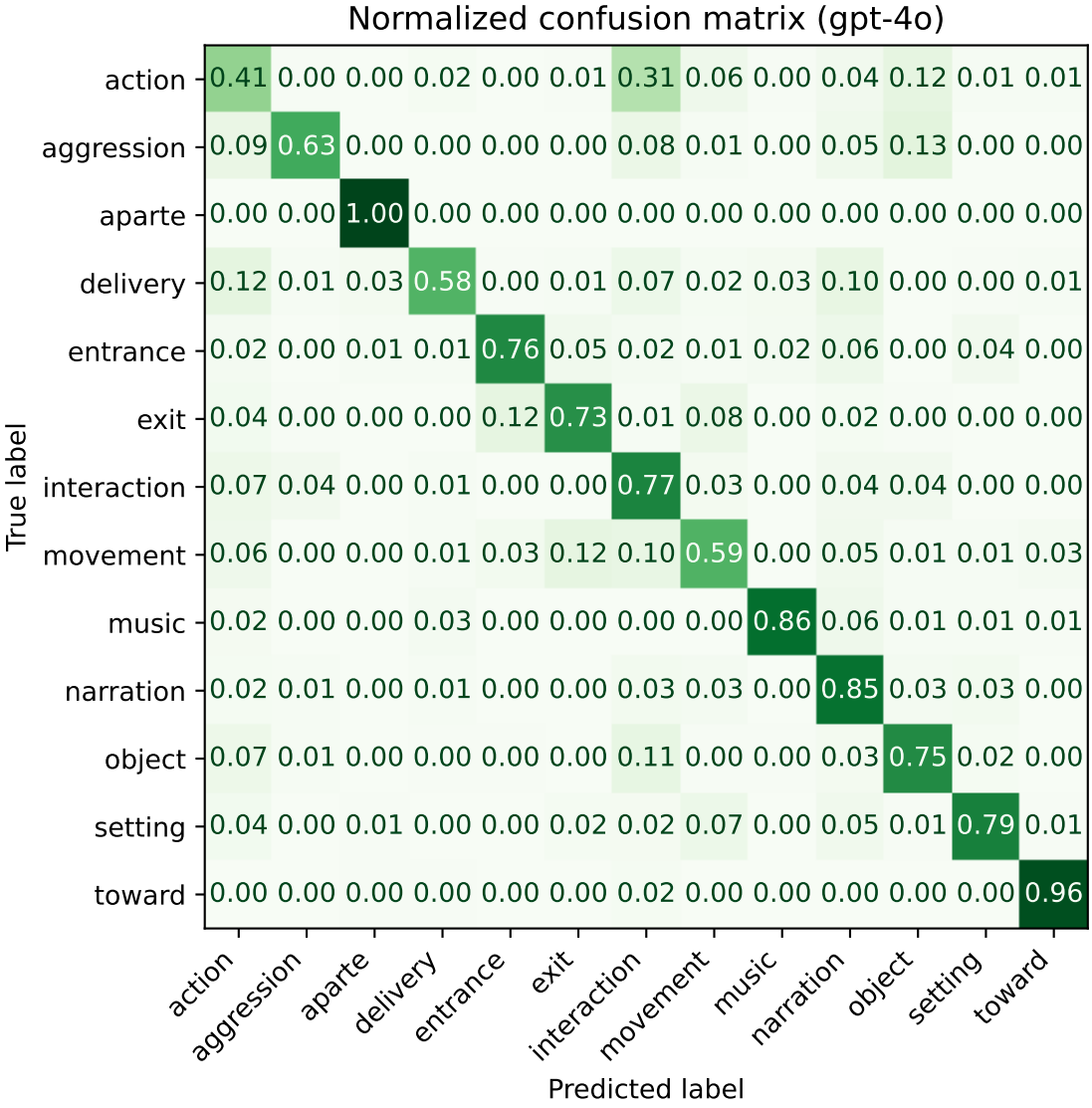
Normalized confusion matrix (gpt-4o)
True label
Predicted label
action 0.41 0.00 0.00 0.02 0.00 0.01 0.31 0.06 0.00 0.04 0.12 0.01 0.01
aggression 0.09 0.63 0.00 0.00 0.00 0.00 0.08 0.01 0.00 0.05 0.13 0.00 0.00
aparte 0.00 0.00 1.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
delivery 0.12 0.01 0.03 0.58 0.00 0.01 0.07 0.02 0.03 0.10 0.00 0.00 0.01
entrance 0.02 0.00 0.01 0.01 0.76 0.05 0.02 0.01 0.02 0.06 0.00 0.04 0.00
exit 0.04 0.00 0.00 0.00 0.12 0.73 0.01 0.08 0.00 0.02 0.00 0.00 0.00
interaction 0.07 0.04 0.00 0.01 0.00 0.00 0.77 0.03 0.00 0.04 0.04 0.00 0.00
movement 0.06 0.00 0.00 0.01 0.03 0.12 0.10 0.59 0.00 0.05 0.01 0.01 0.03
music 0.02 0.00 0.00 0.03 0.00 0.00 0.00 0.00 0.86 0.06 0.01 0.01 0.01
narration 0.02 0.01 0.00 0.01 0.00 0.00 0.03 0.03 0.00 0.85 0.03 0.03 0.00
object 0.07 0.01 0.00 0.00 0.00 0.00 0.11 0.00 0.00 0.03 0.75 0.02 0.00
setting 0.04 0.00 0.01 0.00 0.00 0.02 0.02 0.07 0.00 0.05 0.01 0.79 0.01
toward 0.00 0.00 0.00 0.00 0.00 0.00 0.02 0.00 0.00 0.00 0.00 0.00 0.96
action
aggression
aparte
delivery
entrance
exit
interaction
movement
music
narration
object
setting
toward