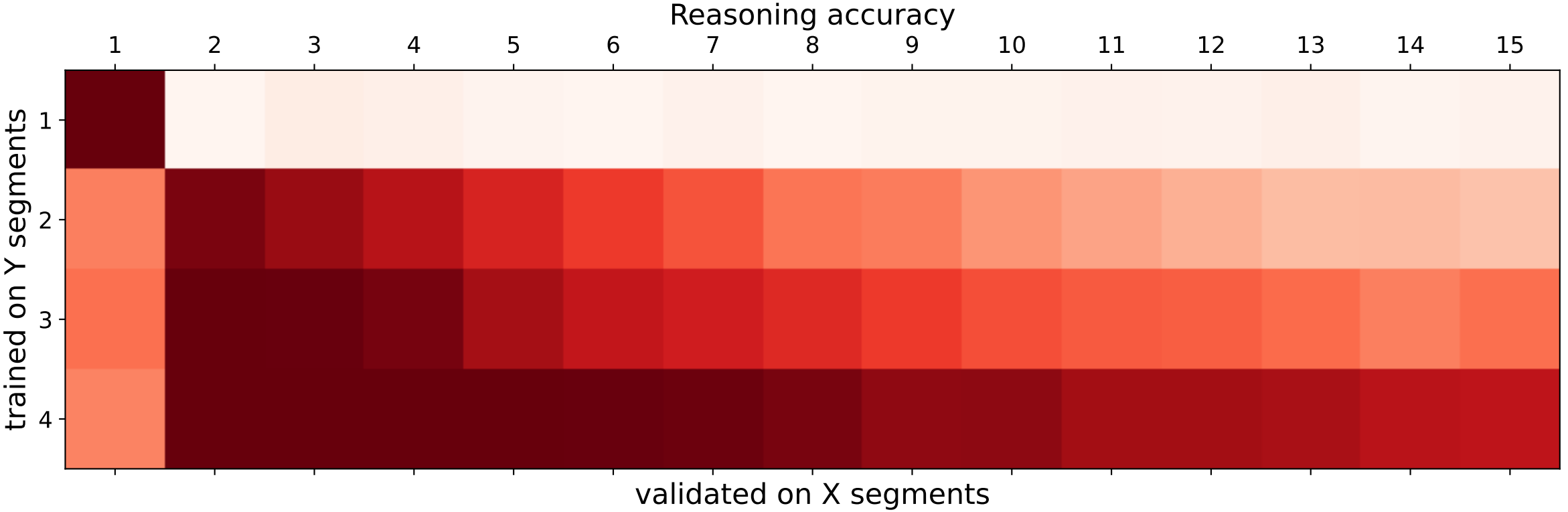
Reasoning accuracy
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
trained on Y segments
1
2
3
4
validated on X segments