 5 R 18 X 24 C 4 J 8 Y 14 V 12 D [SEP] 17.0 [EOS] SOL:C
[EOS]
17.003954088868245
[SEP]
D
12
V
14
Y
8
J
4
C
24
X
18
R
5
K
10
L
16
C
17
[BOS]
[BOS]
17
C
16
L
10
K
5
R
18
X
24
C
4
J
8
Y
14
V
12
D
[SEP]
17.003954088868245
[EOS]
0.8
0.6
0.4
0.2

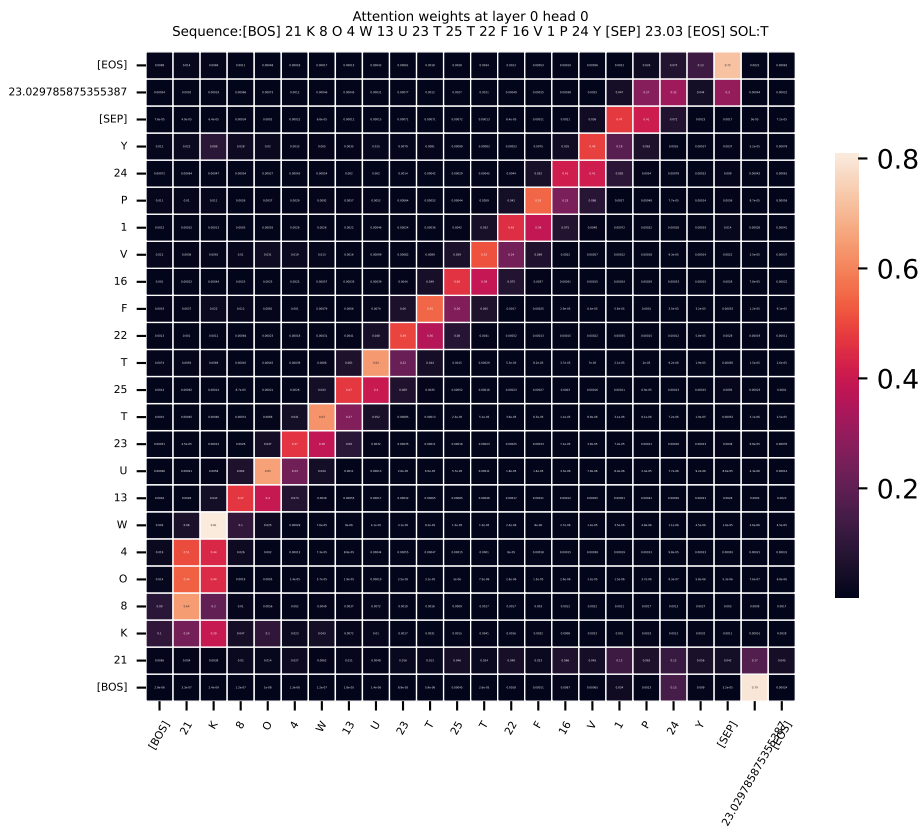

Attention weights at layer 1 head 0
Sequence:[BOS] 17 C 16 L 10 K 5 R 18 X 24 C 4 J 8 Y 14 V 12 D [SEP] 17.0 [EOS] SOL:C
[EOS]
17.003954088868245
[SEP]
D
12
V
14
Y
8
J
4
C
24
X
18
R
5
K
10
L
16
C
17
[BOS]
[BOS]
17
C
16
L
10
K
5
R
18
X
24
C
4
J
8
Y
14
V
12
D
[SEP]
17.003954088868245
[EOS]
0.8
0.6
0.4
0.2

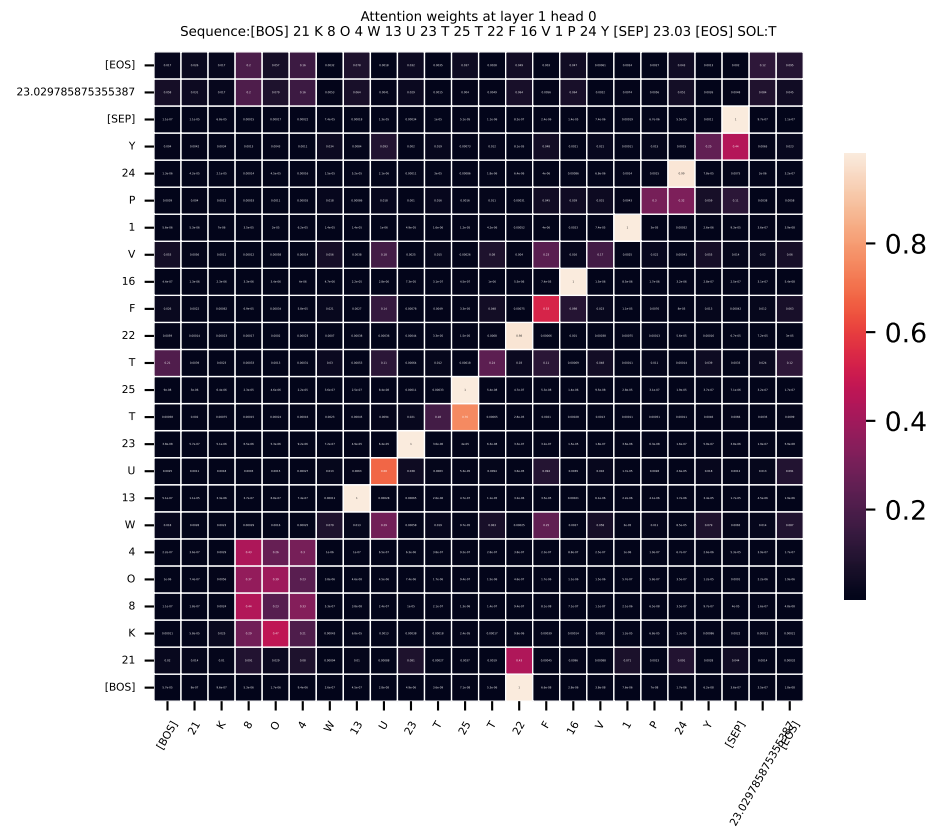

Attention weights at layer 0 head 0
Sequence:[BOS] 14 Y 9 E 6 G 11 D 21 D 17 J 3 T 12 B 4 V 26 X [SEP] 17.17 [EOS] SOL:J
[EOS]
17.174940852589696
[SEP]
X
26
V
4
B
12
T
3
J
17
D
21
D
11
G
6
E
9
Y
14
[BOS]
[BOS] 14 Y 9 E 6 G 11 D 21 D 17 J 3 T 12 B 4 V 26 X [SEP] 17.17 [EOS] [BOS]
0.8
0.6
0.4
0.2

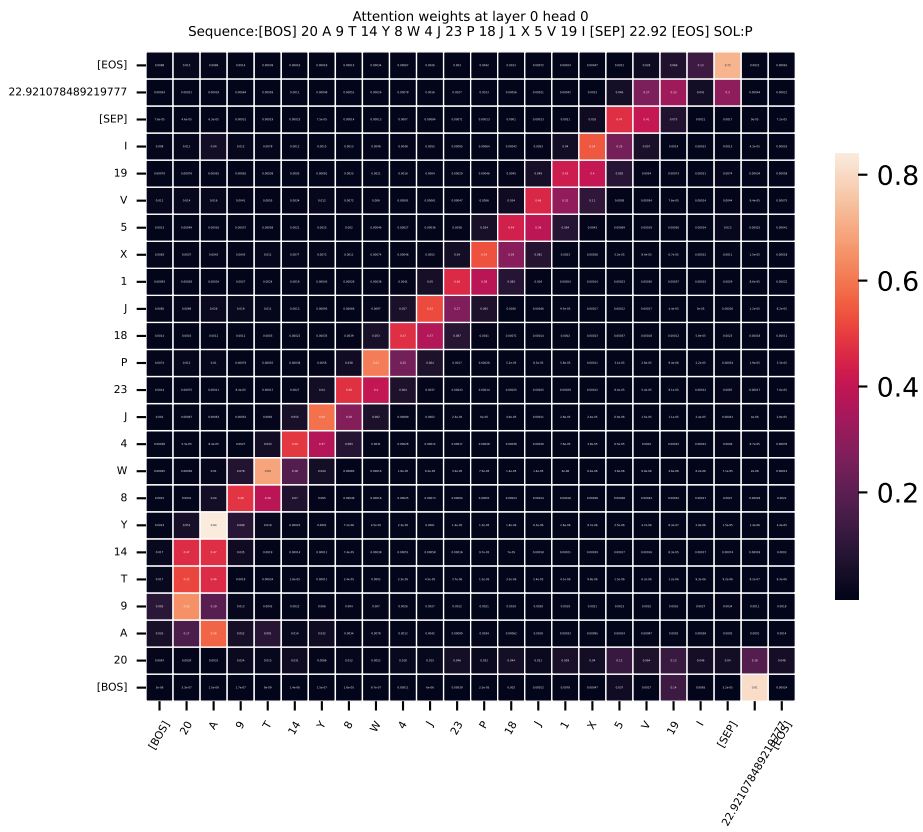

Attention weights at layer 1 head 0
Sequence:[BOS] 14 Y 9 E 6 G 11 D 21 D 17 J 3 T 12 B 4 V 26 X [SEP] 17.17 [EOS] SOL:J
[EOS]
17.174940852589696
[SEP]
X
26
V
4
B
12
T
3
J
17
D
21
D
11
G
6
E
9
Y
14
[BOS]
[BOS] 14 Y 9 E 6 G 11 D 21 D 17 J 3 T 12 B 4 V 26 X [SEP] 17.17 [EOS] [BOS]
0.8
0.6
0.4
0.2

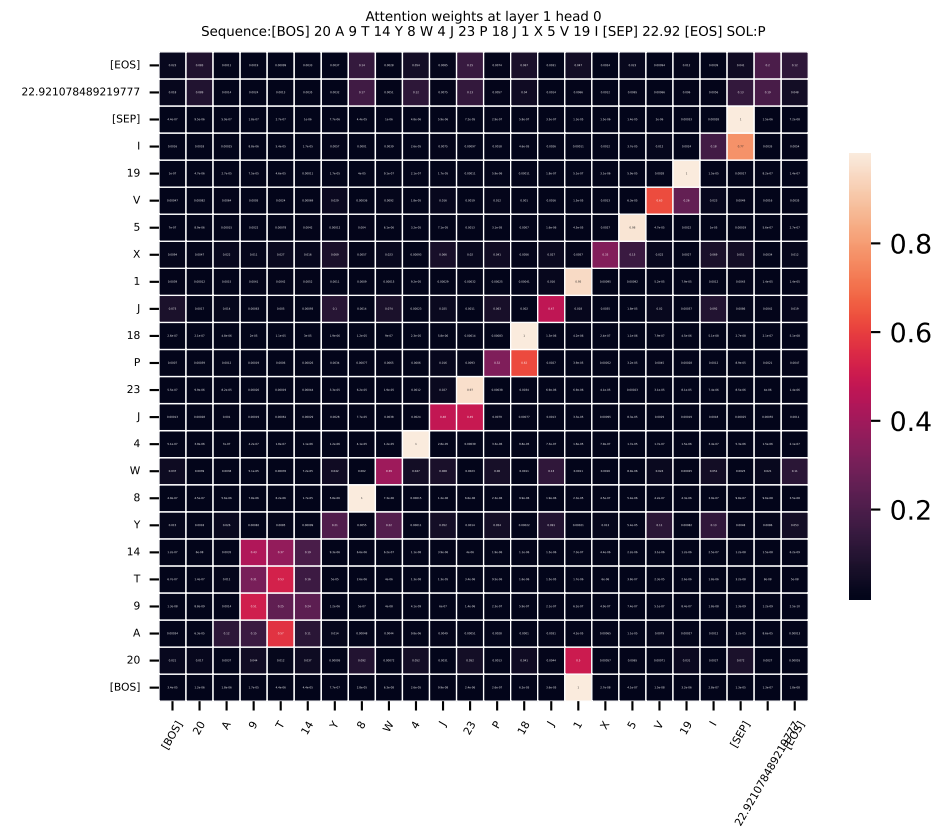

Attention weights at layer 0 head 0
Sequence:[BOS] 15 Q 19 P 20 Q 13 L 23 B 11 I 7 V 2 Q 17 C 18 J [SEP] 16.89 [EOS] SOL:C
[EOS]
16.887394461700218
[SEP]
J
18
C
17
Q
2
V
7
I
11
B
23
L
13
Q
20
P
19
Q
15
[BOS]
[BOS]
15
Q
19
P
20
Q
13
L
23
B
11
I
7
V
2
Q
17
C
18
J
[SEP]
16.887394461700218
[EOS]
0.8
0.6
0.4
0.2

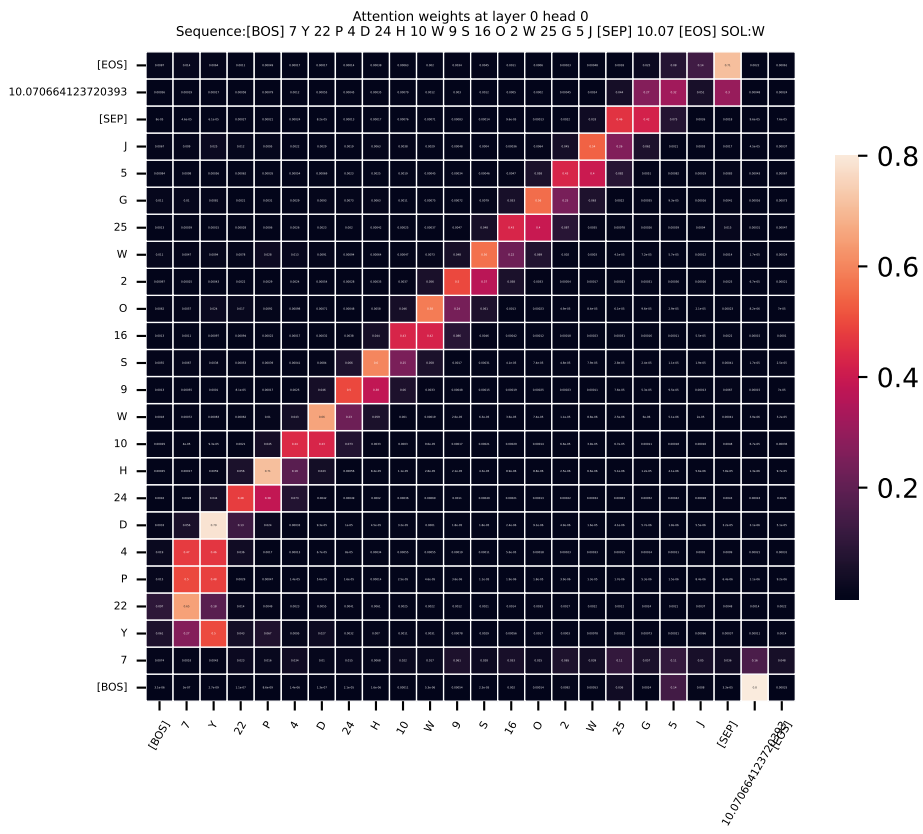

Attention weights at layer 1 head 0
Sequence:[BOS] 15 Q 19 P 20 Q 13 L 23 B 11 I 7 V 2 Q 17 C 18 J [SEP] 16.89 [EOS] SOL:C
[EOS]
16.887394461700218
[SEP]
J
18
C
17
Q
2
V
7
I
11
B
23
L
13
Q
20
P
19
Q
15
[BOS]
[BOS]
15
Q
19
P
20
Q
13
L
23
B
11
I
7
V
2
Q
17
C
18
J
[SEP]
16.887394461700218
[EOS]
0.8
0.6
0.4
0.2

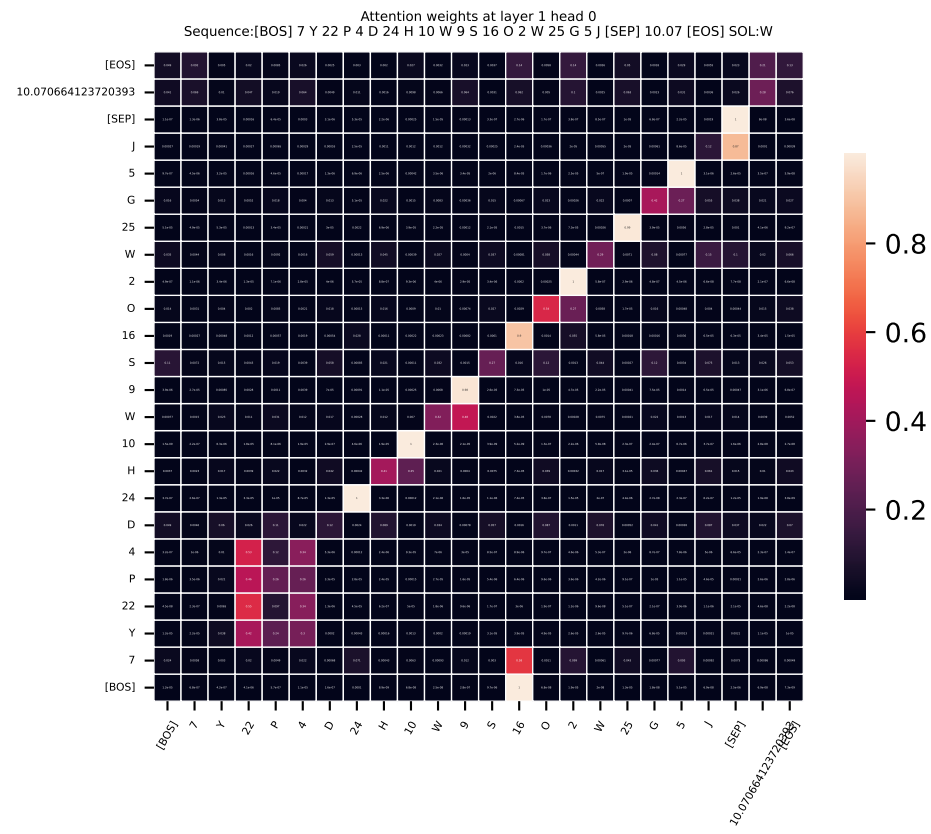

Attention weights at layer 0 head 0
Sequence:[BOS] 21 K 8 O 4 W 13 U 23 T 25 T 22 F 16 V 1 P 24 Y [SEP] 23.03 [EOS] SOL:T
[EOS]
23.029785875355387
[SEP]
Y
24
P
1
V
16
F
22
T
25
T
23
U
13
W
4
O
8
K
21
[BOS]
0.8
0.6
0.4
0.2

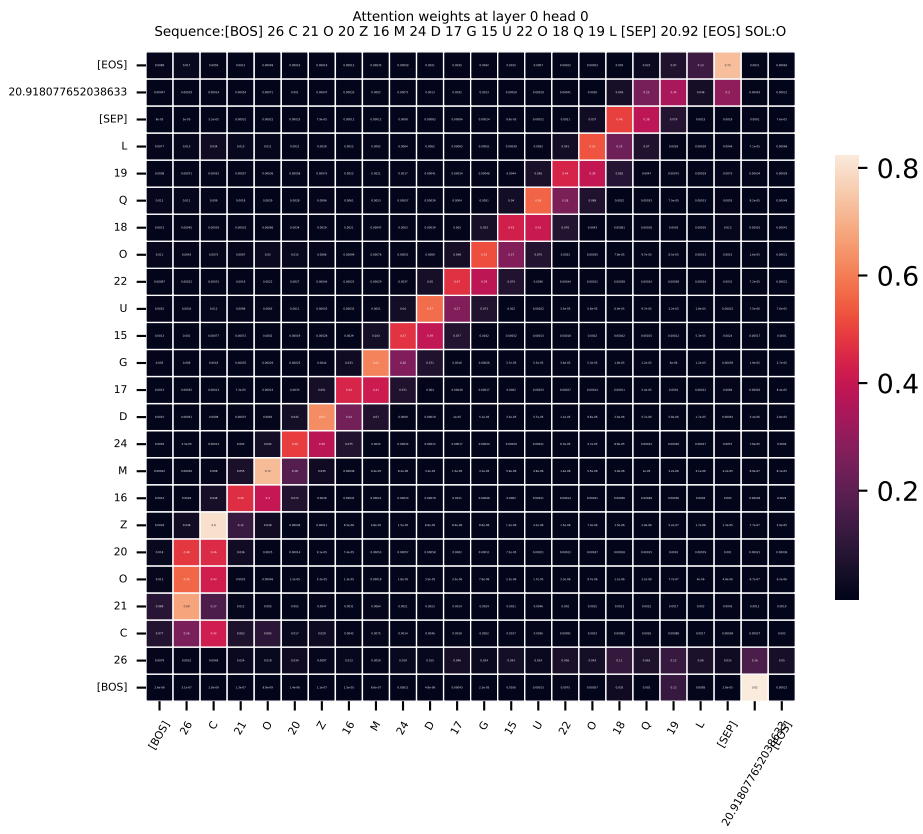

Attention weights at layer 1 head 0
Sequence:[BOS] 21 K 8 O 4 W 13 U 23 T 25 T 22 F 16 V 1 P 24 Y [SEP] 23.03 [EOS] SOL:T
[EOS]
23.029785875355387
[SEP]
Y
24
P
1
V
16
F
22
T
25
T
23
U
13
W
4
O
8
K
21
[BOS]
0.8
0.6
0.4
0.2

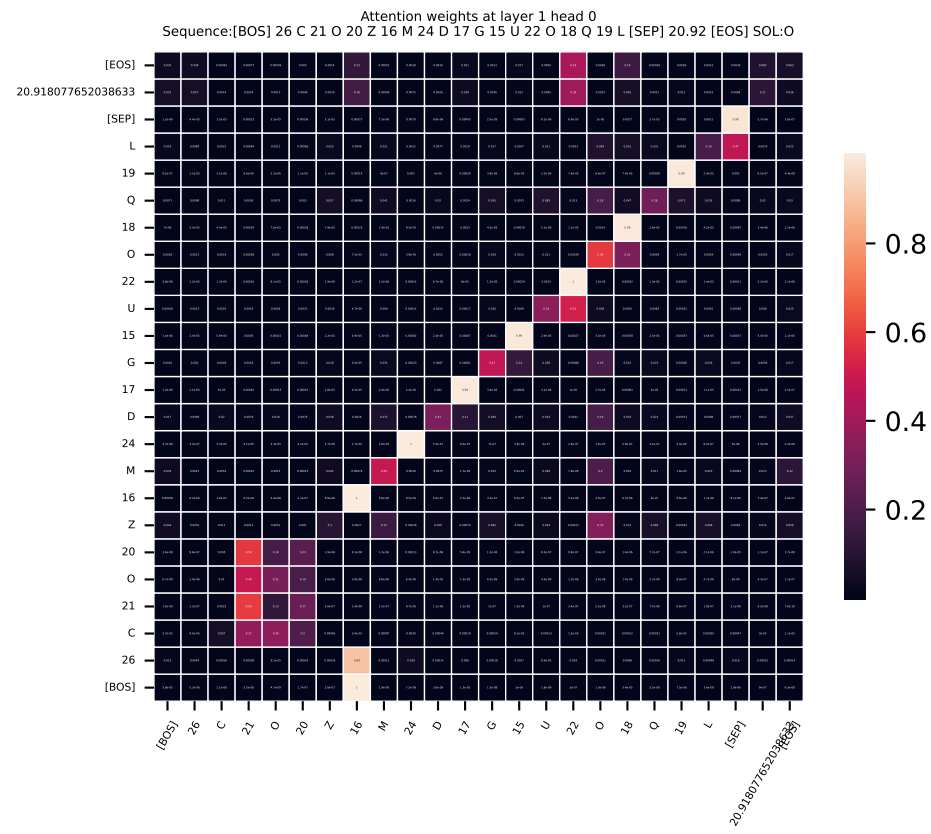

Attention weights at layer 0 head 0
Sequence:[BOS] 20 A 9 T 14 Y 8 W 4 J 23 P 18 J 1 X 5 V 19 I [SEP] 22.92 [EOS] SOL:P
[EOS]
22.921078489219777
[SEP]
I
19
V
5
X
1
J
18
P
23
J
4
W
8
Y
14
T
9
A
20
[BOS]
[BOS] 20 A 9 T 14 Y 8 W 4 J 23 P 18 J 1 X 5 V 19 I [SEP] 22.92 [EOS]
0.8
0.6
0.4
0.2

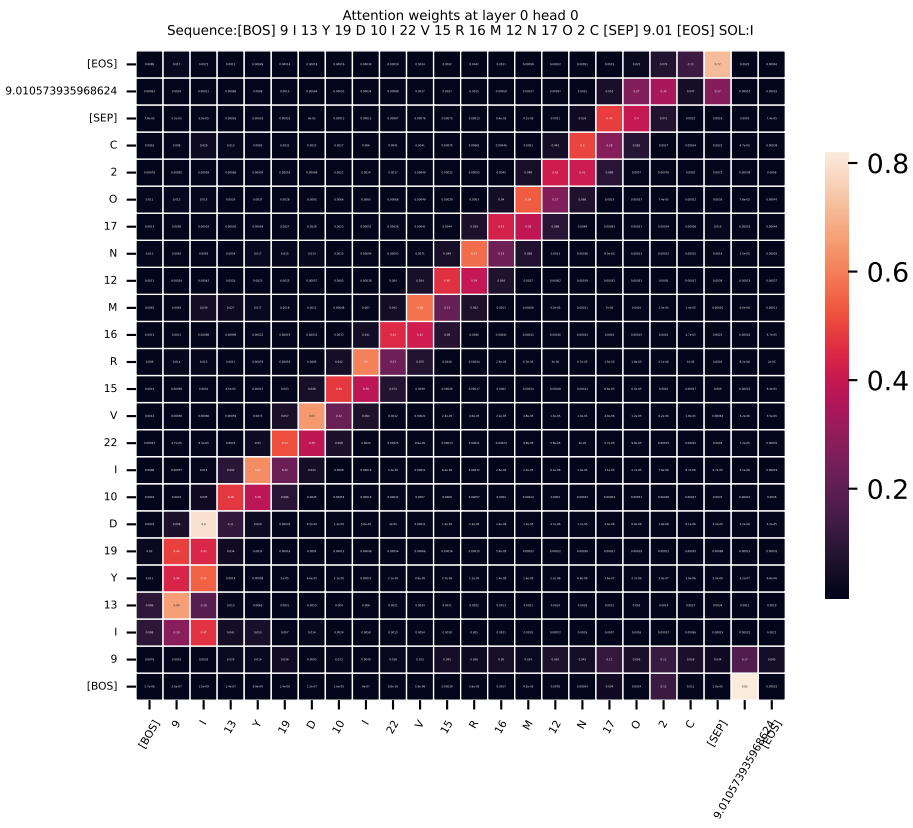

Attention weights at layer 1 head 0
Sequence:[BOS] 20 A 9 T 14 Y 8 W 4 J 23 P 18 J 1 X 5 V 19 I [SEP] 22.92 [EOS] SOL:P
[EOS]
22.921078489219777
[SEP]
I
19
V
5
X
1
J
18
P
23
J
4
W
8
Y
14
T
9
A
20
[BOS]
[BOS] 20 A 9 T 14 Y 8 W 4 J 23 P 18 J 1 X 5 V 19 I [SEP] 22.92 [EOS]
0.8
0.6
0.4
0.2

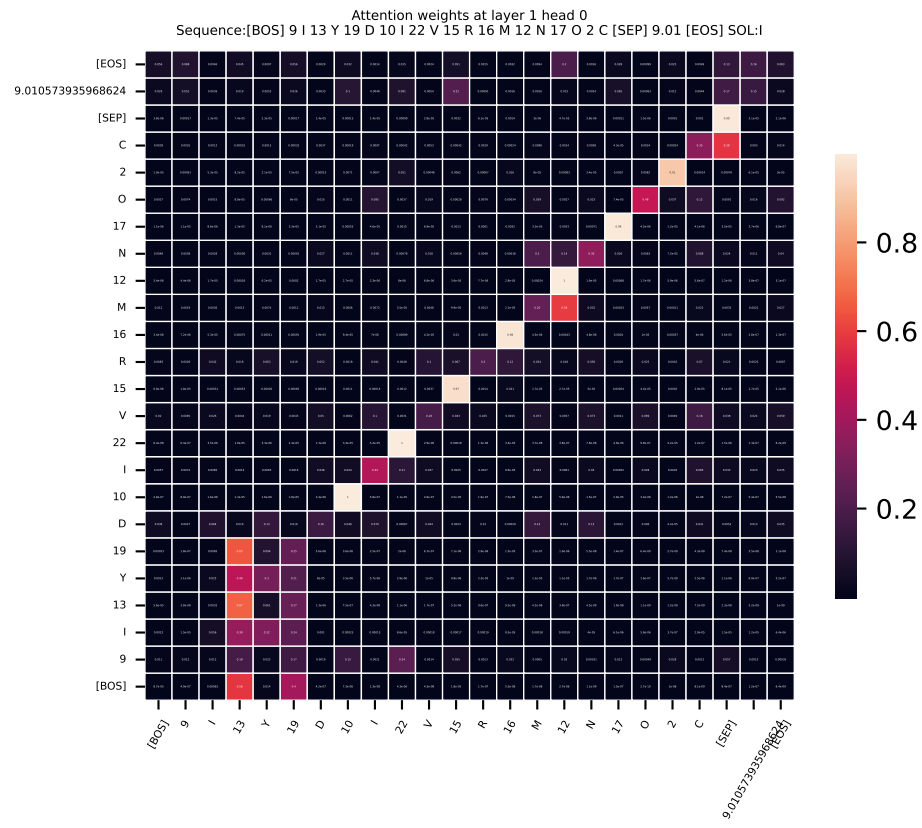

Attention weights at layer 0 head 0
Sequence:[BOS] 7 Y 22 P 4 D 24 H 10 W 9 S 16 O 2 W 25 G 5 J [SEP] 10.07 [EOS] SOL:W
[EOS]
10.070664123720393
[SEP]
J
5
G
25
W
2
O
16
S
9
W
10
H
24
D
4
P
22
Y
7
[BOS]
[BOS] 7 Y 22 P 4 D 24 H 10 W 9 S 16 O 2 W 25 G 5 J [SEP] 10.07 [EOS] SOL:W
0.8
0.6
0.4
0.2

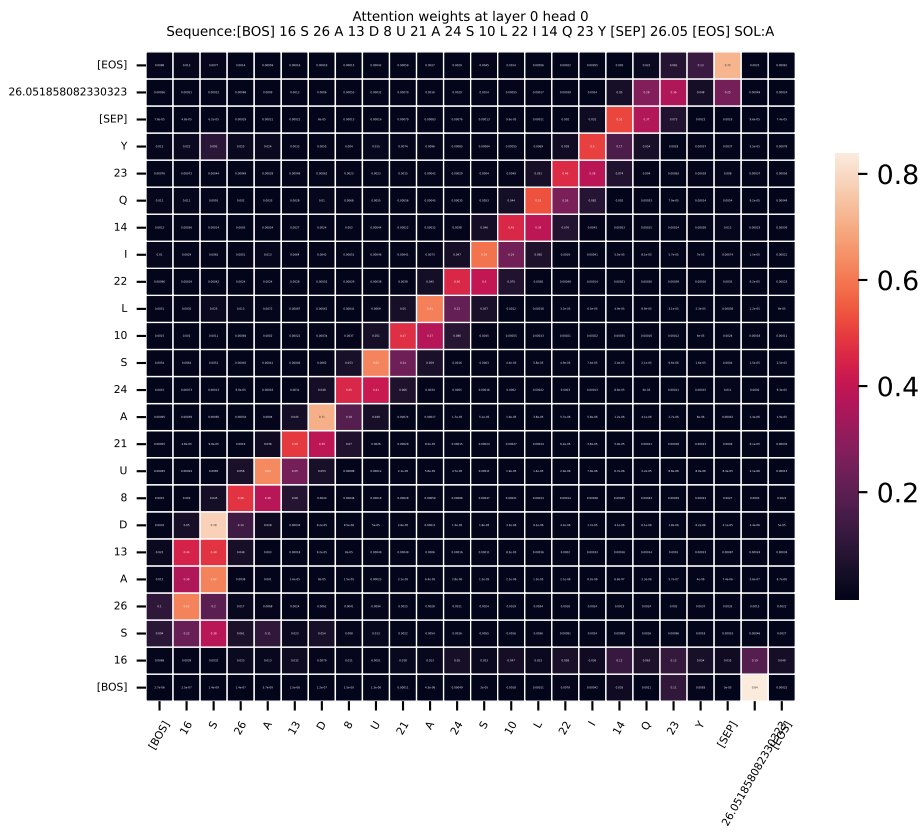

Attention weights at layer 1 head 0
Sequence:[BOS] 7 Y 22 P 4 D 24 H 10 W 9 S 16 O 2 W 25 G 5 J [SEP] 10.07 [EOS] SOL:W
[EOS]
10.070664123720393
[SEP]
J
5
G
25
W
2
O
16
S
9
W
10
H
24
D
4
P
22
Y
7
[BOS]
[BOS] 7 Y 22 P 4 D 24 H 10 W 9 S 16 O 2 W 25 G 5 J [SEP] 10.07 [EOS] SOL:W
0.8
0.6
0.4
0.2

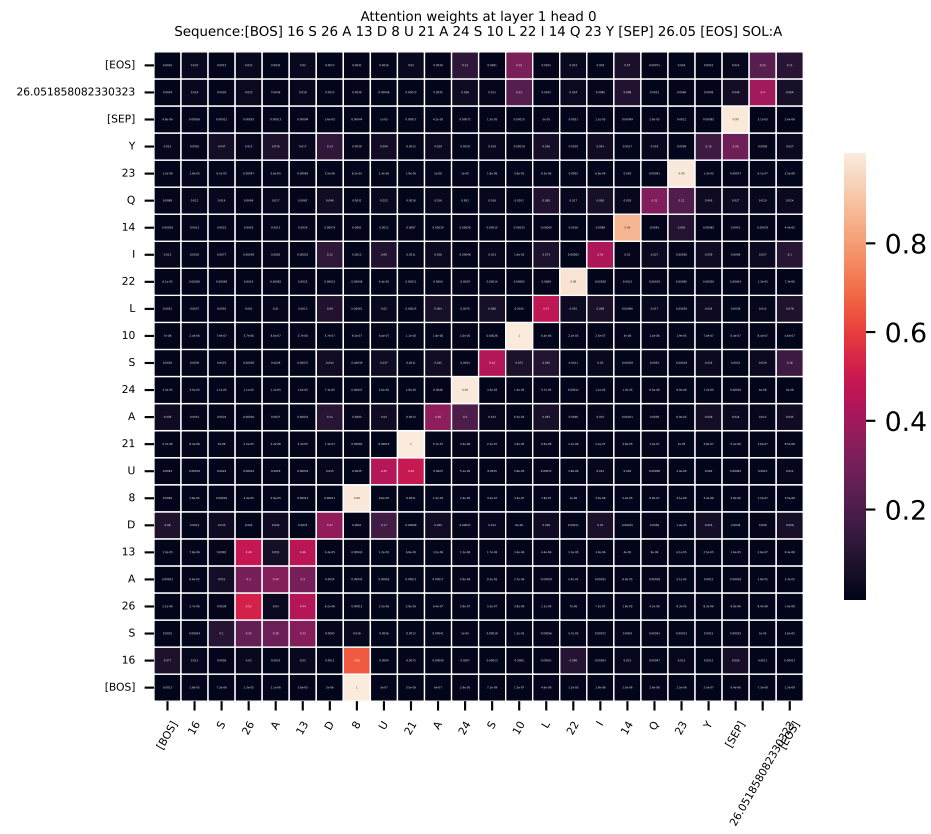

Attention weights at layer 0 head 0
Sequence:[BOS] 26 C 21 O 20 Z 16 M 24 D 17 G 15 U 22 O 18 Q 19 L [SEP] 20.92 [EOS] SOL:O
[EOS]
20.918077652038633
[SEP]
L
19
Q
18
O
22
U
15
G
17
D
24
M
16
Z
20
O
21
C
26
[BOS]
[BOS]
26
C
21
O
20
Z
16
M
24
D
17
G
15
U
22
O
18
Q
19
L
[SEP]
20.918077652038633
[EOS]
0.8
0.6
0.4
0.2

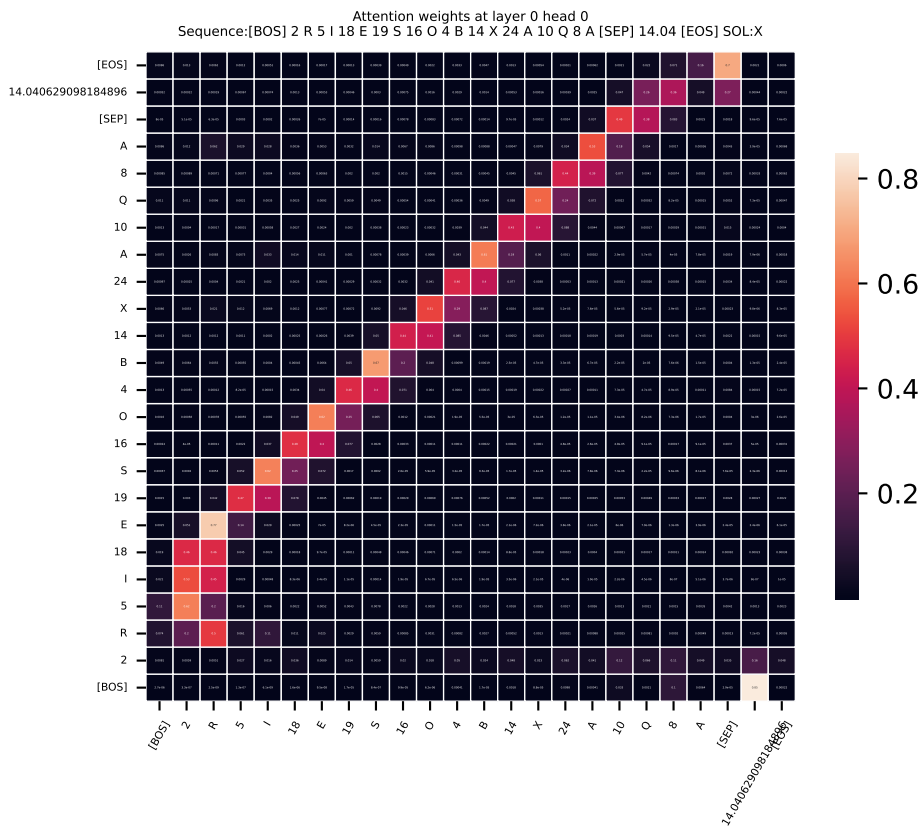

Attention weights at layer 1 head 0
Sequence:[BOS] 26 C 21 O 20 Z 16 M 24 D 17 G 15 U 22 O 18 Q 19 L [SEP] 20.92 [EOS] SOL:O
[EOS]
20.918077652038633
[SEP]
L
19
Q
18
O
22
U
15
G
17
D
24
M
16
Z
20
O
21
C
26
[BOS]
[BOS]
26
C
21
O
20
Z
16
M
24
D
17
G
15
U
22
O
18
Q
19
L
[SEP]
20.918077652038633
[EOS]
0.8
0.6
0.4
0.2

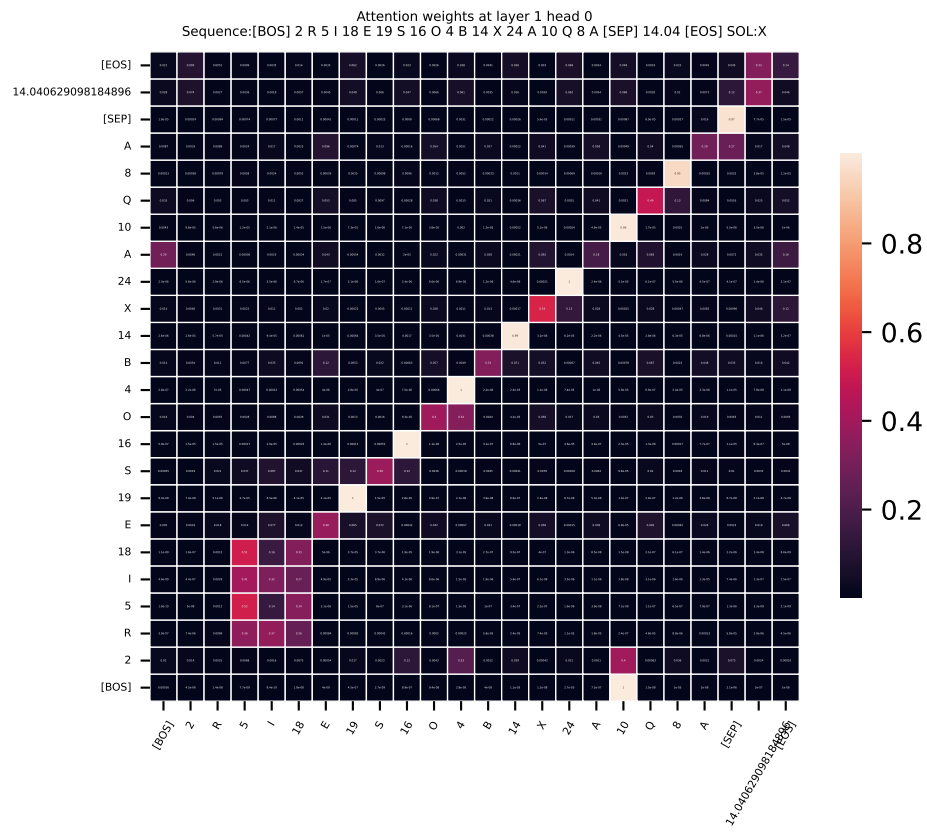

Attention weights at layer 0 head 0
Sequence:[BOS] 9 I 13 Y 19 D 10 I 22 V 15 R 16 M 12 N 17 O 2 C [SEP] 9.01 [EOS] SOL:I
[EOS]
9.010573935968624
[SEP]
C
2
O
17
N
12
M
16
R
15
V
22
I
10
D
19
Y
13
I
9
[BOS]
[BOS] 9 I 13 Y 19 D 10 I 22 V 15 R 16 M 12 N 17 O 2 C [SEP] 9.01 [EOS] SOL:I
0.8
0.6
0.4
0.2

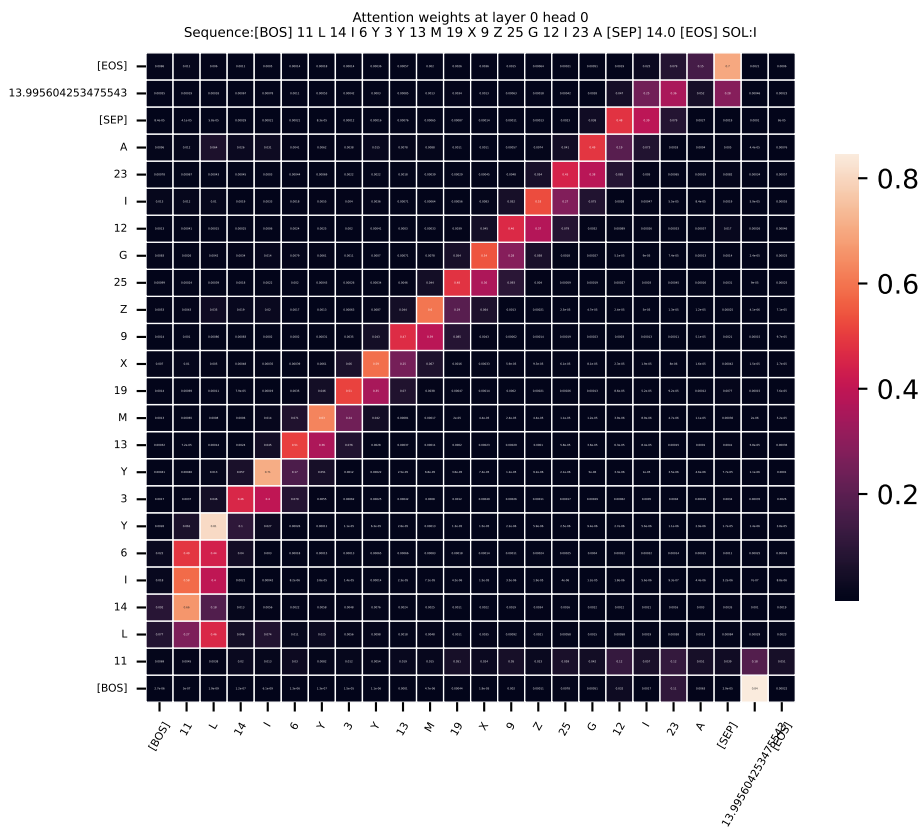

Attention weights at layer 1 head 0
Sequence:[BOS] 9 I 13 Y 19 D 10 I 22 V 15 R 16 M 12 N 17 O 2 C [SEP] 9.01 [EOS] SOL:I
[EOS]
9.010573935968624
[SEP]
C
2
O
17
N
12
M
16
R
15
V
22
I
10
D
19
Y
13
I
9
[BOS]
[BOS] 9 I 13 Y 19 D 10 I 22 V 15 R 16 M 12 N 17 O 2 C [SEP] 9.01 [EOS] SOL:I
0.8
0.6
0.4
0.2

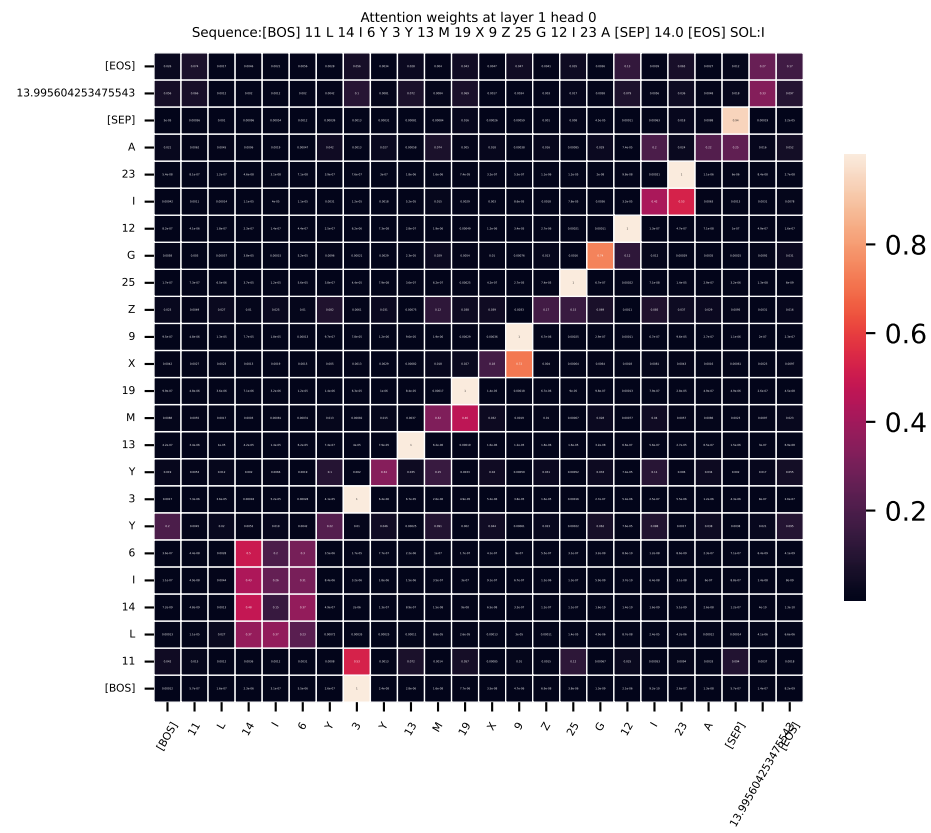

Attention weights at layer 0 head 0
Sequence:[BOS] 16 S 26 A 13 D 8 U 21 A 24 S 10 L 22 I 14 Q 23 Y [SEP] 26.05 [EOS] SOL:A
[EOS]
26.051858082330323
[SEP]
Y
23
Q
14
I
22
L
10
S
24
A
21
U
8
D
13
A
26
S
16
[BOS]
[BOS] 16 S 26 A 13 D 8 U 21 A 24 S 10 L 22 I 14 Q 23 Y [SEP] 26.05 [EOS] SOL:A
0.8
0.6
0.4
0.2

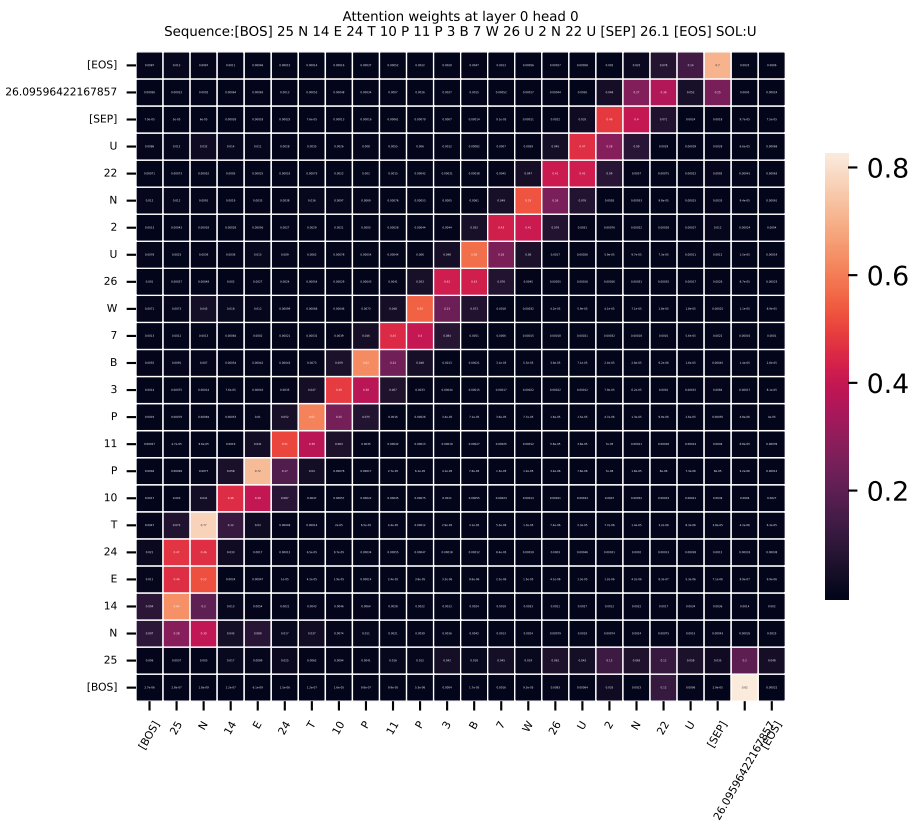

Attention weights at layer 1 head 0
Sequence:[BOS] 16 S 26 A 13 D 8 U 21 A 24 S 10 L 22 I 14 Q 23 Y [SEP] 26.05 [EOS] SOL:A
[EOS]
26.051858082330323
[SEP]
Y
23
Q
14
I
22
L
10
S
24
A
21
U
8
D
13
A
26
S
16
[BOS]
[BOS] 16 S 26 A 13 D 8 U 21 A 24 S 10 L 22 I 14 Q 23 Y [SEP] 26.05 [EOS] SOL:A
0.8
0.6
0.4
0.2

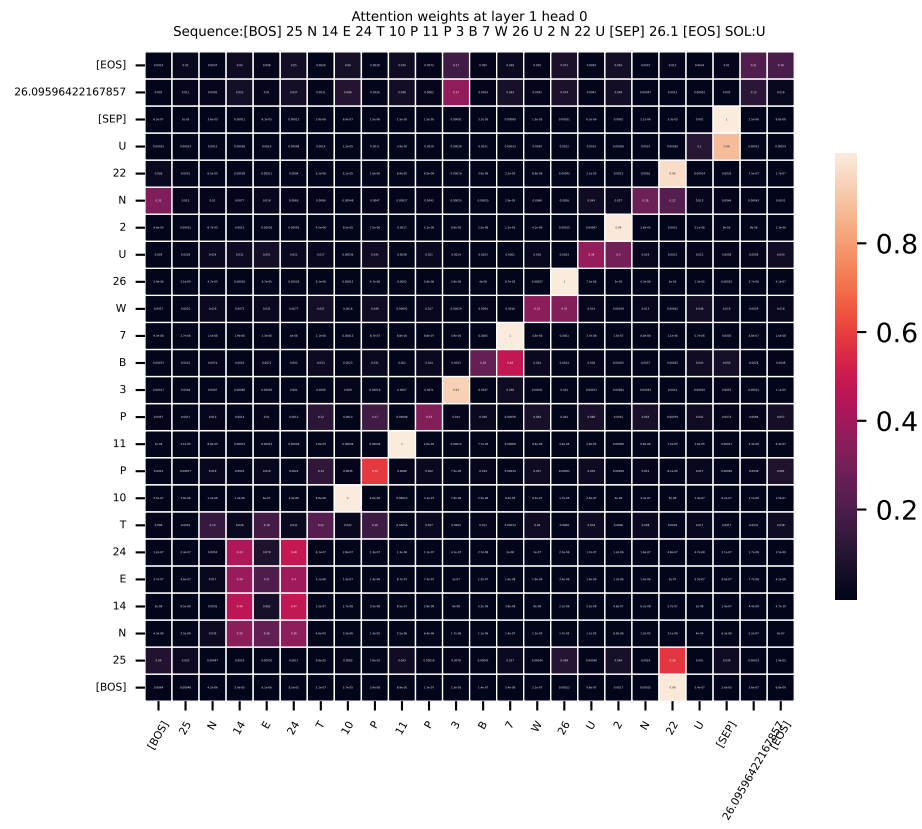

Attention weights at layer 0 head 0
Sequence:[BOS] 2 R 5 I 18 E 19 S 16 O 4 B 14 X 24 A 10 Q 8 A [SEP] 14.04 [EOS] SOL:X
[EOS]
14.040629098184896
[SEP]
A
8
Q
10
A
24
X
14
B
4
O
16
S
19
E
18
I
5
R
2
[BOS]
0.8
0.6
0.4
0.2

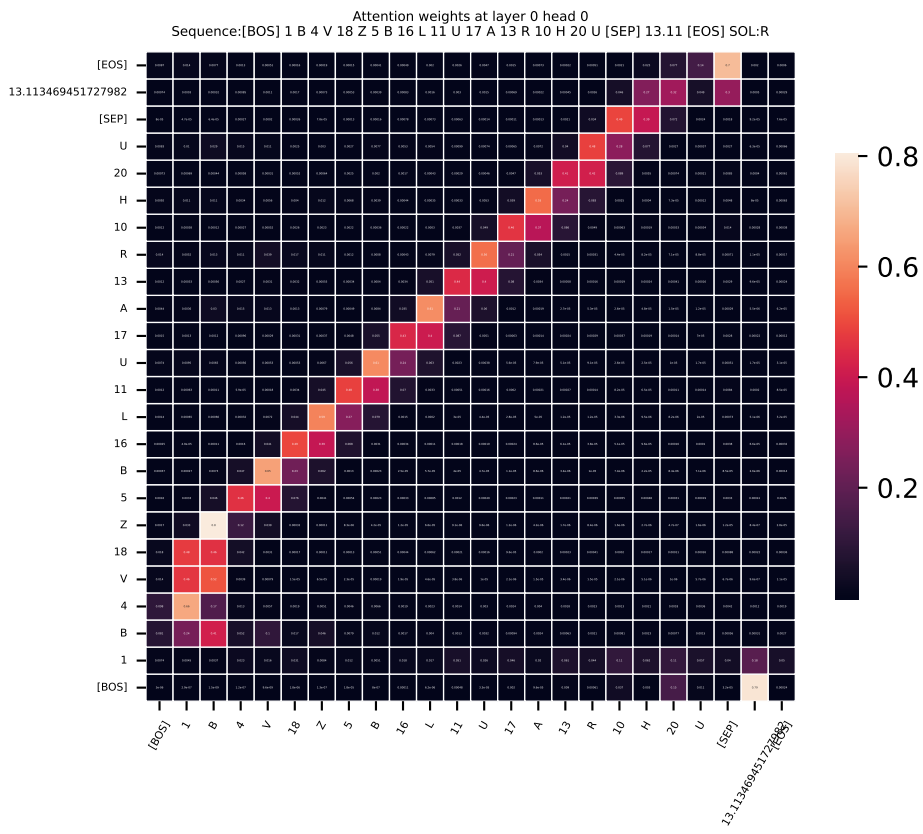

Attention weights at layer 1 head 0
Sequence:[BOS] 2 R 5 I 18 E 19 S 16 O 4 B 14 X 24 A 10 Q 8 A [SEP] 14.04 [EOS] SOL:X
[EOS]
14.040629098184896
[SEP]
A
8
Q
10
A
24
X
14
B
4
O
16
S
19
E
18
I
5
R
2
[BOS]
0.8
0.6
0.4
0.2

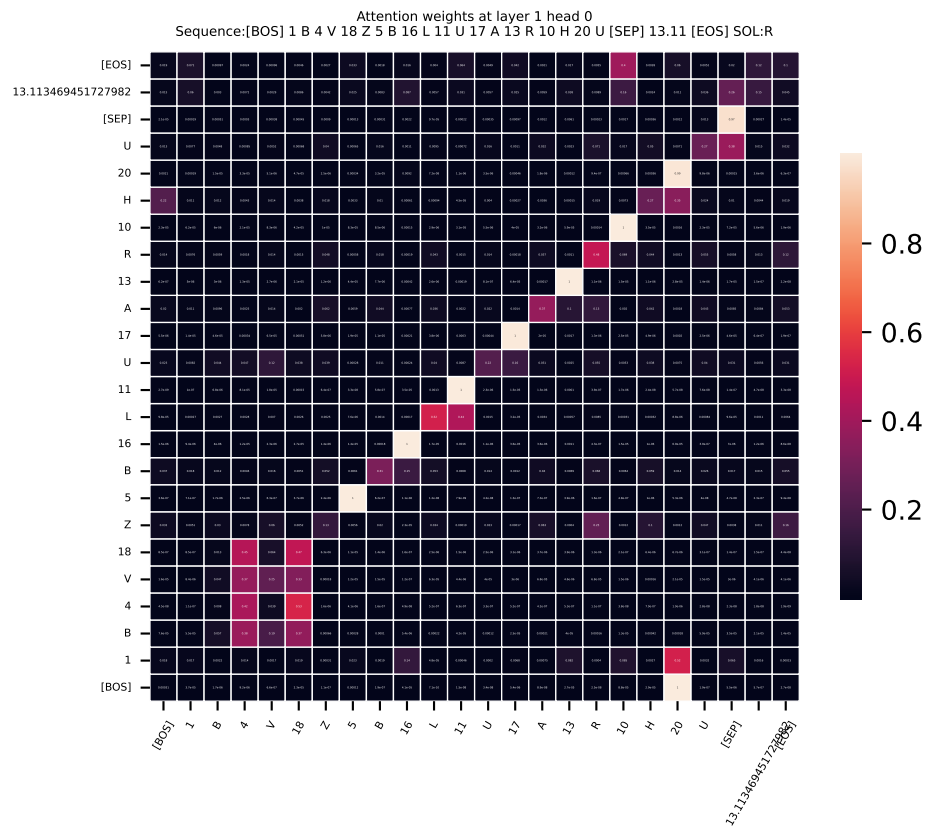

Attention weights at layer 0 head 0
Sequence:[BOS] 11 L 14 I 6 Y 3 Y 13 M 19 X 9 Z 25 G 12 I 23 A [SEP] 14.0 [EOS] SOL:I
[EOS]
13.995604253475543
[SEP]
A
23
I
12
G
25
Z
9
X
19
M
13
Y
3
Y
6
I
14
L
11
[BOS]
0.8
0.6
0.4
0.2

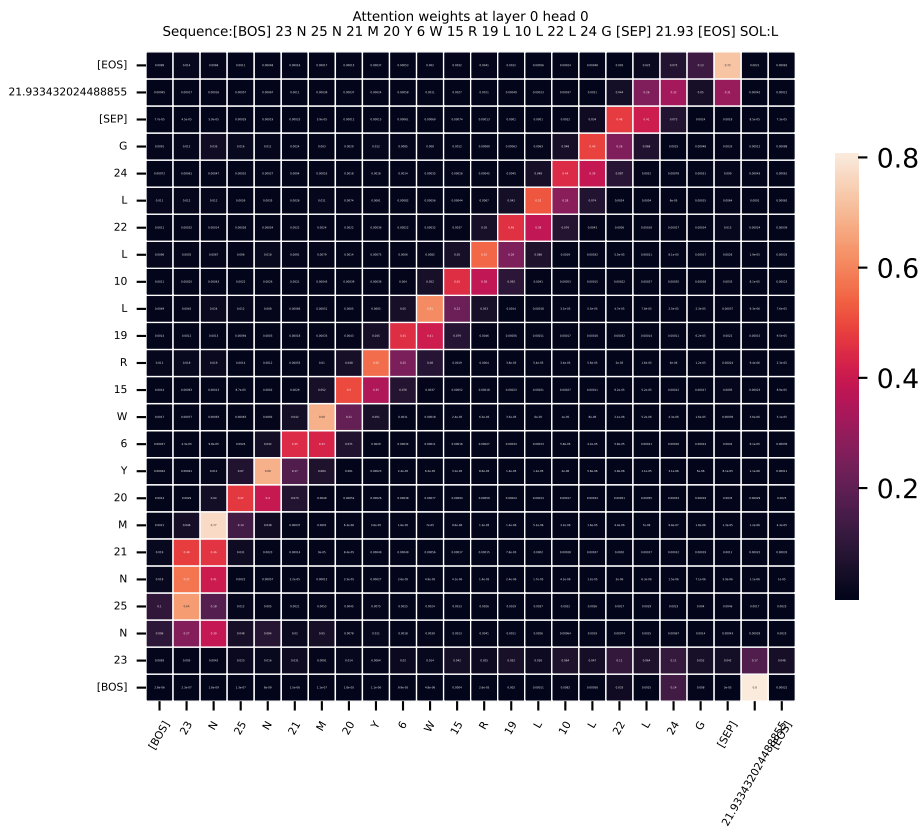

Attention weights at layer 1 head 0
Sequence:[BOS] 11 L 14 I 6 Y 3 Y 13 M 19 X 9 Z 25 G 12 I 23 A [SEP] 14.0 [EOS] SOL:I
[EOS]
13.995604253475543
[SEP]
A
23
I
12
G
25
Z
9
X
19
M
13
Y
3
Y
6
I
14
L
11
[BOS]
0.8
0.6
0.4
0.2

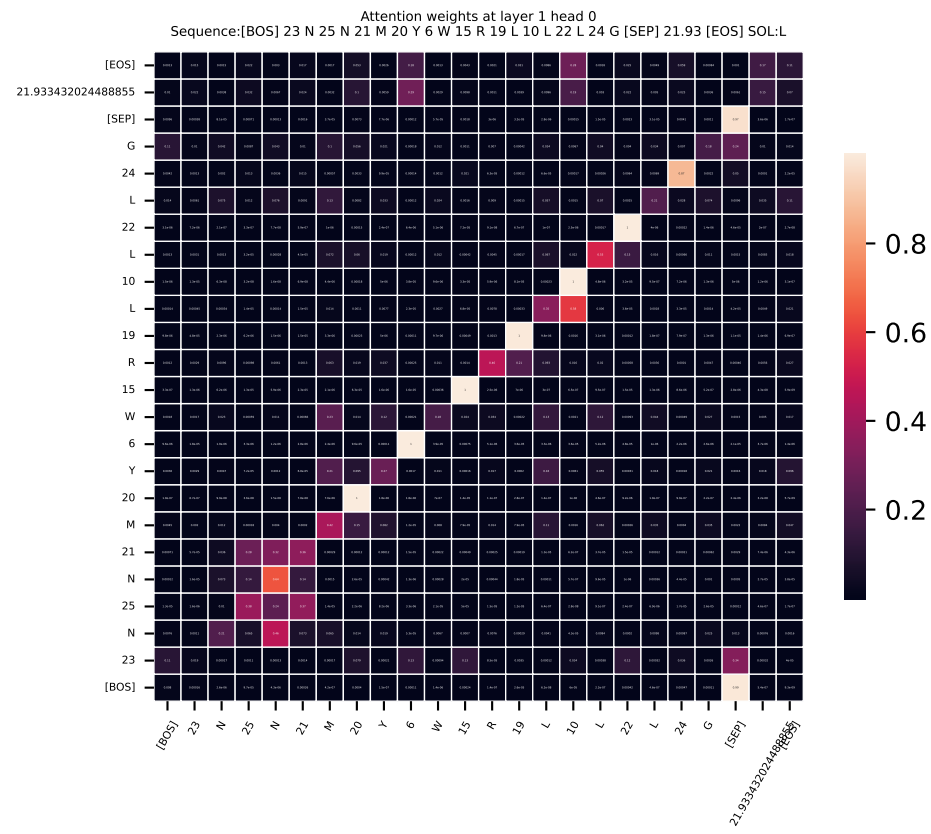

Attention weights at layer 0 head 0
Sequence:[BOS] 25 N 14 E 24 T 10 P 11 P 3 B 7 W 26 U 2 N 22 U [SEP] 26.1 [EOS] SOL:U
[EOS]
26.09596422167857
[SEP]
U
22
N
2
U
26
W
7
B
3
P
11
P
10
T
24
E
14
N
25
[BOS]
[BOS] 25 N 14 E 24 T 10 P 11 P 3 B 7 W 26 U 2 N 22 U [SEP] 26.1 [EOS] SOL:U
0.8
0.6
0.4
0.2

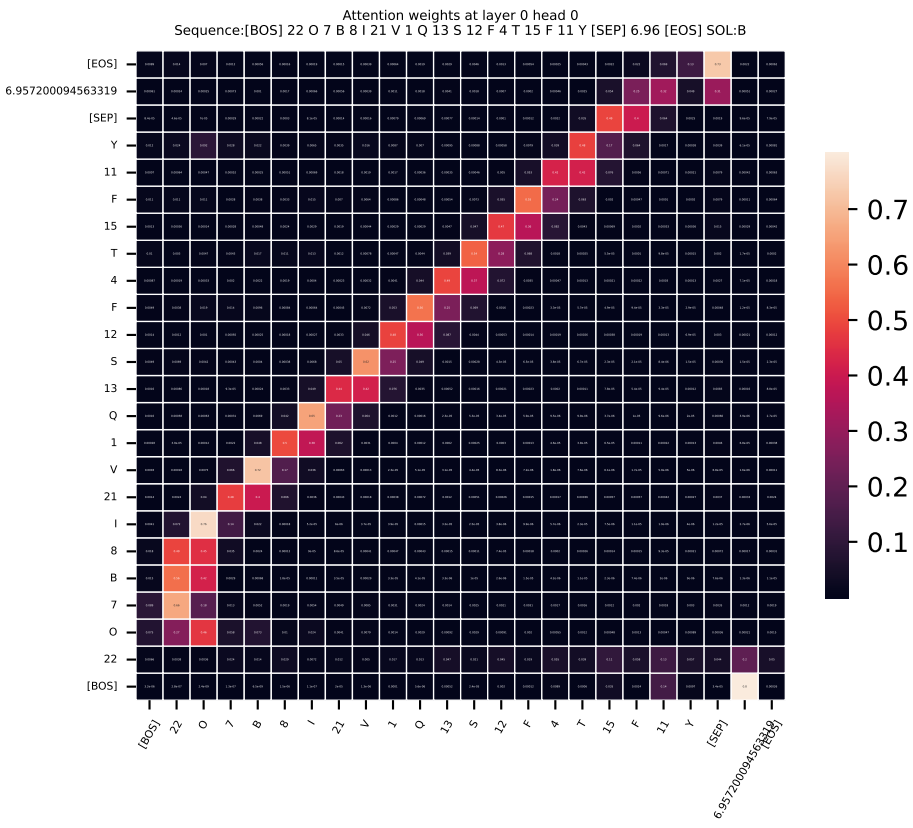

Attention weights at layer 1 head 0
Sequence:[BOS] 25 N 14 E 24 T 10 P 11 P 3 B 7 W 26 U 2 N 22 U [SEP] 26.1 [EOS] SOL:U
[EOS]
26.09596422167857
[SEP]
U
22
N
2
U
26
W
7
B
3
P
11
P
10
T
24
E
14
N
25
[BOS]
[BOS] 25 N 14 E 24 T 10 P 11 P 3 B 7 W 26 U 2 N 22 U [SEP] 26.1 [EOS] SOL:U
0.8
0.6
0.4
0.2

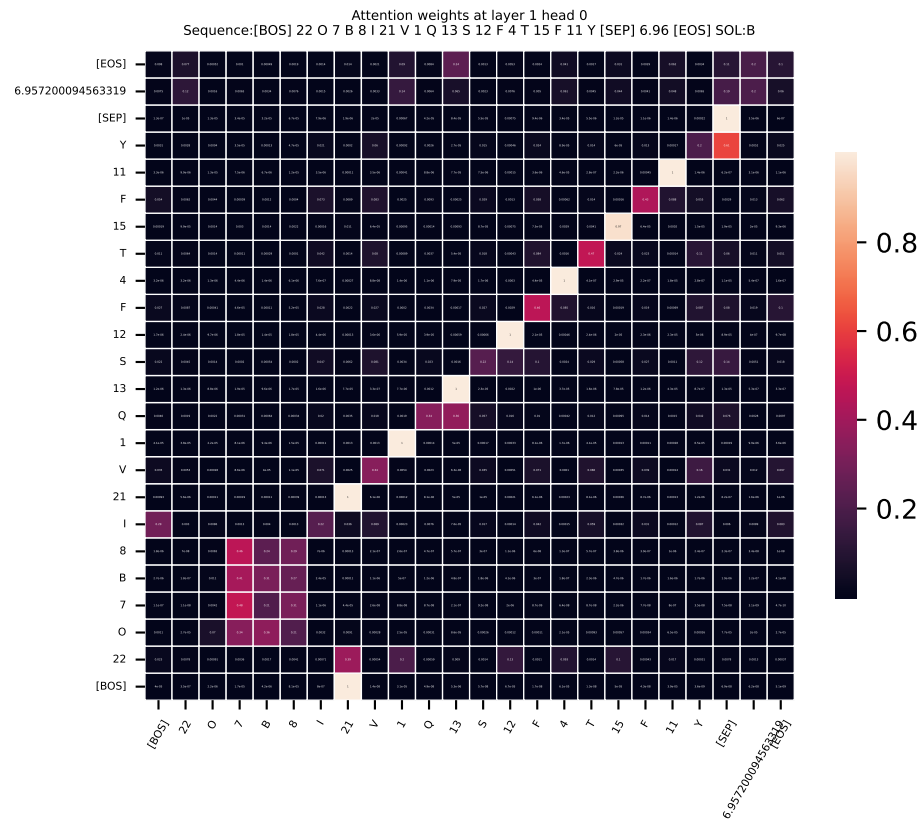

Attention weights at layer 0 head 0
Sequence:[BOS] 1 B 4 V 18 Z 5 B 16 L 11 U 17 A 13 R 10 H 20 U [SEP] 13.11 [EOS] SOL:R
[EOS]
13.113469451727982
[SEP]
U
20
H
10
R
13
A
17
U
11
L
16
B
5
Z
18
V
4
B
1
[BOS]
[BOS] 1 B 4 V 18 Z 5 B 16 L 11 U 17 A 13 R 10 H 20 U [SEP] 13.11 [EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 1 head 0
Sequence:[BOS] 1 B 4 V 18 Z 5 B 16 L 11 U 17 A 13 R 10 H 20 U [SEP] 13.11 [EOS] SOL:R
[EOS]
13.113469451727982
[SEP]
U
20
H
10
R
13
A
17
U
11
L
16
B
5
Z
18
V
4
B
1
[BOS]
[BOS] 1 B 4 V 18 Z 5 B 16 L 11 U 17 A 13 R 10 H 20 U [SEP] 13.11 [EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 0 head 0
Sequence:[BOS] 23 N 25 N 21 M 20 Y 6 W 15 R 19 L 10 L 22 L 24 G [SEP] 21.93 [EOS] SOL:L
[EOS]
21.933432024488855
[SEP]
G
24
L
22
L
10
L
19
R
15
W
6
Y
20
M
21
N
25
N
23
[BOS]
[BOS] 23 N 25 N 21 M 20 Y 6 W 15 R 19 L 10 L 22 L 24 G [SEP]
0.8
0.6
0.4
0.2


Attention weights at layer 1 head 0
Sequence:[BOS] 23 N 25 N 21 M 20 Y 6 W 15 R 19 L 10 L 22 L 24 G [SEP] 21.93 [EOS] SOL:L
[EOS]
21.933432024488855
[SEP]
G
24
L
22
L
10
L
19
R
15
W
6
Y
20
M
21
N
25
N
23
[BOS]
[BOS] 23 N 25 N 21 M 20 Y 6 W 15 R 19 L 10 L 22 L 24 G [SEP]
0.8
0.6
0.4
0.2


Attention weights at layer 0 head 0
Sequence:[BOS] 22 O 7 B 8 I 21 V 1 Q 13 S 12 F 4 T 15 F 11 Y [SEP] 6.96 [EOS] SOL:B
[EOS]
6.957200094563319
[SEP]
Y
11
F
15
T
4
F
12
S
13
Q
1
V
21
I
8
B
7
O
22
[BOS]
[BOS]
22
O
7
B
8
I
21
V
1
Q
13
S
12
F
4
T
15
F
11
Y
[SEP]
6.957200094563319
[EOS]
0.7
0.6
0.5
0.4
0.3
0.2
0.1


Attention weights at layer 1 head 0
Sequence:[BOS] 22 O 7 B 8 I 21 V 1 Q 13 S 12 F 4 T 15 F 11 Y [SEP] 6.96 [EOS] SOL:B
[EOS]
6.957200094563319
[SEP]
Y
11
F
15
T
4
F
12
S
13
Q
1
V
21
I
8
B
7
O
22
[BOS]
[BOS]
22
O
7
B
8
I
21
V
1
Q
13
S
12
F
4
T
15
F
11
Y
[SEP]
6.957200094563319
[EOS]
0.8
0.6
0.4
0.2