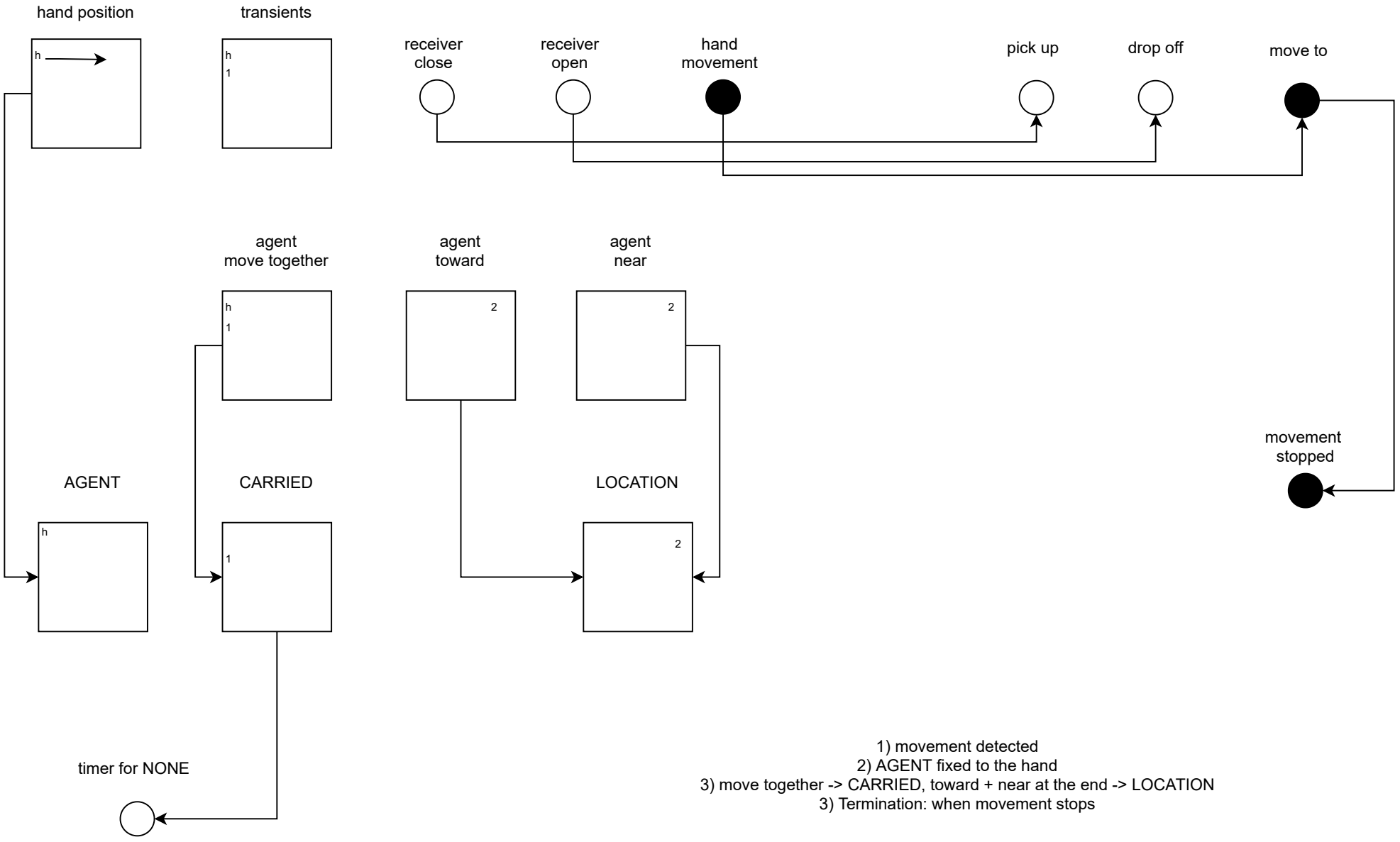
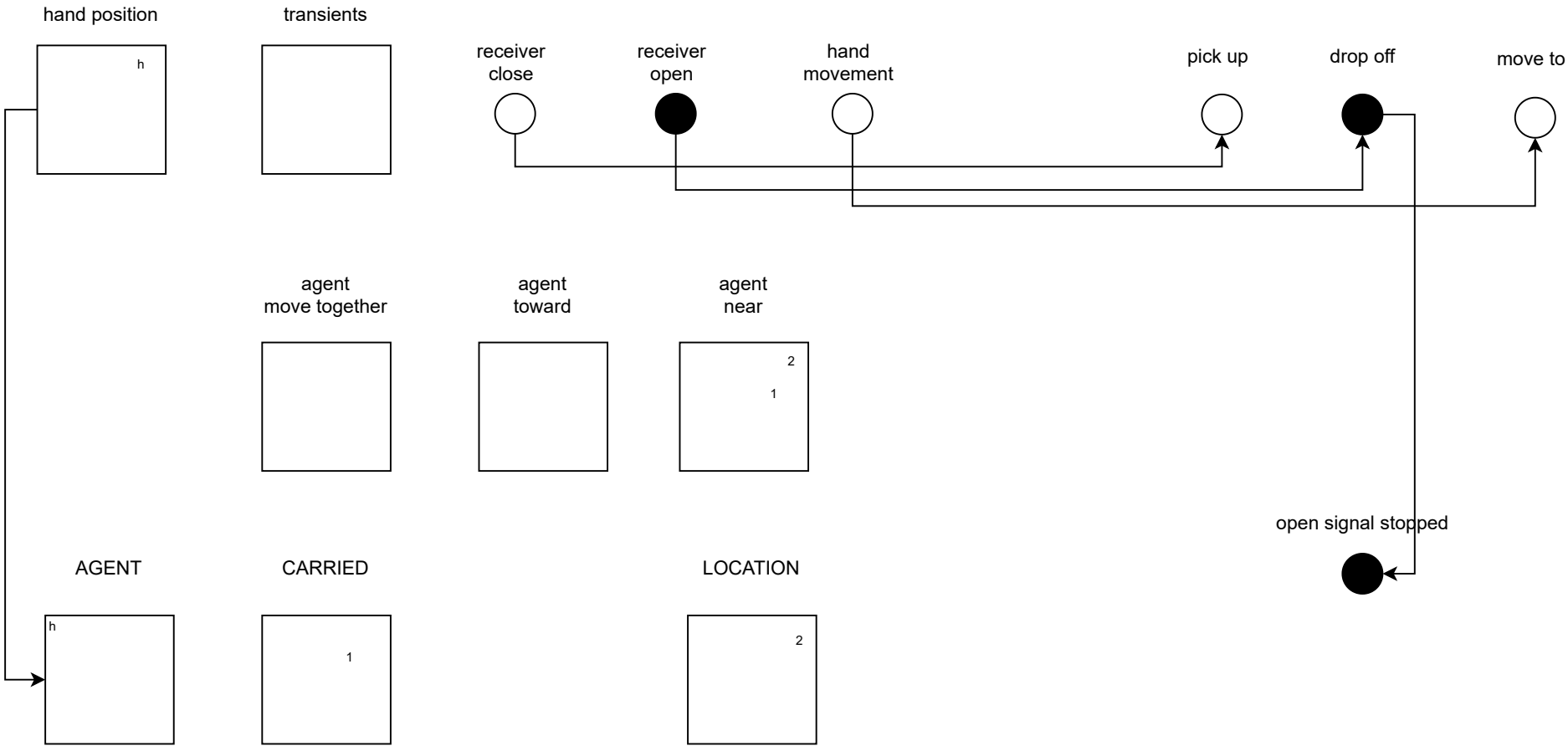
Scenario 1: move_to



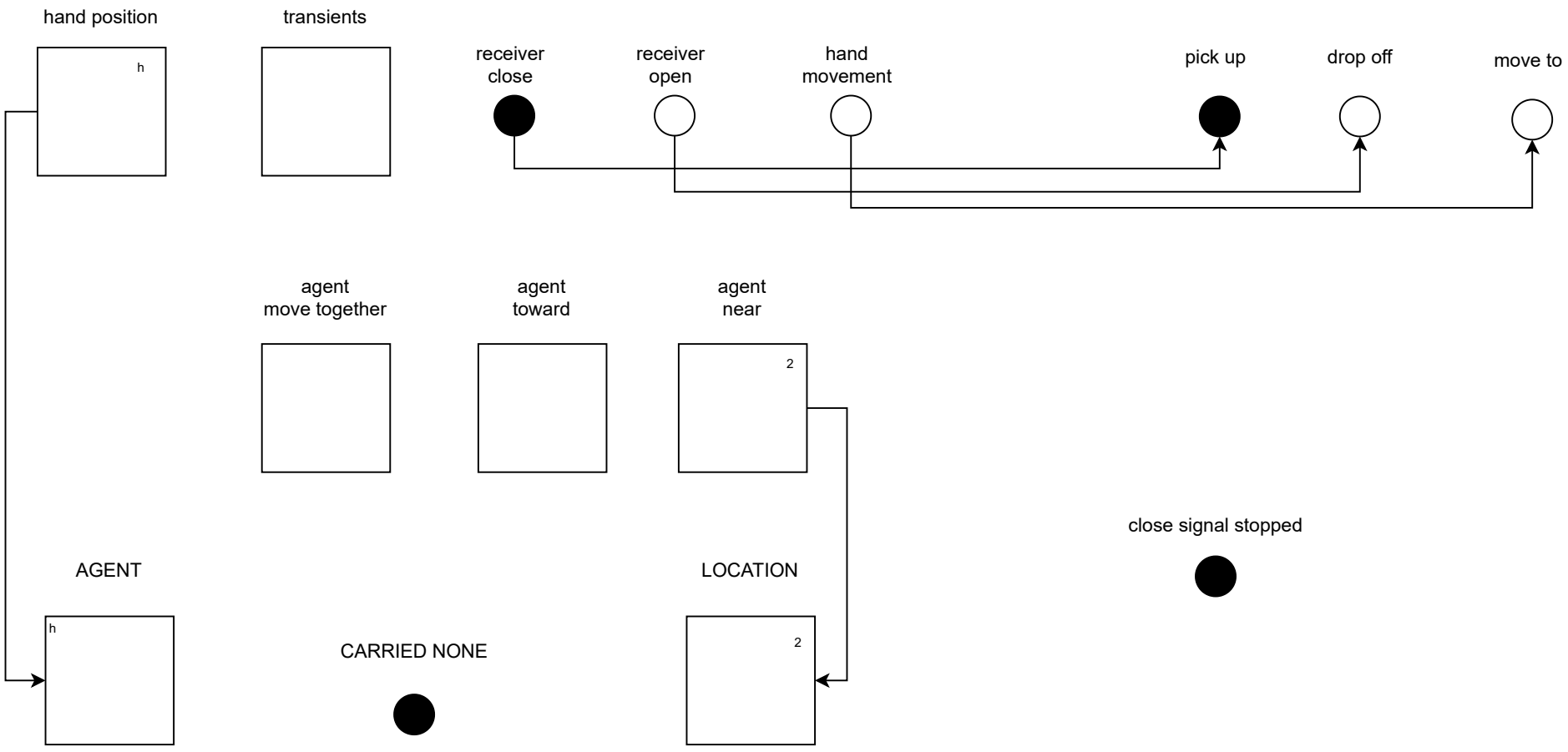
1) movement detected
2) AGENT fixed to the hand
3) move together -> CARRIED, toward + near at the end -> LOCATION
3) Termination: when movement stops

Scenario 2: drop_off



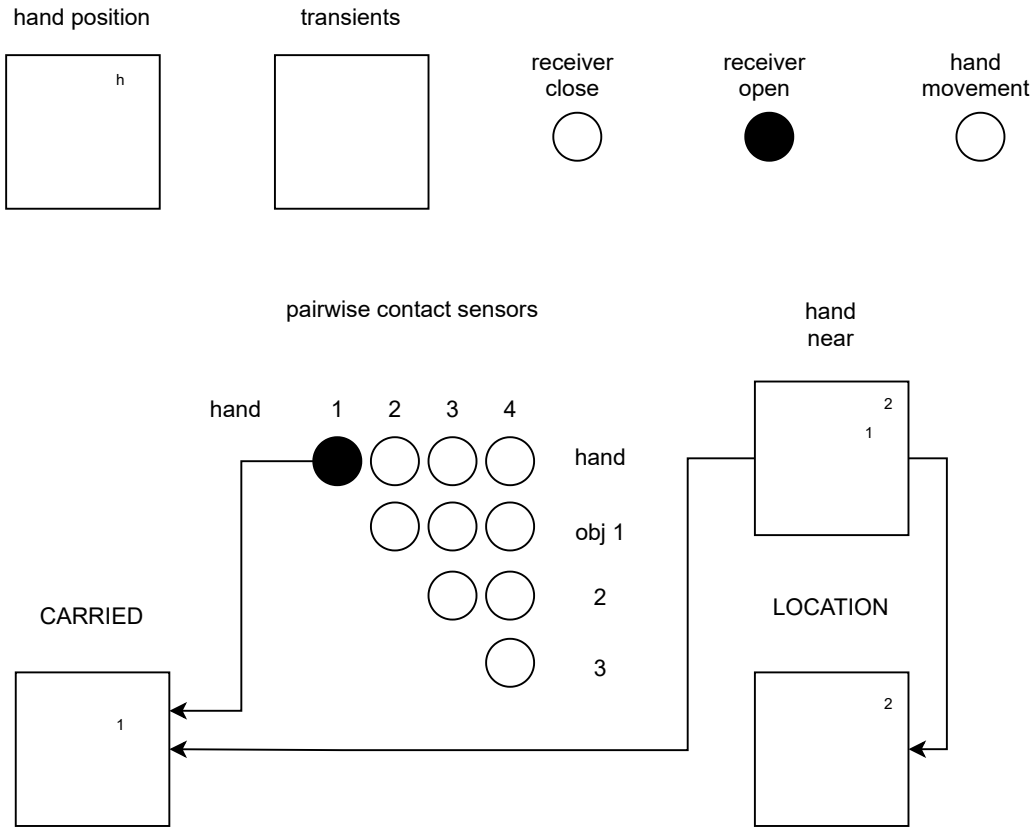
1) hand opening detected (simulated by boost)
2) AGENT fixed to the hand
3) CARRIED / LOCATION just depends on memory from the previous action
3) Termination: when opening motion stops

Scenario 3: pick_up



1) hand opening detected (simulated by boost)
2) AGENT fixed to the hand
3) CARRIED fixed to NONE, LOCATION is an object that is near
3) Termination: when closing motion stops

Scenario 2: drop_off using contact sensor



1) emitter signal decide primitive
2) object that had contact initially -> carried, another object that is near -> location
3) termination when signal stops

PROBLEM: using contact information requires object segmentation in all 2D fields