Let Trajectories Tell Your Intentions

Inferring Human Intentions from Trajectories

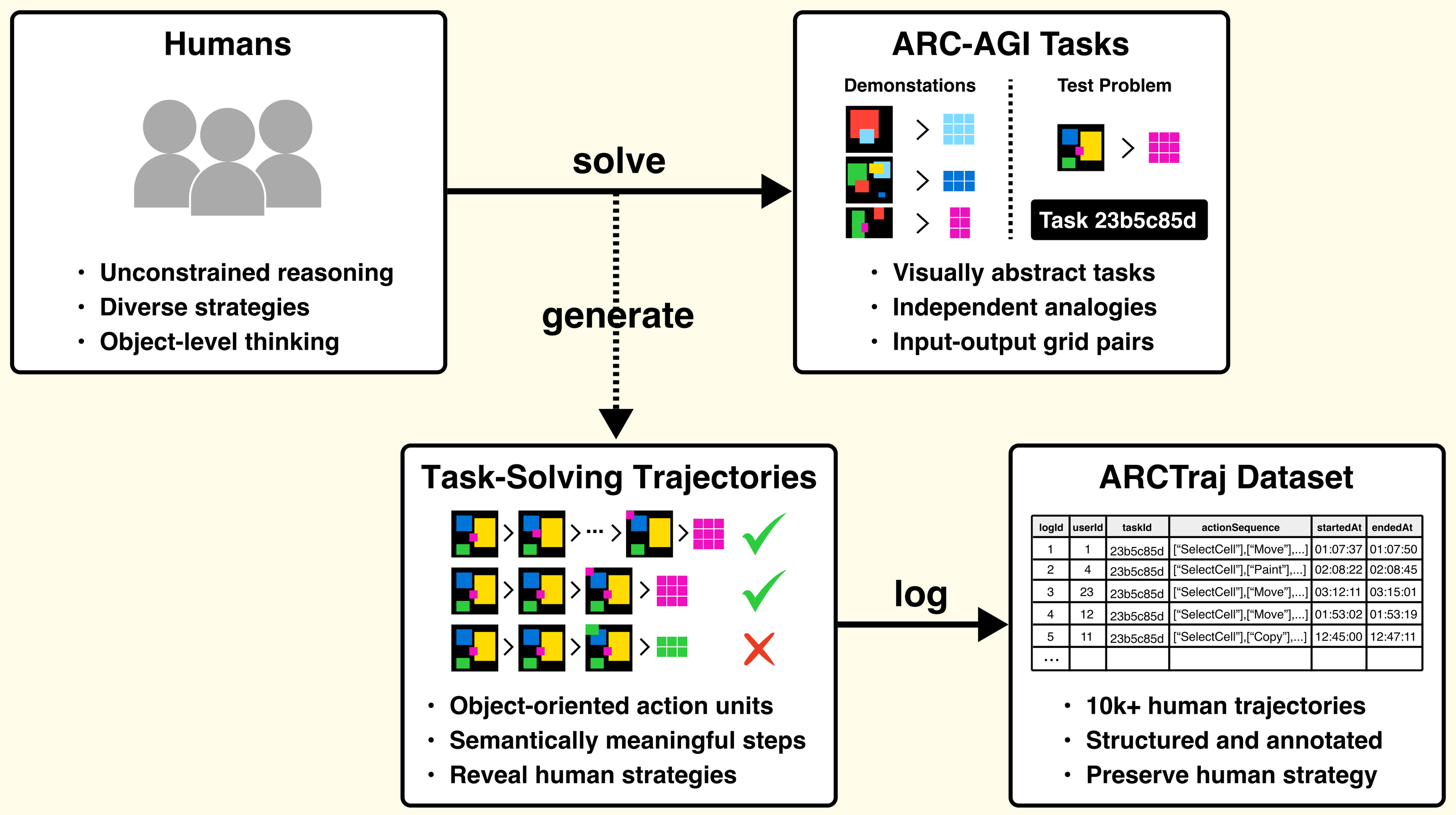
Sejin Kim, Hosung Lee, Sundong Kim

ARCTraj paper

TL;DR Human actions may deviate, but **intentions** can still be *inferred*.

ARCTraj: More Than Action Logs

ARC Tasks + Human Solution → ARCTraj



$$\tau := (s_0, a_0, s_1, \dots, a_{l-1}, s_l) \in \texttt{ARCTraj}$$

- τ is a sequence of **states and actions**
- τ is logged from humans solving ARC tasks
- Rich supervision, but **intentions not explicit** in τ

Why Intention Estimation Matter

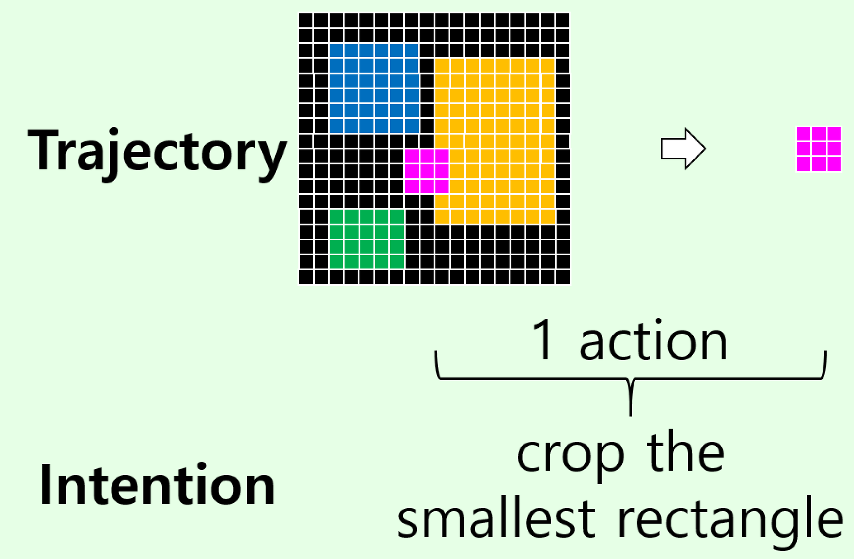
→ Actions → ✗ *(Diverge)*

→ Intentions → ✓ *(Stable)*

- Inconsistent actions may reflect consistent goals
- Inferring **intent patterns** aids understanding

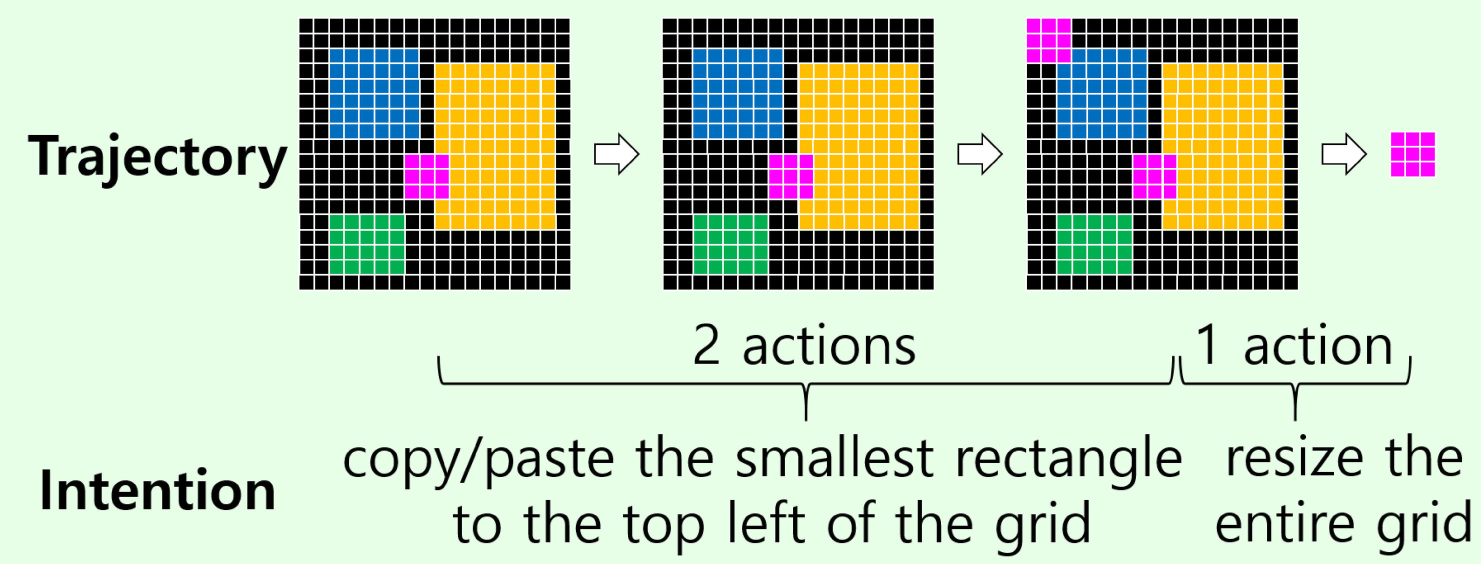
Observed Actions $\nRightarrow$ User Intentions

Intention ≡ Action

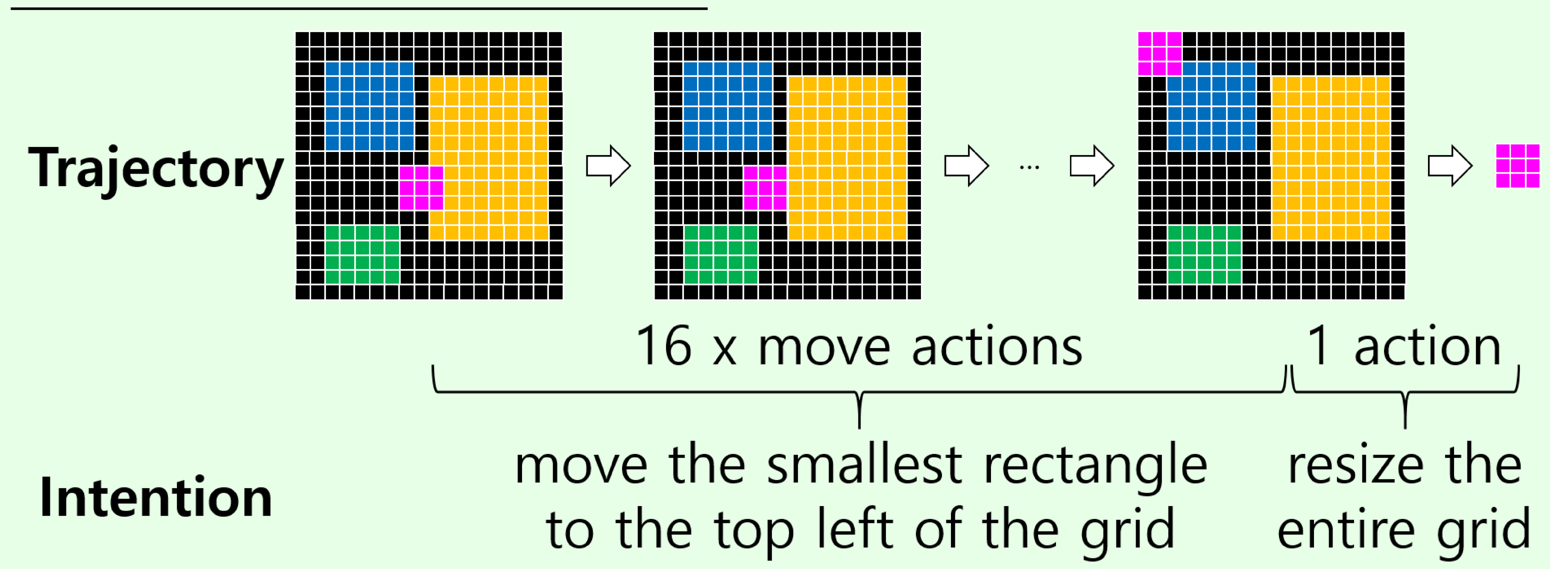


Intention ≡ Action Sequence

Case 1: Tool Limitation



Case 2: User Unfamiliarity



(*Caveat: Sometimes, intention may be* ***spoiled by confusion***)

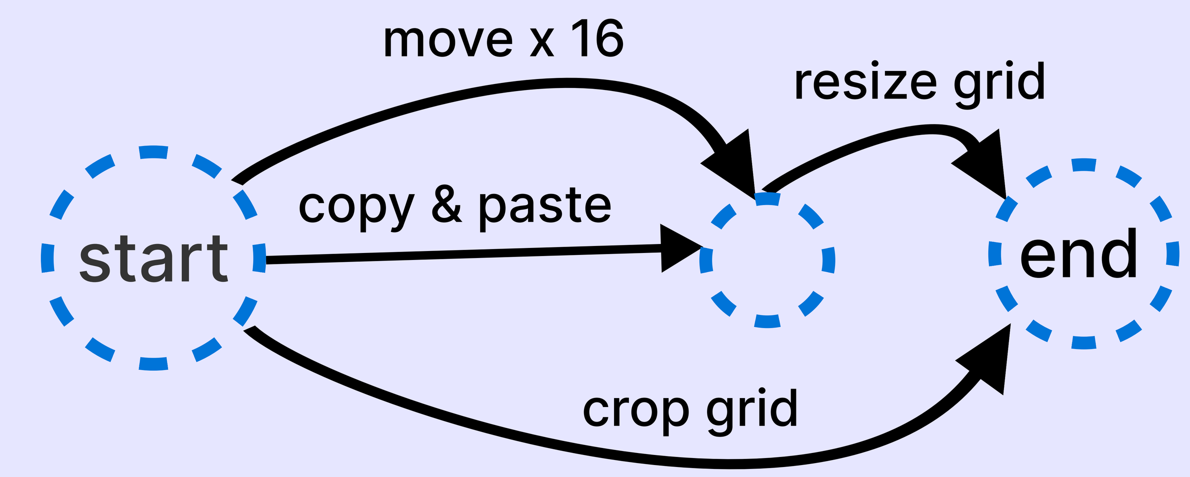
Same Intention, Diverging Paths

$$\tau_i \neq \tau_j, \quad \text{but possibly} \quad \mathcal{I}(\tau_i) = \mathcal{I}(\tau_j)$$

(An intention can be represented through various trajectories)

Inferring Intentions from ARCTraj

Step 1: Extracting Popular States ($\mathcal{P}$)



- $\mathcal{P}$ are extracted **intra-task** by frequency

Step 2: Segmenting Trajectories (τ)

$$\tau \longrightarrow \tau_1 \circ \tau_2 \circ \cdots \circ \tau_n$$

$$\text{where } \tau_i = (s_i^{\text{in}}, \dots, s_i^{\text{out}}), \quad s_i^{\text{in}}, s_i^{\text{out}} \in \mathcal{P}$$

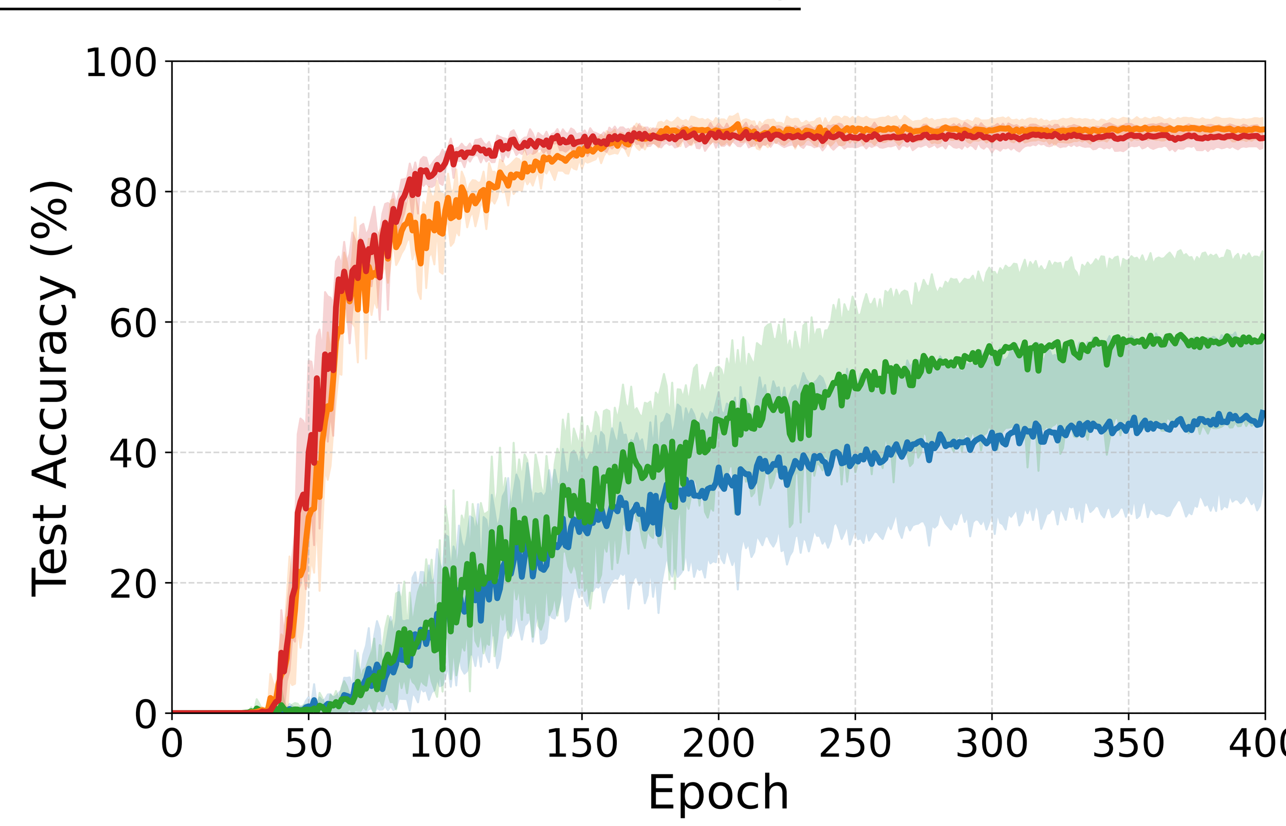
- Each τ is split into **segments** (τ_i)
- Splits occur at **popular state pairs** ($s_i^{\text{in}}, s_i^{\text{out}}$)

Step 3: Representing Intention ($\mathcal{I}(\tau_i)$)

$$\mathcal{I}(\tau_i) := (s_i^{\text{in}}, s_i^{\text{out}})$$

- Segment is interpreted as a distinct **intent transition**

Effect of Intention-Guided Learning



- Faster convergence (● > ●), Better accuracy (● > ●)
- Human intent improves both **efficiency & generalization**

Addressing and Visualizing Misalignments in Human Task-Solving Trajectories

sejinkim@gist.ac.kr