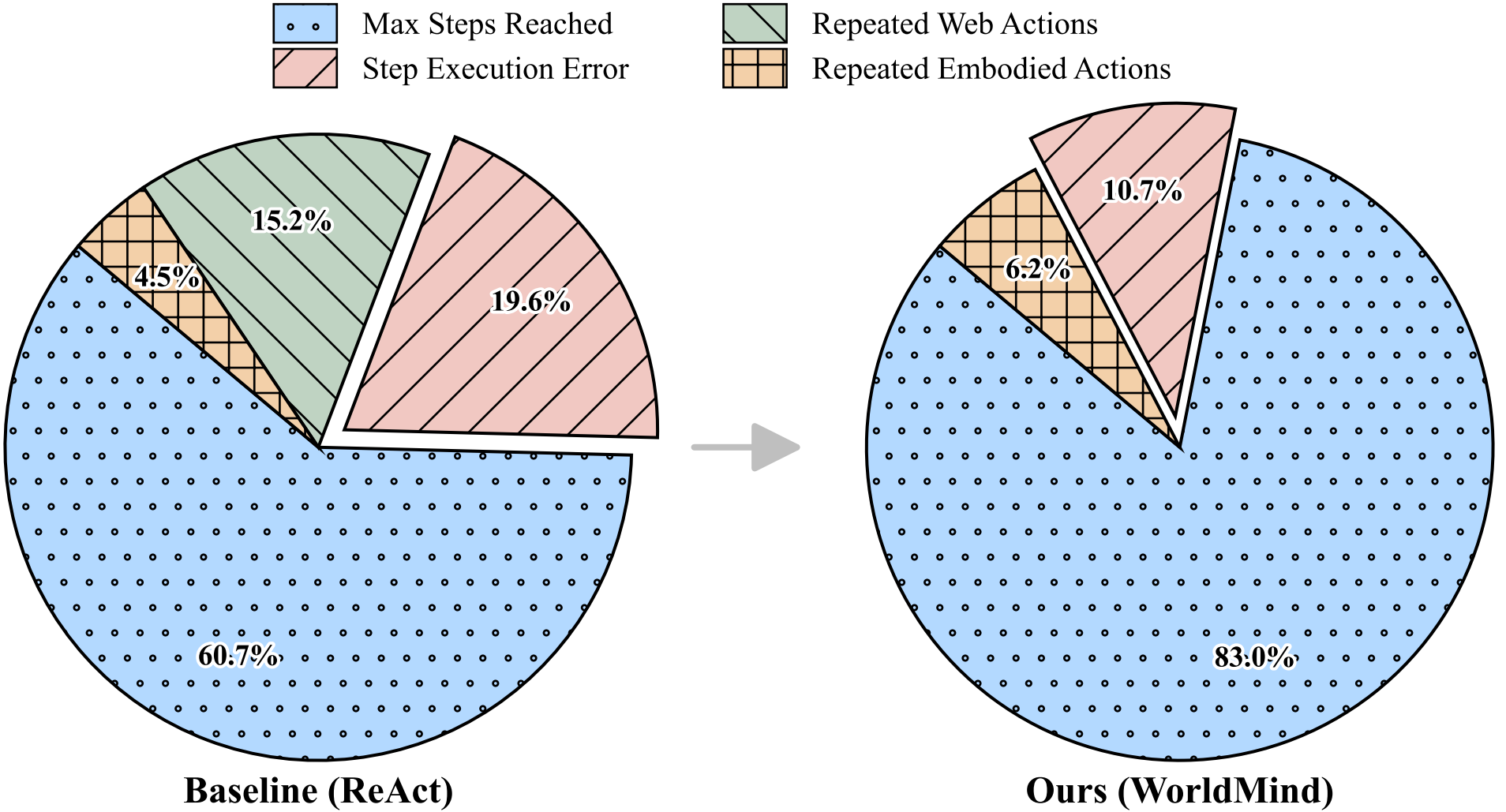
Max Steps Reached
Step Execution Error
Repeated Web Actions
Repeated Embodied Actions
15.2%
19.6%
4.5%
60.7%
10.7%
6.2%
83.0%
Baseline (ReAct)
Ours (WorldMind)