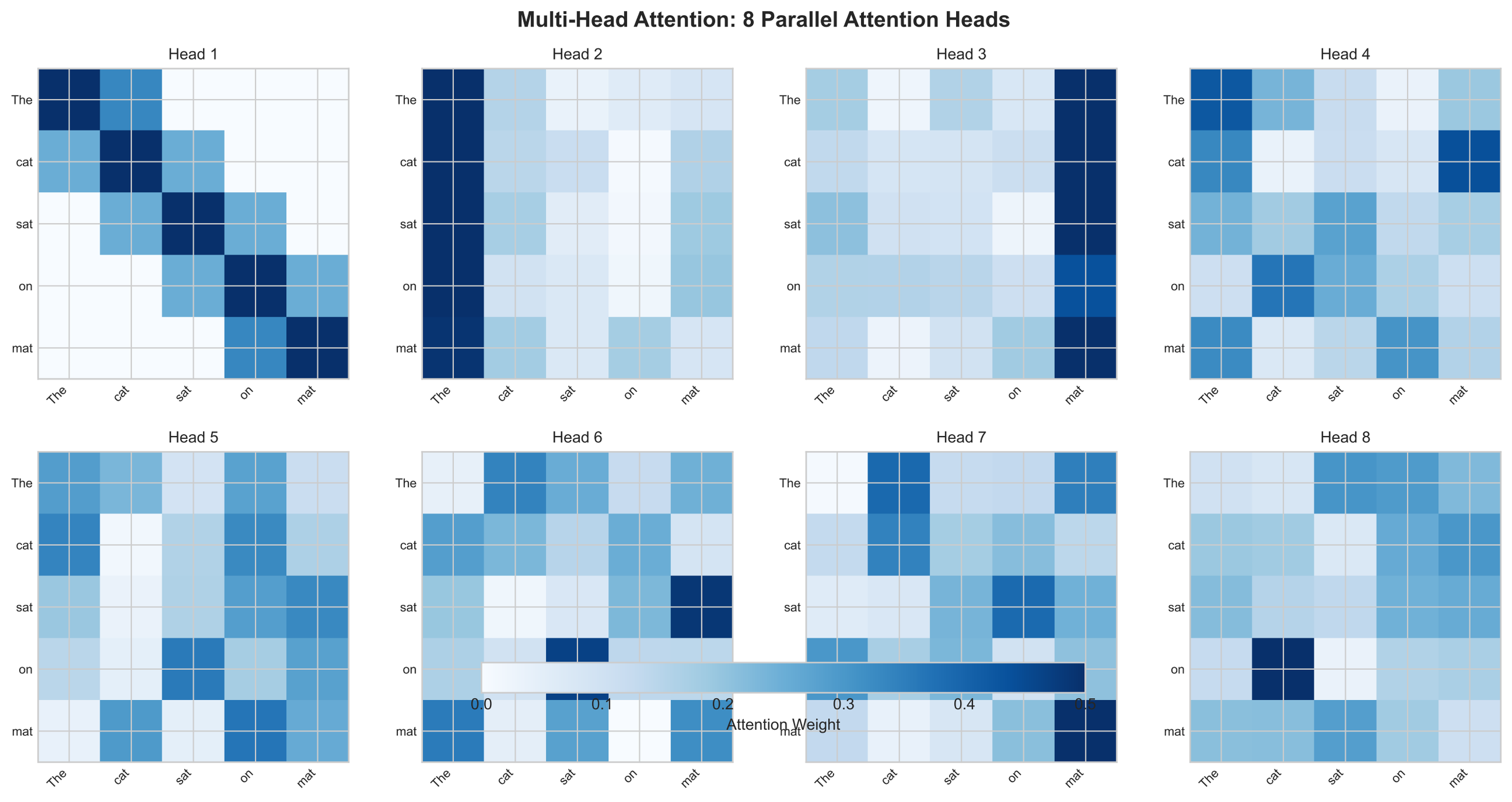
Multi-Head Attention: 8 Parallel Attention Heads
Head 1
Head 2
Head 3
Head 4
Head 5
Head 6
Head 7
Head 8
The
cat
sat
on
mat
0.0
0.1
0.2
0.3
0.4
0.5
Attention Weight