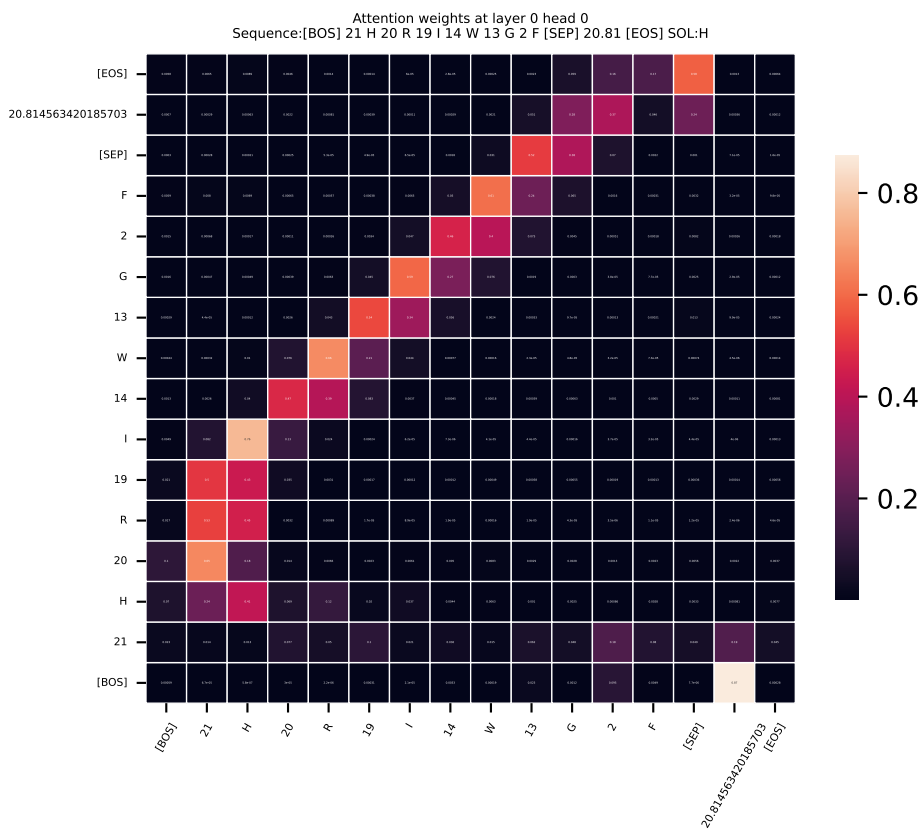
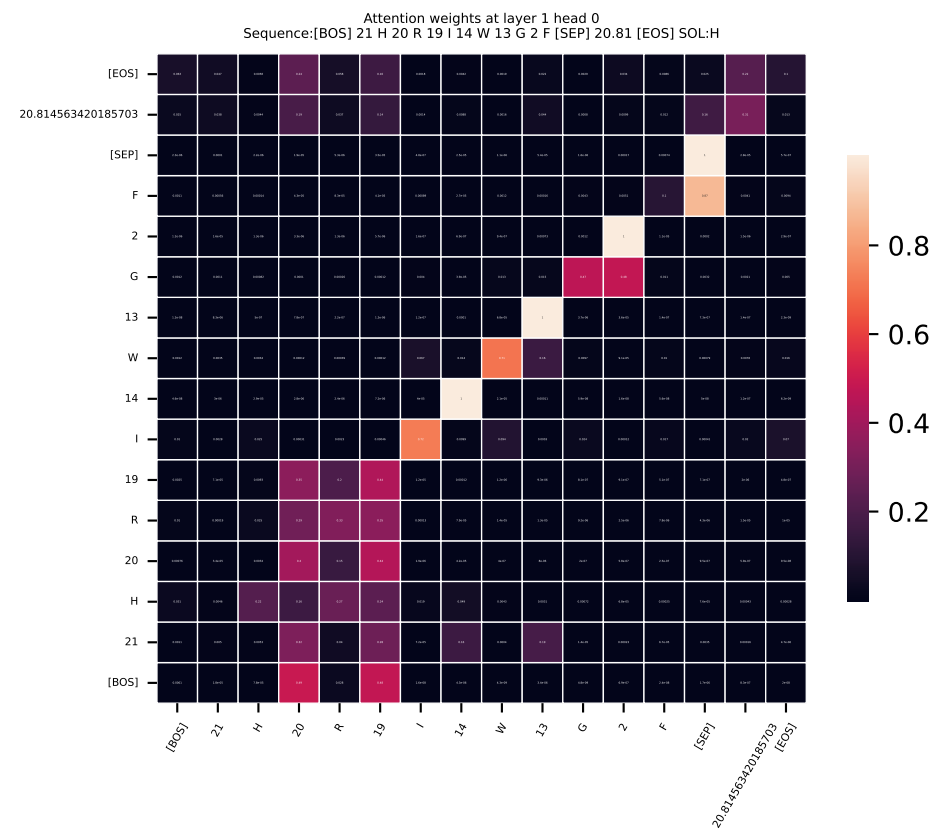
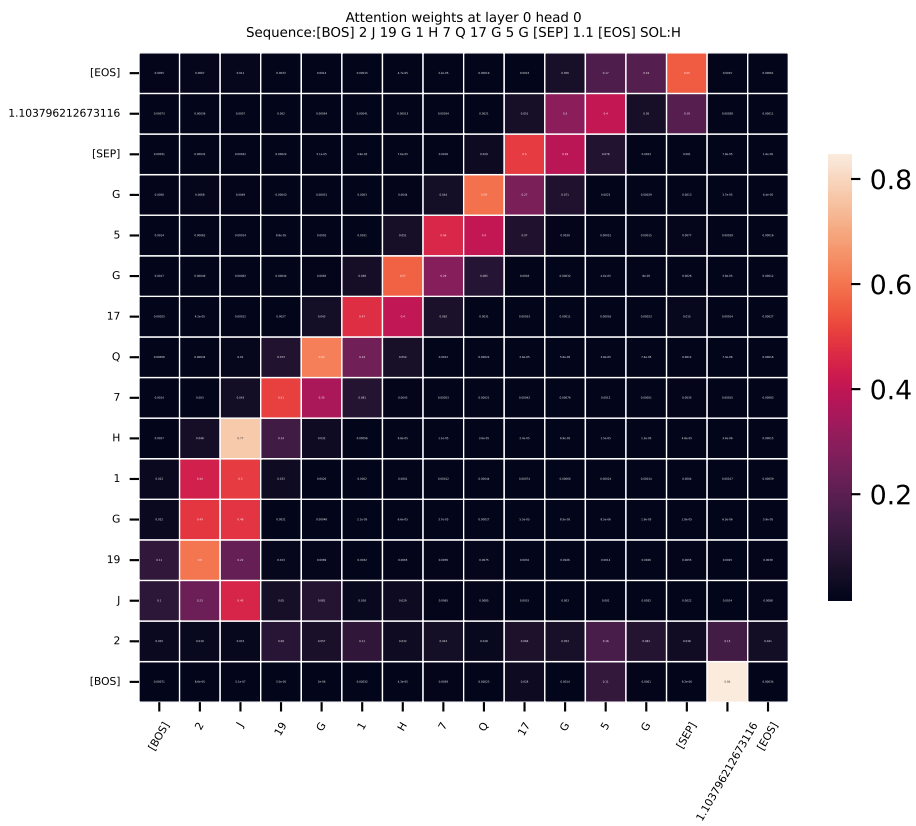
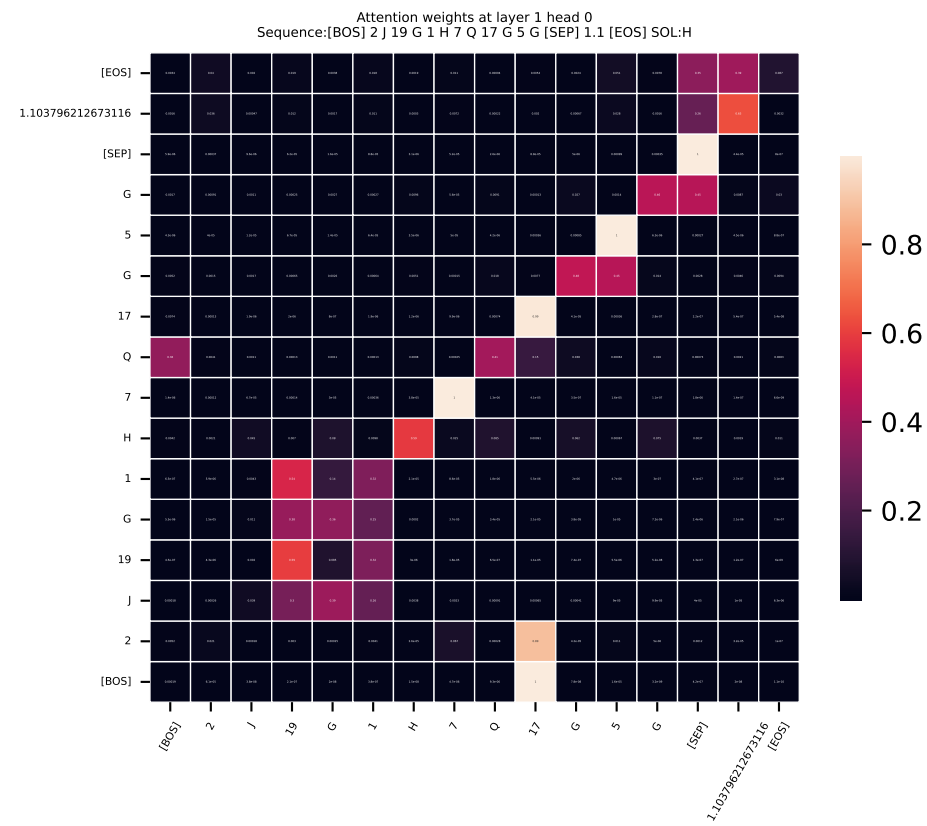
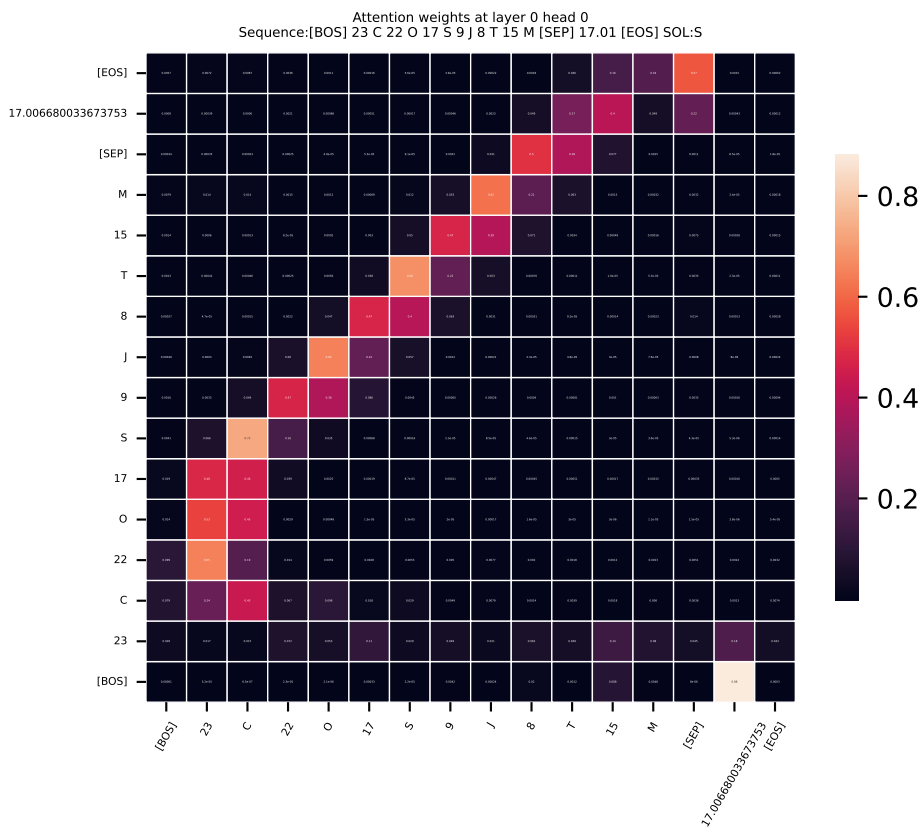
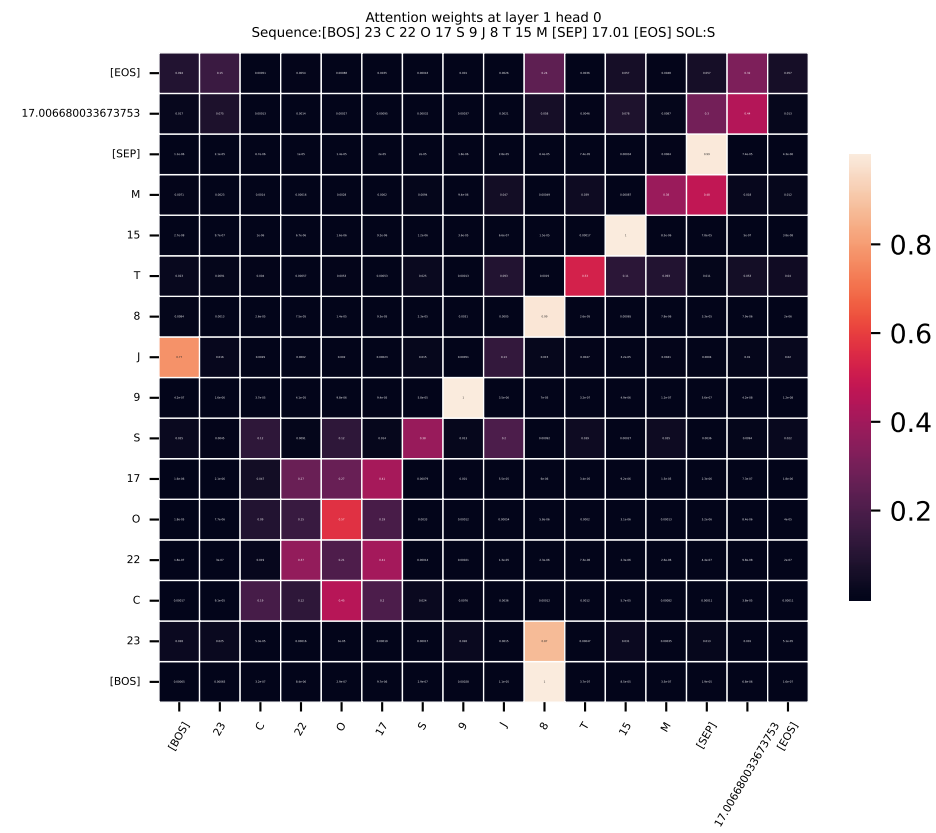

Attention weights at layer 0 head 0
Sequence:[BOS] 21 H 20 R 19 I 14 W 13 G 2 F [SEP] 20.81 [EOS] SOL:H
[EOS]
20.814563420185703
[SEP]
F
2
G
13
W
14
I
19
R
20
H
21
[BOS]
0.8
0.6
0.4
0.2

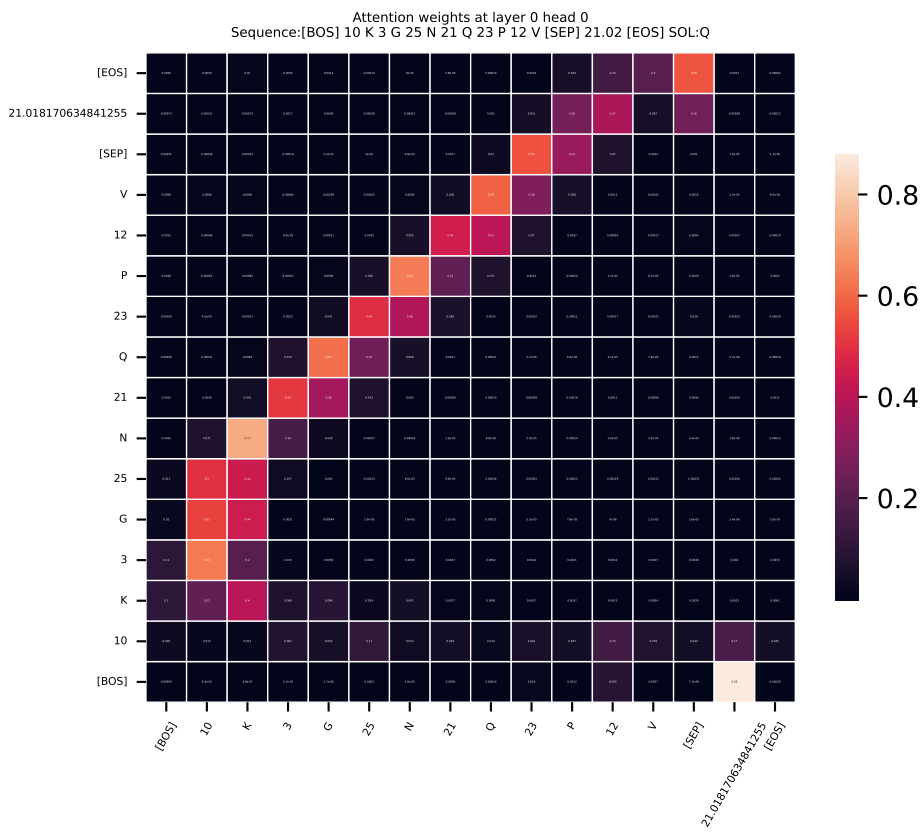

Attention weights at layer 1 head 0
Sequence:[BOS] 21 H 20 R 19 I 14 W 13 G 2 F [SEP] 20.81 [EOS] SOL:H
[EOS]
20.814563420185703
[SEP]
F
2
G
13
W
14
I
19
R
20
H
21
[BOS]
0.8
0.6
0.4
0.2

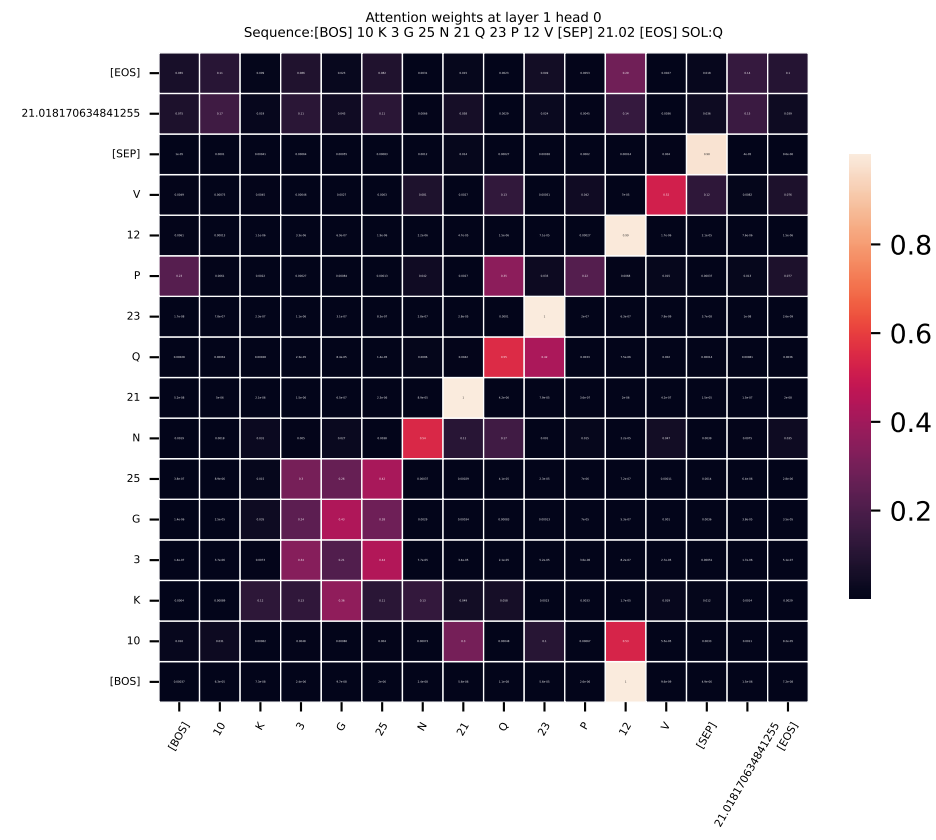

Attention weights at layer 0 head 0
Sequence:[BOS] 2 J 19 G 1 H 7 Q 17 G 5 G [SEP] 1.1 [EOS] SOL:H
[EOS]
1.103796212673116
[SEP]
G
5
G
17
Q
7
H
1
G
19
J
2
[BOS]
[BOS]
2
J
19
G
1
H
7
Q
17
G
5
G
[SEP]
1.103796212673116
[EOS]
0.8
0.6
0.4
0.2

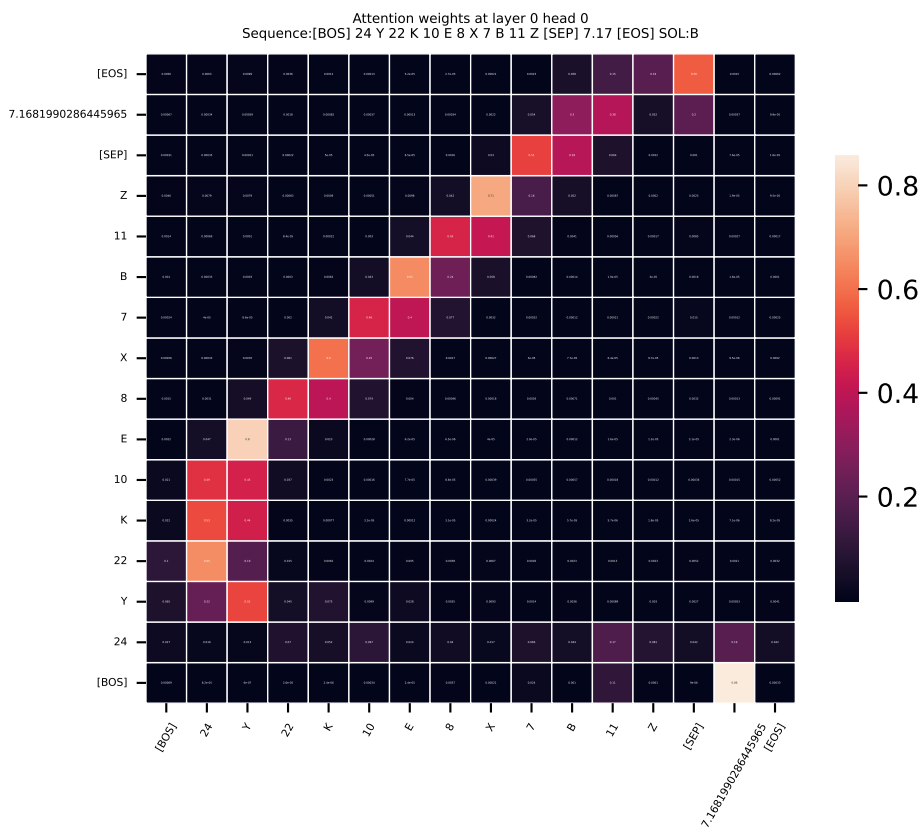

Attention weights at layer 1 head 0
Sequence:[BOS] 2 J 19 G 1 H 7 Q 17 G 5 G [SEP] 1.1 [EOS] SOL:H
[EOS]
1.103796212673116
[SEP]
G
5
G
17
Q
7
H
1
G
19
J
2
[BOS]
[BOS]
2
J
19
G
1
H
7
Q
17
G
5
G
[SEP]
1.103796212673116
[EOS]
0.8
0.6
0.4
0.2

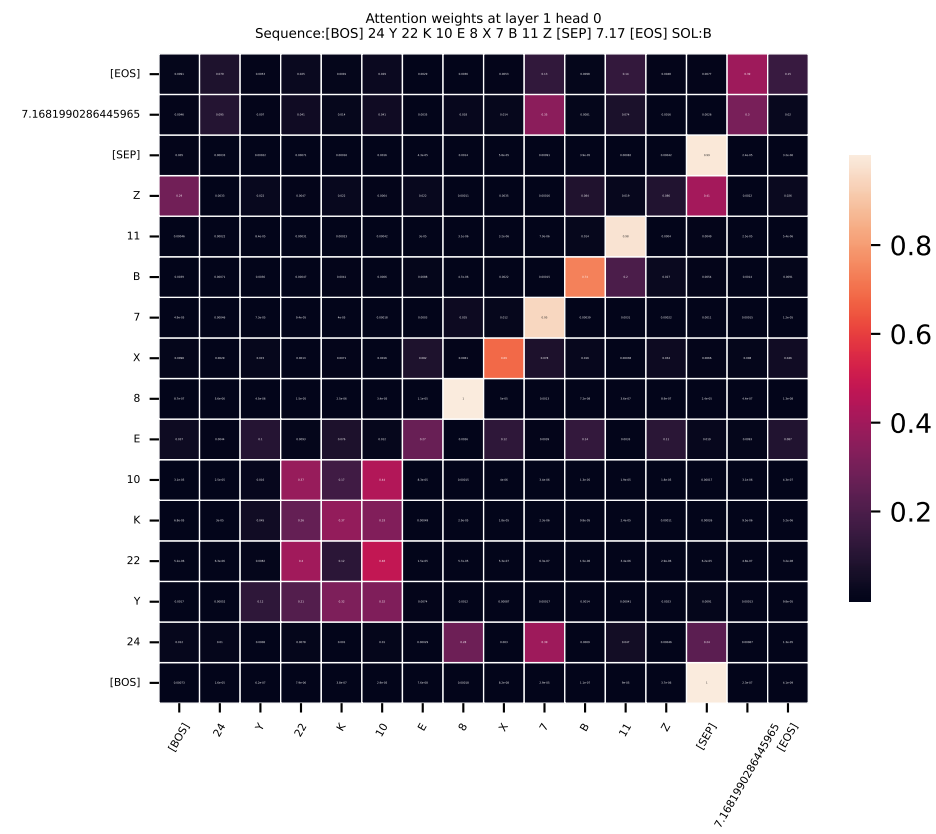

Attention weights at layer 0 head 0
Sequence:[BOS] 23 C 22 O 17 S 9 J 8 T 15 M [SEP] 17.01 [EOS] SOL:S
[EOS]
17.006680033673753
[SEP]
M
15
T
8
J
9
S
17
O
22
C
23
[BOS]
[BOS]
23
C
22
O
17
S
9
J
8
T
15
M
[SEP]
17.006680033673753
[EOS]
0.8
0.6
0.4
0.2

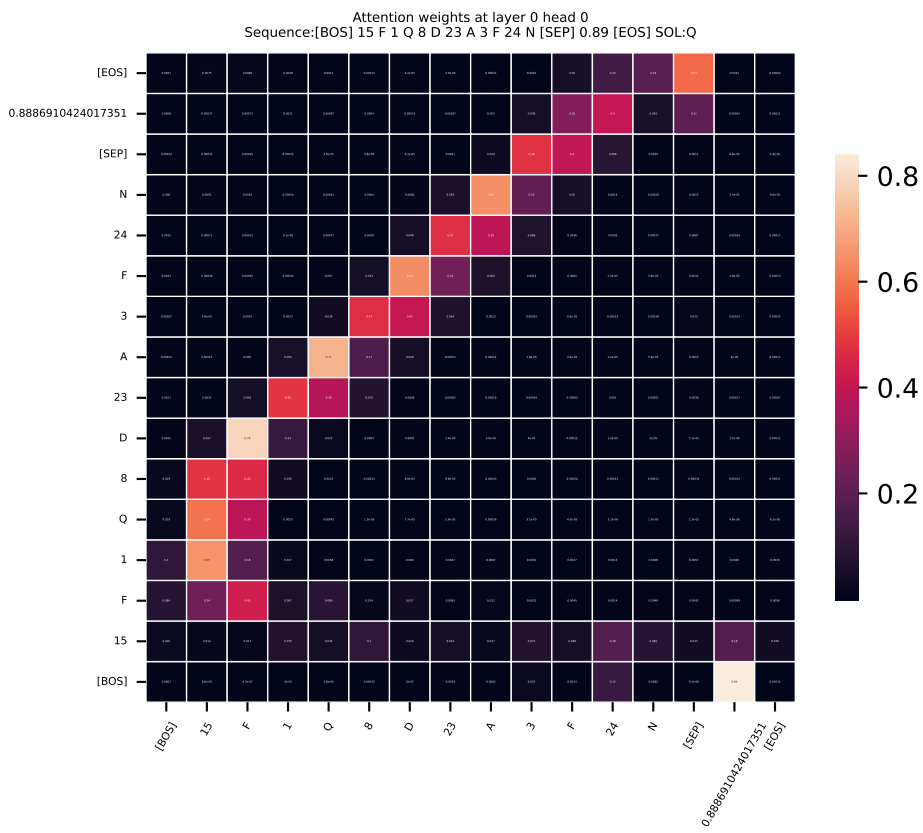

Attention weights at layer 1 head 0
Sequence:[BOS] 23 C 22 O 17 S 9 J 8 T 15 M [SEP] 17.01 [EOS] SOL:S
[EOS]
17.006680033673753
[SEP]
M
15
T
8
J
9
S
17
O
22
C
23
[BOS]
[BOS]
23
C
22
O
17
S
9
J
8
T
15
M
[SEP]
17.006680033673753
[EOS]
0.8
0.6
0.4
0.2

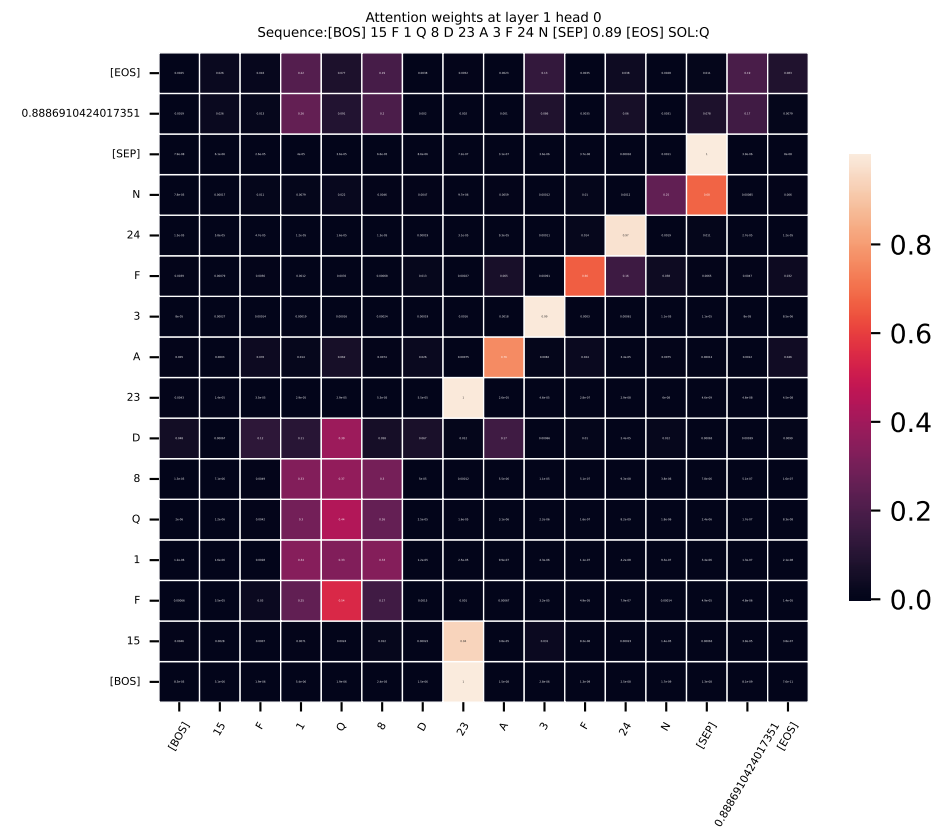

Attention weights at layer 0 head 0
Sequence:[BOS] 10 K 3 G 25 N 21 Q 23 P 12 V [SEP] 21.02 [EOS] SOL:Q
[EOS]
21.018170634841255
[SEP]
V
12
P
23
Q
21
N
25
G
3
K
10
[BOS]
[BOS] 10 K 3 G 25 N 21 Q 23 P 12 V [SEP] 21.018170634841255 [EOS]
0.8
0.6
0.4
0.2

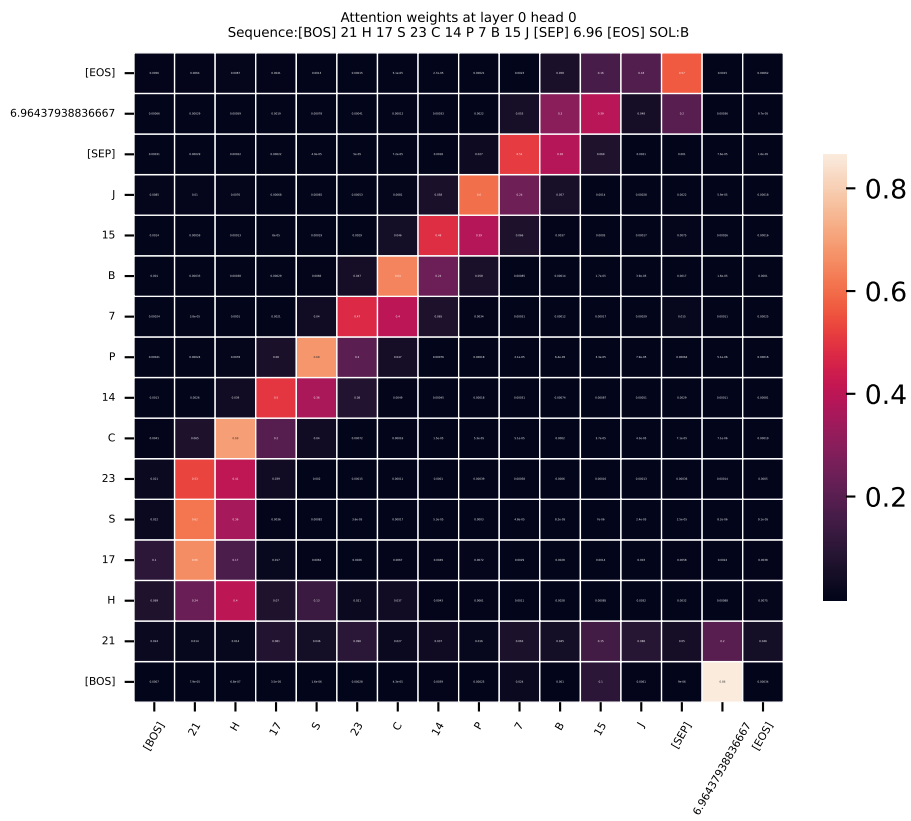

Attention weights at layer 1 head 0
Sequence:[BOS] 10 K 3 G 25 N 21 Q 23 P 12 V [SEP] 21.02 [EOS] SOL:Q
[EOS]
21.018170634841255
[SEP]
V
12
P
23
Q
21
N
25
G
3
K
10
[BOS]
[BOS] 10 K 3 G 25 N 21 Q 23 P 12 V [SEP] 21.018170634841255 [EOS]
0.8
0.6
0.4
0.2

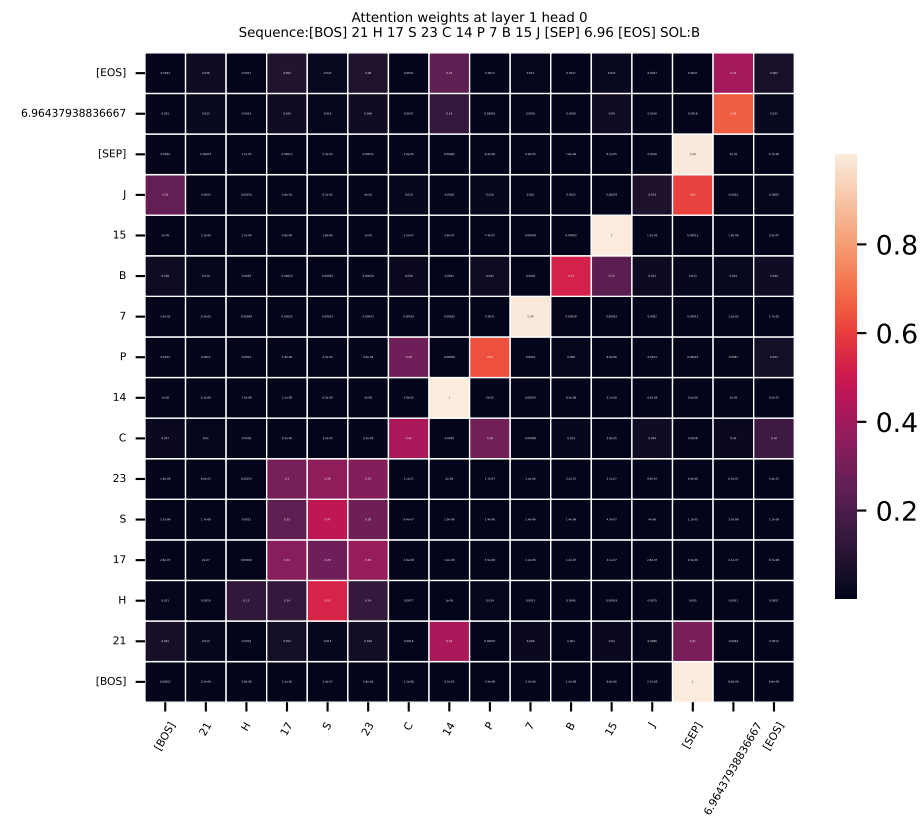

Attention weights at layer 0 head 0
Sequence:[BOS] 24 Y 22 K 10 E 8 X 7 B 11 Z [SEP] 7.17 [EOS] SOL:B
[EOS]
7.1681990286445965
[SEP]
Z
11
B
7
X
8
E
10
K
22
Y
24
[BOS]
[BOS]
24
Y
22
K
10
E
8
X
7
B
11
Z
[SEP]
7.1681990286445965
[EOS]
0.8
0.6
0.4
0.2

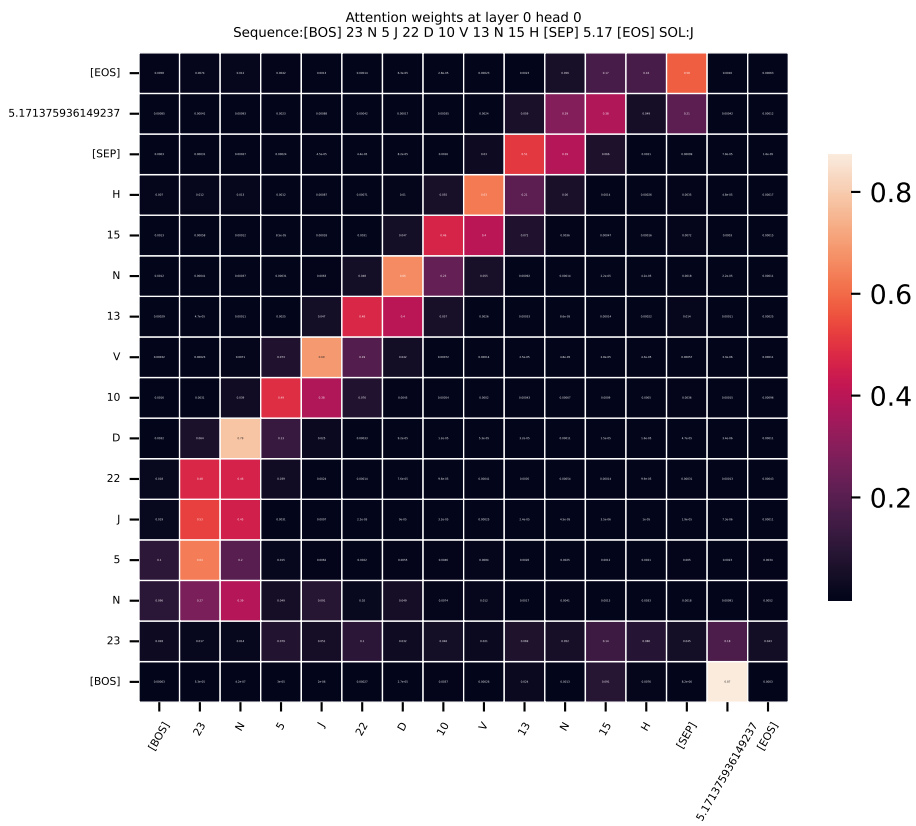

Attention weights at layer 1 head 0
Sequence:[BOS] 24 Y 22 K 10 E 8 X 7 B 11 Z [SEP] 7.17 [EOS] SOL:B
[EOS]
7.1681990286445965
[SEP]
Z
11
B
7
X
8
E
10
K
22
Y
24
[BOS]
[BOS]
24
Y
22
K
10
E
8
X
7
B
11
Z
[SEP]
7.1681990286445965
[EOS]
0.8
0.6
0.4
0.2

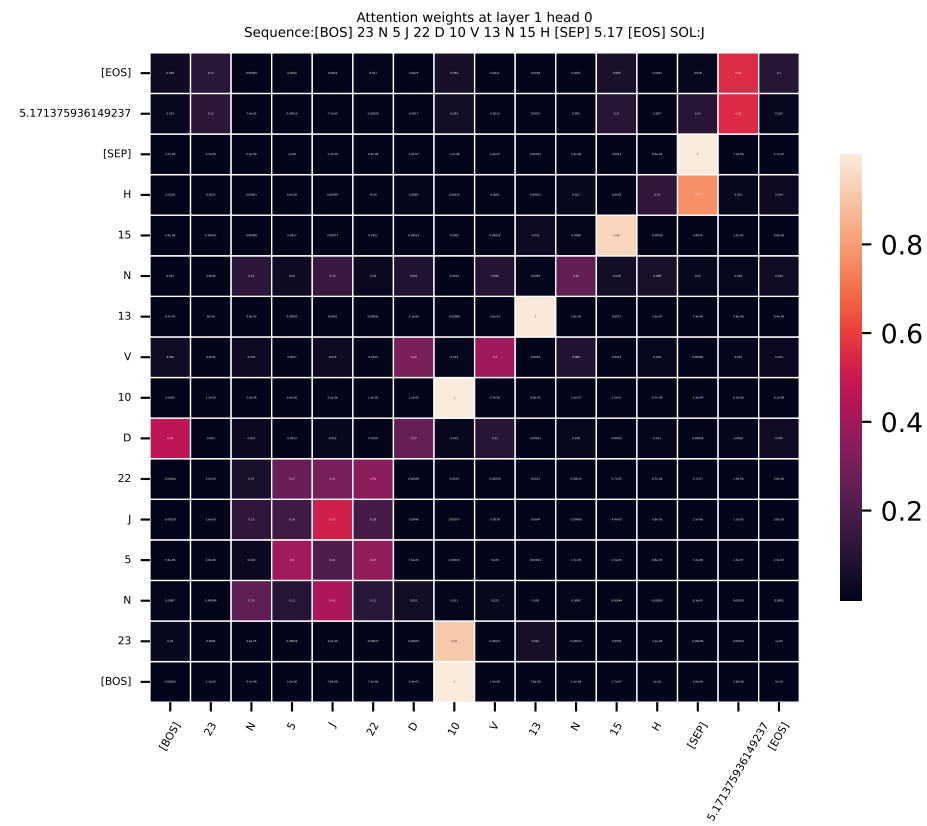

Attention weights at layer 0 head 0
Sequence:[BOS] 15 F 1 Q 8 D 23 A 3 F 24 N [SEP] 0.89 [EOS] SOL:Q
[EOS]
0.8886910424017351
[SEP]
N
24
F
3
A
23
D
8
Q
1
F
15
[BOS]
[BOS]
15
F
1
Q
8
D
23
A
3
F
24
N
[SEP]
0.8886910424017351
[EOS]
0.8
0.6
0.4
0.2

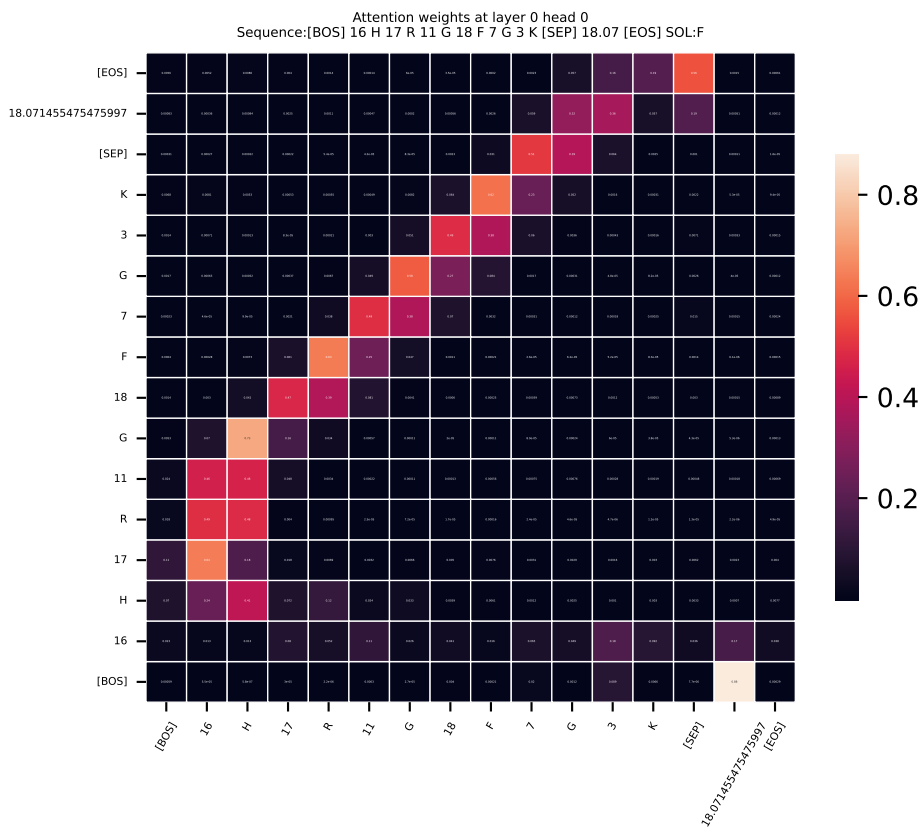

Attention weights at layer 1 head 0
Sequence:[BOS] 15 F 1 Q 8 D 23 A 3 F 24 N [SEP] 0.89 [EOS] SOL:Q
[EOS]
0.8886910424017351
[SEP]
N
24
F
3
A
23
D
8
Q
1
F
15
[BOS]
[BOS]
15
F
1
Q
8
D
23
A
3
F
24
N
[SEP]
0.8886910424017351
[EOS]
0.8
0.6
0.4
0.2
0.0

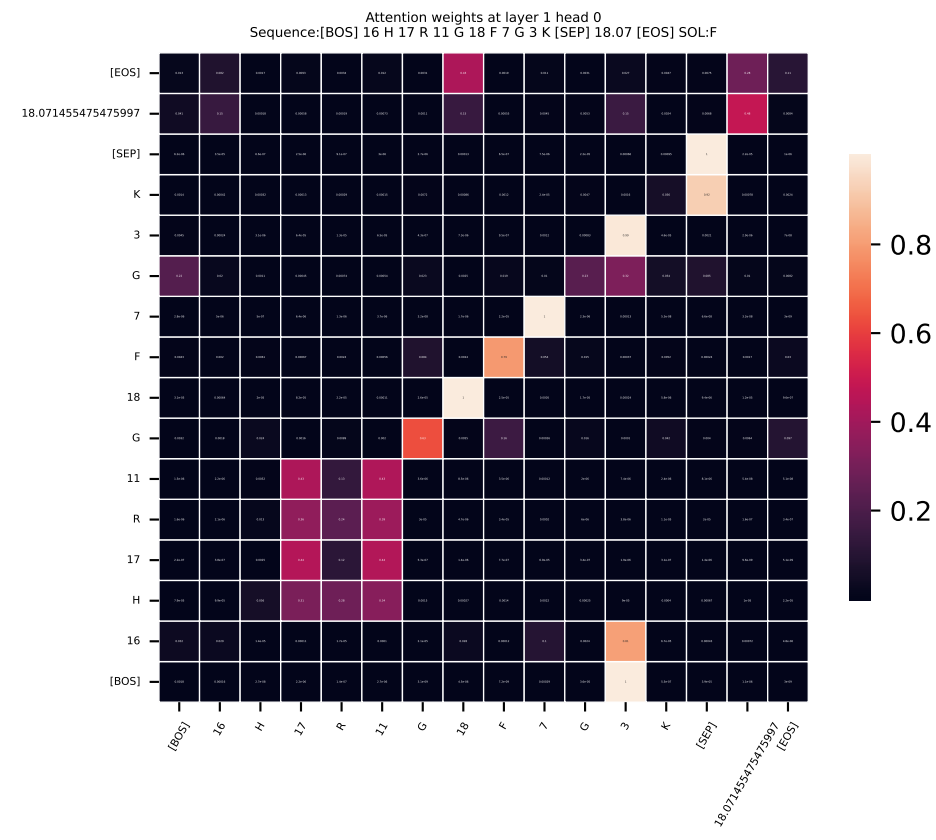

Attention weights at layer 0 head 0
Sequence:[BOS] 21 H 17 S 23 C 14 P 7 B 15 J [SEP] 6.96 [EOS] SOL:B
[EOS]
6.96437938836667
[SEP]
J
15
B
7
P
14
C
23
S
17
H
21
[BOS]
[BOS]
21
H
17
S
23
C
14
P
7
B
15
J
[SEP]
6.96437938836667
[EOS]
0.8
0.6
0.4
0.2

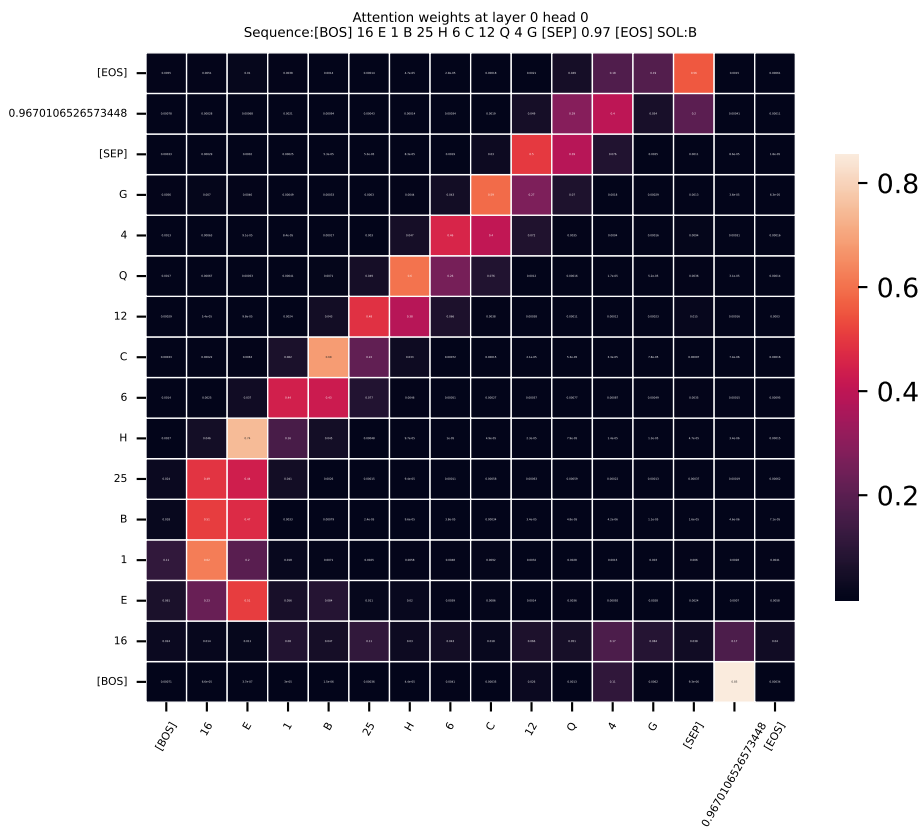

Attention weights at layer 1 head 0
Sequence:[BOS] 21 H 17 S 23 C 14 P 7 B 15 J [SEP] 6.96 [EOS] SOL:B
[EOS]
6.96437938836667
[SEP]
J
15
B
7
P
14
C
23
S
17
H
21
[BOS]
[BOS]
21
H
17
S
23
C
14
P
7
B
15
J
[SEP]
6.96437938836667
[EOS]
0.8
0.6
0.4
0.2

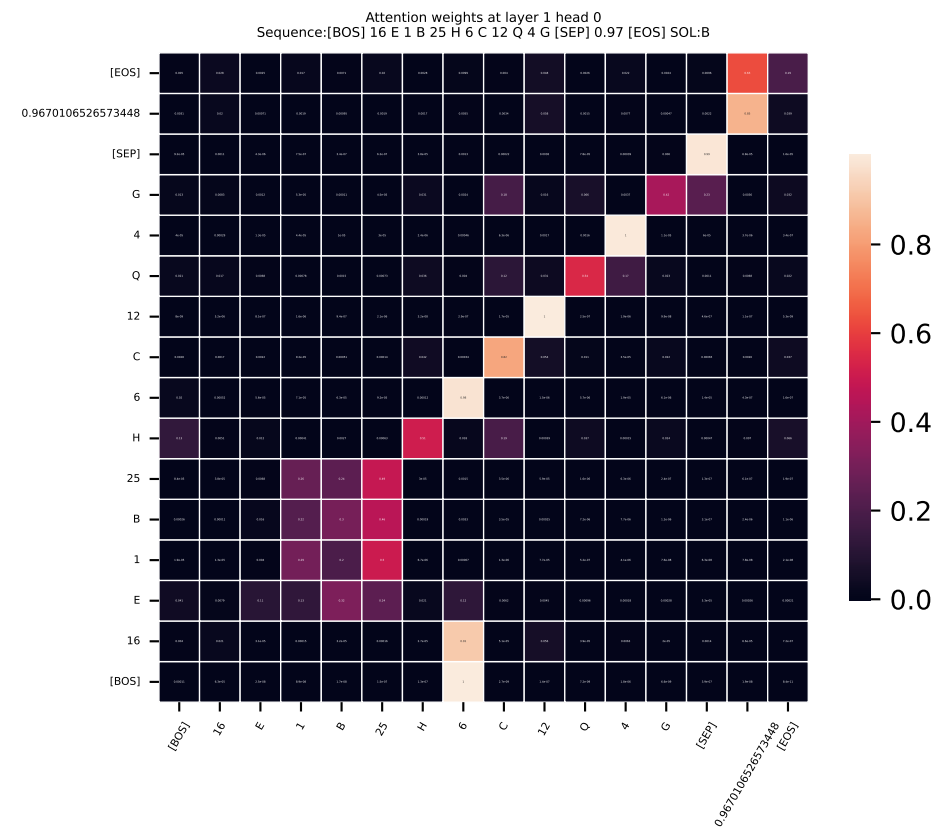

Attention weights at layer 0 head 0
Sequence:[BOS] 23 N 5 J 22 D 10 V 13 N 15 H [SEP] 5.17 [EOS] SOL:J
[EOS]
5.171375936149237
[SEP]
H
15
N
13
V
10
D
22
J
5
N
23
[BOS]
[BOS]
23
N
5
J
22
D
10
V
13
N
15
H
[SEP]
5.171375936149237
[EOS]
0.8
0.6
0.4
0.2

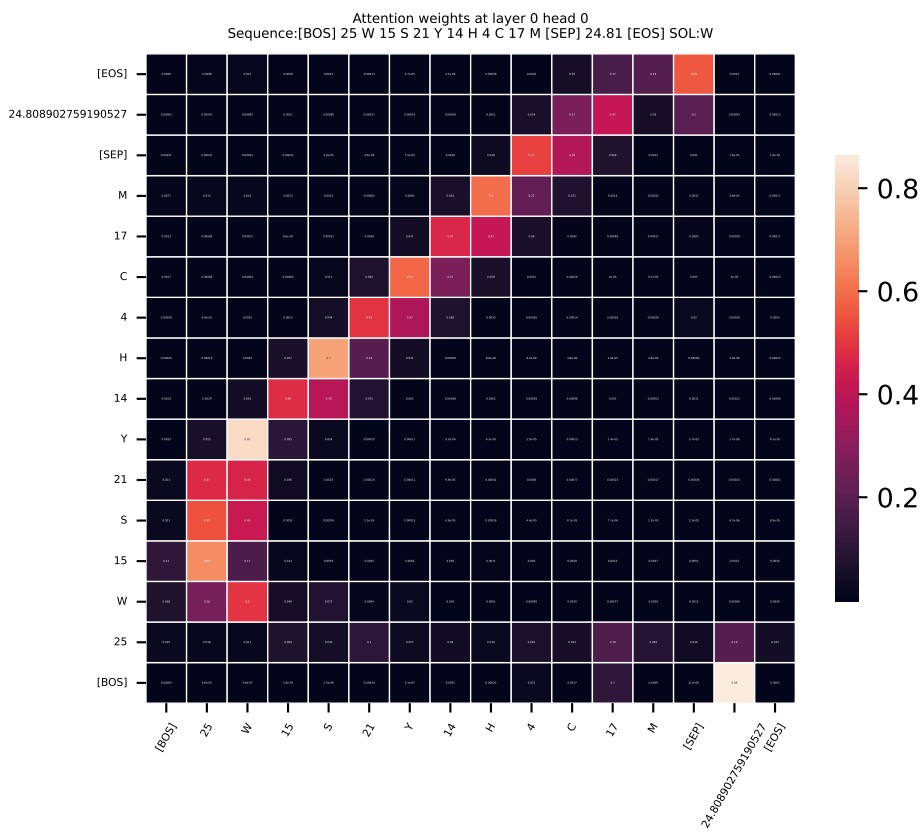

Attention weights at layer 1 head 0
Sequence:[BOS] 23 N 5 J 22 D 10 V 13 N 15 H [SEP] 5.17 [EOS] SOL:J
[EOS]
5.171375936149237
[SEP]
H
15
N
13
V
10
D
22
J
5
N
23
[BOS]
[BOS]
23
N
5
J
22
D
10
V
13
N
15
H
[SEP]
5.171375936149237
[EOS]
0.8
0.6
0.4
0.2

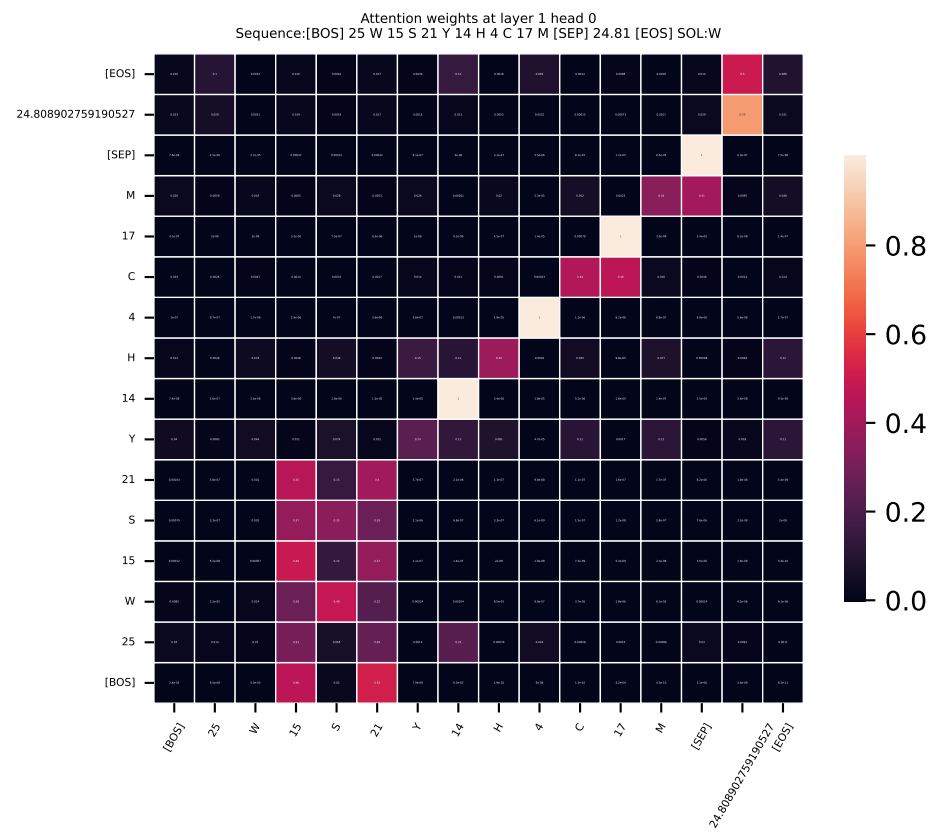

Attention weights at layer 0 head 0
Sequence:[BOS] 16 H 17 R 11 G 18 F 7 G 3 K [SEP] 18.07 [EOS] SOL:F
[EOS]
18.071455475475997
[SEP]
K
3
G
7
F
18
G
11
R
17
H
16
[BOS]
[BOS]
16
H
17
R
11
G
18
F
7
G
3
K
[SEP]
18.071455475475997
[EOS]
0.8
0.6
0.4
0.2

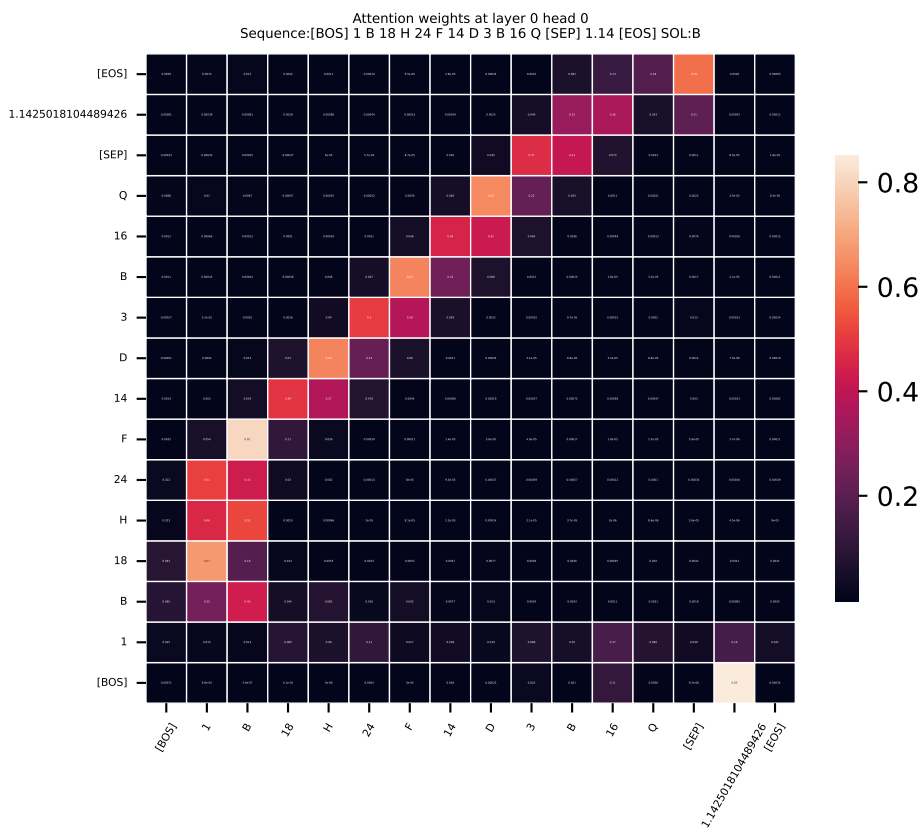

Attention weights at layer 1 head 0
Sequence:[BOS] 16 H 17 R 11 G 18 F 7 G 3 K [SEP] 18.07 [EOS] SOL:F
[EOS]
18.071455475475997
[SEP]
K
3
G
7
F
18
G
11
R
17
H
16
[BOS]
[BOS]
16
H
17
R
11
G
18
F
7
G
3
K
[SEP]
18.071455475475997
[EOS]
0.8
0.6
0.4
0.2

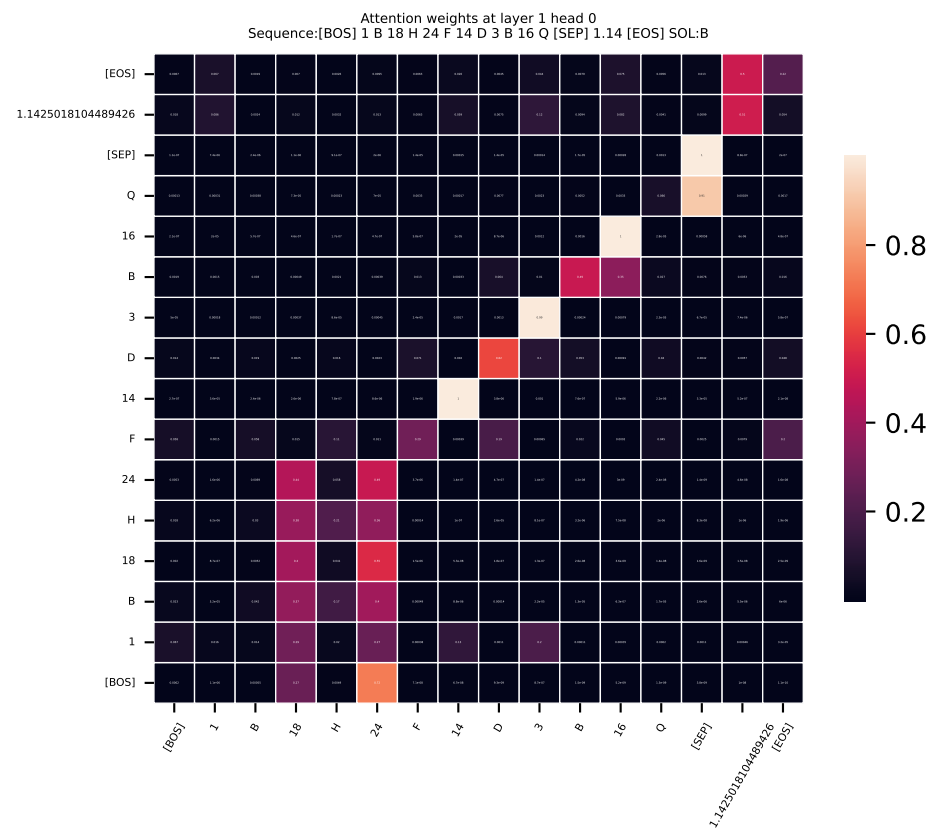

Attention weights at layer 0 head 0
Sequence:[BOS] 16 E 1 B 25 H 6 C 12 Q 4 G [SEP] 0.97 [EOS] SOL:B
[EOS]
0.9670106526573448
[SEP]
G
4
Q
12
C
6
H
25
B
1
E
16
[BOS]
0.8
0.6
0.4
0.2

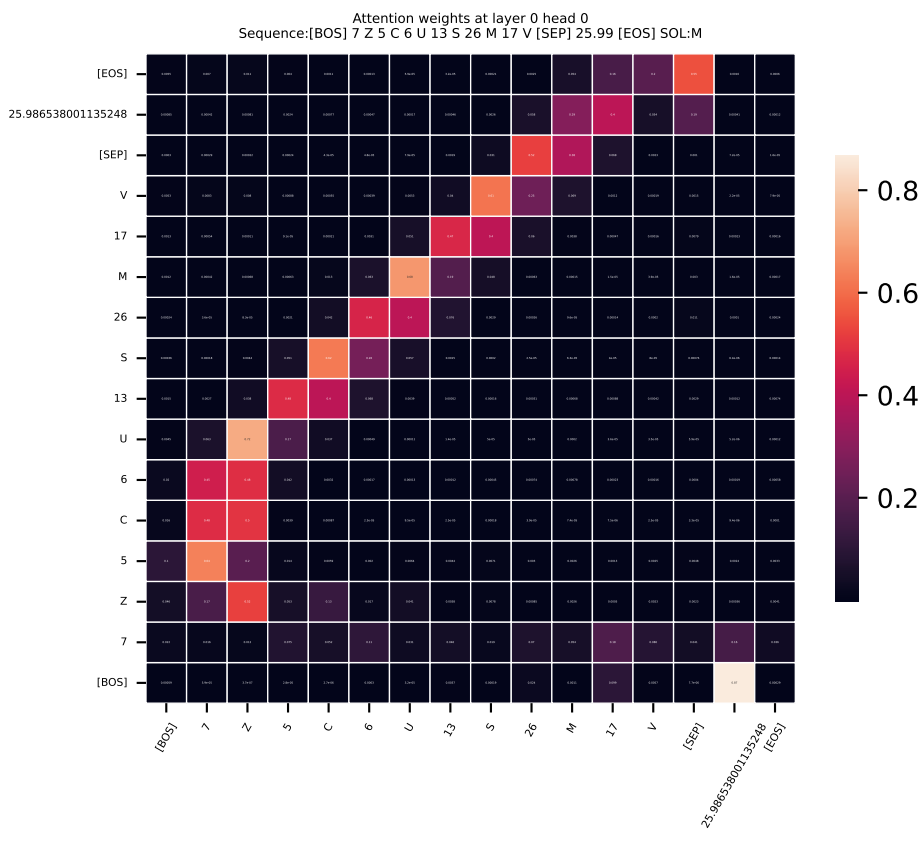

Attention weights at layer 1 head 0
Sequence:[BOS] 16 E 1 B 25 H 6 C 12 Q 4 G [SEP] 0.97 [EOS] SOL:B
[EOS]
0.9670106526573448
[SEP]
G
4
Q
12
C
6
H
25
B
1
E
16
[BOS]
0.8
0.6
0.4
0.2
0.0

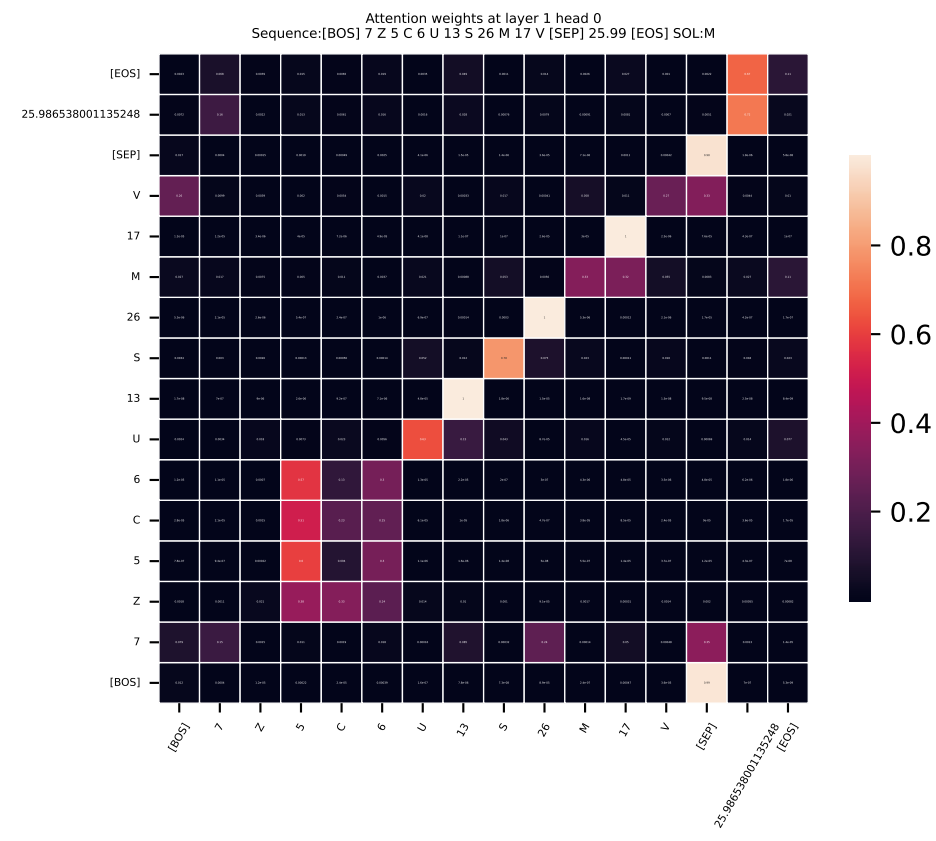

Attention weights at layer 0 head 0
Sequence:[BOS] 25 W 15 S 21 Y 14 H 4 C 17 M [SEP] 24.81 [EOS] SOL:W
[EOS]
24.808902759190527
[SEP]
M
17
C
4
H
14
Y
21
S
15
W
25
[BOS]
0.8
0.6
0.4
0.2

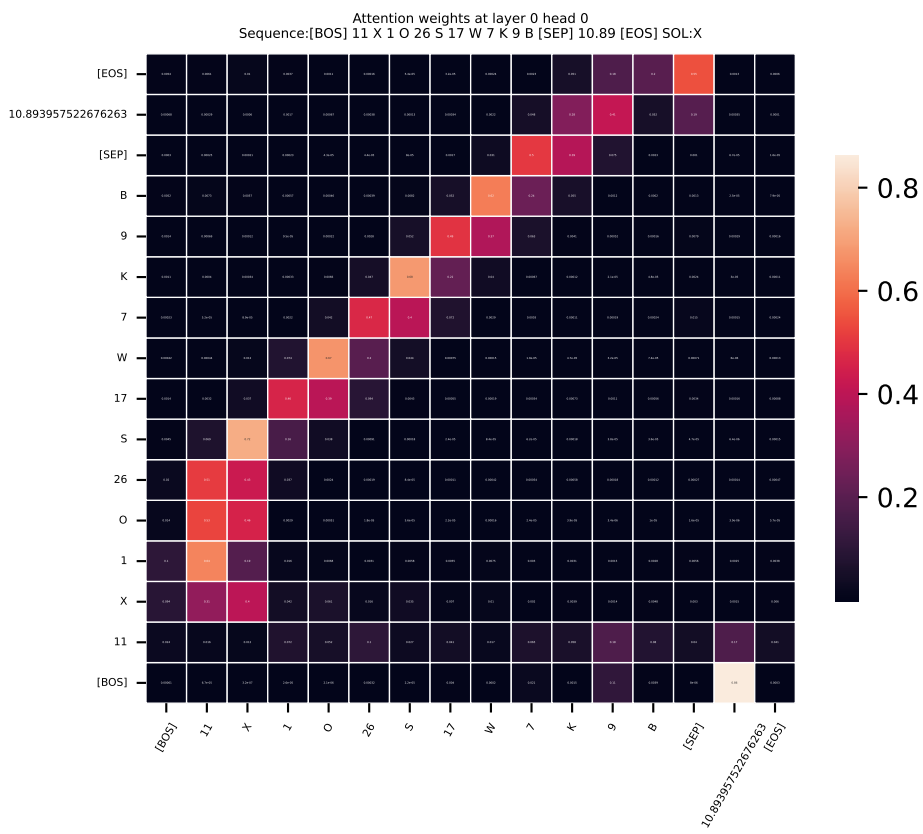

Attention weights at layer 1 head 0
Sequence:[BOS] 25 W 15 S 21 Y 14 H 4 C 17 M [SEP] 24.81 [EOS] SOL:W
[EOS]
24.808902759190527
[SEP]
M
17
C
4
H
14
Y
21
S
15
W
25
[BOS]
0.8
0.6
0.4
0.2
0.0

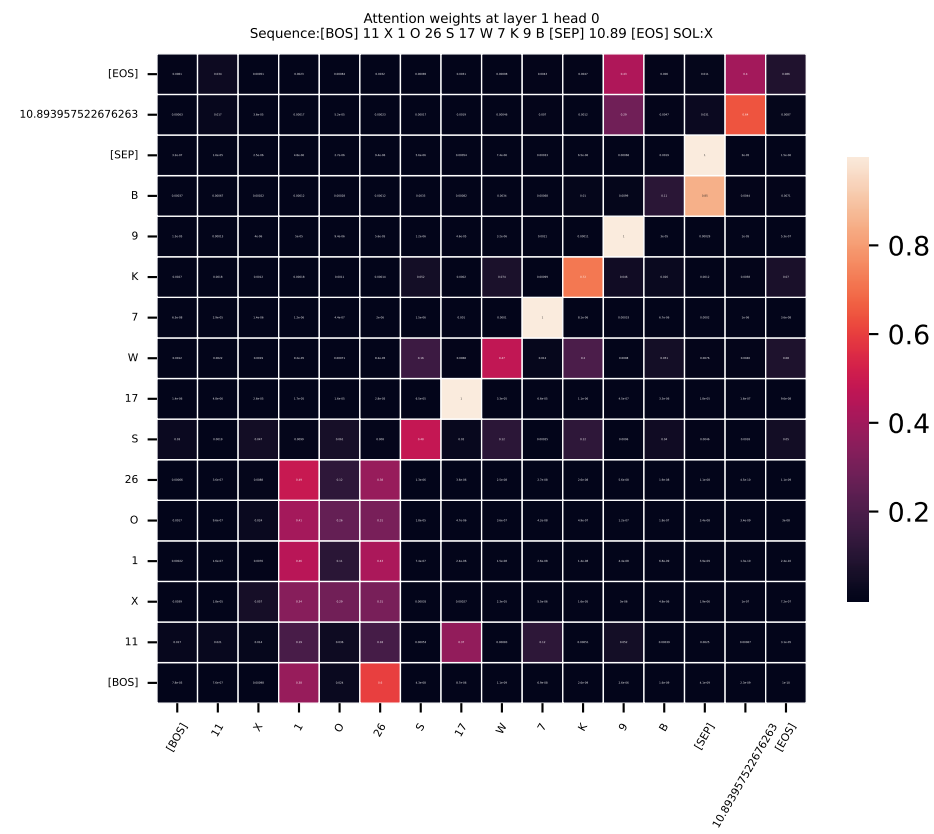

Attention weights at layer 0 head 0
Sequence:[BOS] 1 B 18 H 24 F 14 D 3 B 16 Q [SEP] 1.14 [EOS] SOL:B
[EOS]
1.1425018104489426
[SEP]
Q
16
B
3
D
14
F
24
H
18
B
1
[BOS]
[BOS] 1 B 18 H 24 F 14 D 3 B 16 Q [SEP] 1.1425018104489426 [EOS]
0.8
0.6
0.4
0.2

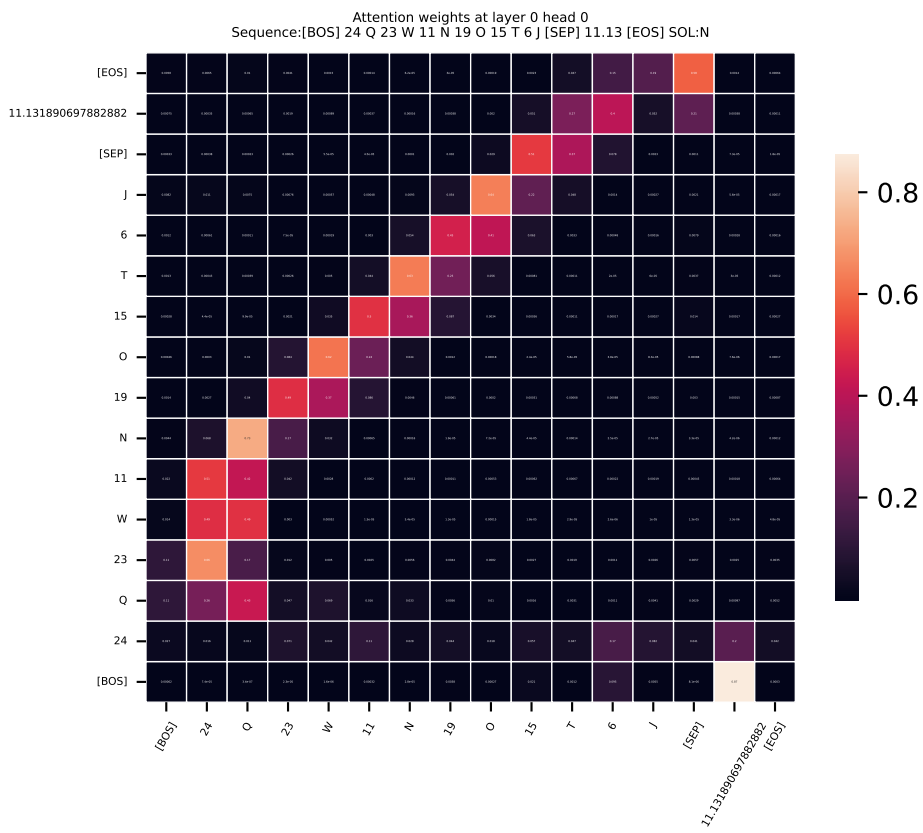

Attention weights at layer 1 head 0
Sequence:[BOS] 1 B 18 H 24 F 14 D 3 B 16 Q [SEP] 1.14 [EOS] SOL:B
[EOS]
1.1425018104489426
[SEP]
Q
16
B
3
D
14
F
24
H
18
B
1
[BOS]
[BOS] 1 B 18 H 24 F 14 D 3 B 16 Q [SEP] 1.1425018104489426 [EOS]
0.8
0.6
0.4
0.2

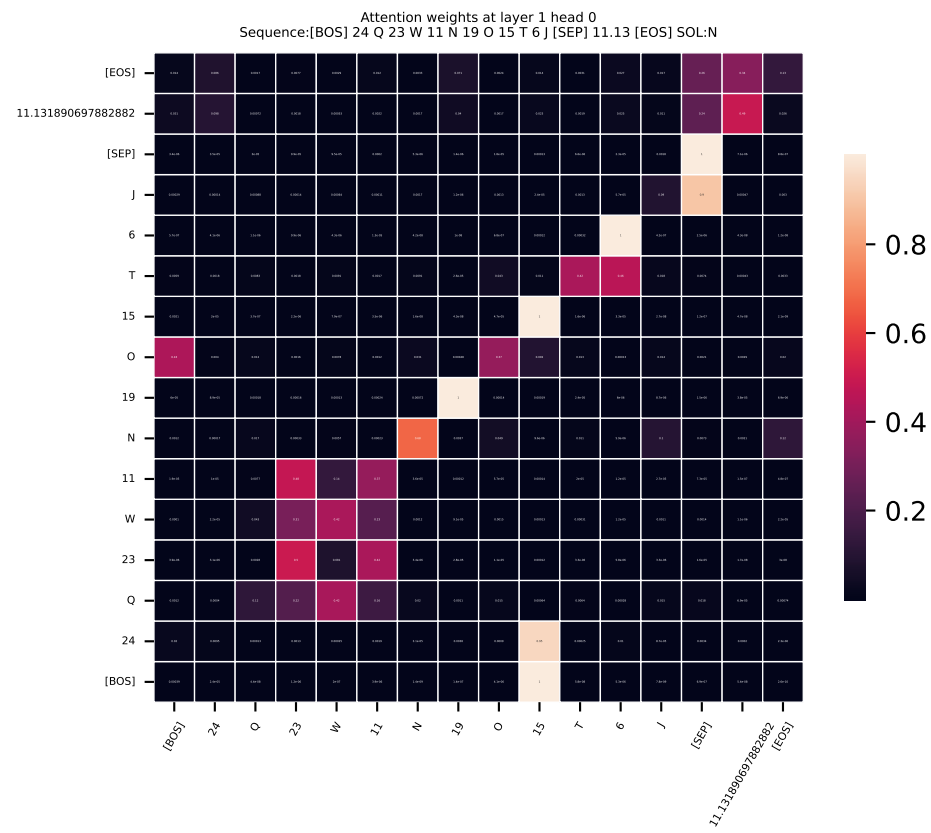

Attention weights at layer 0 head 0
Sequence:[BOS] 7 Z 5 C 6 U 13 S 26 M 17 V [SEP] 25.99 [EOS] SOL:M
[EOS]
25.986538001135248
[SEP]
V
17
M
26
S
13
U
6
C
5
Z
7
[BOS]
[BOS] 7 Z 5 C 6 U 13 S 26 M 17 V [SEP] 25.986538001135248 [EOS]
0.8
0.6
0.4
0.2

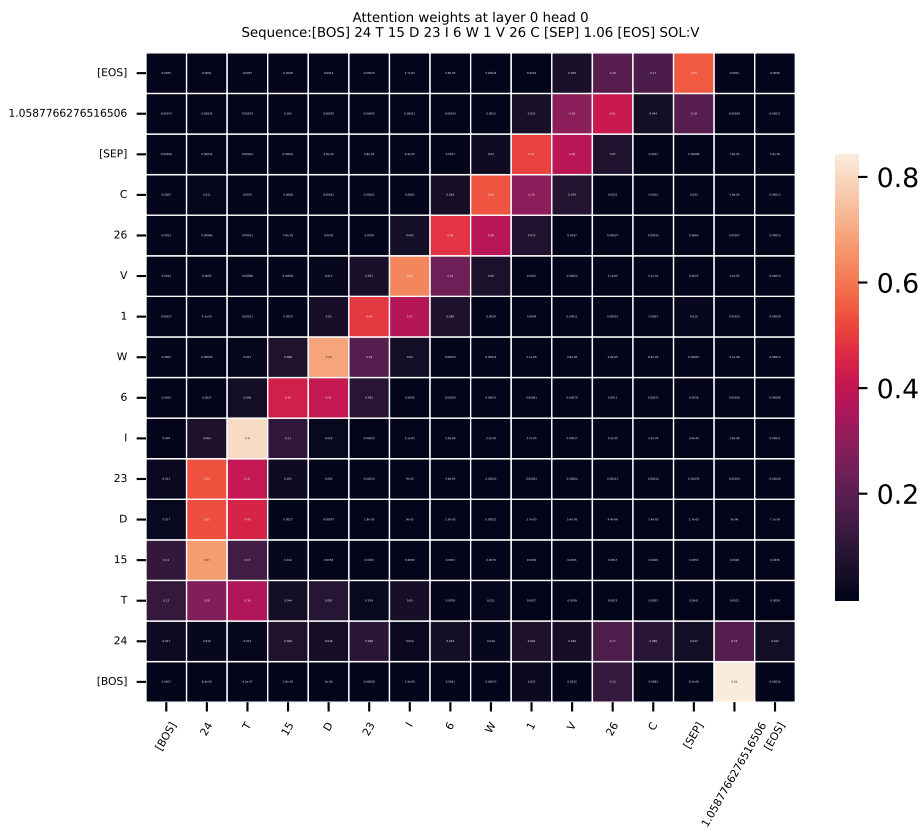

Attention weights at layer 1 head 0
Sequence:[BOS] 7 Z 5 C 6 U 13 S 26 M 17 V [SEP] 25.99 [EOS] SOL:M
[EOS]
25.986538001135248
[SEP]
V
17
M
26
S
13
U
6
C
5
Z
7
[BOS]
[BOS] 7 Z 5 C 6 U 13 S 26 M 17 V [SEP] 25.986538001135248 [EOS]
0.8
0.6
0.4
0.2

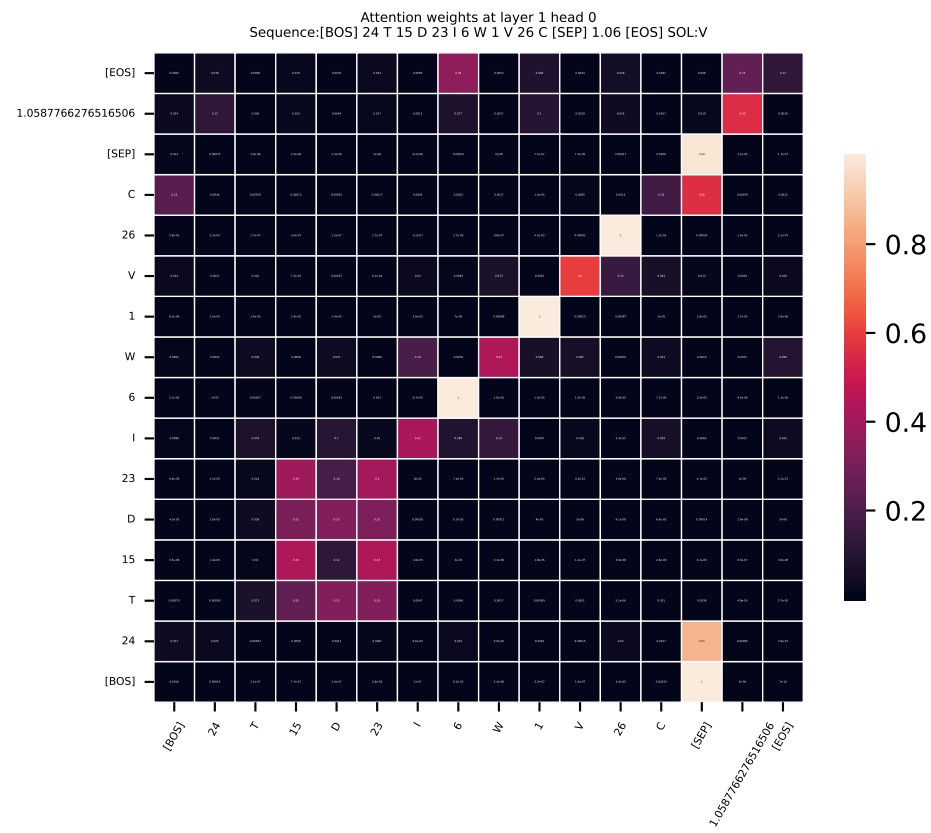

Attention weights at layer 0 head 0
Sequence:[BOS] 11 X 1 O 26 S 17 W 7 K 9 B [SEP] 10.89 [EOS] SOL:X
[EOS]
10.893957522676263
[SEP]
B
9
K
7
W
17
S
26
O
1
X
11
[BOS]
0.8
0.6
0.4
0.2

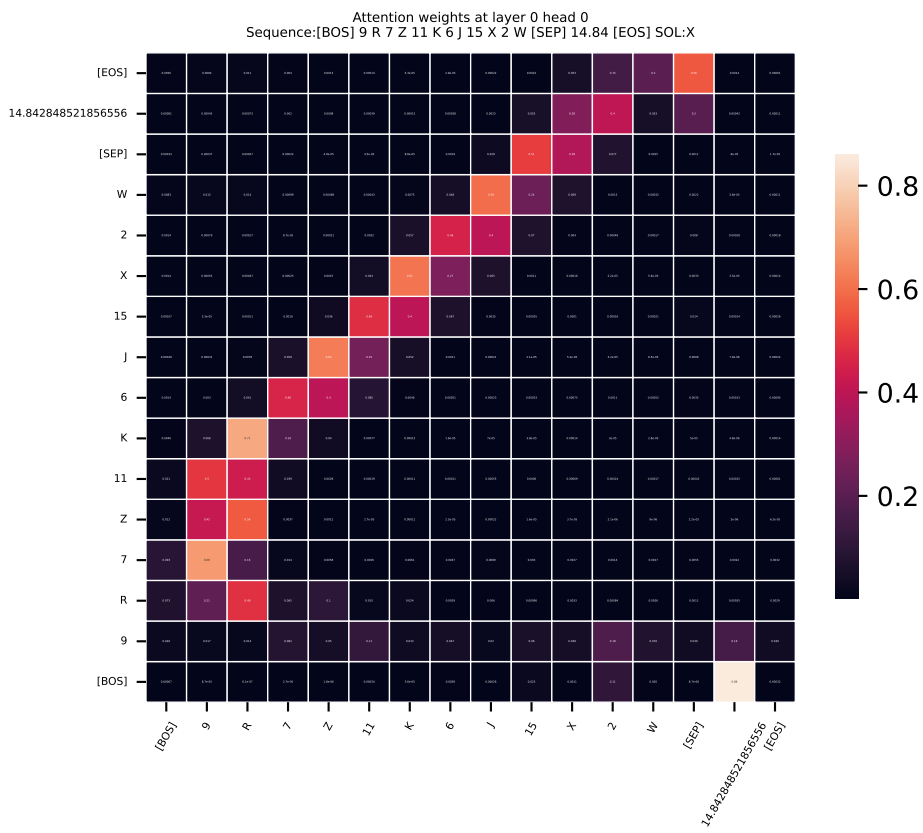

Attention weights at layer 1 head 0
Sequence:[BOS] 11 X 1 O 26 S 17 W 7 K 9 B [SEP] 10.89 [EOS] SOL:X
[EOS]
10.893957522676263
[SEP]
B
9
K
7
W
17
S
26
O
1
X
11
[BOS]
0.8
0.6
0.4
0.2

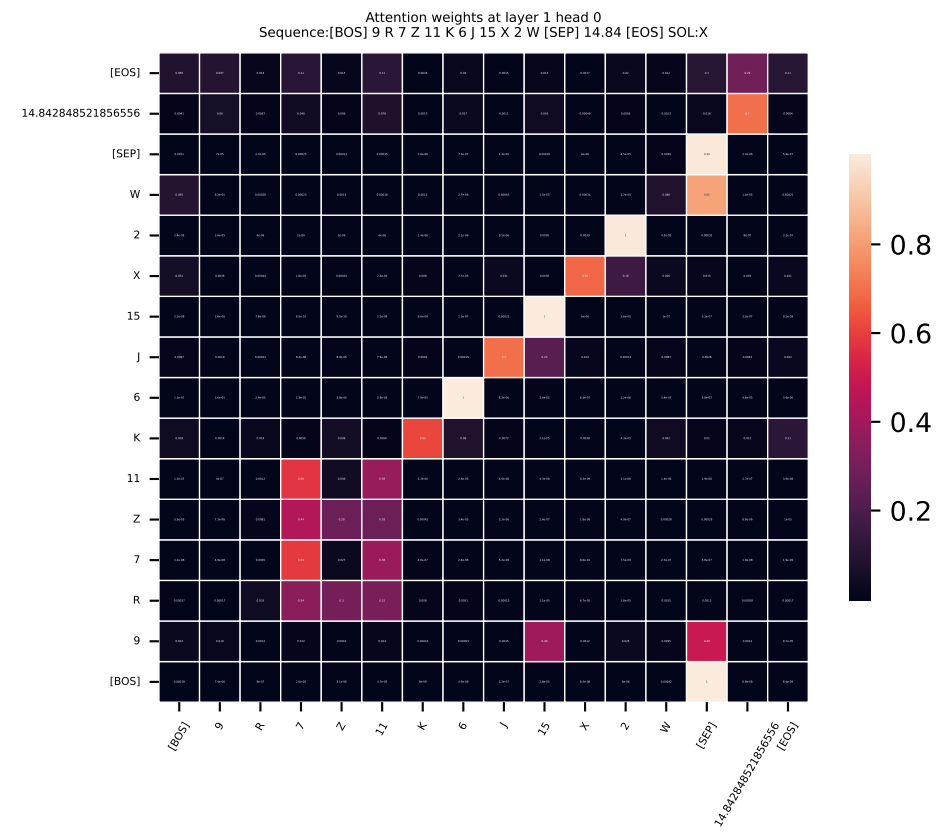

Attention weights at layer 0 head 0
Sequence:[BOS] 24 Q 23 W 11 N 19 O 15 T 6 J [SEP] 11.13 [EOS] SOL:N
[EOS]
11.131890697882882
[SEP]
J
6
T
15
O
19
N
11
W
23
Q
24
[BOS]
0.8
0.6
0.4
0.2

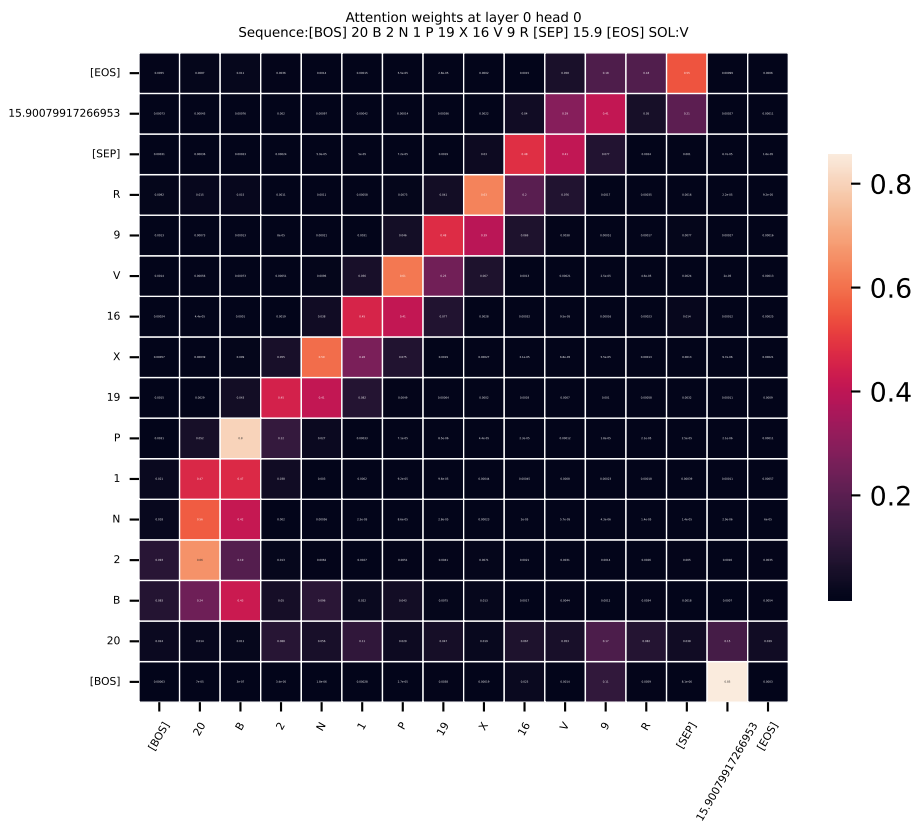

Attention weights at layer 1 head 0
Sequence:[BOS] 24 Q 23 W 11 N 19 O 15 T 6 J [SEP] 11.13 [EOS] SOL:N
[EOS]
11.131890697882882
[SEP]
J
6
T
15
O
19
N
11
W
23
Q
24
[BOS]
0.8
0.6
0.4
0.2

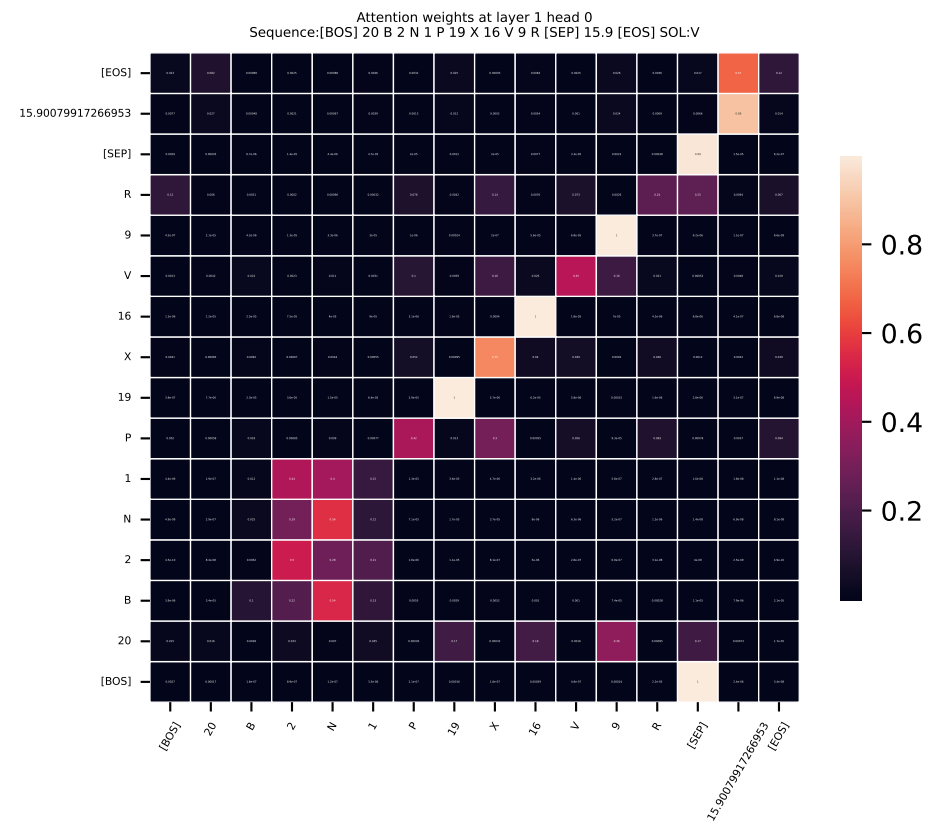

Attention weights at layer 0 head 0
Sequence:[BOS] 24 T 15 D 23 I 6 W 1 V 26 C [SEP] 1.06 [EOS] SOL:V
[EOS]
1.0587766276516506
[SEP]
C
26
V
1
W
6
I
23
D
15
T
24
[BOS]
0.8
0.6
0.4
0.2

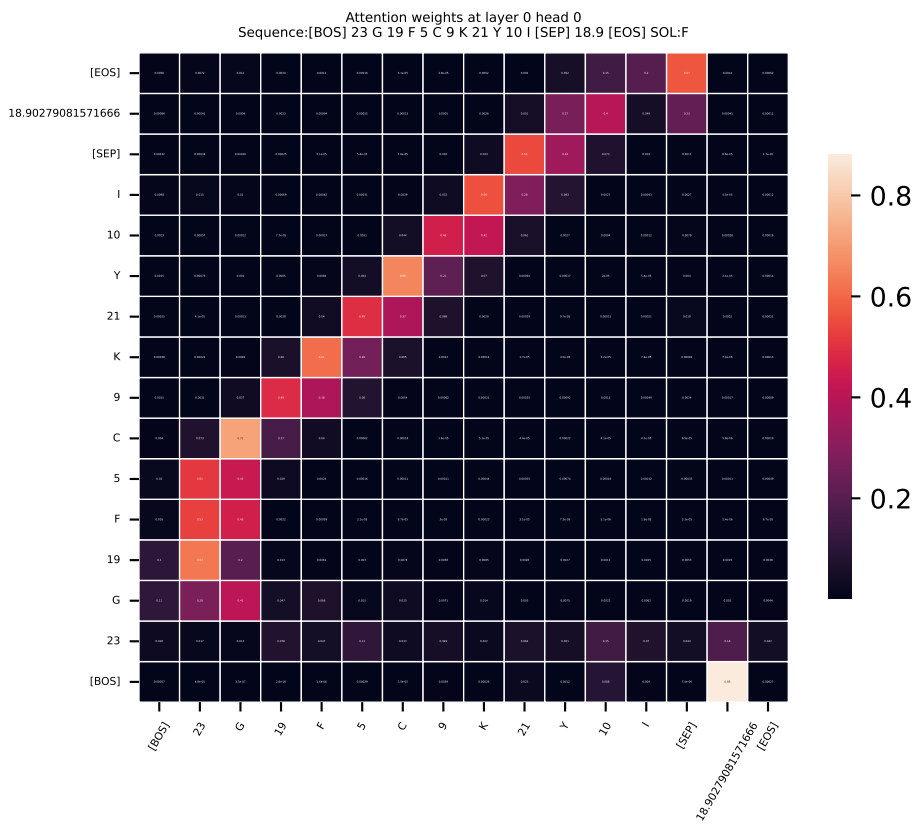

Attention weights at layer 1 head 0
Sequence:[BOS] 24 T 15 D 23 I 6 W 1 V 26 C [SEP] 1.06 [EOS] SOL:V
[EOS]
1.0587766276516506
[SEP]
C
26
V
1
W
6
I
23
D
15
T
24
[BOS]
0.8
0.6
0.4
0.2

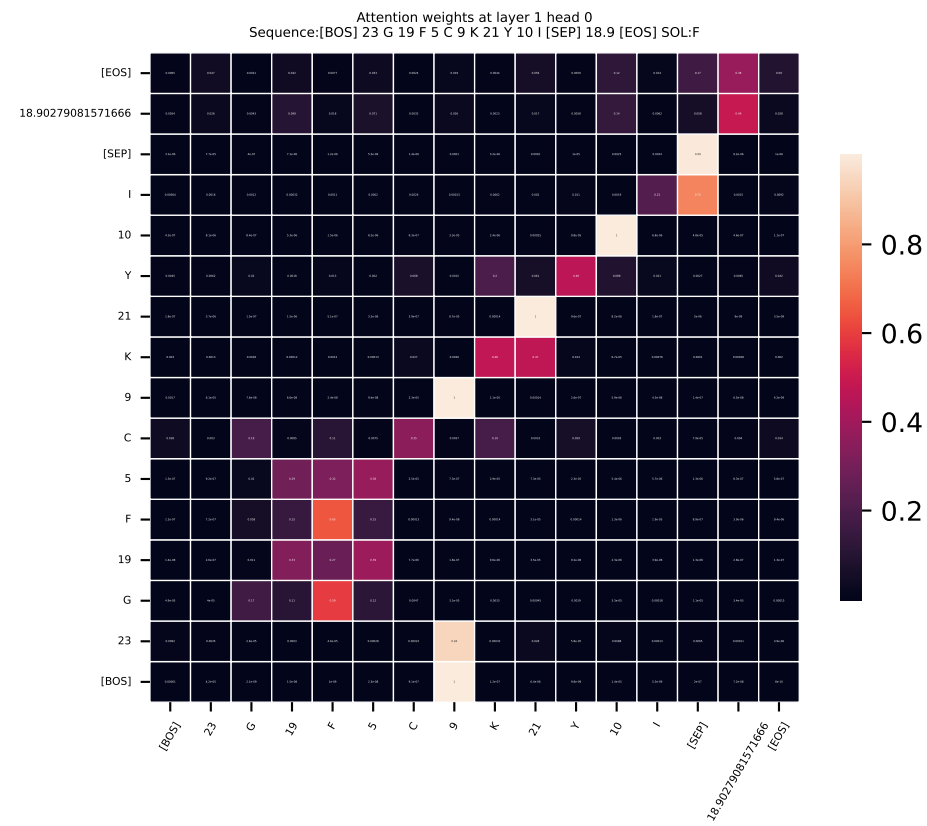

Attention weights at layer 0 head 0
Sequence:[BOS] 9 R 7 Z 11 K 6 J 15 X 2 W [SEP] 14.84 [EOS] SOL:X
[EOS]
14.842848521856556
[SEP]
W
2
X
15
J
6
K
11
Z
7
R
9
[BOS]
[BOS] 9 R 7 Z 11 K 6 J 15 X 2 W [SEP] 14.842848521856556 [EOS]
0.8
0.6
0.4
0.2

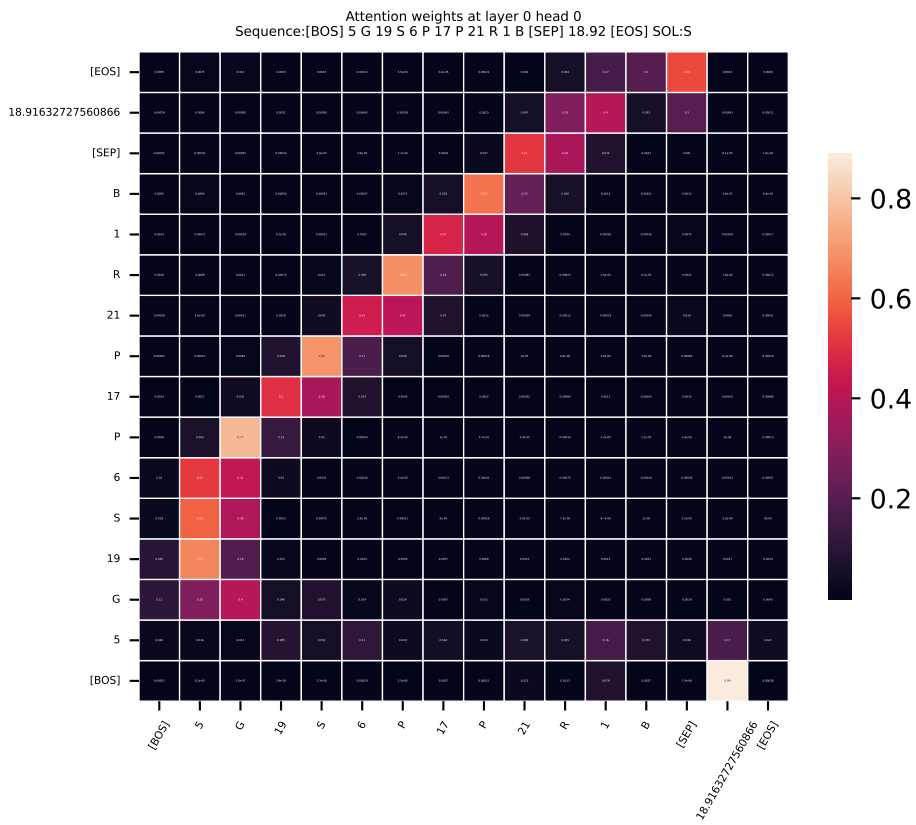

Attention weights at layer 1 head 0
Sequence:[BOS] 9 R 7 Z 11 K 6 J 15 X 2 W [SEP] 14.84 [EOS] SOL:X
[EOS]
14.842848521856556
[SEP]
W
2
X
15
J
6
K
11
Z
7
R
9
[BOS]
[BOS] 9 R 7 Z 11 K 6 J 15 X 2 W [SEP] 14.842848521856556 [EOS]
0.8
0.6
0.4
0.2

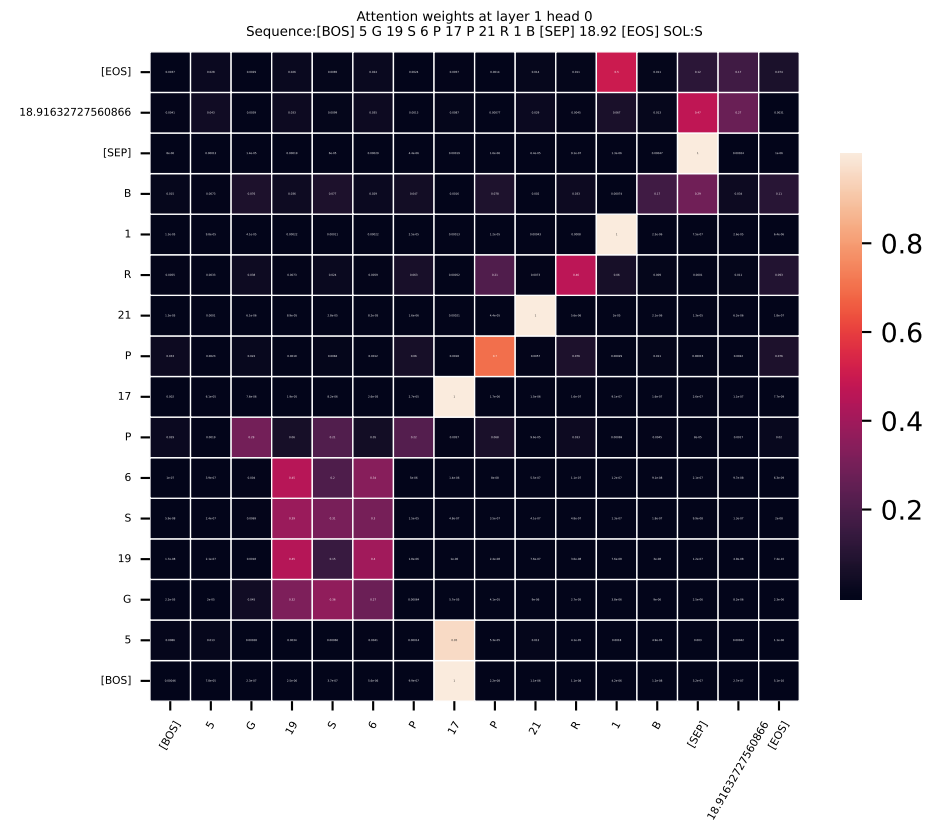

Attention weights at layer 0 head 0
Sequence:[BOS] 20 B 2 N 1 P 19 X 16 V 9 R [SEP] 15.9 [EOS] SOL:V
[EOS]
15.90079917266953
[SEP]
R
9
V
16
X
19
P
1
N
2
B
20
[BOS]
0.8
0.6
0.4
0.2

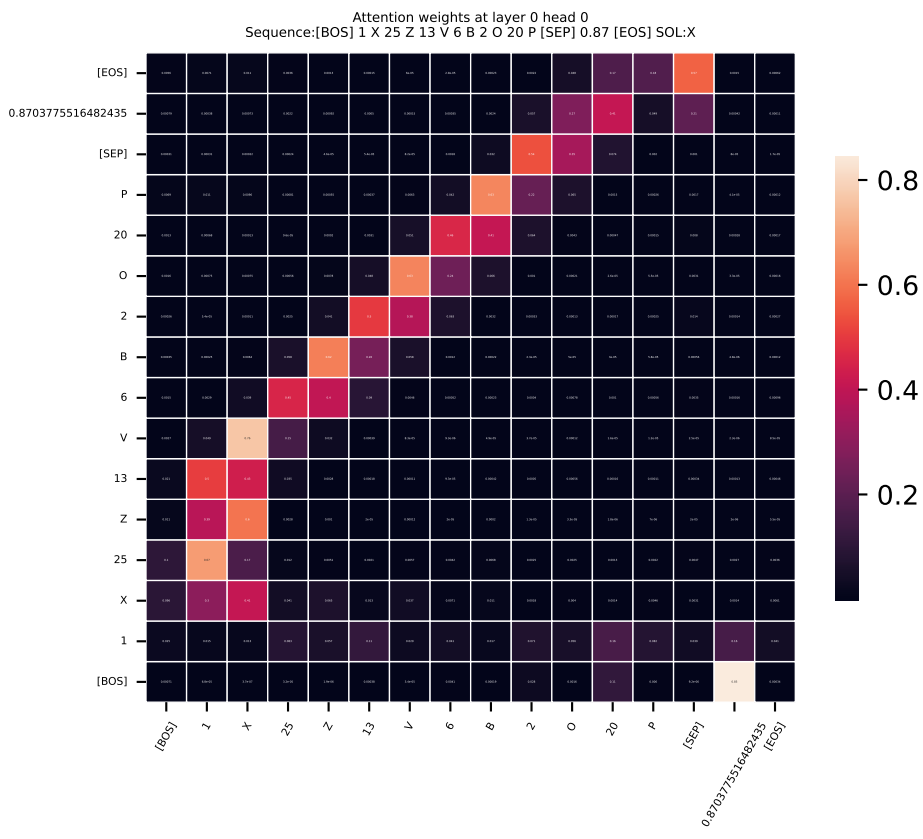

Attention weights at layer 1 head 0
Sequence:[BOS] 20 B 2 N 1 P 19 X 16 V 9 R [SEP] 15.9 [EOS] SOL:V
[EOS]
15.90079917266953
[SEP]
R
9
V
16
X
19
P
1
N
2
B
20
[BOS]
0.8
0.6
0.4
0.2

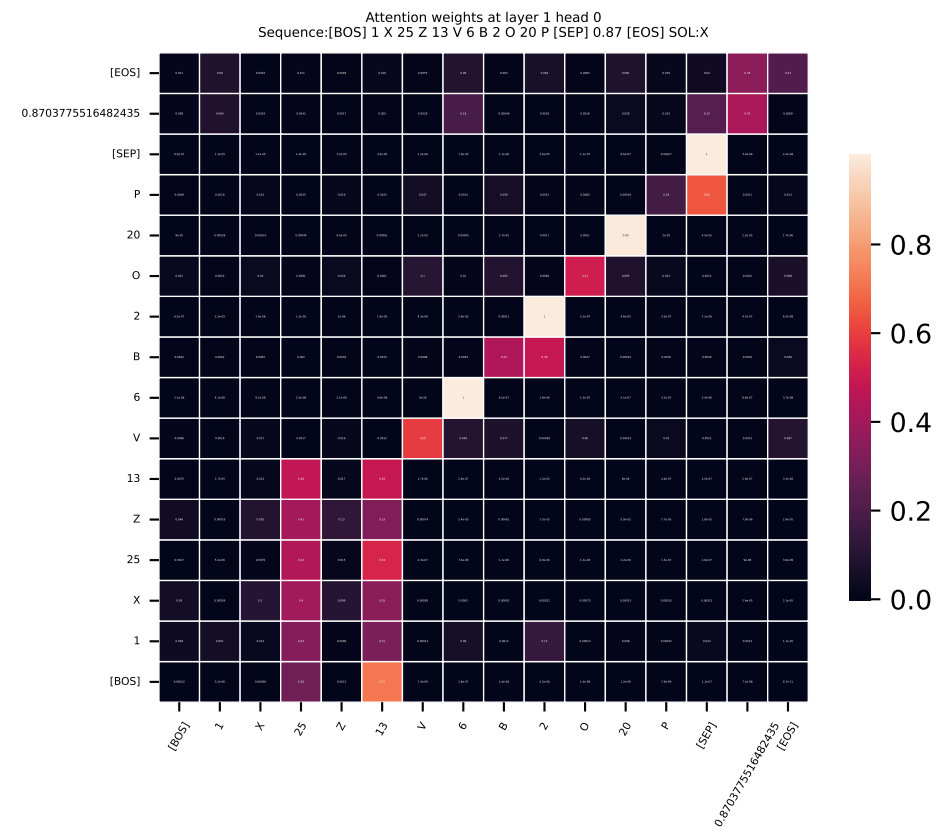

Attention weights at layer 0 head 0
Sequence:[BOS] 23 G 19 F 5 C 9 K 21 Y 10 I [SEP] 18.9 [EOS] SOL:F
[EOS]
18.90279081571666
[SEP]
I
10
Y
21
K
9
C
5
F
19
G
23
[BOS]
0.8
0.6
0.4
0.2

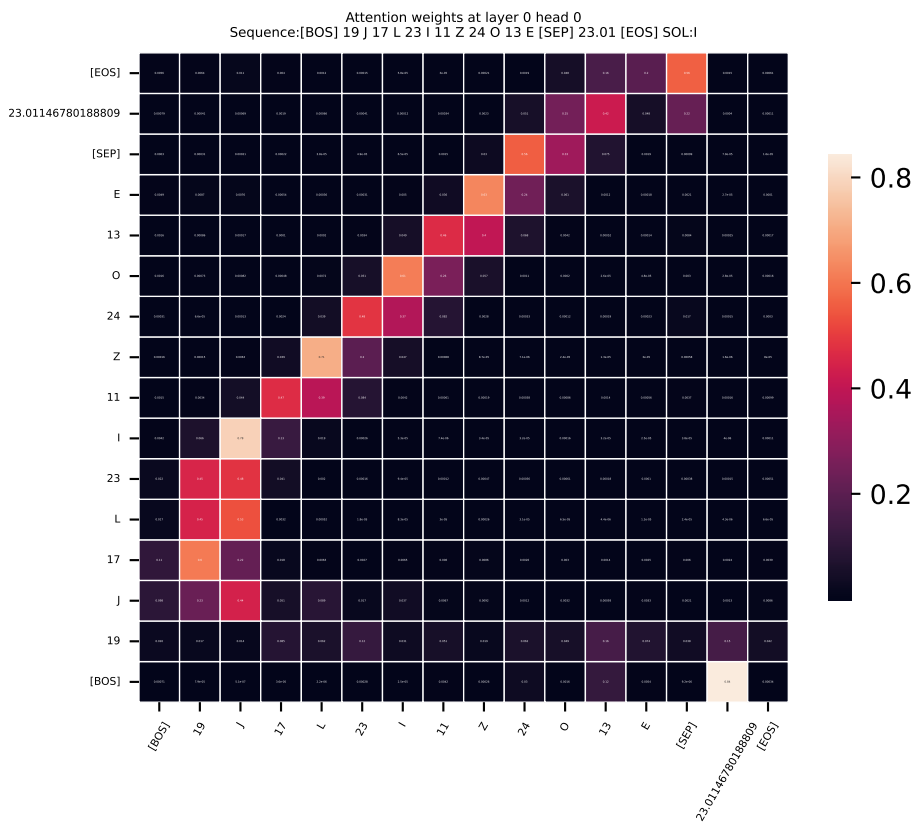

Attention weights at layer 1 head 0
Sequence:[BOS] 23 G 19 F 5 C 9 K 21 Y 10 I [SEP] 18.9 [EOS] SOL:F
[EOS]
18.90279081571666
[SEP]
I
10
Y
21
K
9
C
5
F
19
G
23
[BOS]
0.8
0.6
0.4
0.2

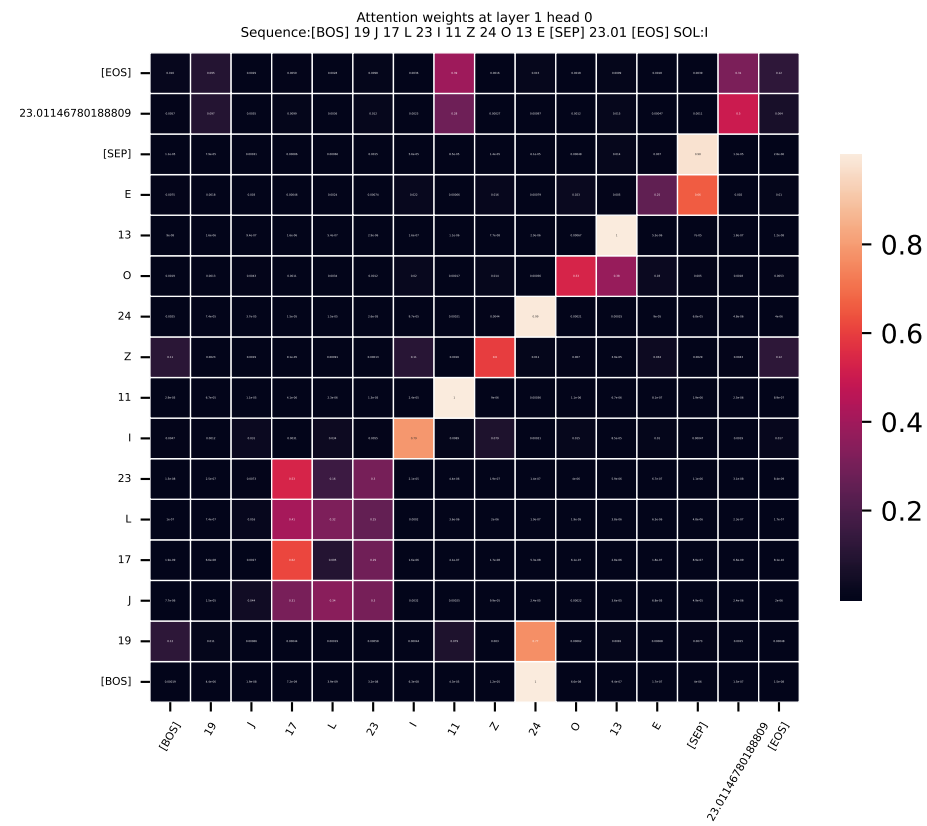

Attention weights at layer 0 head 0
Sequence:[BOS] 5 G 19 S 6 P 17 P 21 R 1 B [SEP] 18.92 [EOS] SOL:S
[EOS]
18.91632727560866
[SEP]
B
1
R
21
P
17
P
6
S
19
G
5
[BOS]
[BOS]
5
G
19
S
6
P
17
P
21
R
1
B
[SEP]
18.91632727560866
[EOS]
0.8
0.6
0.4
0.2

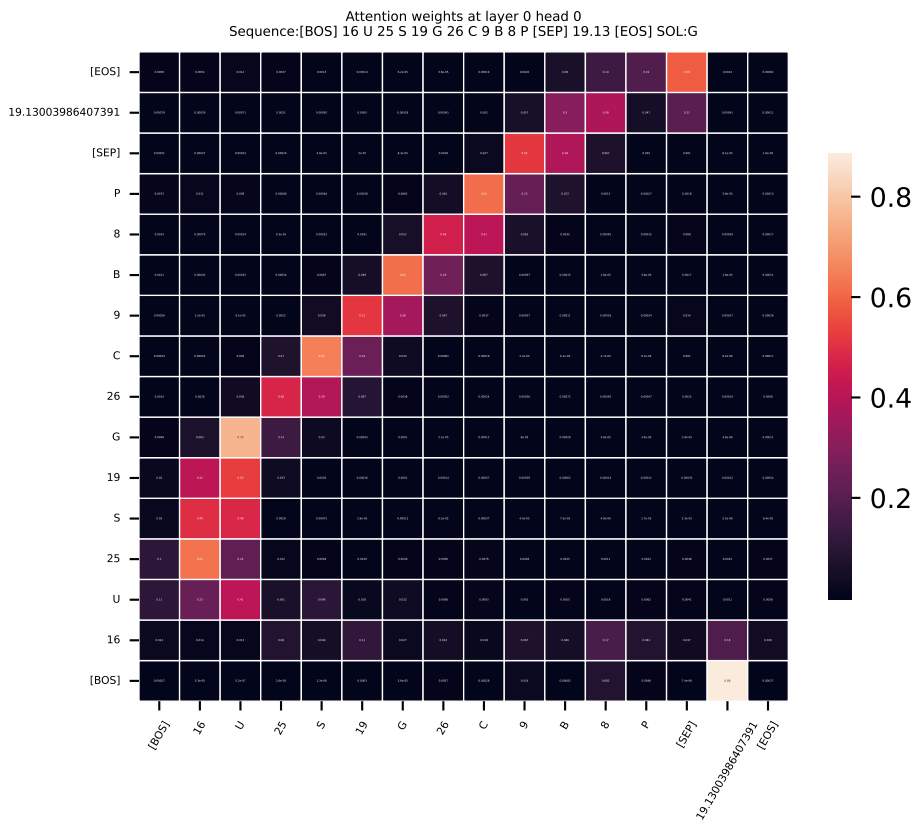

Attention weights at layer 1 head 0
Sequence:[BOS] 5 G 19 S 6 P 17 P 21 R 1 B [SEP] 18.92 [EOS] SOL:S
[EOS]
18.91632727560866
[SEP]
B
1
R
21
P
17
P
6
S
19
G
5
[BOS]
[BOS]
5
G
19
S
6
P
17
P
21
R
1
B
[SEP]
18.91632727560866
[EOS]
0.8
0.6
0.4
0.2

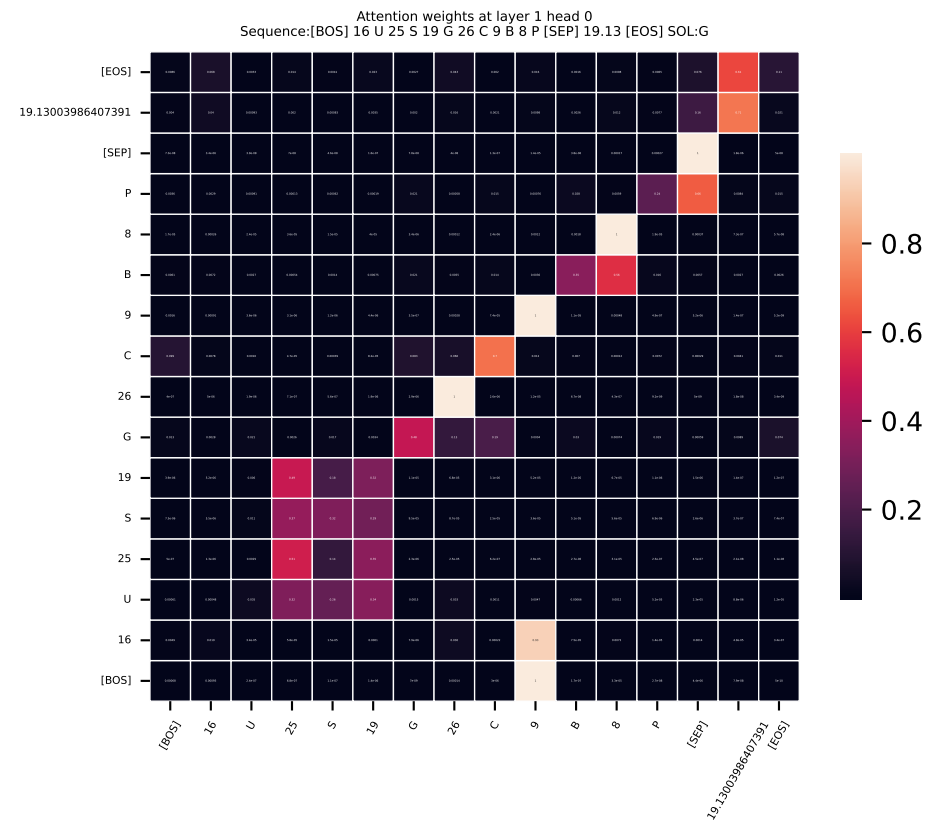

Attention weights at layer 0 head 0
Sequence:[BOS] 1 X 25 Z 13 V 6 B 2 O 20 P [SEP] 0.87 [EOS] SOL:X
[EOS]
0.8703775516482435
[SEP]
P
20
O
2
B
6
V
13
Z
25
X
1
[BOS]
[BOS] 1 X 25 Z 13 V 6 B 2 O 20 P [SEP] 0.8703775516482435 [EOS]
0.8
0.6
0.4
0.2

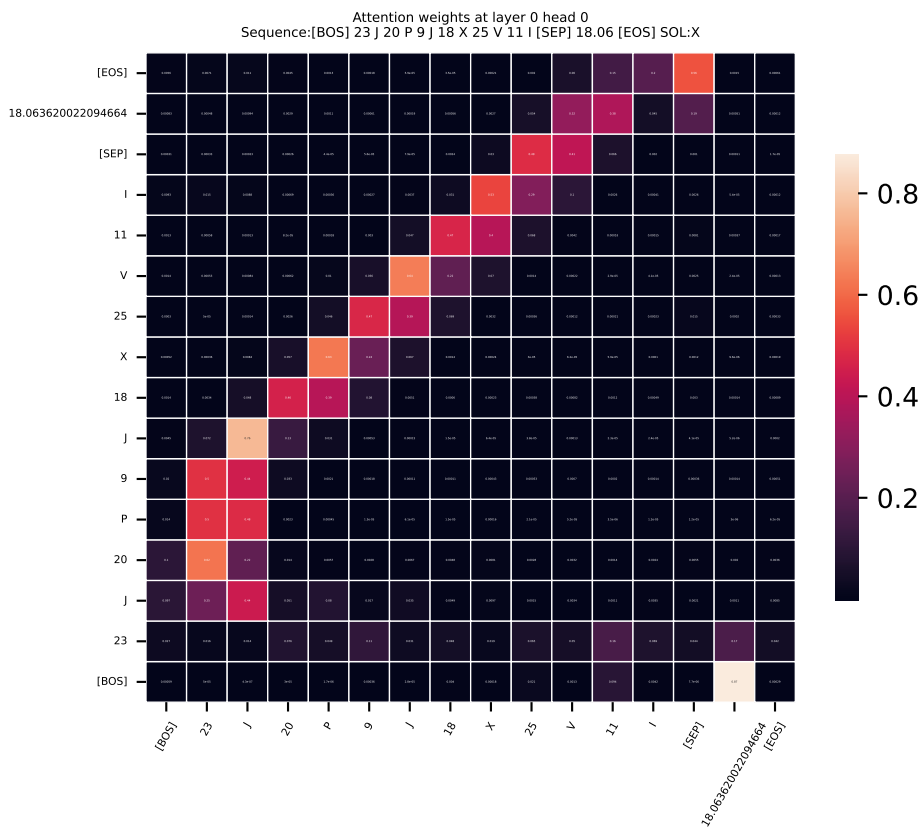

Attention weights at layer 1 head 0
Sequence:[BOS] 1 X 25 Z 13 V 6 B 2 O 20 P [SEP] 0.87 [EOS] SOL:X
[EOS]
0.8703775516482435
[SEP]
P
20
O
2
B
6
V
13
Z
25
X
1
[BOS]
[BOS] 1 X 25 Z 13 V 6 B 2 O 20 P [SEP] 0.8703775516482435 [EOS]
0.8
0.6
0.4
0.2
0.0

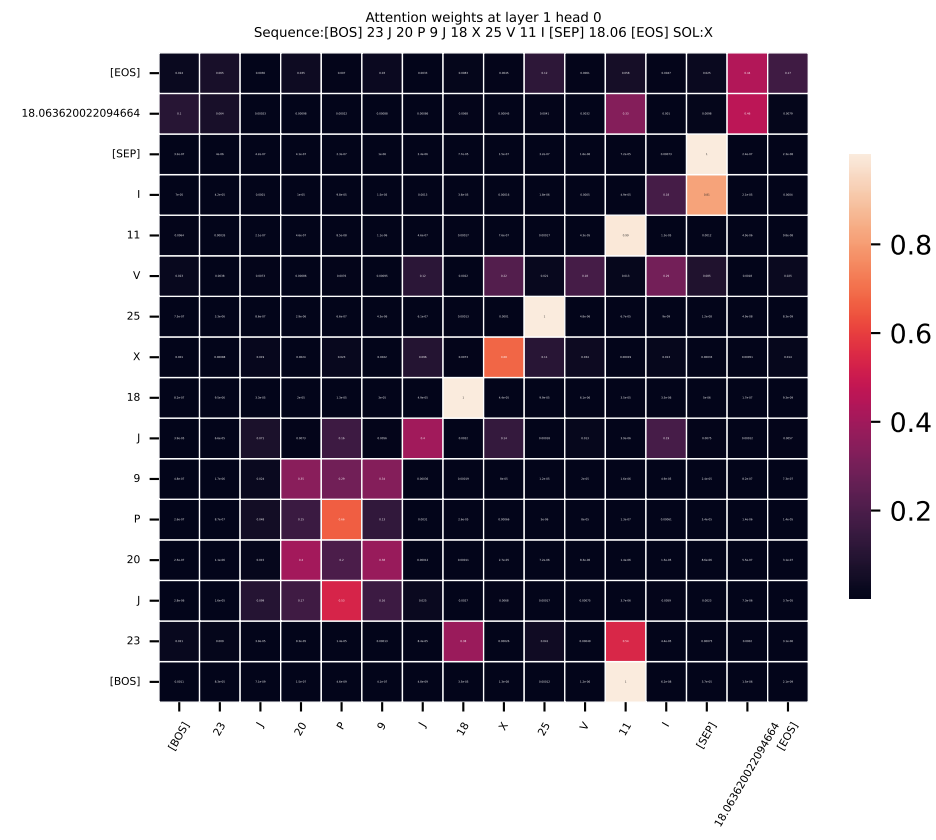

Attention weights at layer 0 head 0
Sequence:[BOS] 19 J 17 L 23 I 11 Z 24 O 13 E [SEP] 23.01 [EOS] SOL:I
[EOS]
23.01146780188809
[SEP]
E
13
O
24
Z
11
I
23
L
17
J
19
[BOS]
[BOS]
19
J
17
L
23
I
11
Z
24
O
13
E
[SEP]
23.01146780188809
[EOS]
0.8
0.6
0.4
0.2

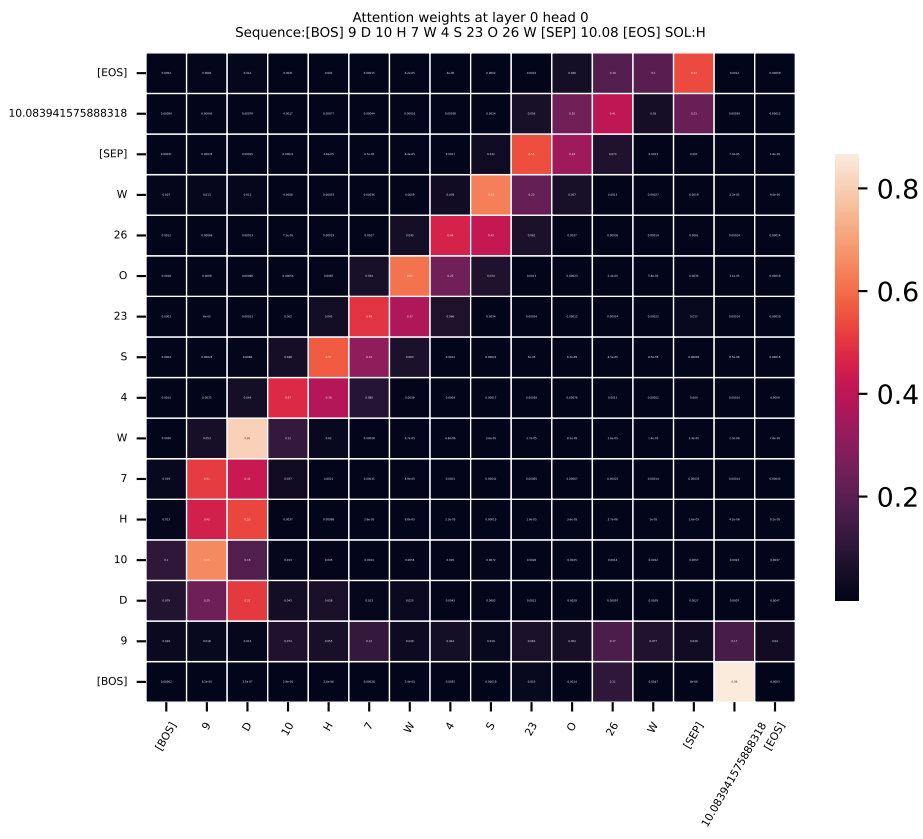

Attention weights at layer 1 head 0
Sequence:[BOS] 19 J 17 L 23 I 11 Z 24 O 13 E [SEP] 23.01 [EOS] SOL:I
[EOS]
23.01146780188809
[SEP]
E
13
O
24
Z
11
I
23
L
17
J
19
[BOS]
[BOS]
19
J
17
L
23
I
11
Z
24
O
13
E
[SEP]
23.01146780188809
[EOS]
0.8
0.6
0.4
0.2

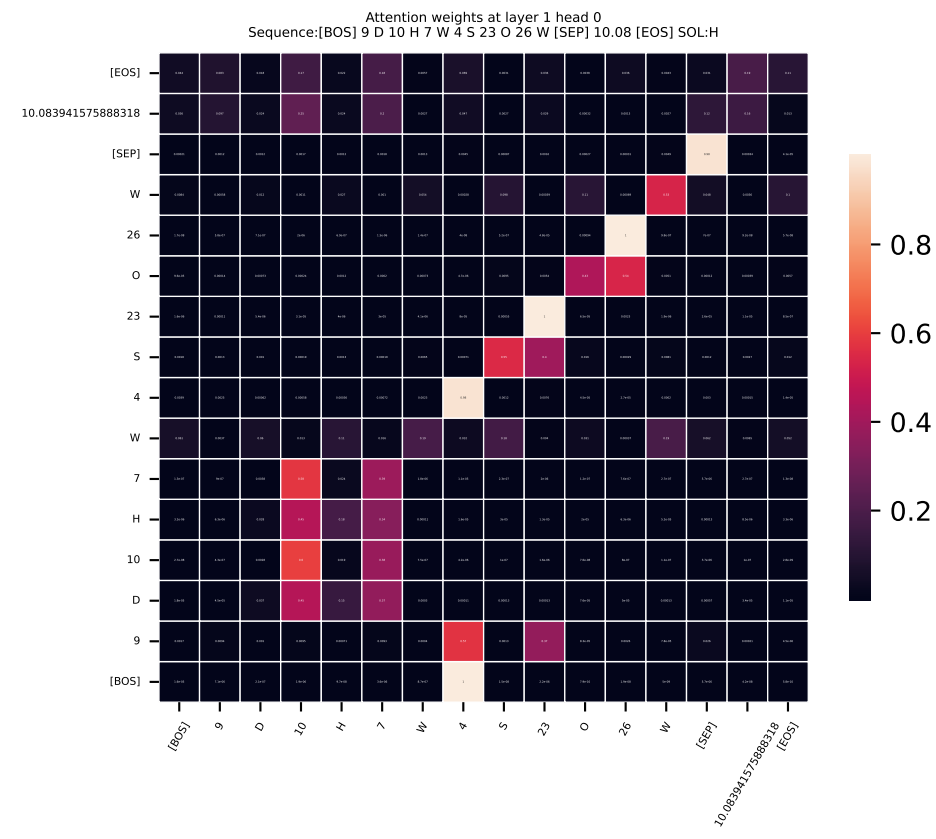

Attention weights at layer 0 head 0
Sequence:[BOS] 16 U 25 S 19 G 26 C 9 B 8 P [SEP] 19.13 [EOS] SOL:G
[EOS]
19.13003986407391
[SEP]
P
8
B
9
C
26
G
19
S
25
U
16
[BOS]
0.8
0.6
0.4
0.2

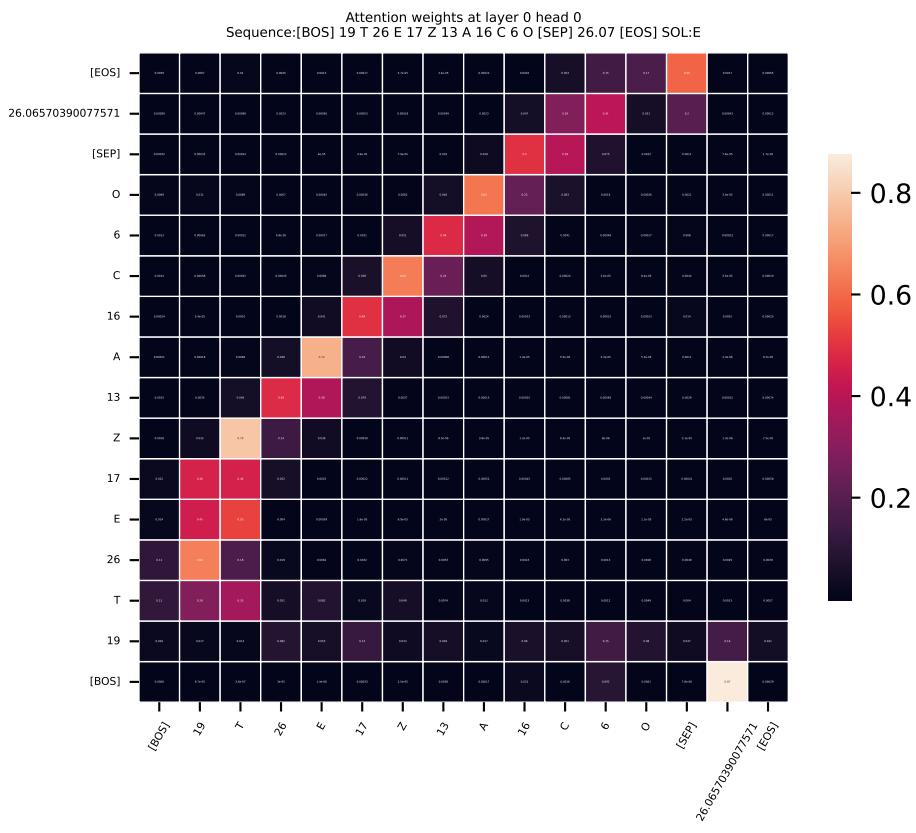

Attention weights at layer 1 head 0
Sequence:[BOS] 16 U 25 S 19 G 26 C 9 B 8 P [SEP] 19.13 [EOS] SOL:G
[EOS]
19.13003986407391
[SEP]
P
8
B
9
C
26
G
19
S
25
U
16
[BOS]
0.8
0.6
0.4
0.2

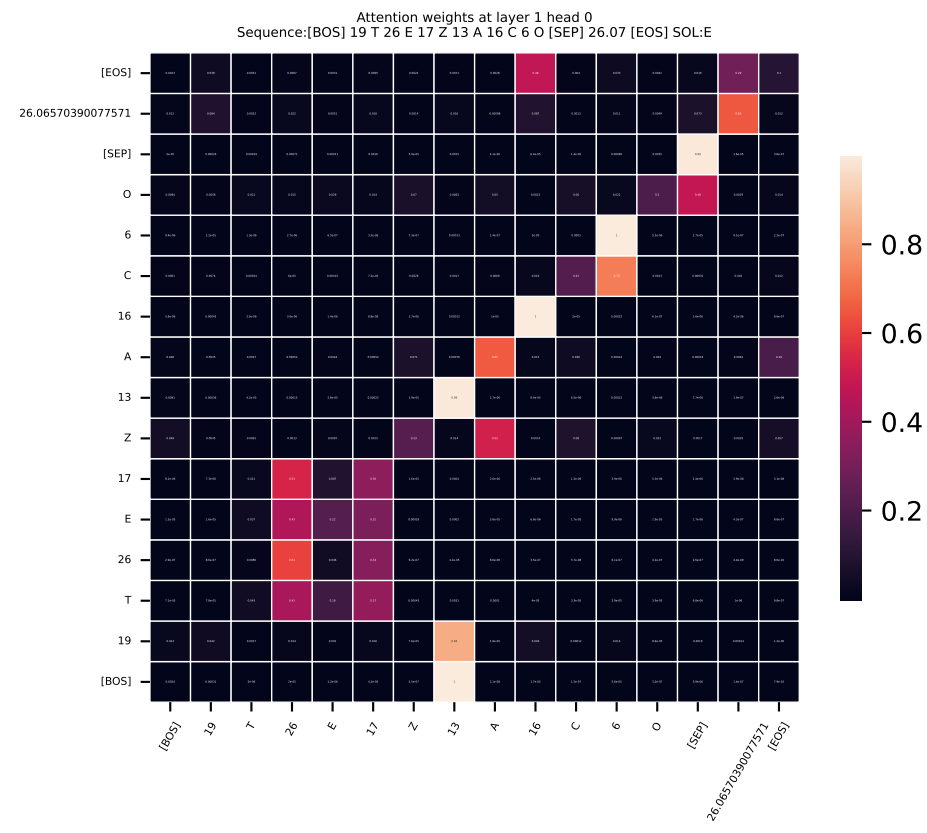

Attention weights at layer 0 head 0
Sequence:[BOS] 23 J 20 P 9 J 18 X 25 V 11 I [SEP] 18.06 [EOS] SOL:X
[EOS]
18.063620022094664
[SEP]
I
11
V
25
X
18
J
9
P
20
J
23
[BOS]
[BOS]
23
J
20
P
9
J
18
X
25
V
11
I
[SEP]
18.063620022094664
[EOS]
0.8
0.6
0.4
0.2

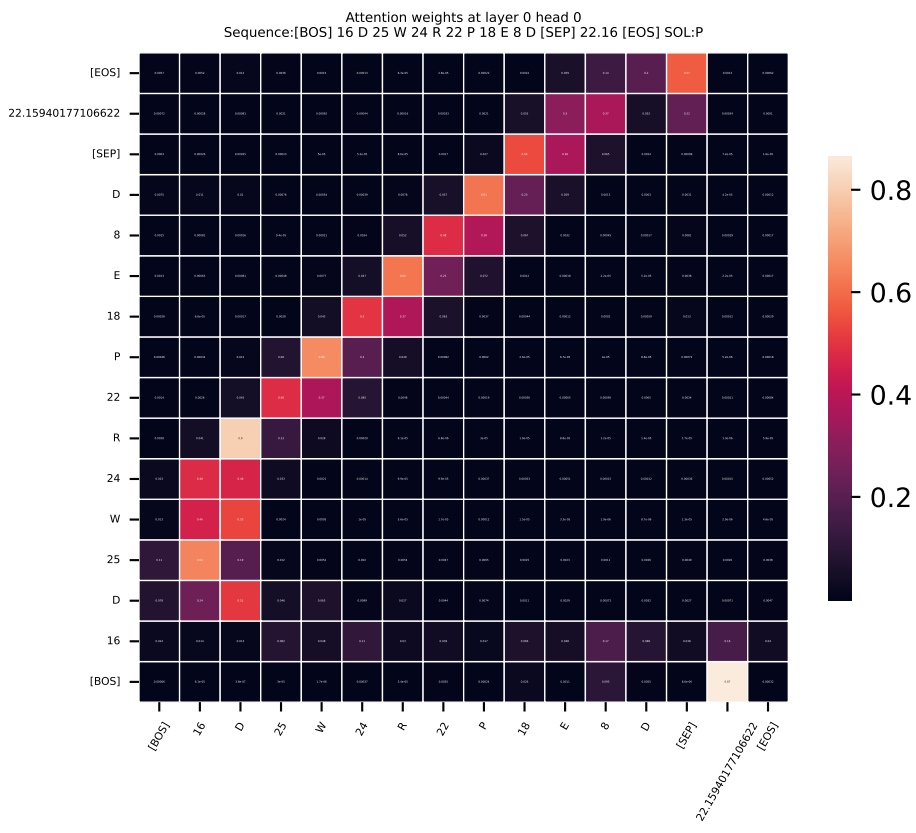

Attention weights at layer 1 head 0
Sequence:[BOS] 23 J 20 P 9 J 18 X 25 V 11 I [SEP] 18.06 [EOS] SOL:X
[EOS]
18.063620022094664
[SEP]
I
11
V
25
X
18
J
9
P
20
J
23
[BOS]
[BOS]
23
J
20
P
9
J
18
X
25
V
11
I
[SEP]
18.063620022094664
[EOS]
0.8
0.6
0.4
0.2

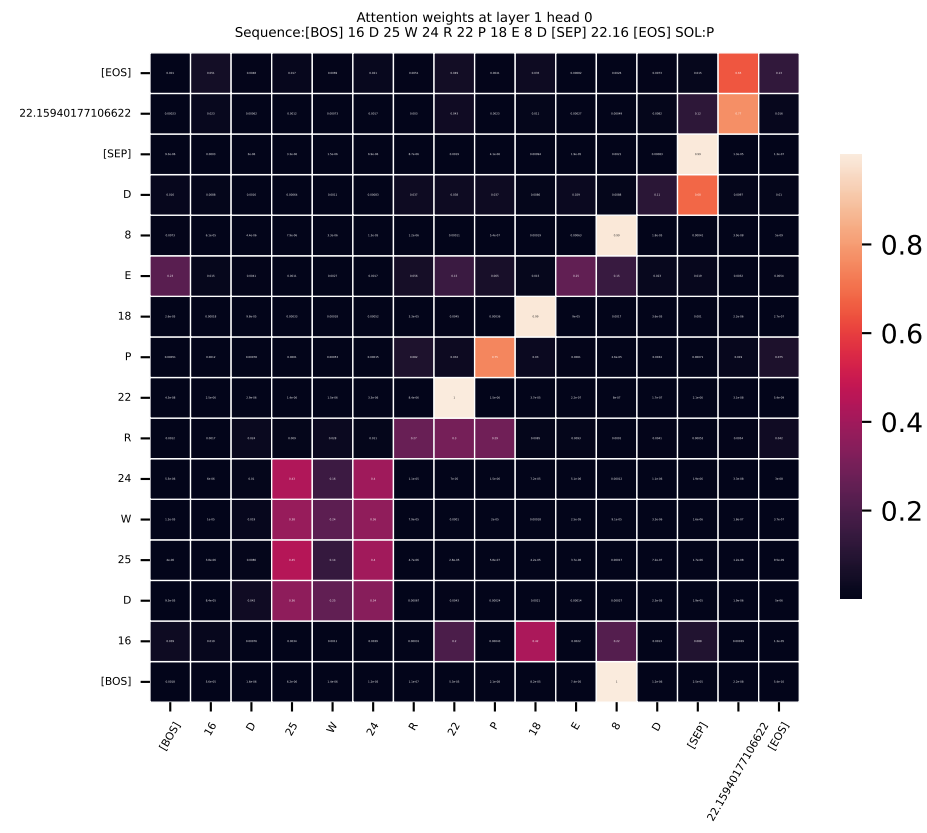

Attention weights at layer 0 head 0
Sequence:[BOS] 9 D 10 H 7 W 4 S 23 O 26 W [SEP] 10.08 [EOS] SOL:H
[EOS]
10.083941575888318
[SEP]
W
26
O
23
S
4
W
7
H
10
D
9
[BOS]
[BOS]
9
D
10
H
7
W
4
S
23
O
26
W
[SEP]
10.083941575888318
[EOS]
0.8
0.6
0.4
0.2

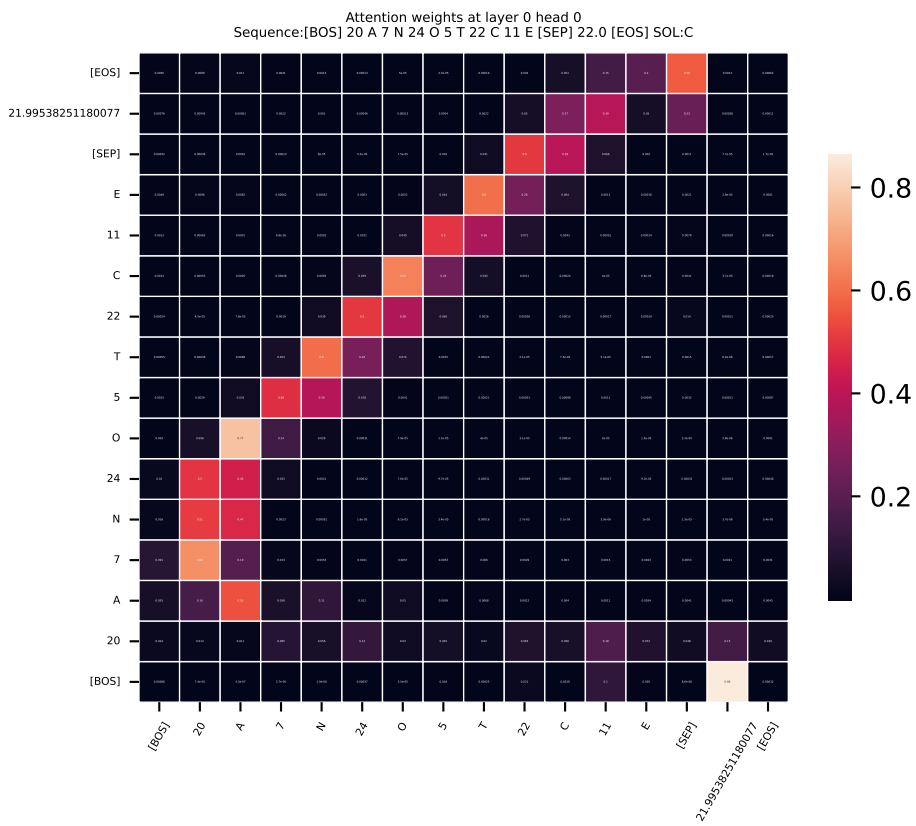

Attention weights at layer 1 head 0
Sequence:[BOS] 9 D 10 H 7 W 4 S 23 O 26 W [SEP] 10.08 [EOS] SOL:H
[EOS]
10.083941575888318
[SEP]
W
26
O
23
S
4
W
7
H
10
D
9
[BOS]
[BOS]
9
D
10
H
7
W
4
S
23
O
26
W
[SEP]
10.083941575888318
[EOS]
0.8
0.6
0.4
0.2

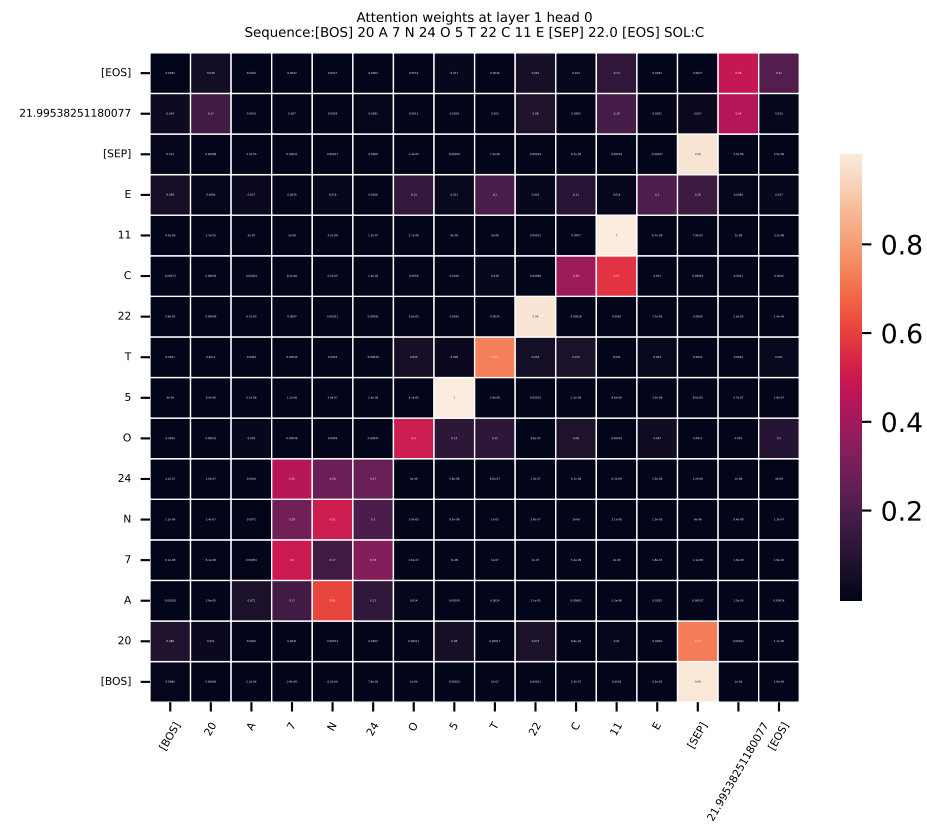

Attention weights at layer 0 head 0
Sequence:[BOS] 19 T 26 E 17 Z 13 A 16 C 6 O [SEP] 26.07 [EOS] SOL:E
[EOS]
26.06570390077571
[SEP]
O
6
C
16
A
13
Z
17
E
26
T
19
[BOS]
[BOS]
19
T
26
E
17
Z
13
A
16
C
6
O
[SEP]
26.06570390077571
[EOS]
0.8
0.6
0.4
0.2

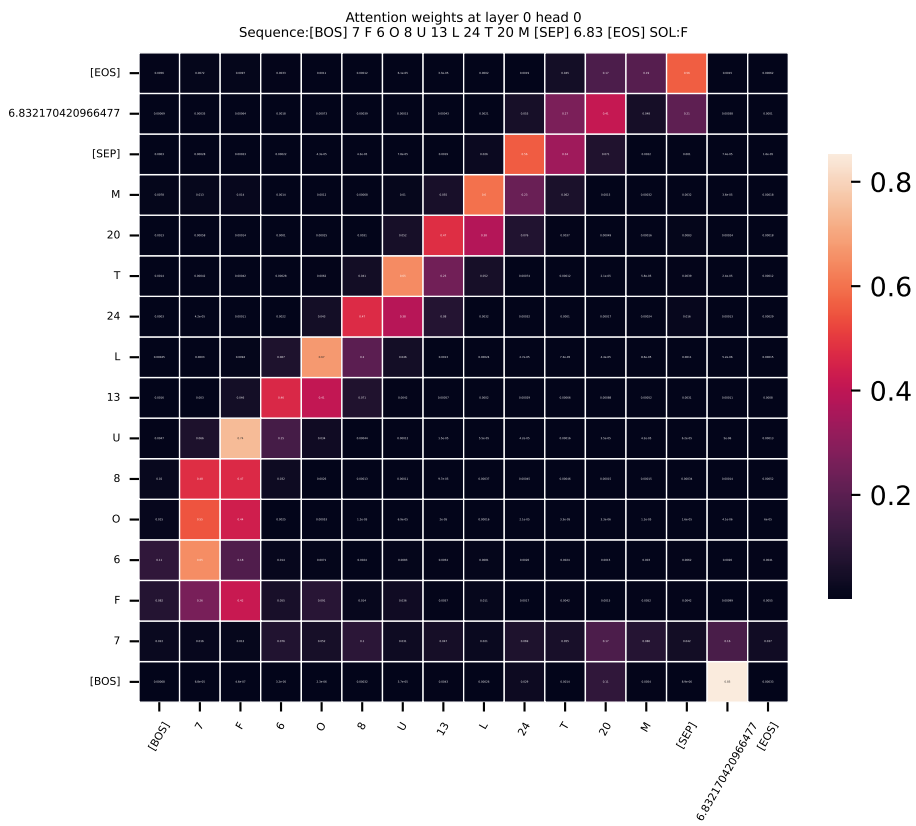

Attention weights at layer 1 head 0
Sequence:[BOS] 19 T 26 E 17 Z 13 A 16 C 6 O [SEP] 26.07 [EOS] SOL:E
[EOS]
26.06570390077571
[SEP]
O
6
C
16
A
13
Z
17
E
26
T
19
[BOS]
[BOS]
19
T
26
E
17
Z
13
A
16
C
6
O
[SEP]
26.06570390077571
[EOS]
0.8
0.6
0.4
0.2

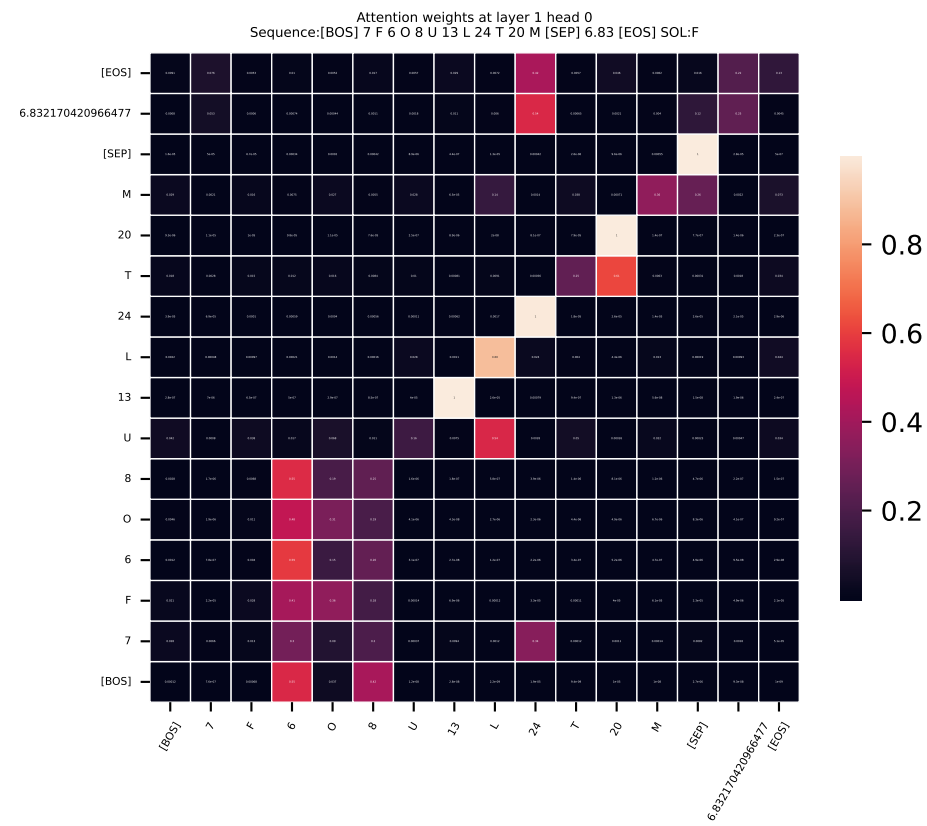

Attention weights at layer 0 head 0
Sequence:[BOS] 16 D 25 W 24 R 22 P 18 E 8 D [SEP] 22.16 [EOS] SOL:P
[EOS]
22.15940177106622
[SEP]
D
8
E
18
P
22
R
24
W
25
D
16
[BOS]
[BOS]
16
D
25
W
24
R
22
P
18
E
8
D
[SEP]
22.15940177106622
[EOS]
0.8
0.6
0.4
0.2

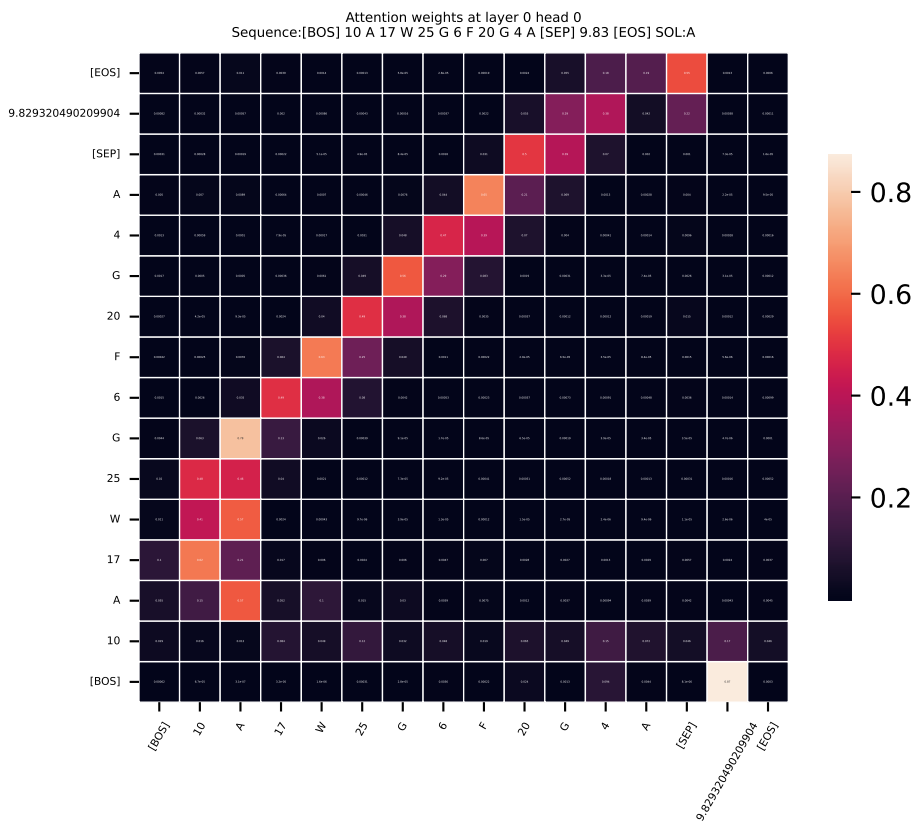

Attention weights at layer 1 head 0
Sequence:[BOS] 16 D 25 W 24 R 22 P 18 E 8 D [SEP] 22.16 [EOS] SOL:P
[EOS]
22.15940177106622
[SEP]
D
8
E
18
P
22
R
24
W
25
D
16
[BOS]
[BOS]
16
D
25
W
24
R
22
P
18
E
8
D
[SEP]
22.15940177106622
[EOS]
0.8
0.6
0.4
0.2

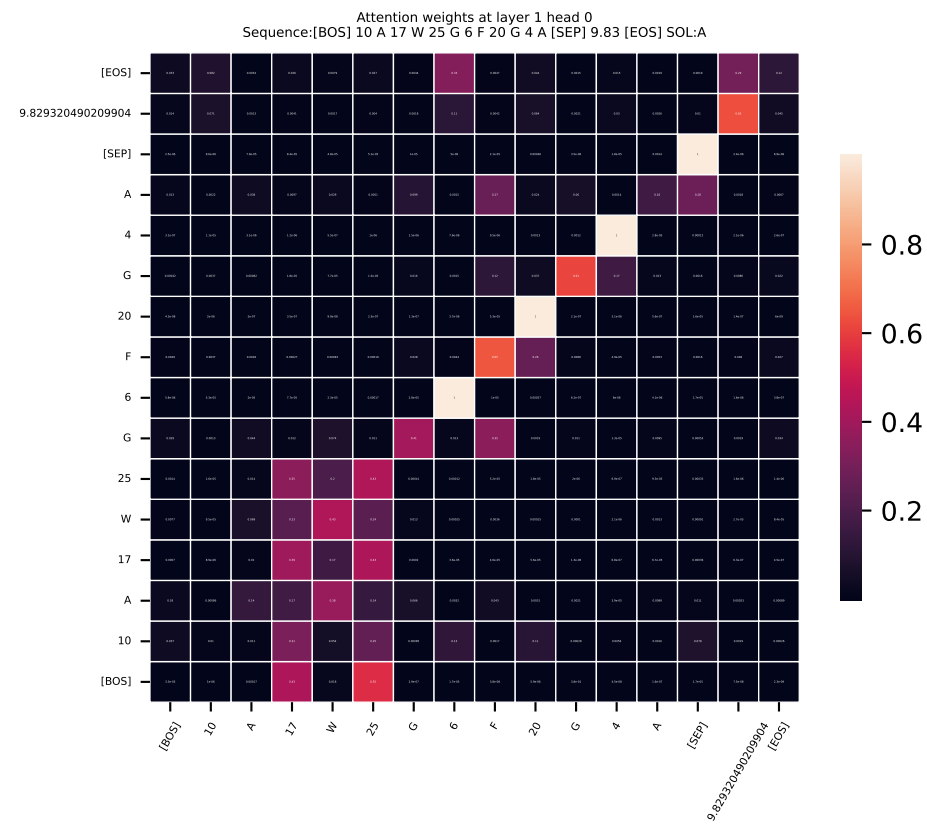

Attention weights at layer 0 head 0
Sequence:[BOS] 20 A 7 N 24 O 5 T 22 C 11 E [SEP] 22.0 [EOS] SOL:C
[EOS]
21.99538251180077
[SEP]
E
11
C
22
T
5
O
24
N
7
A
20
[BOS]
0.8
0.6
0.4
0.2

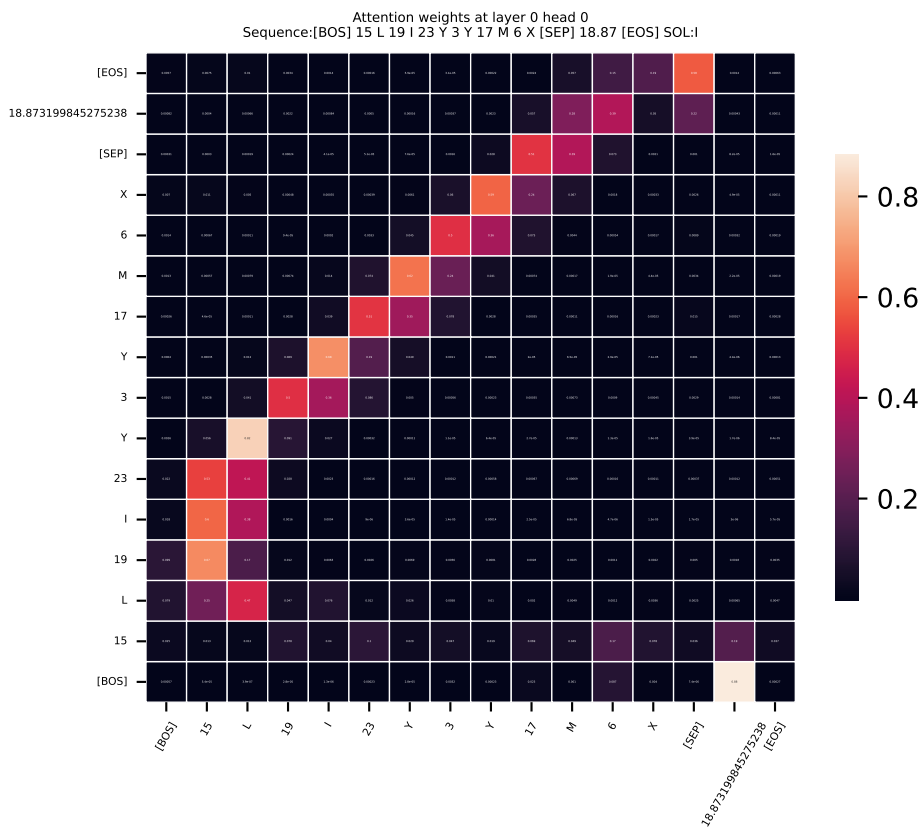

Attention weights at layer 1 head 0
Sequence:[BOS] 20 A 7 N 24 O 5 T 22 C 11 E [SEP] 22.0 [EOS] SOL:C
[EOS]
21.99538251180077
[SEP]
E
11
C
22
T
5
O
24
N
7
A
20
[BOS]
0.8
0.6
0.4
0.2

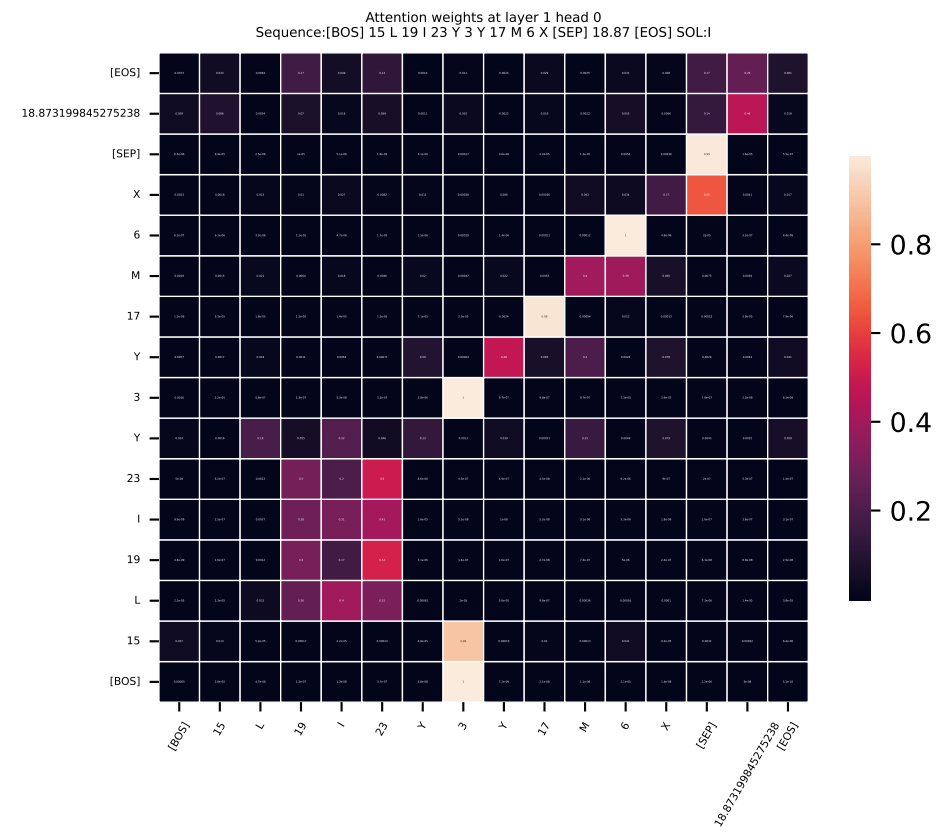

Attention weights at layer 0 head 0
Sequence:[BOS] 7 F 6 O 8 U 13 L 24 T 20 M [SEP] 6.83 [EOS] SOL:F
[EOS]
6.832170420966477
[SEP]
M
20
T
24
L
13
U
8
O
6
F
7
[BOS]
[BOS]
7
F
6
O
8
U
13
L
24
T
20
M
[SEP]
6.832170420966477
[EOS]
0.8
0.6
0.4
0.2

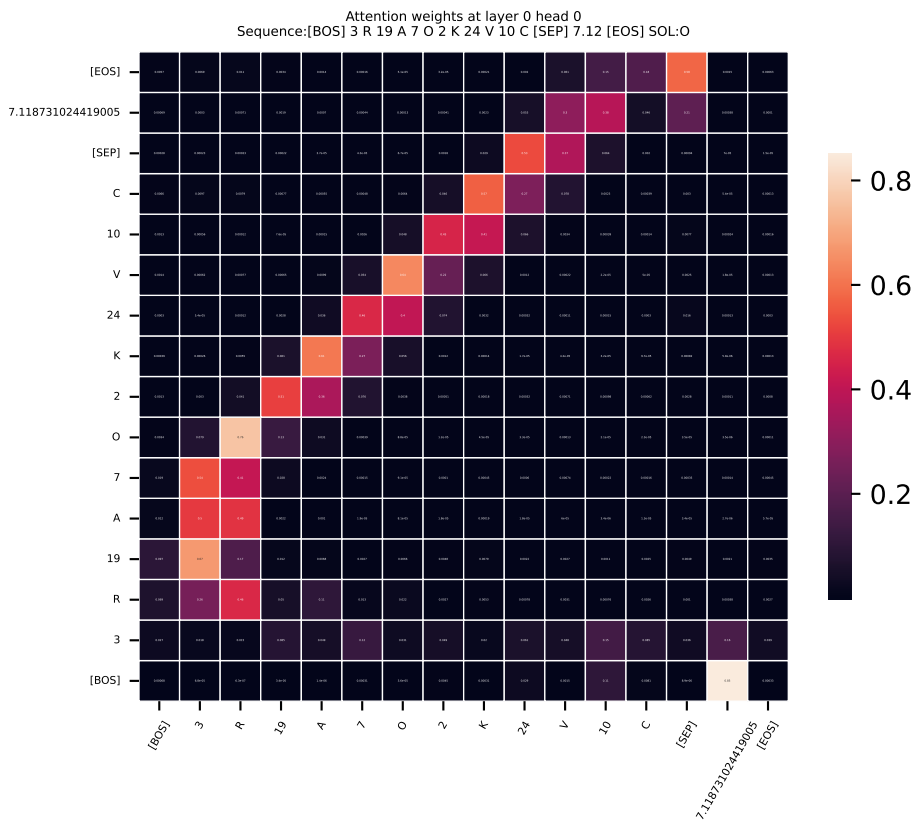

Attention weights at layer 1 head 0
Sequence:[BOS] 7 F 6 O 8 U 13 L 24 T 20 M [SEP] 6.83 [EOS] SOL:F
[EOS]
6.832170420966477
[SEP]
M
20
T
24
L
13
U
8
O
6
F
7
[BOS]
[BOS]
7
F
6
O
8
U
13
L
24
T
20
M
[SEP]
6.832170420966477
[EOS]
0.8
0.6
0.4
0.2

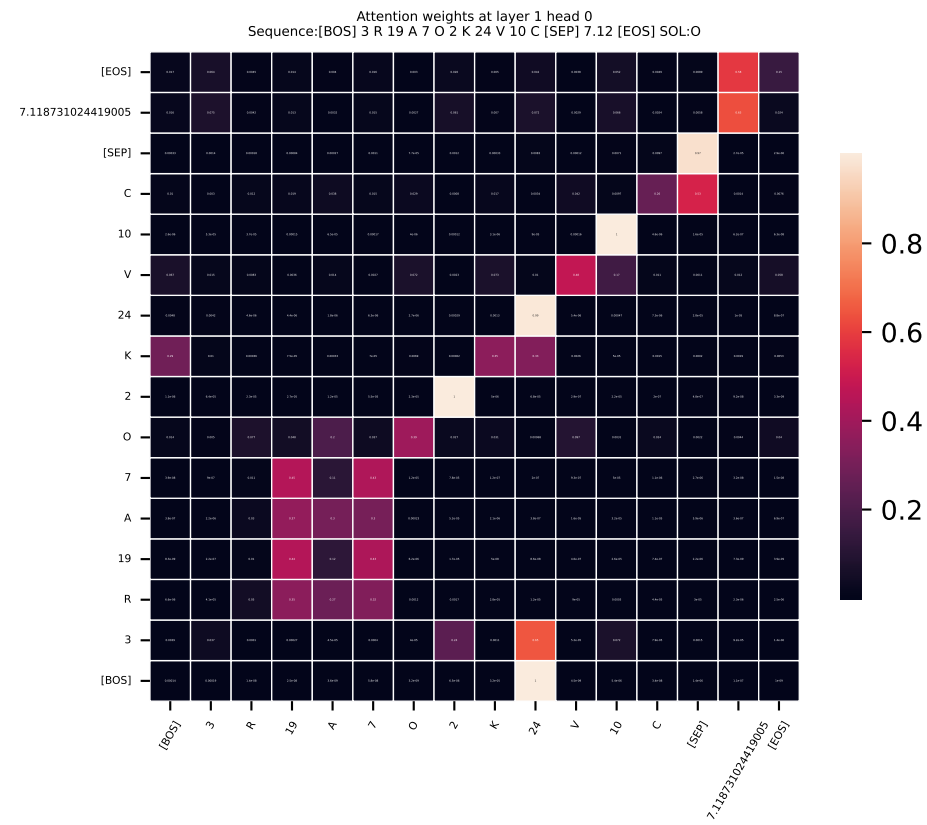

Attention weights at layer 0 head 0
Sequence:[BOS] 10 A 17 W 25 G 6 F 20 G 4 A [SEP] 9.83 [EOS] SOL:A
[EOS]
9.829320490209904
[SEP]
A
4
G
20
F
6
G
25
W
17
A
10
[BOS]
[BOS]
10
A
17
W
25
G
6
F
20
G
4
A
[SEP]
9.829320490209904
[EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 1 head 0
Sequence:[BOS] 10 A 17 W 25 G 6 F 20 G 4 A [SEP] 9.83 [EOS] SOL:A
[EOS]
9.829320490209904
[SEP]
A
4
G
20
F
6
G
25
W
17
A
10
[BOS]
[BOS]
10
A
17
W
25
G
6
F
20
G
4
A
[SEP]
9.829320490209904
[EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 0 head 0
Sequence:[BOS] 15 L 19 I 23 Y 3 Y 17 M 6 X [SEP] 18.87 [EOS] SOL:I
[EOS]
18.873199845275238
[SEP]
X
6
M
17
Y
3
Y
23
I
19
L
15
[BOS]
[BOS]
15
L
19
I
23
Y
3
Y
17
M
6
X
[SEP]
18.873199845275238
[EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 1 head 0
Sequence:[BOS] 15 L 19 I 23 Y 3 Y 17 M 6 X [SEP] 18.87 [EOS] SOL:I
[EOS]
18.873199845275238
[SEP]
X
6
M
17
Y
3
Y
23
I
19
L
15
[BOS]
[BOS]
15
L
19
I
23
Y
3
Y
17
M
6
X
[SEP]
18.873199845275238
[EOS]
0.8
0.6
0.4
0.2


Attention weights at layer 0 head 0
Sequence:[BOS] 3 R 19 A 7 O 2 K 24 V 10 C [SEP] 7.12 [EOS] SOL:O
[EOS]
7.118731024419005
[SEP]
C
10
V
24
K
2
O
7
A
19
R
3
[BOS]
0.8
0.6
0.4
0.2


Attention weights at layer 1 head 0
Sequence:[BOS] 3 R 19 A 7 O 2 K 24 V 10 C [SEP] 7.12 [EOS] SOL:O
[EOS]
7.118731024419005
[SEP]
C
10
V
24
K
2
O
7
A
19
R
3
[BOS]
0.8
0.6
0.4
0.2