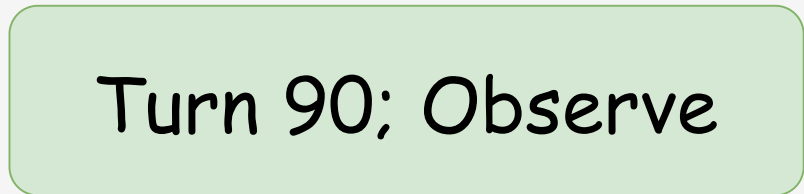
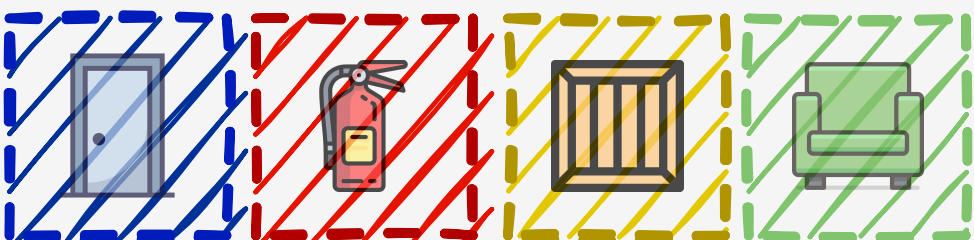
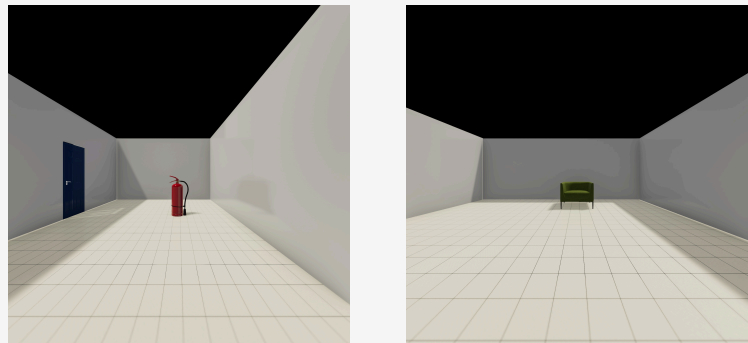
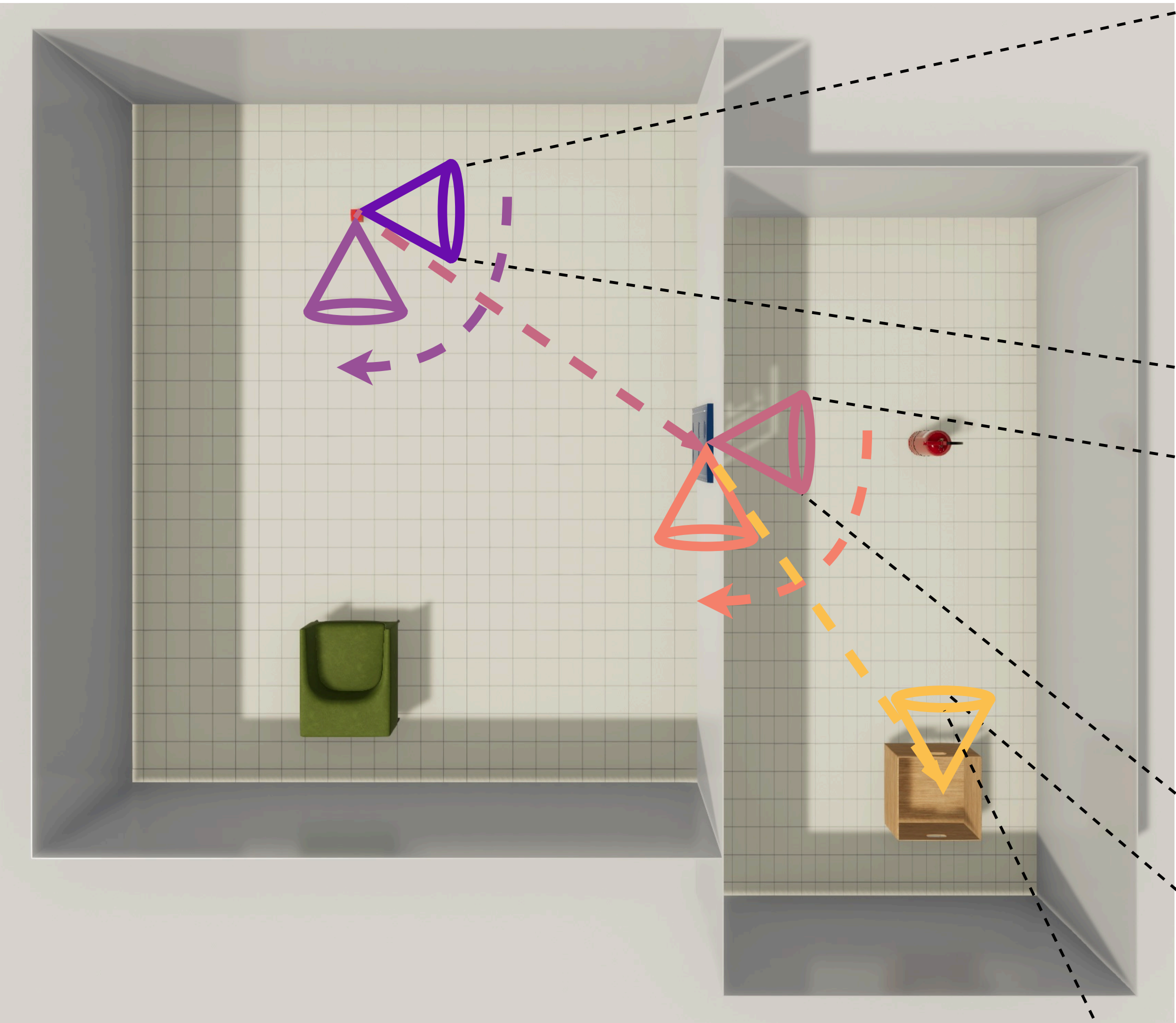


Environment Topdown



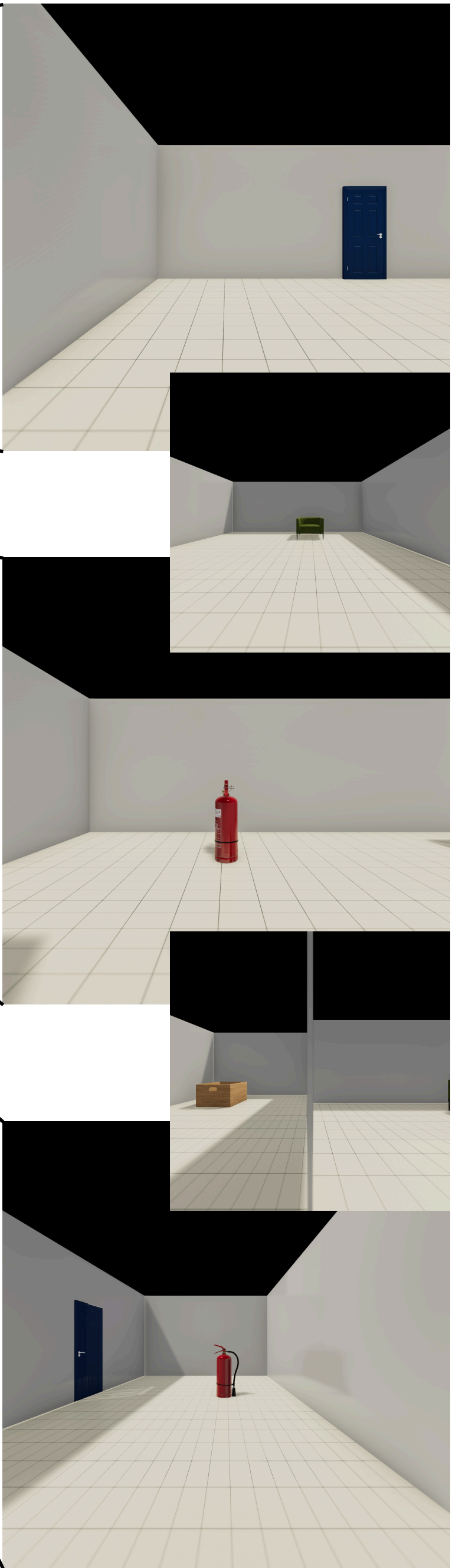
vision observation

text observation

predicted 2D coord

agent action

Active exploration



- 1

front-right, mid distance

Turn 90; Observe
- 2

front, mid distance

Turn 90; Observe
- 3

front, near

Turn 180; Go to ; Observe
- 4

front-left, mid distance

Turn 90; Observe
- 5

front, mid distance
 front-left, mid distance

Go to ; Turn 180; Observe

Spatial Belief

Evaluation (Pairwise):
Where is relative to

✓ Northeast, slightly far

Evaluation (Rotation):
Object order if rotate clockwise

✓

Representation
Map each object to 2D coord

