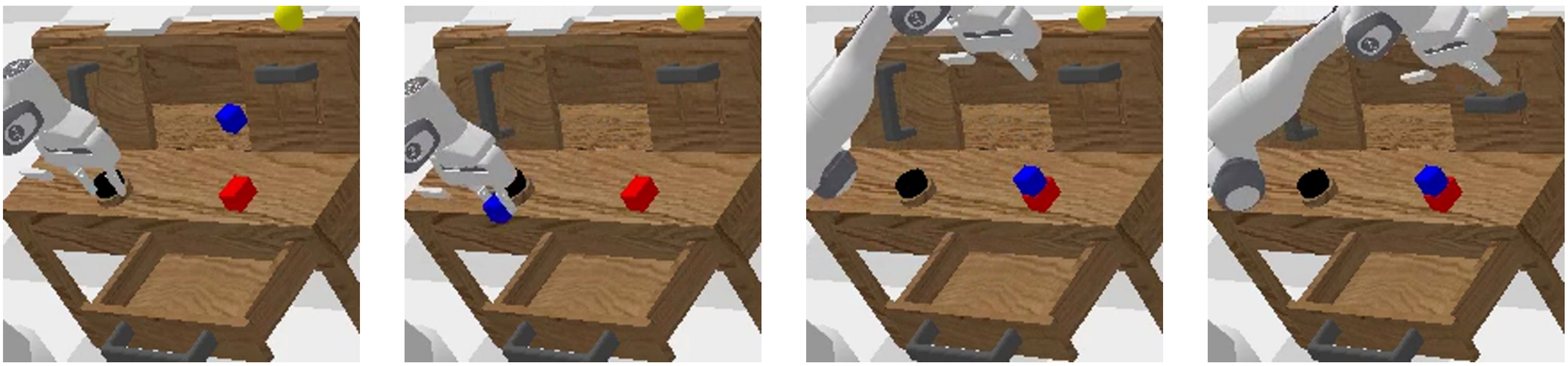
(a) CALVIN | Sequence 0526

Complete the five instructed manipulation subgoals in sequence.

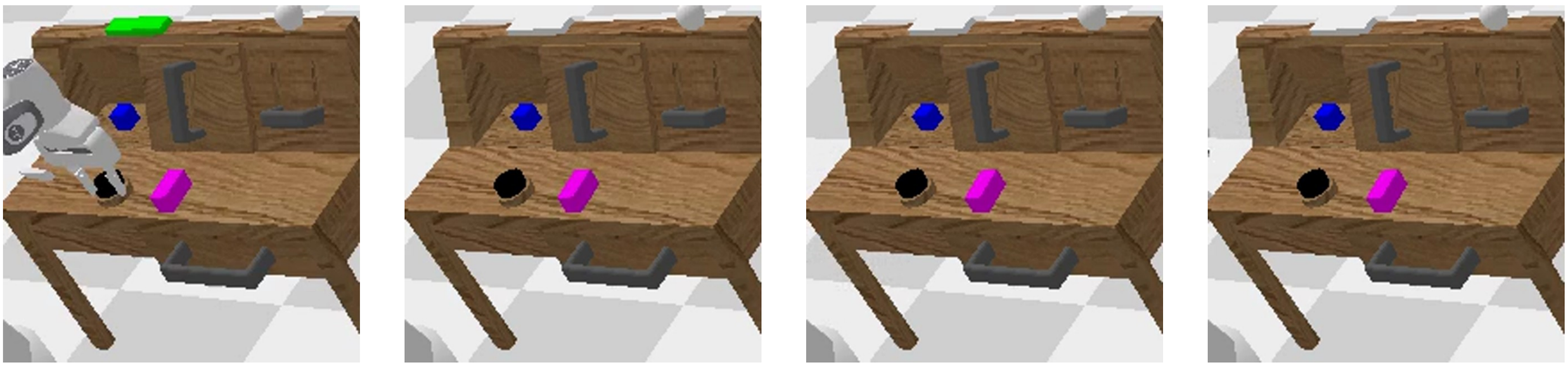


0.0 s | 4.8 s | 12.0 s | 16.8 s

Claude Opus 5.5 Reported success | 40-round ceiling | 229.4 s

(b) CALVIN | Sequence 0557

Complete the five instructed manipulation subgoals in sequence.

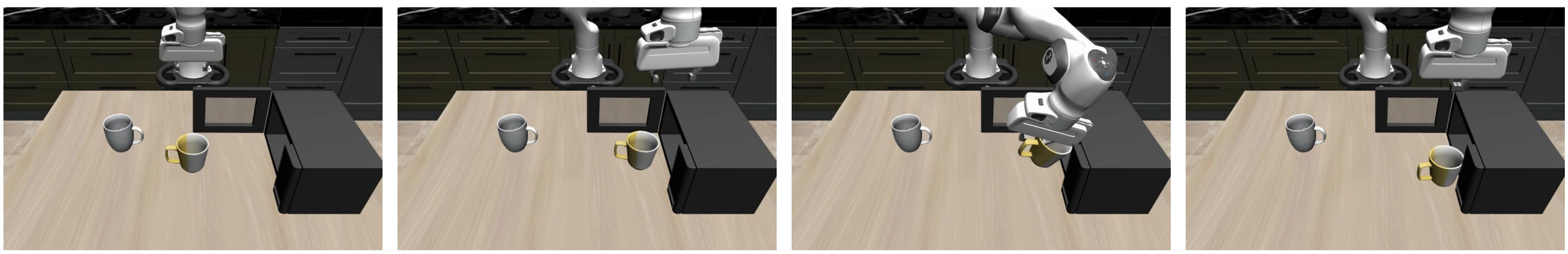


0.0 s | 59.7 s | 149.2 s | 208.9 s

Claude Opus 5.5 Round limit reached | 40 rounds | 978.6 s

(c) LIBERO-PRO | Mug and Microwave

Put the yellow-and-white mug in the microwave and close it.



0.0 s | 70.8 s | 176.9 s | 247.7 s

Claude Opus 5.5 Round limit reached | 24 rounds | 1308.6 s