

Agent's Internal State



Interactive Agent(VLM)

Hypothesis Box(H-Box)

Hypothesis: Pneumothorax?

Belief: 0.3

$$S_t = (HBox_t, t, p_t)$$

Policy π_θ

$a_t = \langle \text{Claim} \rangle$



$a_t = \langle \text{Abstain} \rangle$



$$p \leftarrow \text{SharpenBelief}(p)$$

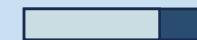
$$p \leftarrow 0.5$$

Belief Update

Hypothesis Box(H-Box)

Hypothesis Confirmed:
Pneumothorax

Belief: 0.3 → 0.8



$$\begin{aligned} p_{t+1} &\leftarrow \text{FuseEvidence}(p_t, p_{\text{evidence}}) \\ \text{probe} &\leftarrow \text{True} \end{aligned}$$



Evidence Score: 0.8

$$ROI, p_{\text{evidence}} \leftarrow KBCS(x, c)$$

$a_t = \langle \text{Probe \& Ground (P\&G)} \rangle$

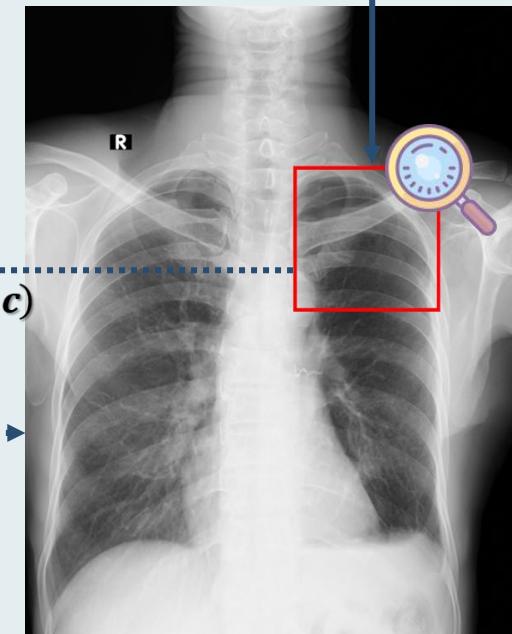
$\langle \text{Stop} \rangle$

External Evidence Grounding



External Tool
KBCS

Localize ROI
Score Confidence



$t=1 \langle \text{Init} \rangle \rightarrow t=2 \dots \rightarrow t=T_{\max}$