Human-Aware Epistemic Task Planning for Human-Robot Collaboration

The workshop on Human-Aware and Explainable Planning *(HAXP) at* ICAPS 2024

Shashank Shekhar, Anthony Favier, and Rachid Alami

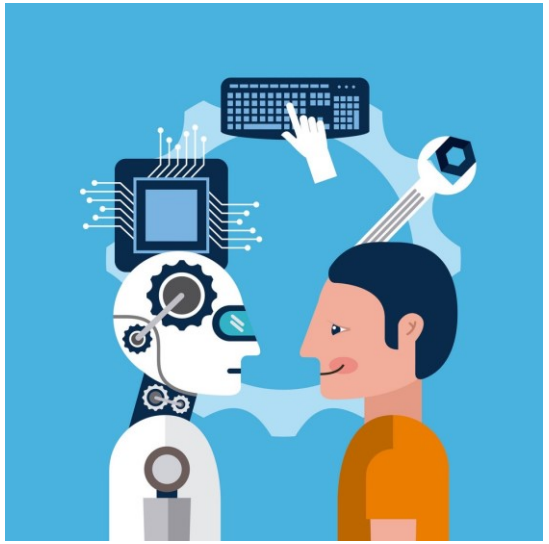
LAAS-CNRS, Toulouse, France, ANITI

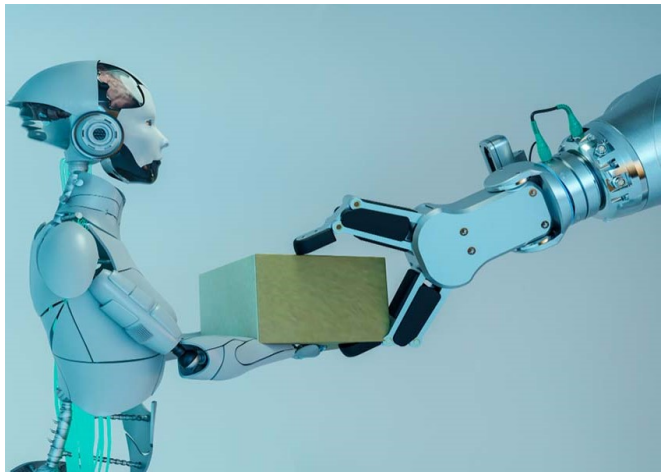
3 June 2024

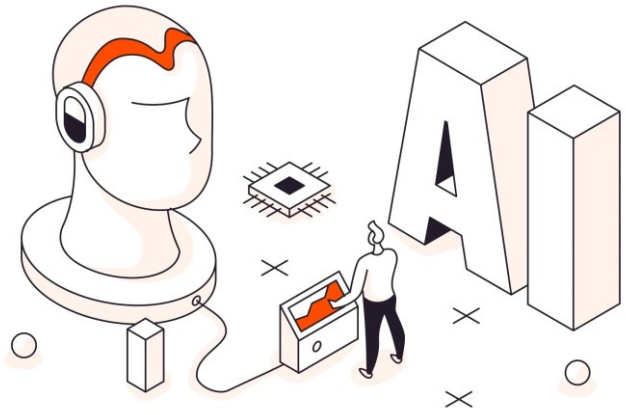
Our Focus

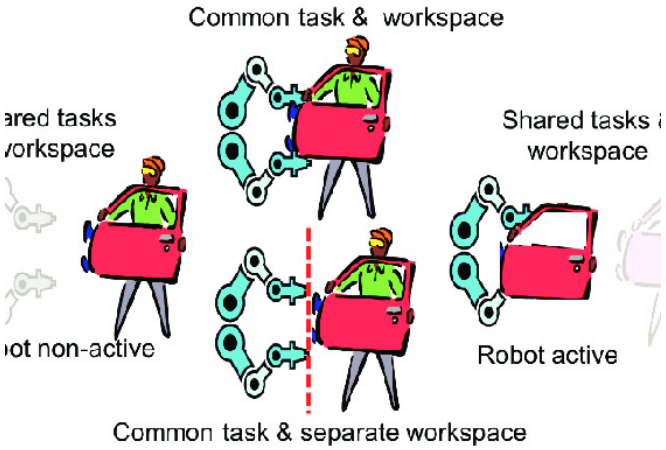
decisional aspects of HR collaboration

- Artificial Intelligence (AI) is required to equip robots with the **ability to make decisions** in settings like home, workshop, etc.

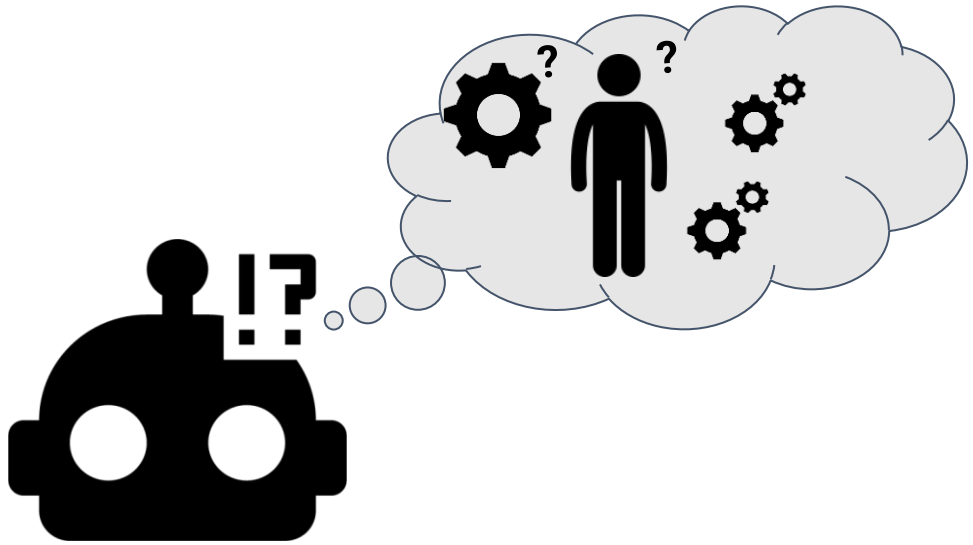








The presence of **human** in the environment is challenging



- **Human-Aware Task Planning (HATP)** takes into consideration a variety of factors, such as developing robot policies that are *proactive*, *safe*, *legible*, *acceptable*, *predictable, etc.* (Alami et al., 2006).

Related Work:
Planning for HR Collaboration

- **Offline Planning** (for a shared goal)
  - Builds joint policies before the execution starts
  - Built joint policies that include *coordinated human robot actions*
  - Policies shared are *assumed* to be followed by **H** & **R** (Alami et al., 2006; Roncone et al., 2017; Lallement et al., 2018)

Related Work:
Planning for HR Collaboration

- **Offline Planning** (for a shared goal)
  - Builds joint policies before the execution starts
  - Built joint policies that include *coordinated human robot actions*
  - Policies shared are *assumed* to be followed by **H** & **R** (Alami et al., 2006; Roncone et al., 2017; Lallement et al., 2018)

- **Online Planning**
  - Online human-aware planners prepare reactive policies for the real-time operation
  - They prioritizing *responsiveness* and *adaptiveness* (Darvish et al., 2021, Ramachandruni et al., 2023)

Related Work: Uncontrollable Human Operator

- Handled by using a **distinct human model**
  - Builds robot's policies by predicting human decisions and actions (Hoffman et al., 2007; Unhelkar 2019, 2020; Buisan et al., 2022)

Related Work: Uncontrollable Human Operator

- Handled by using a **distinct human model**
  - Builds robot's policies by predicting human decisions and actions (Hoffman et al., 2007; Unhelkar 2019, 2020; Buisan et al., 2022)

- **HATP/EHDA** (Buisan et al., 2022)
  - It stands for **H**uman-**A**ware **T**ask **P**lanner which **E**stimates **H**uman **D**ecisions and **A**ctions
  - (Qualitative) Offline planning framework
  - Explicit *task models* for both HUMAN and ROBOT – with (non-) shared goals
  - Supports planning with agent's ***false beliefs***
  - The joint policy is produced in a ***turn-taking*** manner

Rational behind HATP/EHDA

- We consider that the human's goal is known
- We focus on rational choices aligned with that objective

Rational behind HATP/EHDA

- We consider that the human's goal is known
- We focus on rational choices aligned with that objective
- Planning accounts for **non-deterministic** choices available to **H**, allowing **R** to adapt its behavior accordingly
  - We will now delve into HATP/EHDA framework description 

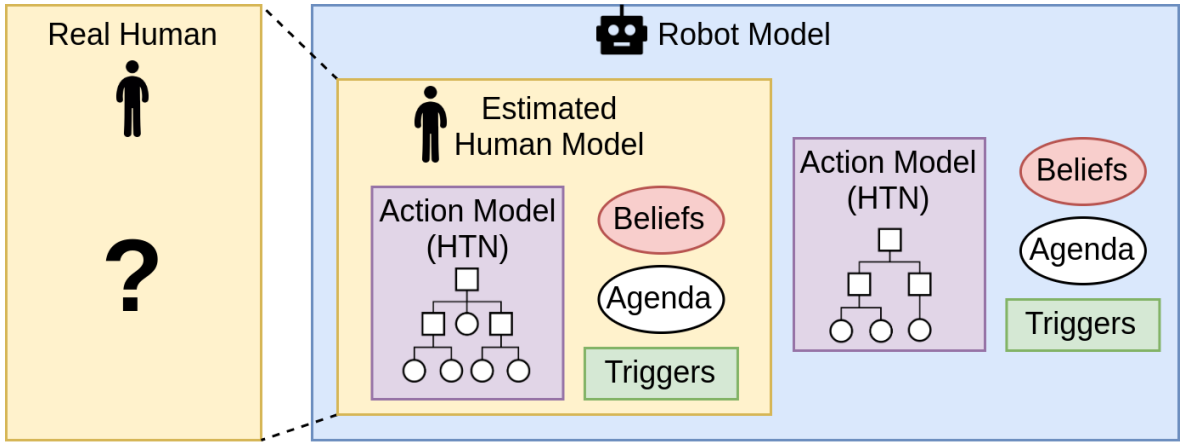
Agents' Models: RM and HM

Beliefs: agent's knowledge from their perspective

Agenda: agent's goals

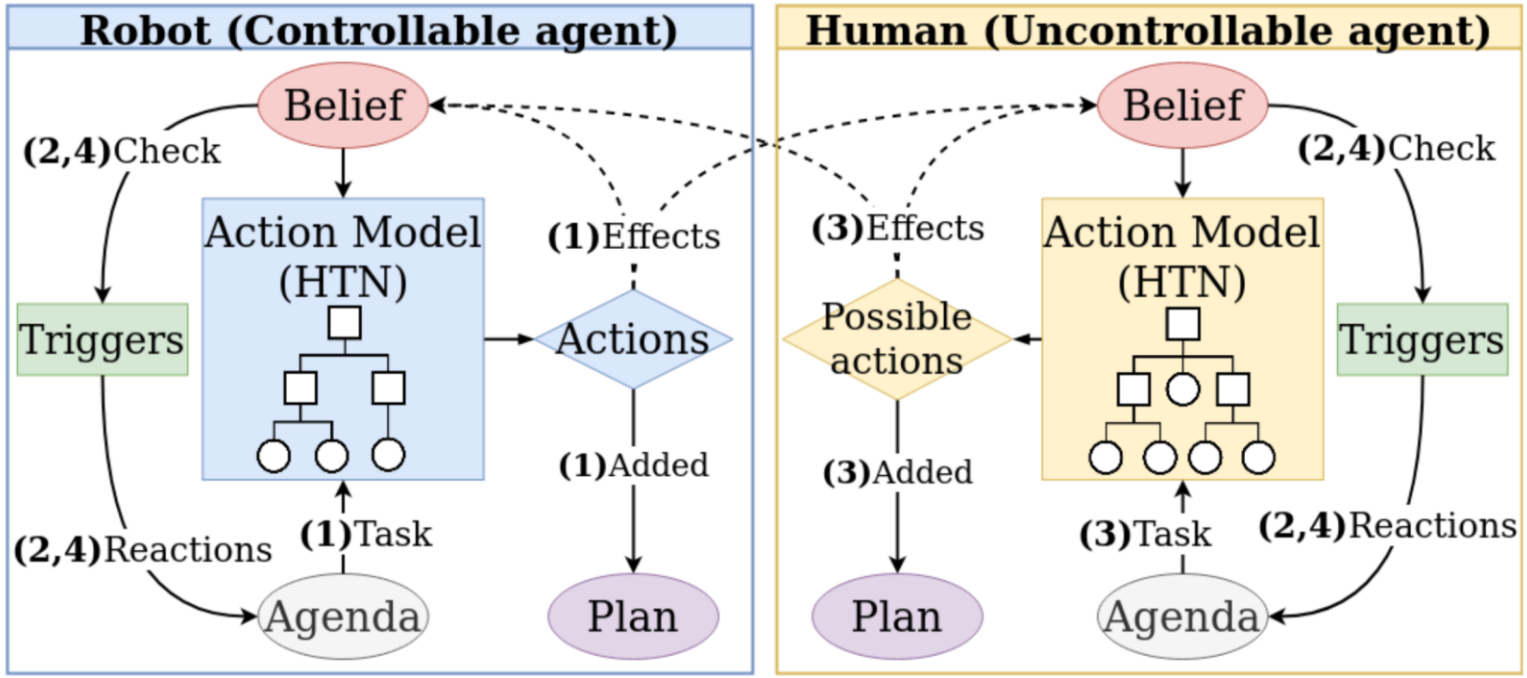
Action Model: agent's capabilities

Triggers: agent's possible reaction

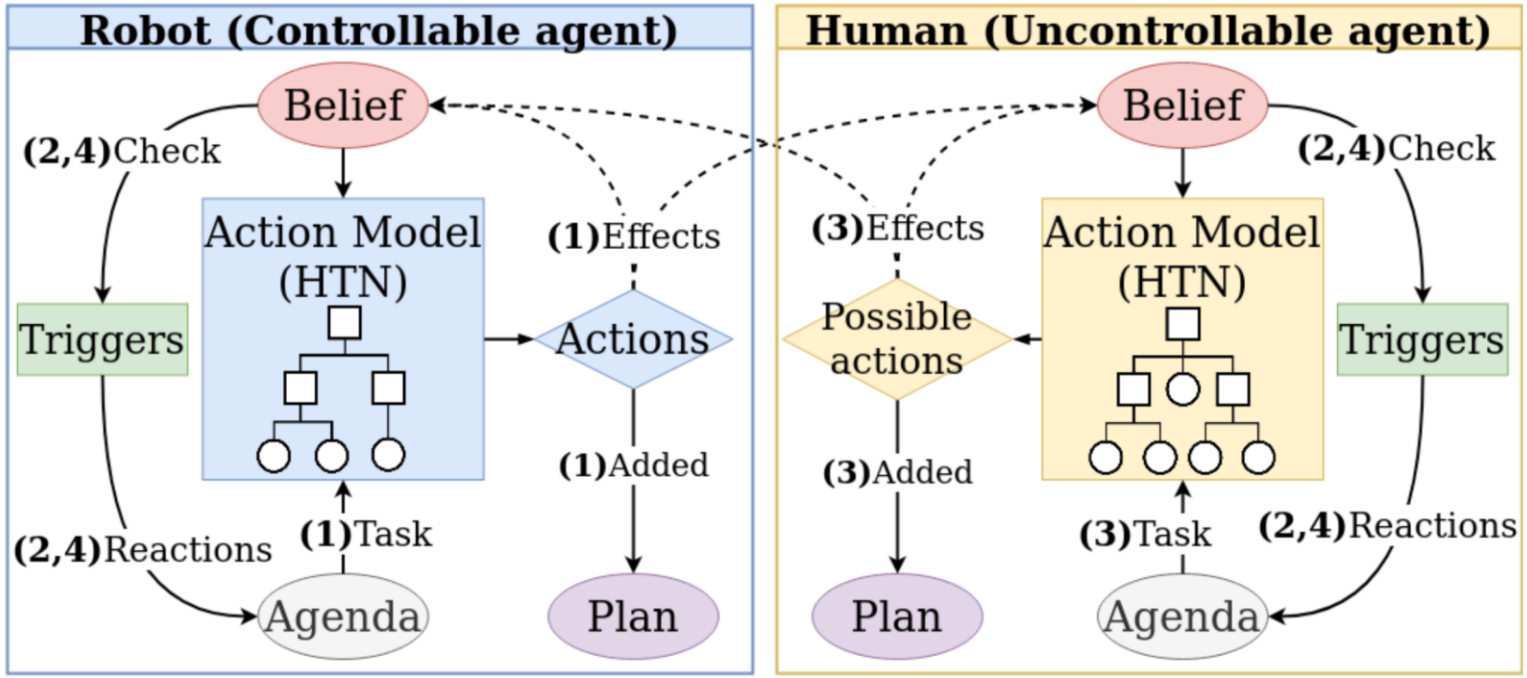


Similar model structures but their purposes are fundamentally different!

HATP/EHDA Exploration

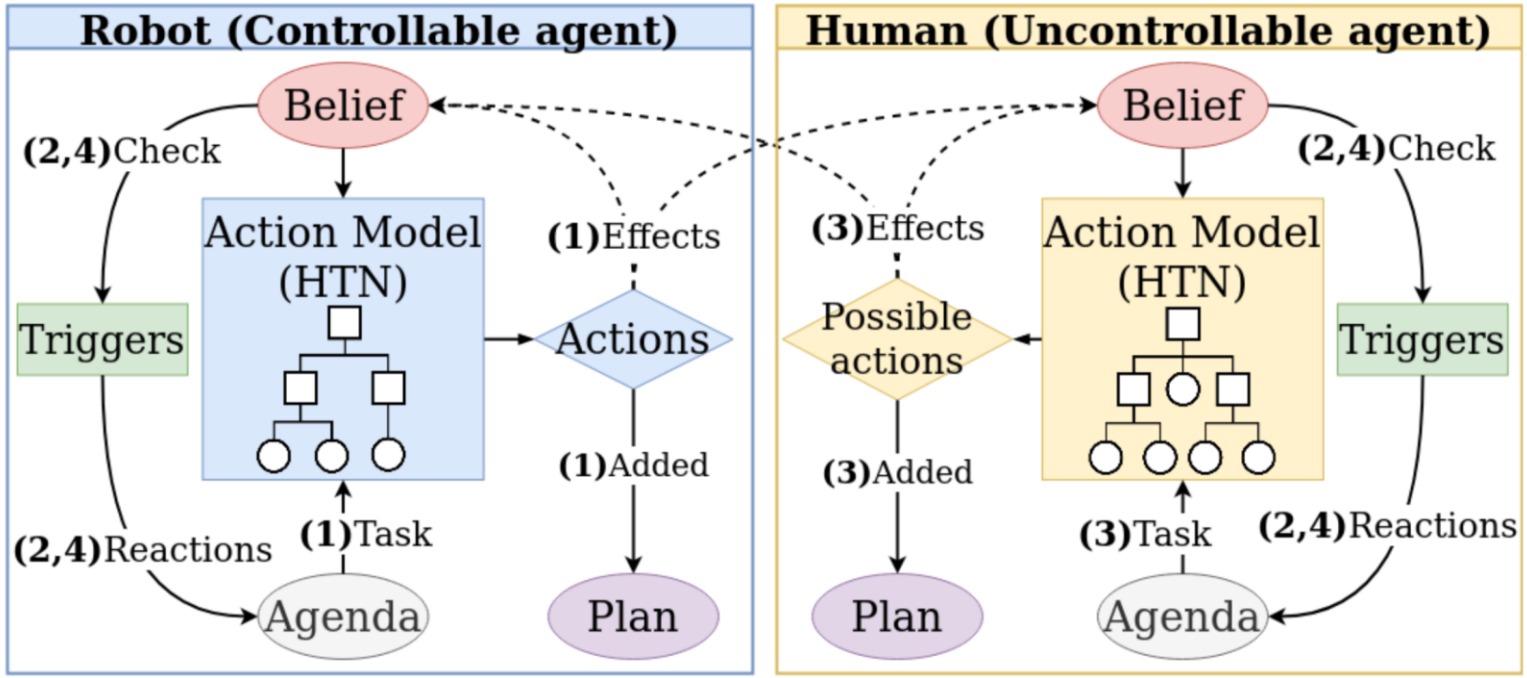


HATP/EHDA Exploration



Step 1. Compute the robot's next action – append it into the plan, update agents' beliefs

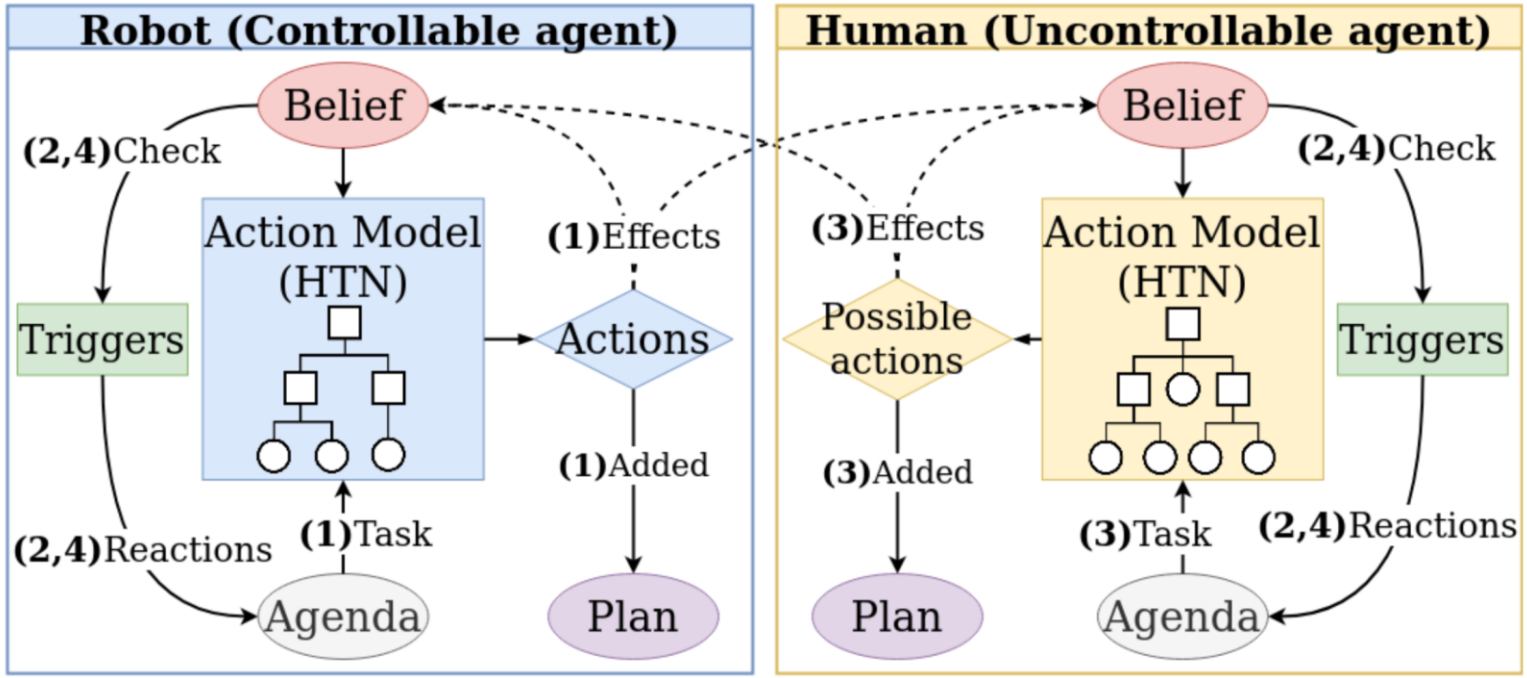
HATP/EHDA Exploration



Step 1. Compute the robot's next action – append it into the plan, update agents' beliefs

Step 3. Get possible human actions based on their updated beliefs – append them in the plan and apply the effects on agents' beliefs

HATP/EHDA Exploration



Step 1. Compute the robot's next action – append it into the plan, update agents' beliefs

Step 3. Get possible human actions based on their updated beliefs – append them in the plan and apply the effects on agents' beliefs

Steps 2. & 4. Check the triggers to append reactions in agents' agendas

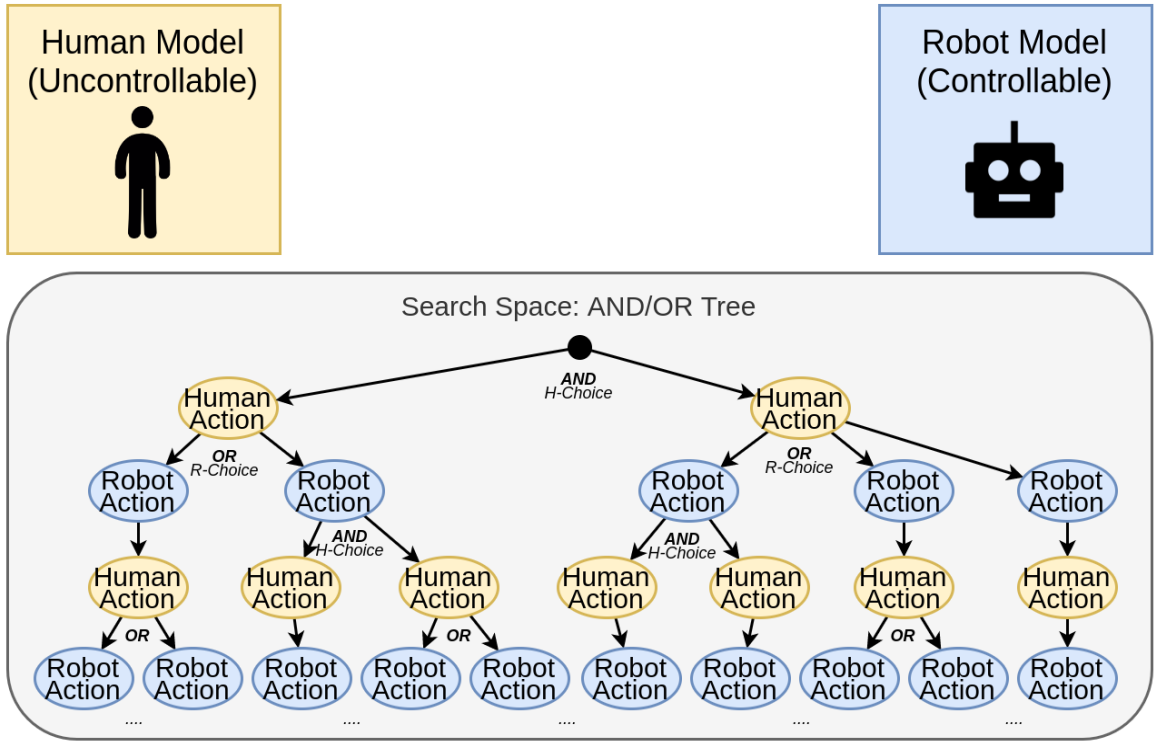
HATP/EHDA – The Planning Process

Human Actions

- Estimated with Human Model
- Non-deterministic (**AND**)

Robot Actions

- **Computed** with Robot Model
- Best choice (**OR**)



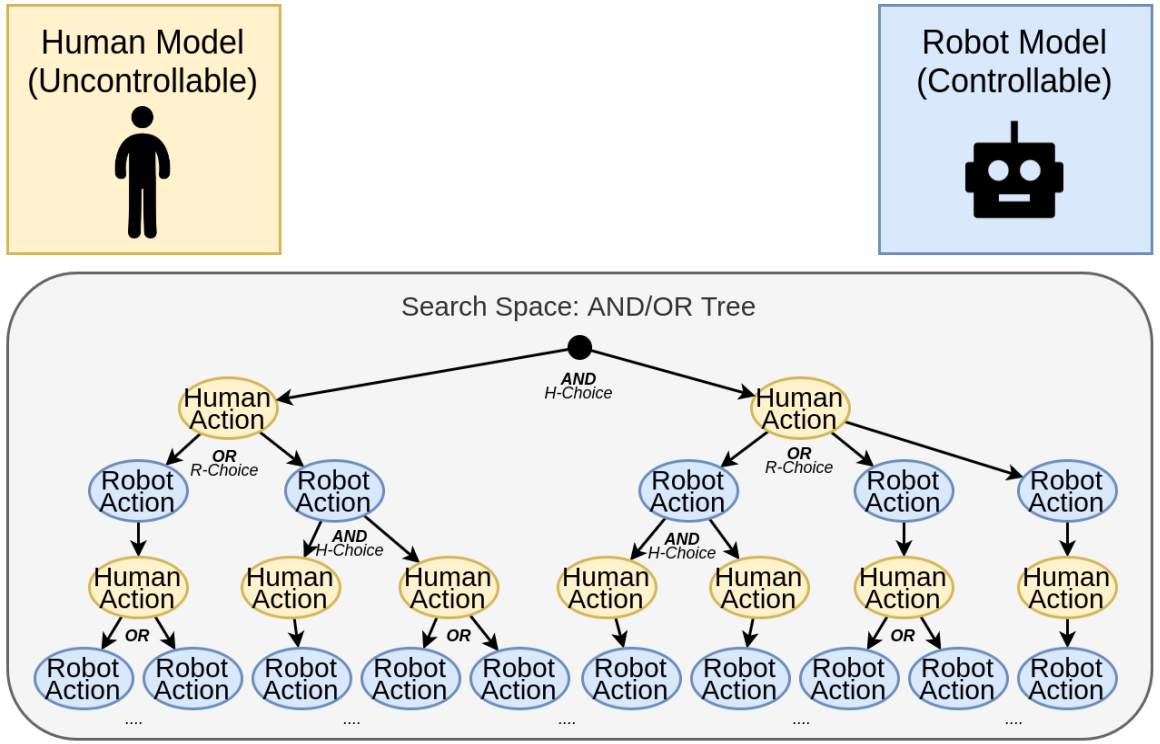
HATP/EHDA – The Planning Process

Human Actions

- Estimated with Human Model
- Non-deterministic (**AND**)

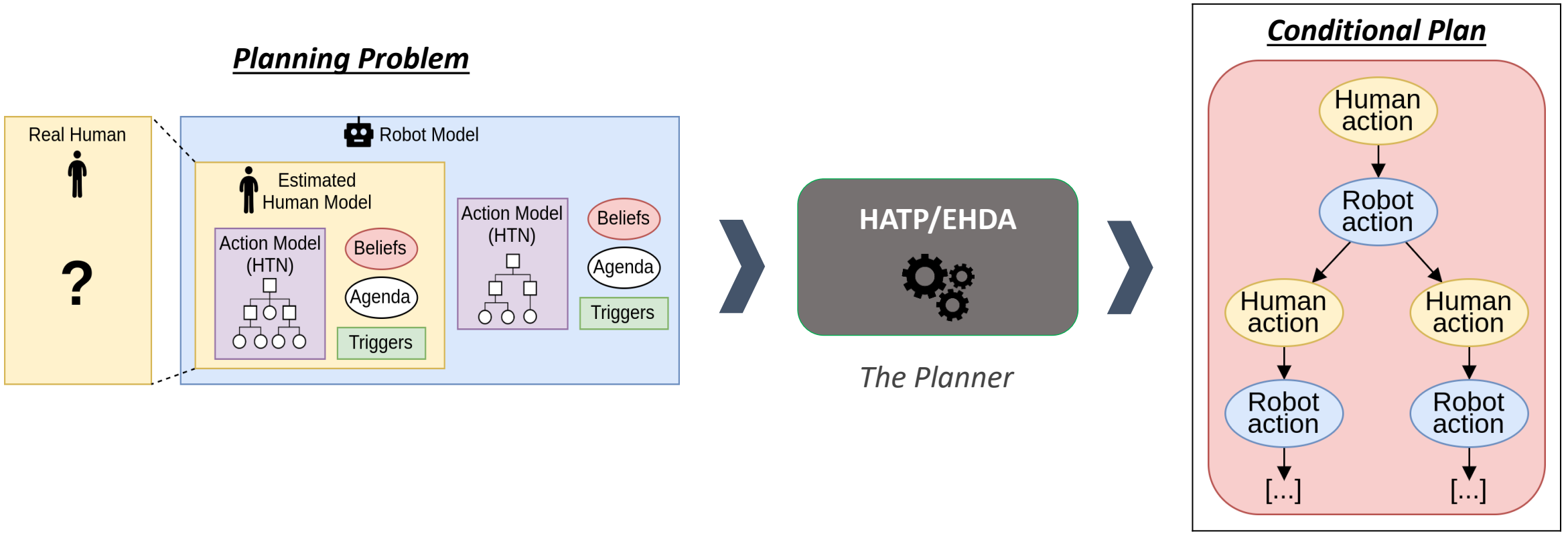
Robot Actions

- **Computed** with Robot Model
- Best choice (**OR**)



Key Point: The planner constructs the robot's policy by interleaving anticipated human actions to accomplish the objective.

HATP/EHDA – The Planning Process [1]



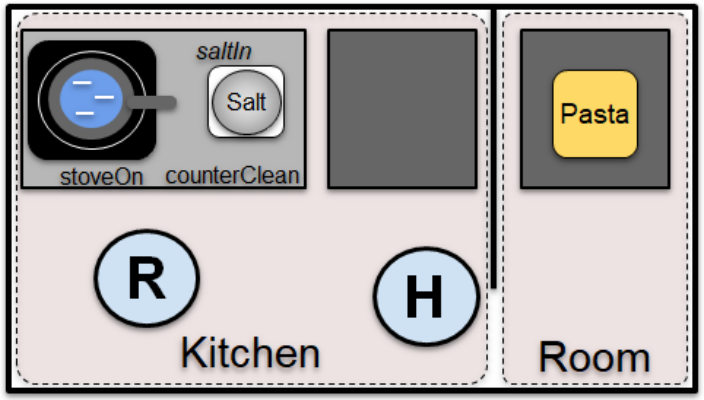
[1] Guilhem Buisan, Anthony Favier, Amandine Mayima, Rachid Alami. HATP/EHDA: A Robot Task Planner Anticipating and Eliciting Human Decisions and Actions. In the proc. of *IEEE International Conference On Robotics and Automation (ICRA 2022)*.

Execution experiences are not always shared!

- We assumes that the human's goal is known
- We focus on rational choices aligned with that objective
- Planning accounts for **non-deterministic** choices available to **H**, allowing **R** to adapt its behavior accordingly
- ***(novelty and contribution)*** Building robot's policy by **anticipating periods** when the human may not actively participate in the task or is away for finishing a subtask
  - **Agents have (non-) shared execution experience**
  - **Agents can have divergent beliefs**

Illustrative Example

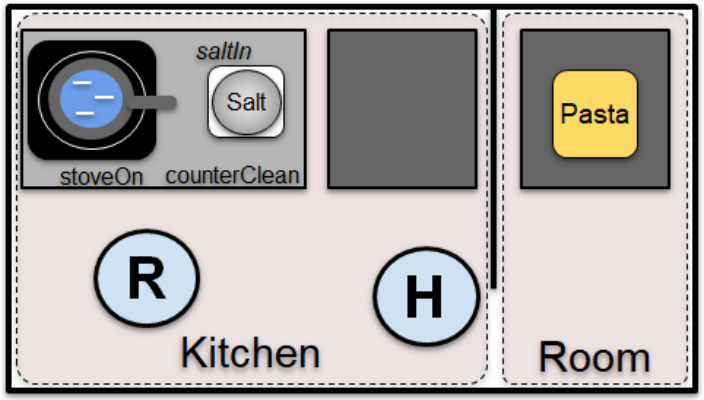
- **Consider the following scenario**
  - **Human** and **robot** team up to achieve a task together
    - For example: **cooking pasta**



Pasta cooking shared task

Illustrative Example

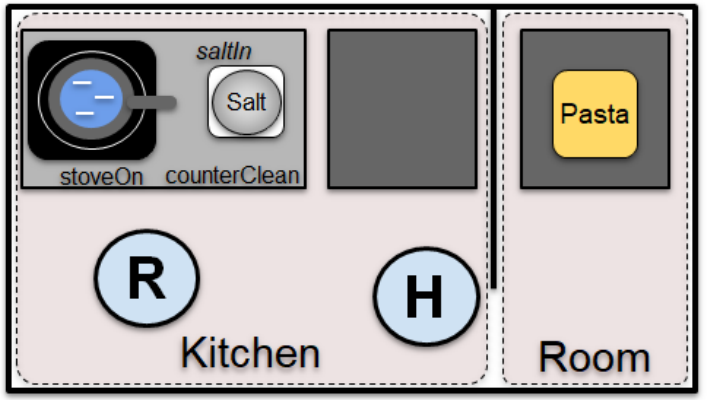
- **<u>Consider the following scenario</u>**
  - **Human** and **robot** team up to achieve a task together
    - For example: **<u>cooking pasta</u>**
  - **Human** needs to leave the *Kitchen*, but…
    - **(*Non-determinism*) Humans** cannot be controlled and **choose** when to leave for fetching the pasta packet
    - **Humans** are rational agents



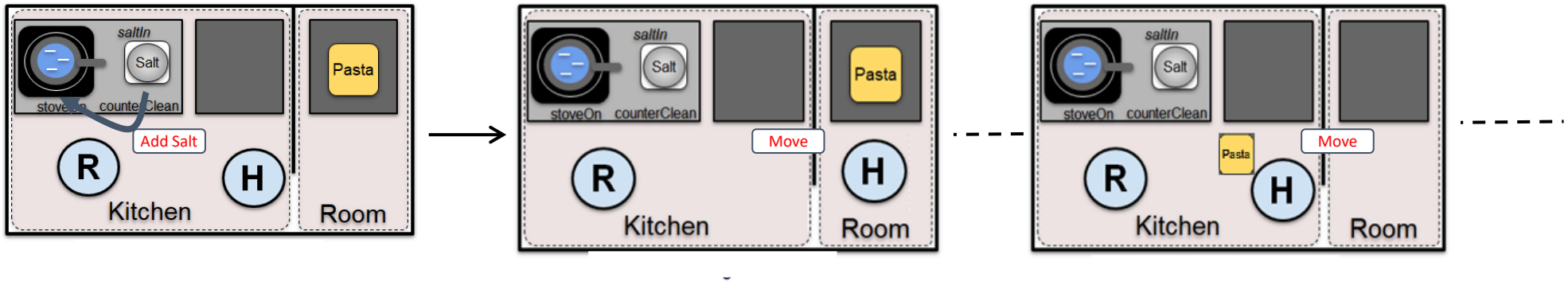
Pasta cooking shared task

Illustrative Example: Human Choice 1

- **Consider the following scenario**
  - **Human** and **robot** team up to achieve a task together
    - For example: **cooking pasta**
  - **Human** needs to leave the *Kitchen*, but…
    - ***(Non-determinism)*** **Humans** cannot be controlled and **choose** when to leave for fetching the pasta packet
    - **Humans** are rational agents

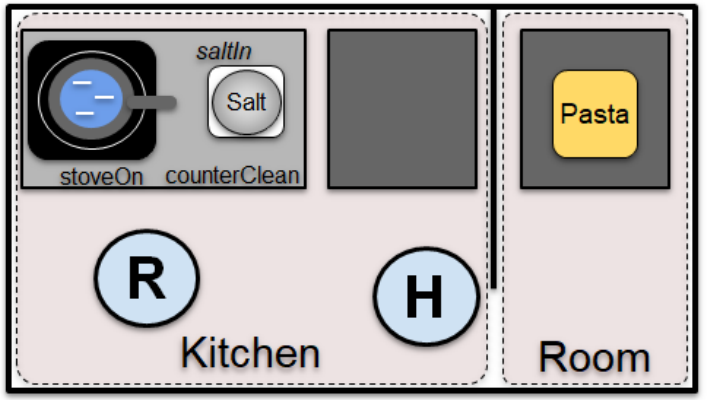


Pasta cooking shared task

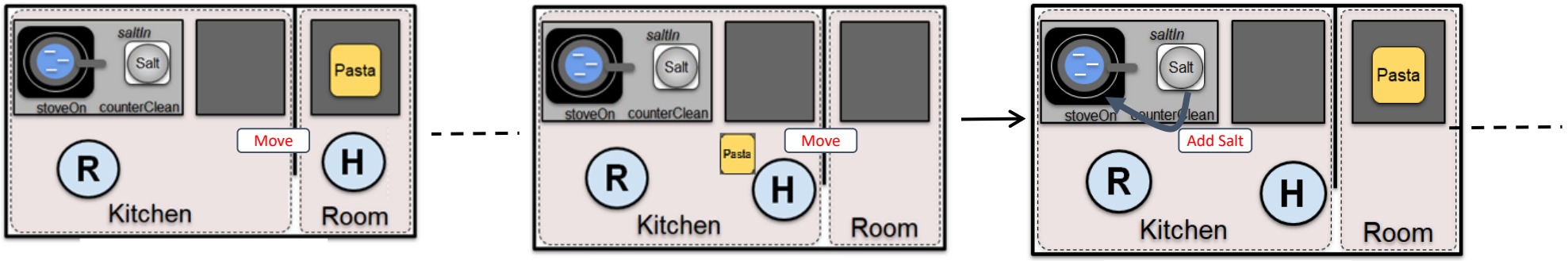


Illustrative Example: Human Choice 2

- **Consider the following scenario**
  - **Human** and **robot** team up to achieve a task together
    - For example: **cooking pasta**
  - **Human** needs to leave the *Kitchen*, but…
    - ***(Non-determinism)*** **Humans** cannot be controlled and **choose** when to leave for fetching the pasta packet
    - **Humans** are rational agents

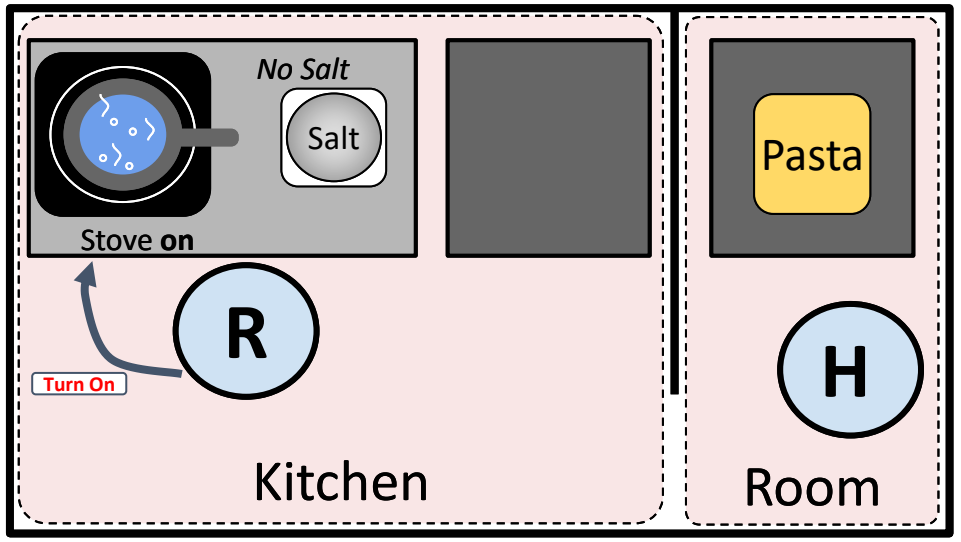


Pasta cooking shared task

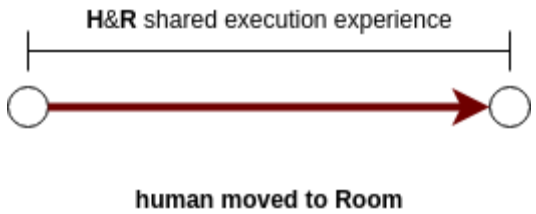


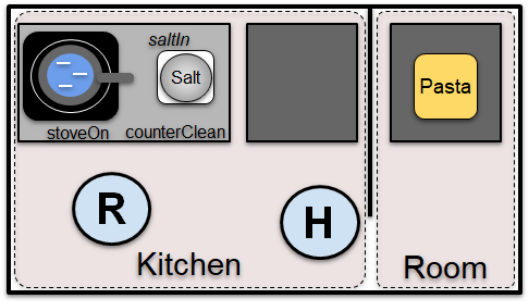
Illustrative Example: Suppose human picked the 2nd option

- **Consider the following scenario**
  - **Human** and **robot** team up to achieve a task together
    - For example: **cooking pasta**
  - **Human** needs to leave the *Kitchen*, but…
    - (***Non-determinism***) **Humans** cannot be controlled and **choose** when to leave for fetching the pasta packet
    - **Humans** are rational agents
  - Meanwhile, **robot** can progress in their absence towards achieving the task
    - **Humans** missed the real progress achieved by **robot**
    - **Humans** can *estimate* the set of worlds they expect to see one of those when back
      - *We consider that humans carry an estimated model of the robot.*



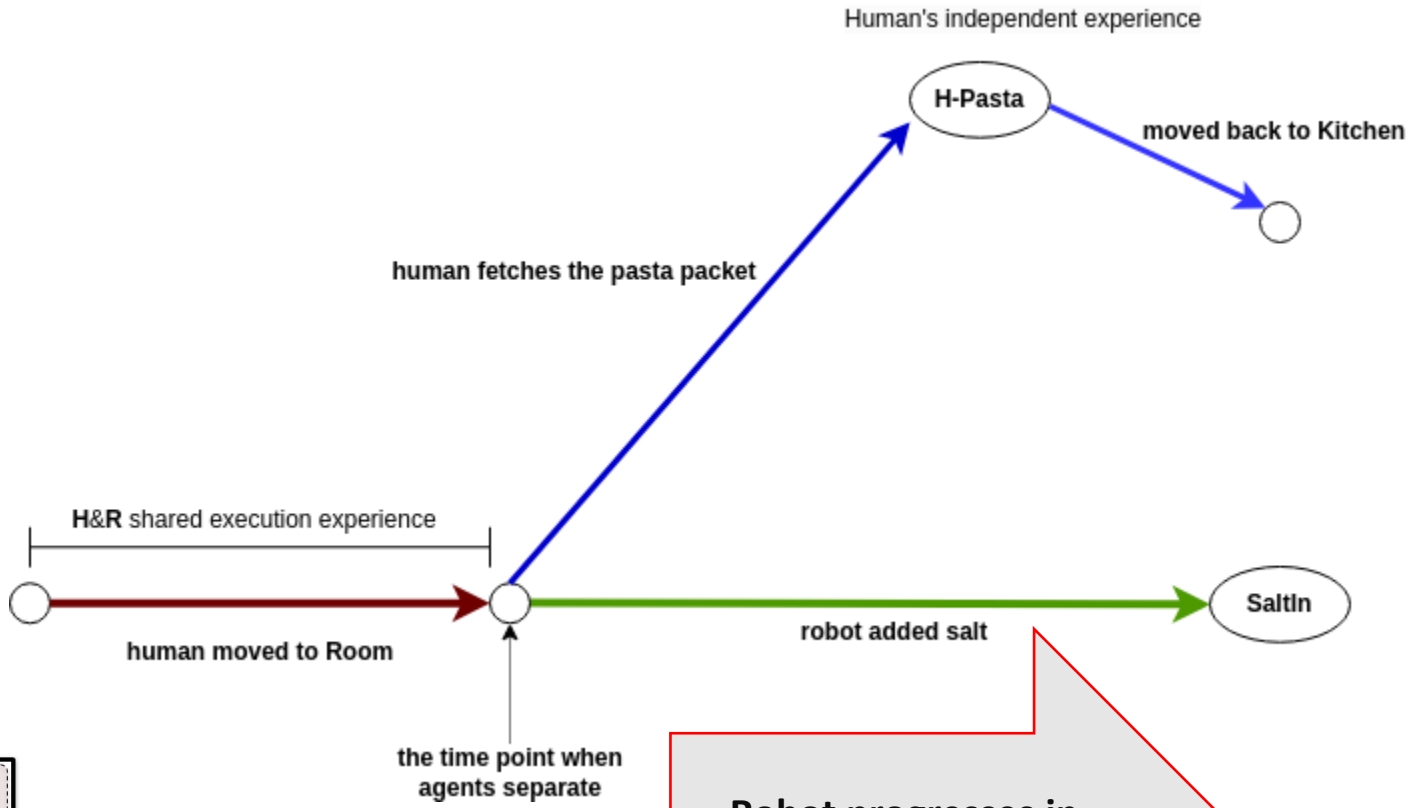
Let us consider a single plan trace to understand the (non-) shared execution experiences

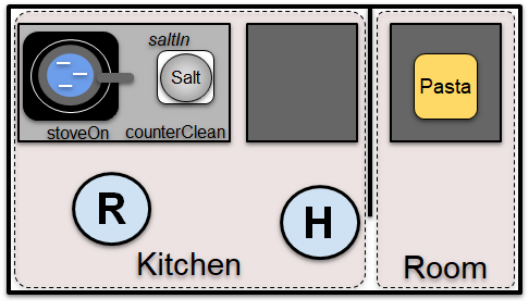




Pasta cooking shared task

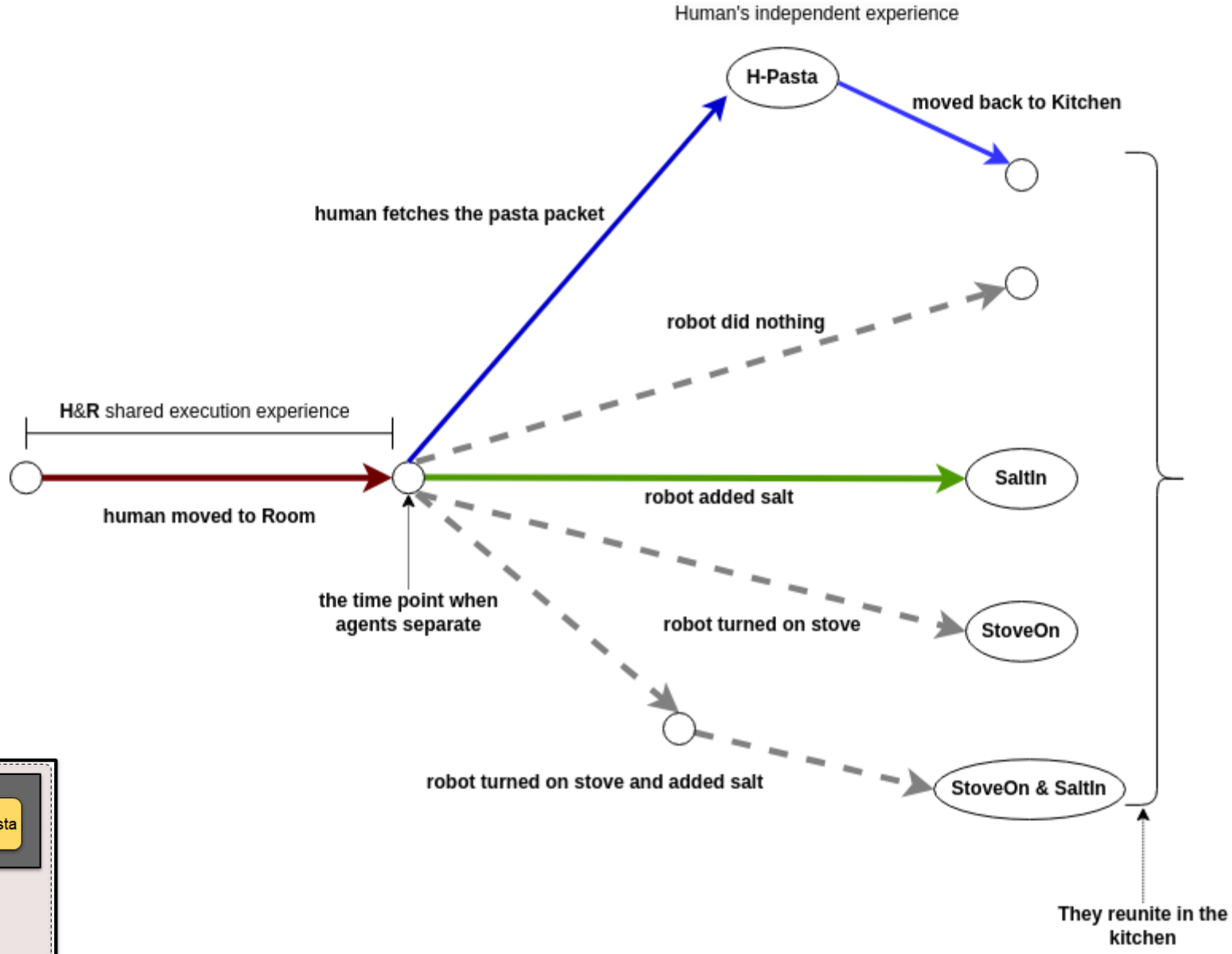
By the time human fetches pasta, robot adds salt to the pan

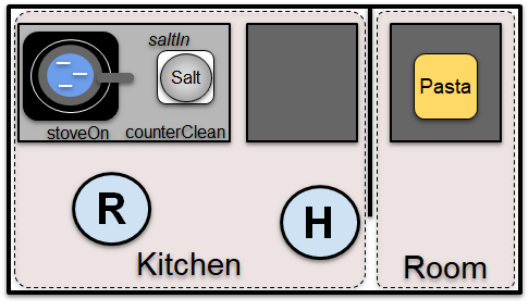




Pasta cooking shared task

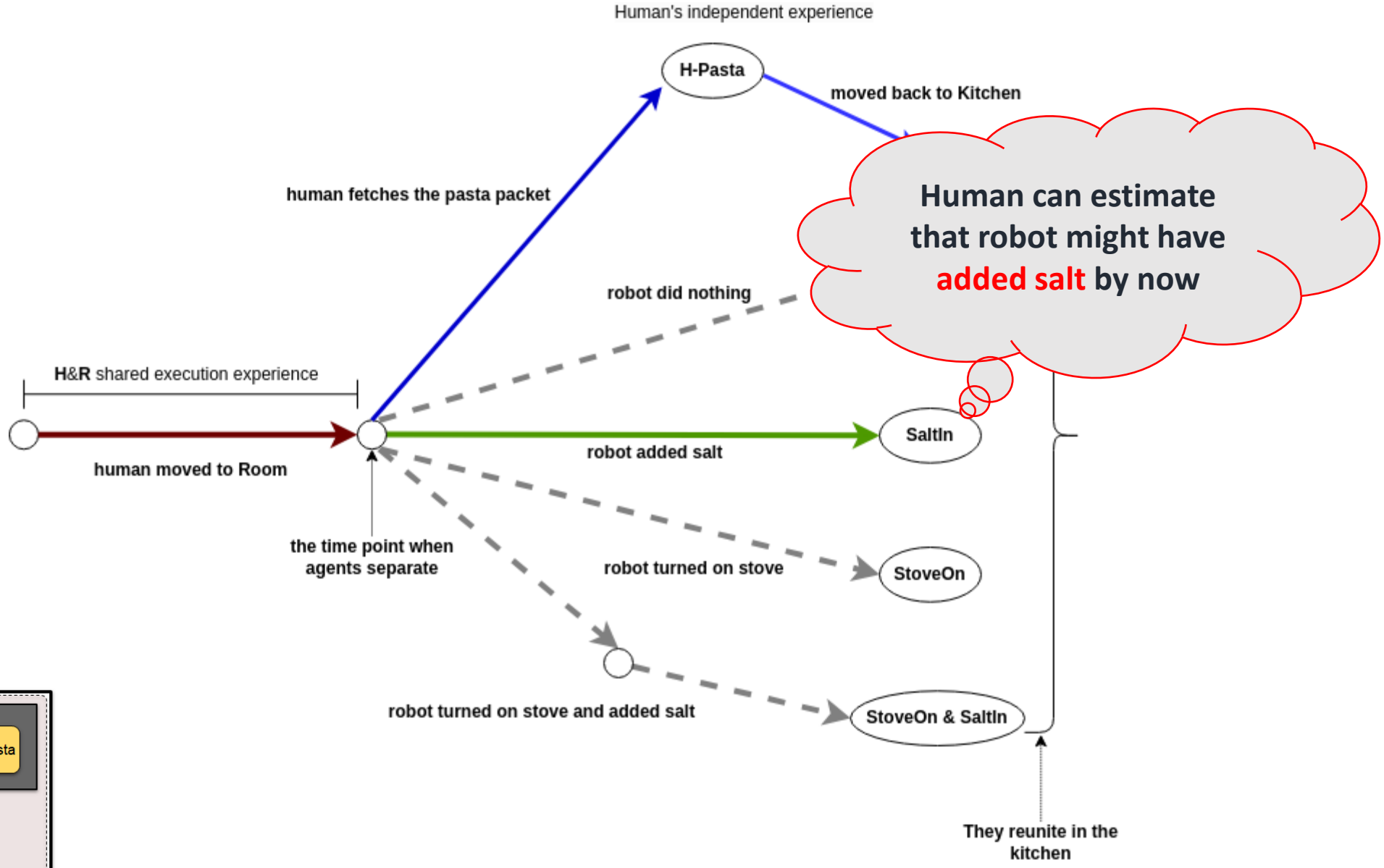
Anticipating Potential Progress

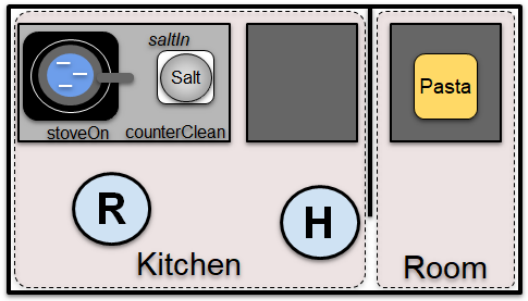




Pasta cooking shared task

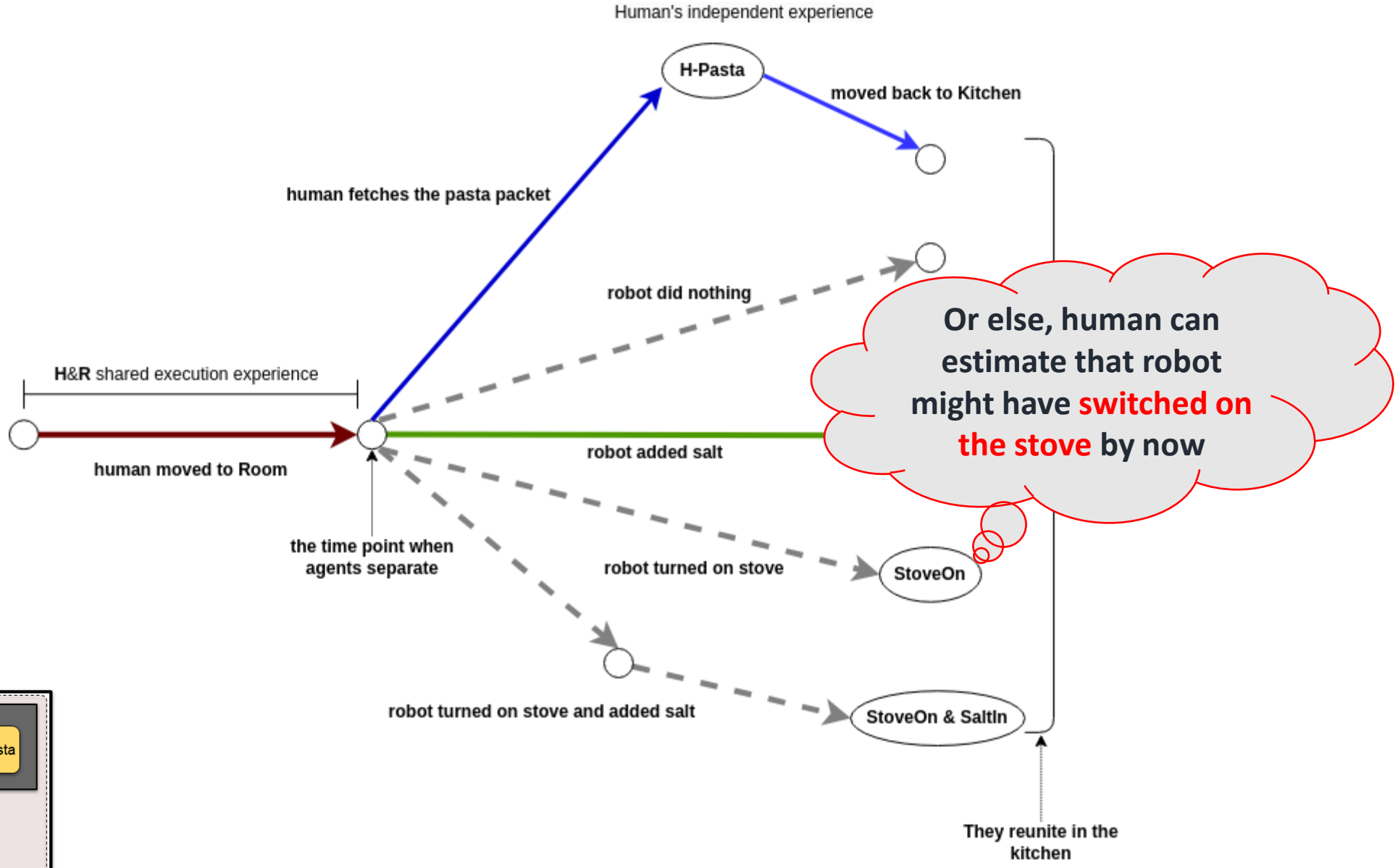
Anticipating Potential Progress

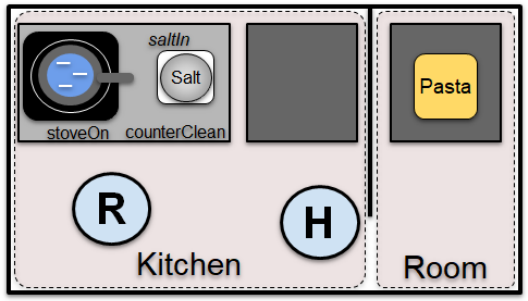




Pasta cooking shared task

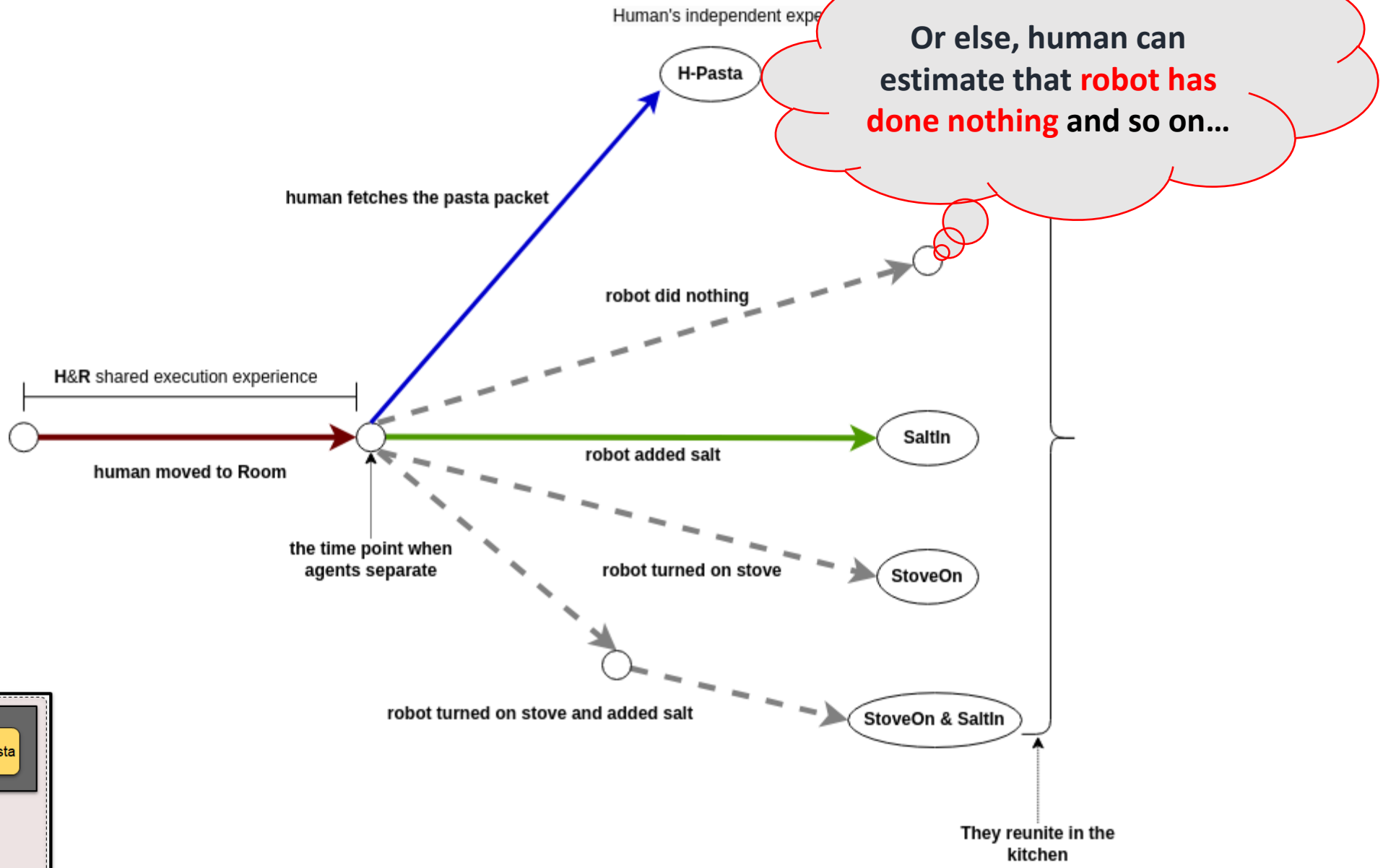
Anticipating Potential Progress

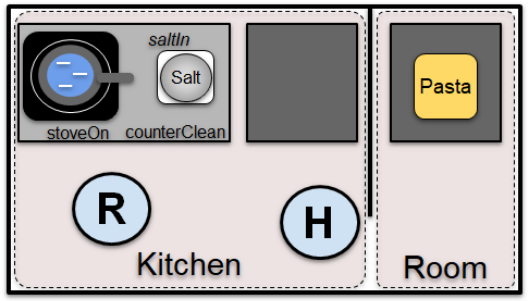




Pasta cooking shared task

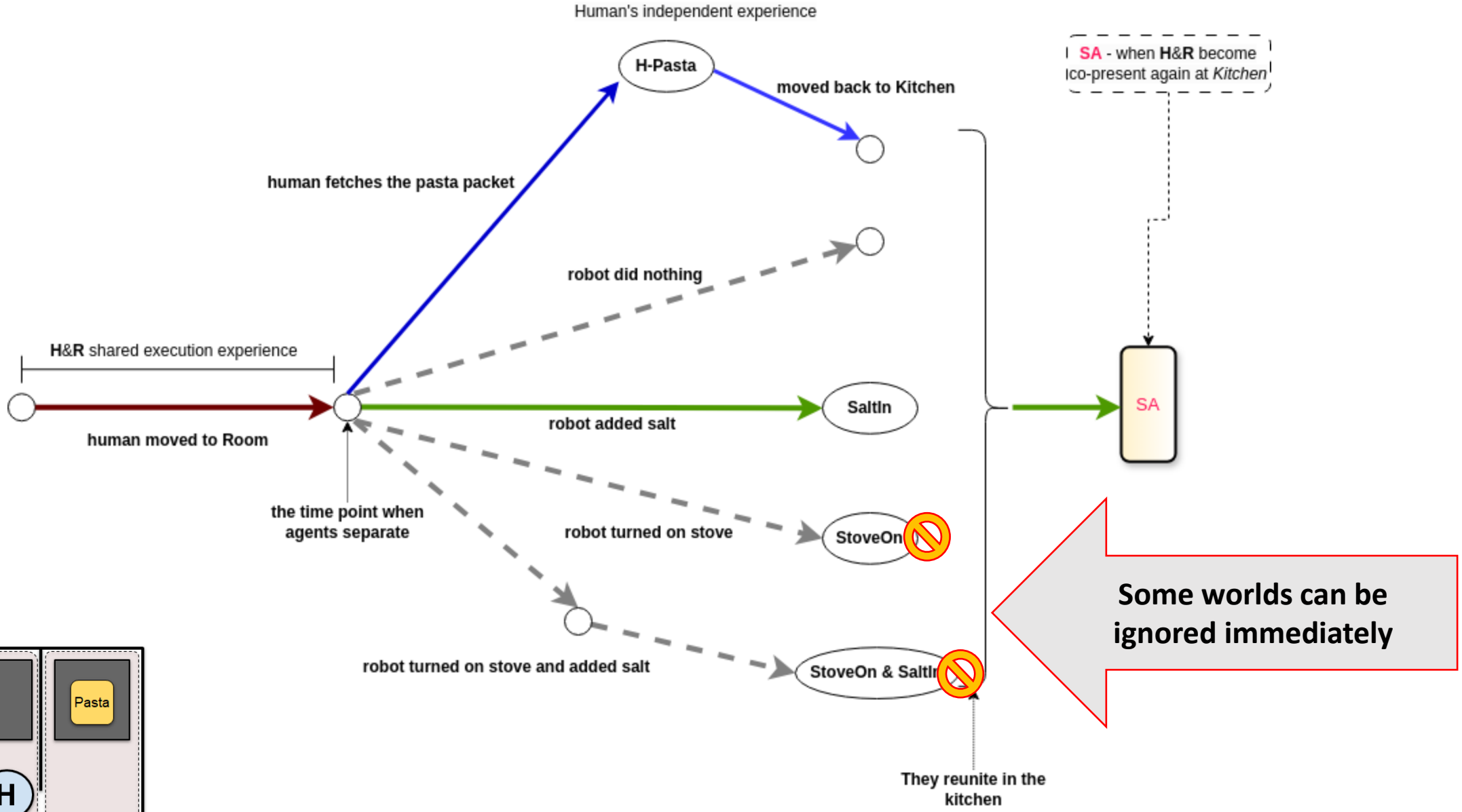
Anticipating Potential Progress

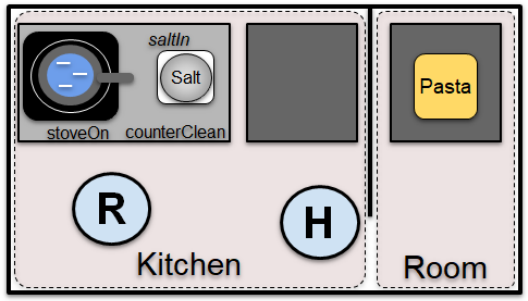




Pasta cooking shared task

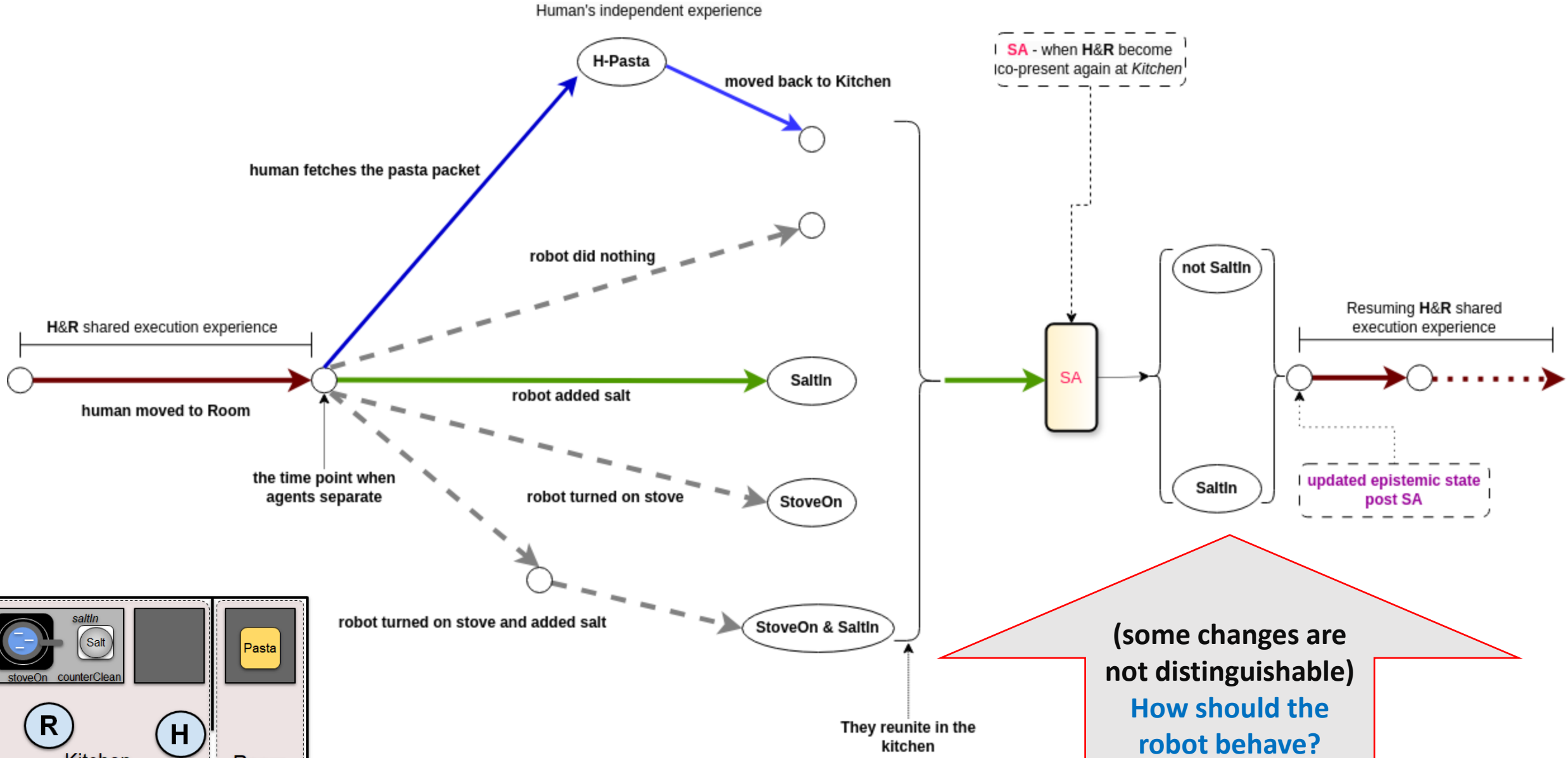
Human Can Assess Changes





Pasta cooking shared task

Human Can Assess Changes



Pasta cooking shared task

Human Aware Epistemic Task planner

- **<u>Our Objectives</u>**
  - Building **robot's** policy that accounts for **human** uncontrollable behaviors

Human Aware Epistemic Task planner

- **<u>Our Objectives</u>**
  - Building **robot's** policy that accounts for **human** uncontrollable behaviors
  - Also, keeping track of potential advancements the **robot** can achieve from the **human's perspective** – that human can estimate to be possible
    - (This is effective when the execution experience is not shared)

Human Aware Epistemic Task planner

- **<u>Our Objectives</u>**
  - Building **robot's** policy that accounts for **human** uncontrollable behaviors
  - Also, keeping track of potential advancements the **robot** can achieve from the **human's perspective** – that human can estimate to be possible
    - (This is effective when the execution experience is not shared)
  - The **robot's** policy has appropriate course of action depending on different situations arise:
    - Responding to the human enquiry, *e.g., whether the salt is added?*
    - Robot communicates, *e.g., salt is added*
    - Robot's (ontic) action, *e.g., adds salt in the presence of human*

Human Aware Epistemic Task planner

- **Consideration**
  - We consider that human has an estimated robot's model (**HuMM**)
    - This is effective in managing potential advancements
  - To plan **robot's actions** the planner uses robot model (RM)
    - Note that **RM ≠ HuMM (disparate model)**
  - And to anticipate and estimate possible **human actions**, human model (HM) is used

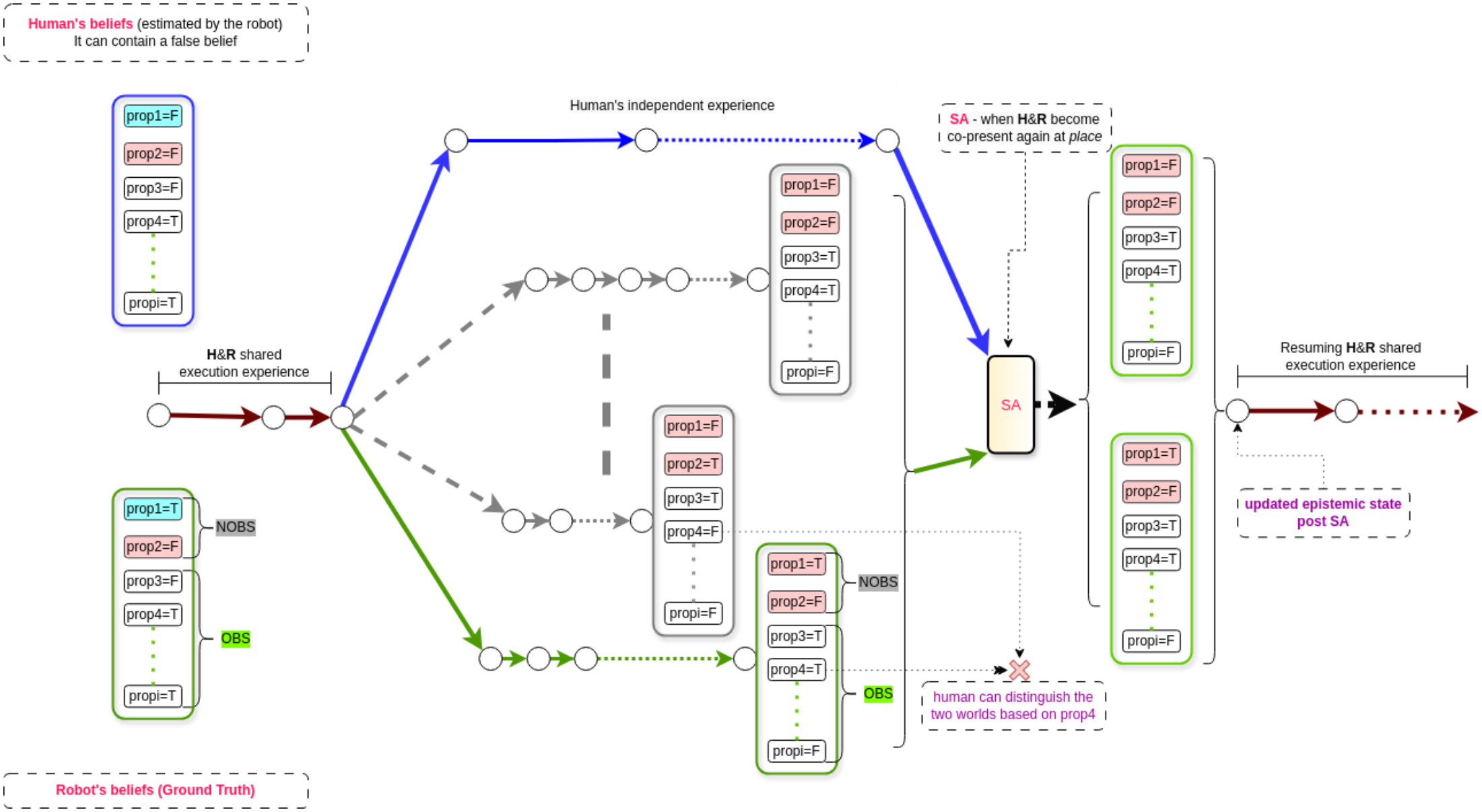
Human Aware Epistemic Task planner

- **Consideration**
  - We consider that human has an estimated robot's model (**HuMM**)
    - This is effective in managing potential advancements
  - To plan **robot's actions** the planner uses robot model (RM)
    - Note that **RM ≠ HuMM (disparate model)**
  - And to anticipate and estimate possible **human actions**, human model (HM) is used

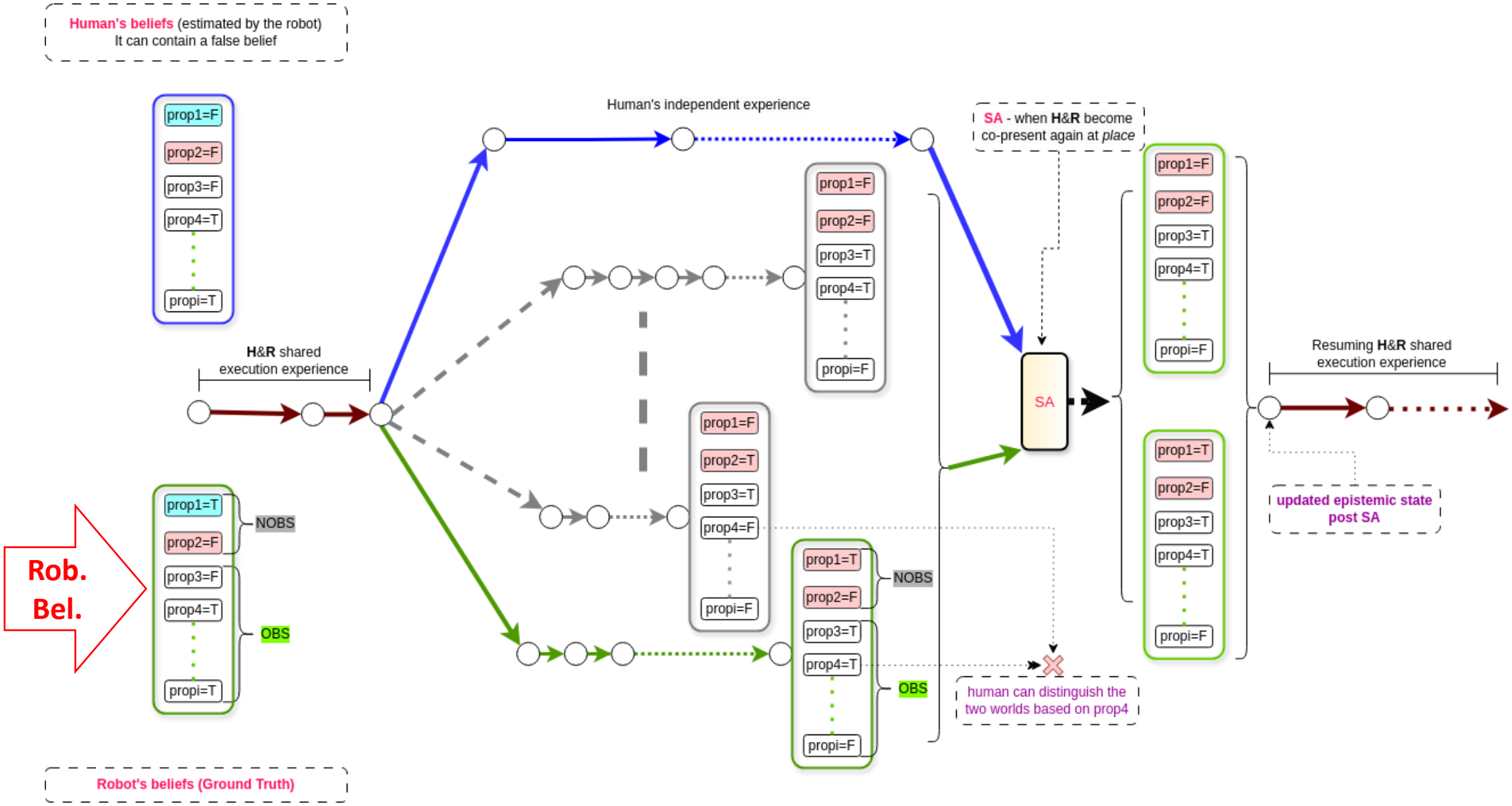
- **Today's focus**
  - The framework we developed is general and works for **RM ≠ HuMM**
  - But to simplify the presentation, we assume that **RM = HuMM**

Modeling Aspects

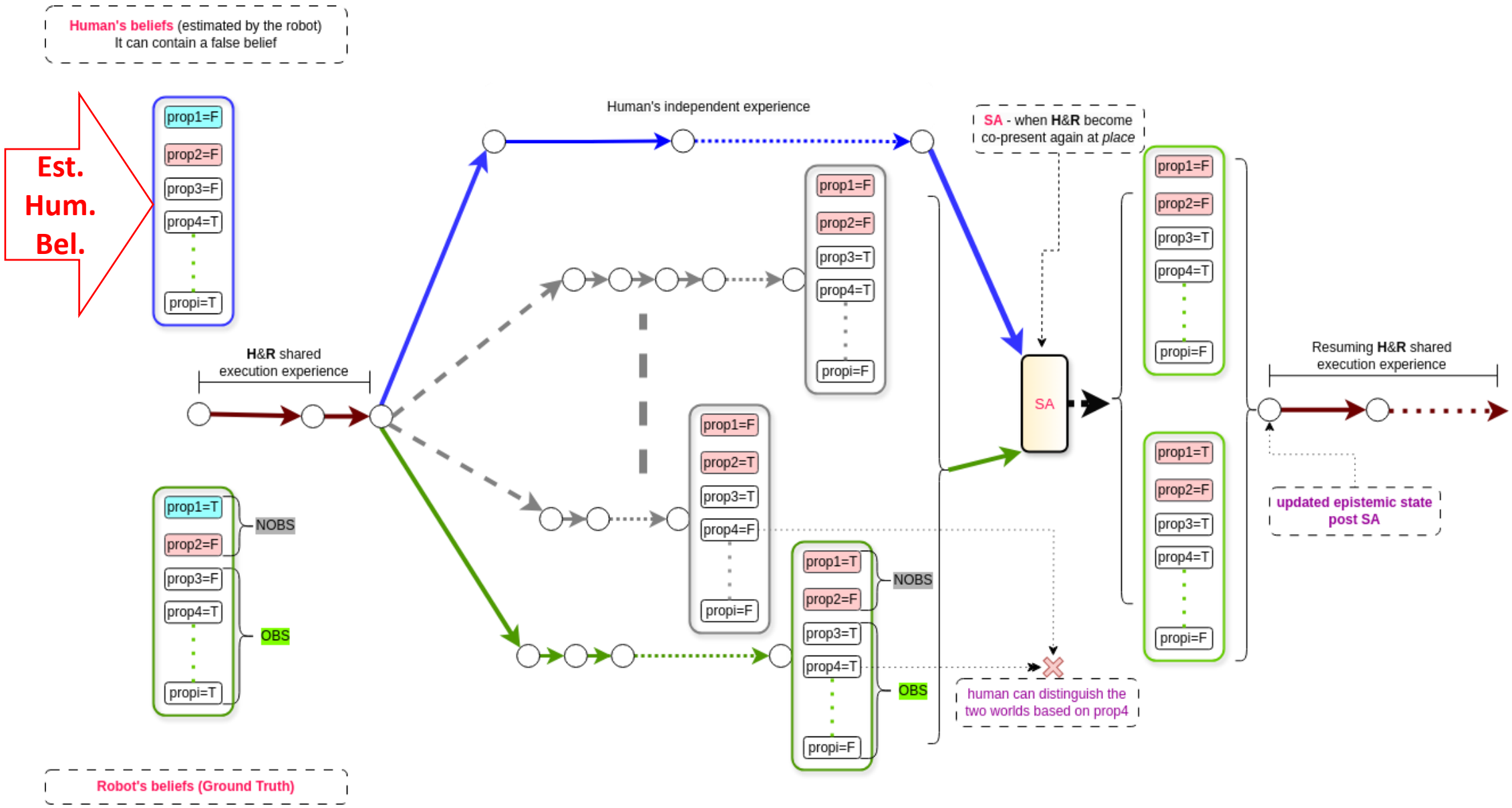
Modeling aspects of our planning framework



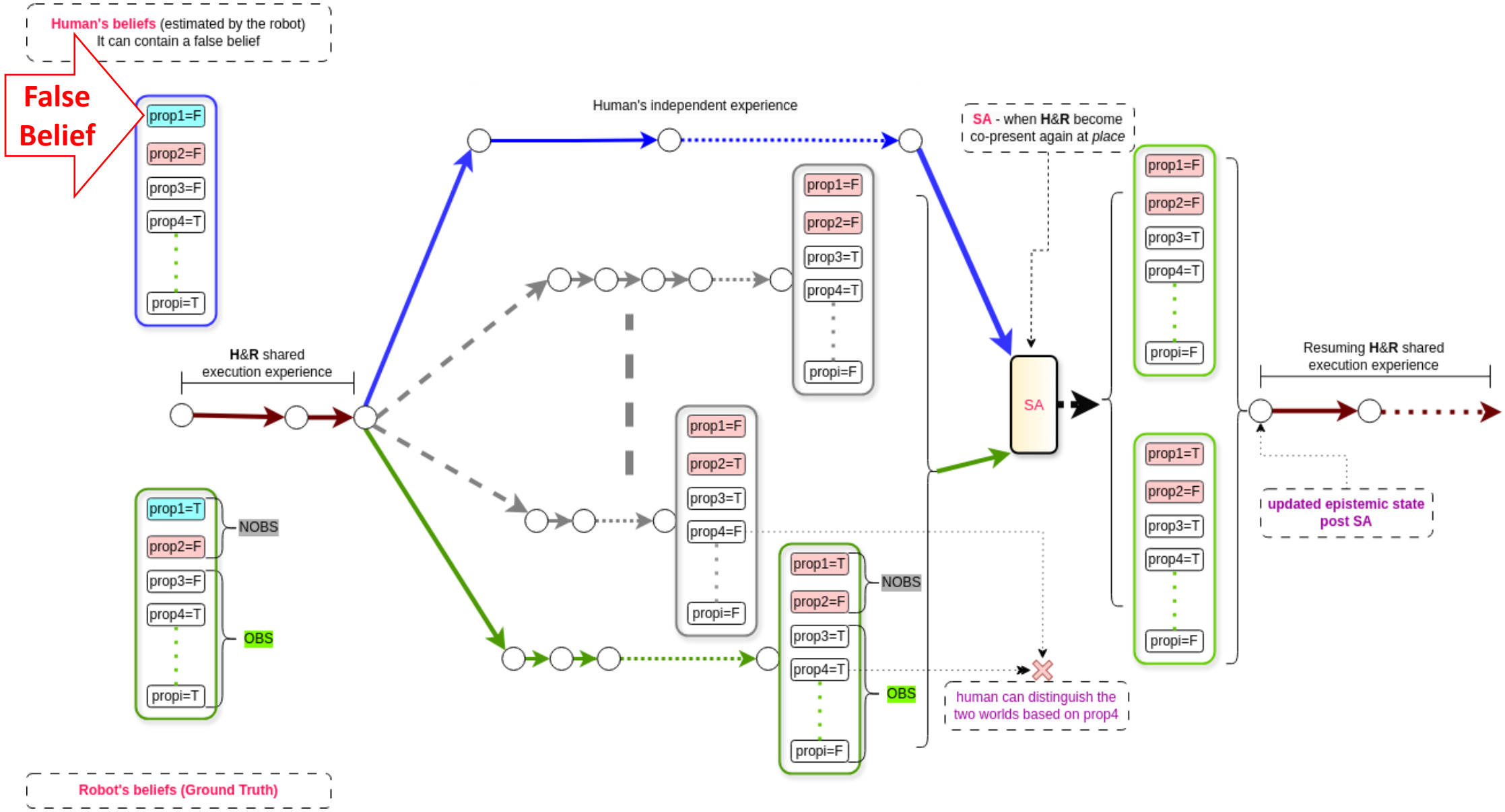
Modeling aspects: **Agents' Beliefs**



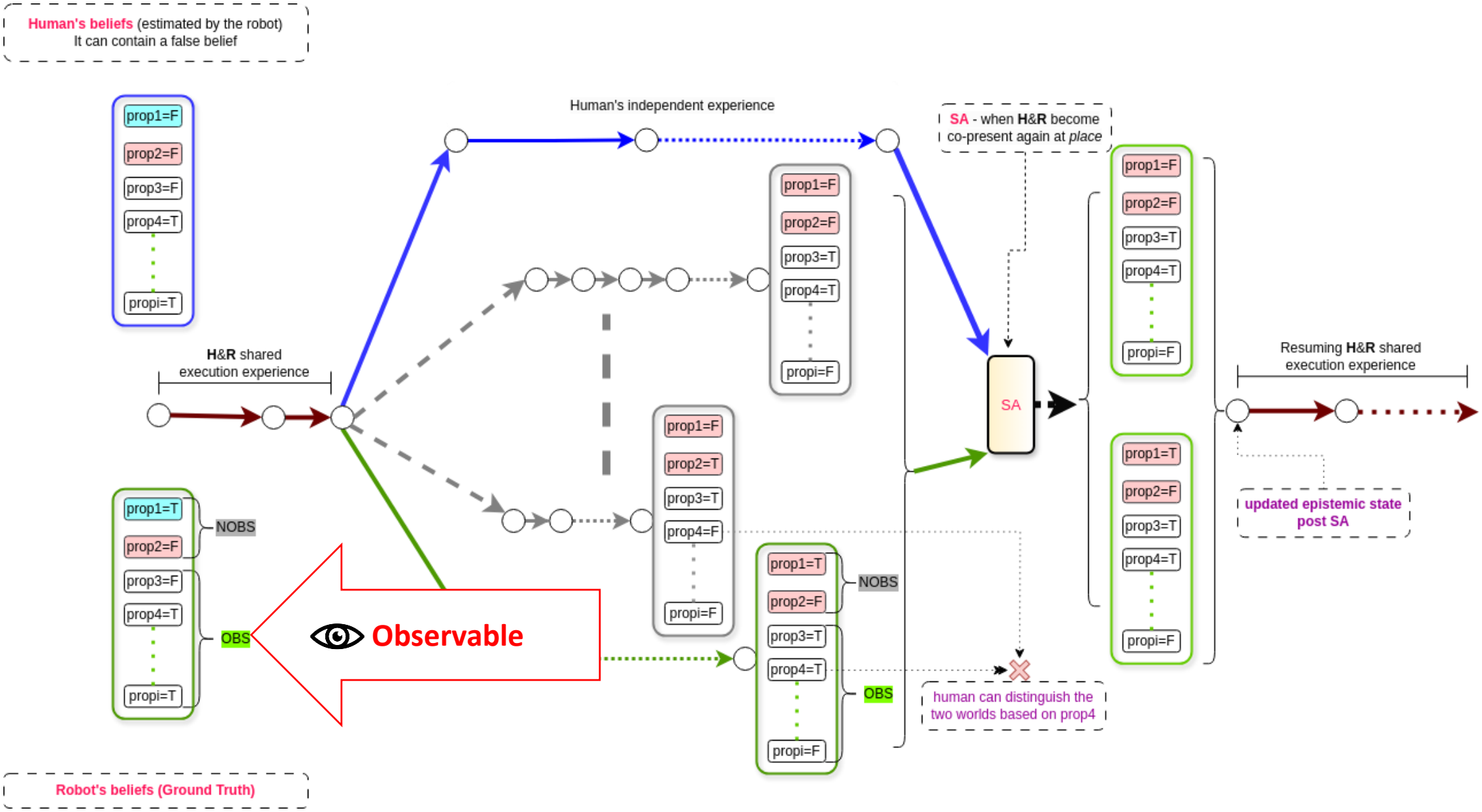
Modeling aspects: **Agents' Beliefs**



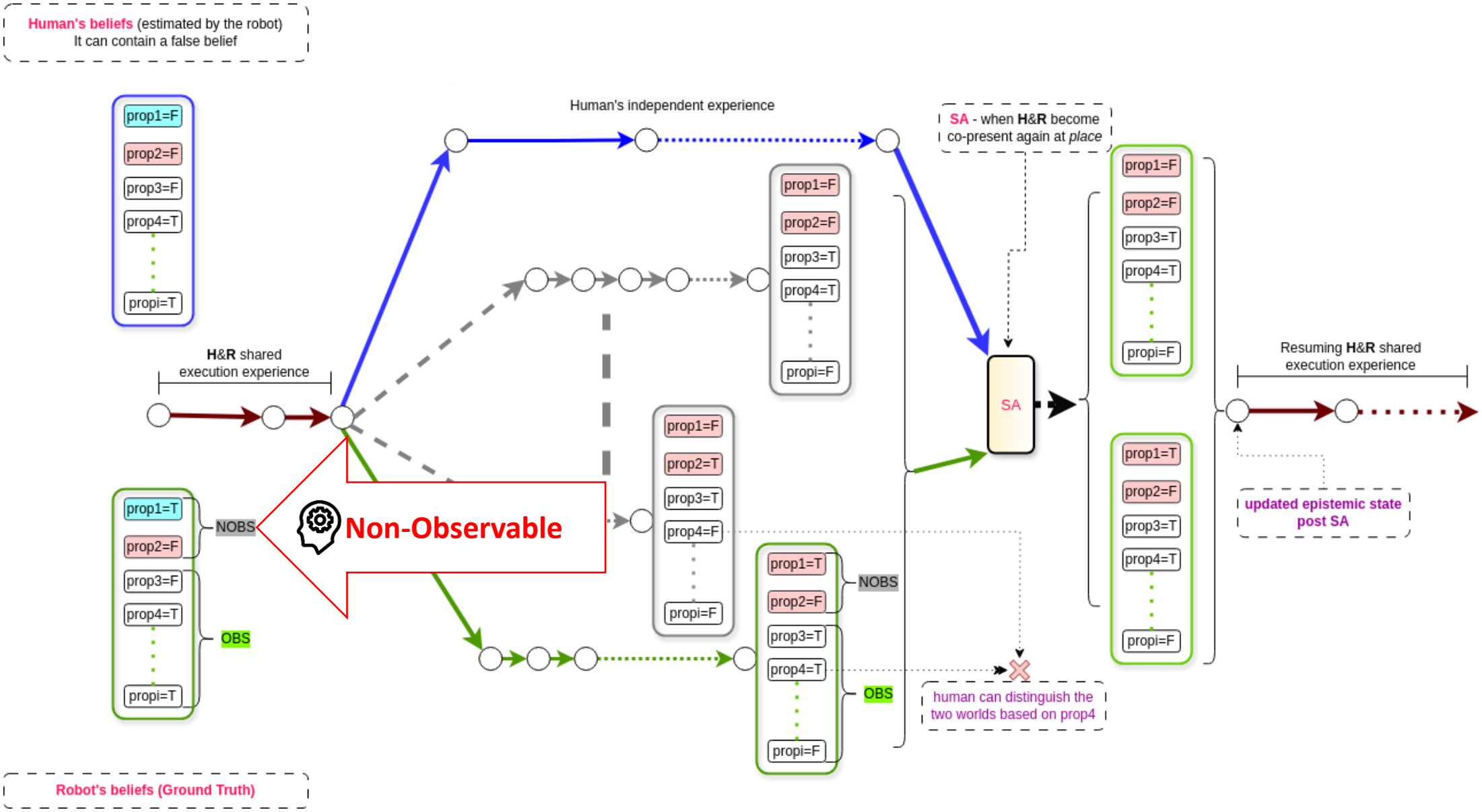
Modeling aspects: **Agents' Beliefs**



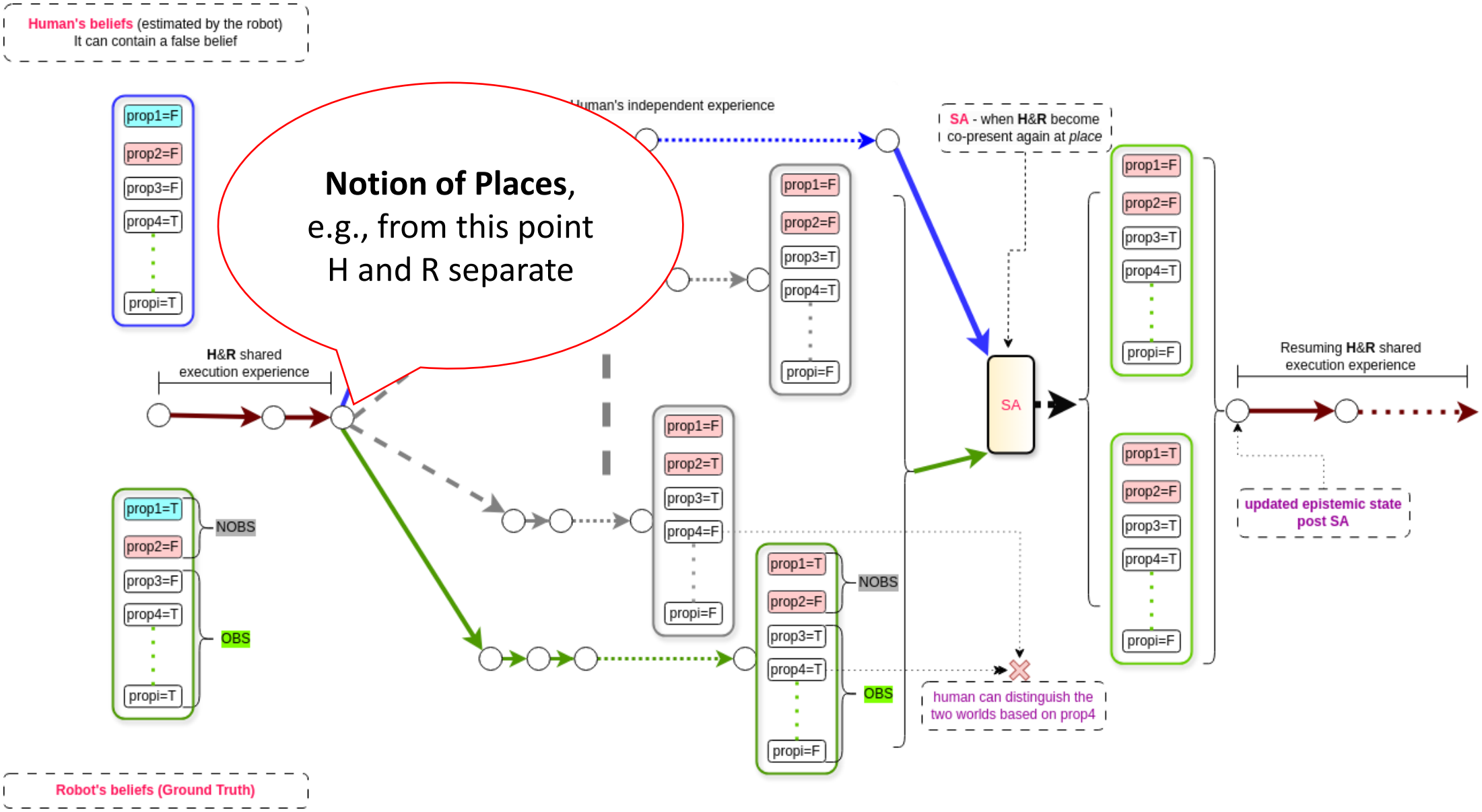
Modeling aspects: **(Non-) Observable Variables**



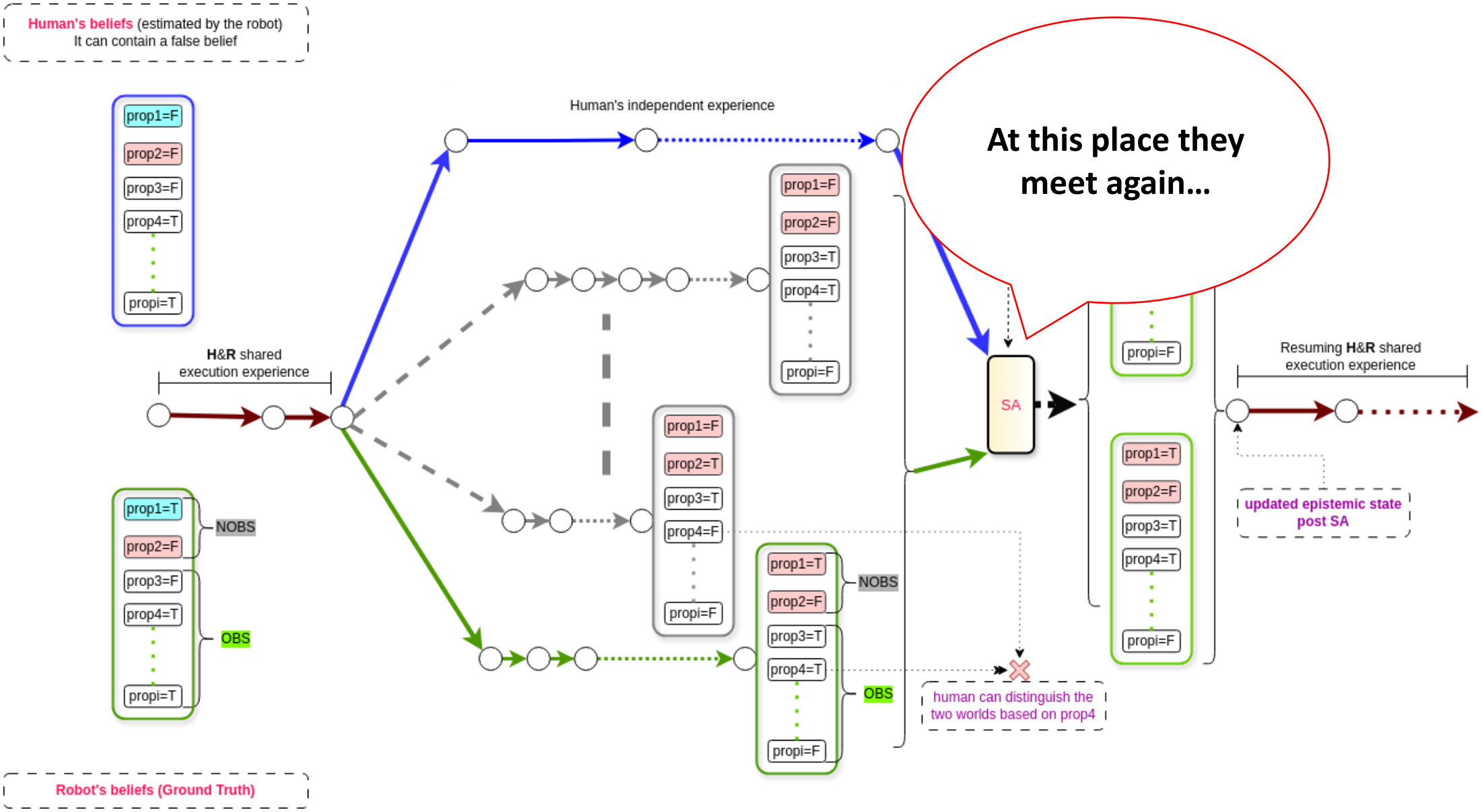
Modeling aspects: **(Non-) Observable Variables**



Modeling aspects: **Notion of Places**



Modeling aspects: **Notion of Places**



Modeling aspects: **Shared Experience**

- **An agent can observe an action execution:**
  - When **observer** is also the **actor**, and
  - If the observer **shares** the same **place** as of the actor when the action is executed (in ***place***)
    - In that case, the actor and observer are also said to be **co-present**

Modeling aspects: **Shared Experience**

- **An agent can observe an action execution:**
  - When **observer** is also the **actor**, and
  - If the observer **shares** the same **place** as of the actor when the action is executed (in ***place***)
    - In that case, the actor and observer are also said to be **co-present**
- **Action's effects:**
  - Effects will be observed by the **actor** and all the **co-present observers**
  - Moreover, changes w.r.t. observable variables can also be observed later when an agent enters **place**
  - But non-co-present agents **cannot assess** the changes w.r.t. non-observable variables
  - Accordingly, agents' beliefs are updated!

Modeling aspects: Situation Assessment

- To systematically manage the **evolution** and **contraction** of estimated beliefs of the agents:
  - our framework implements a **situation assessment** process
  - it utilizes models for co-presence, (non-) observable variables, etc.

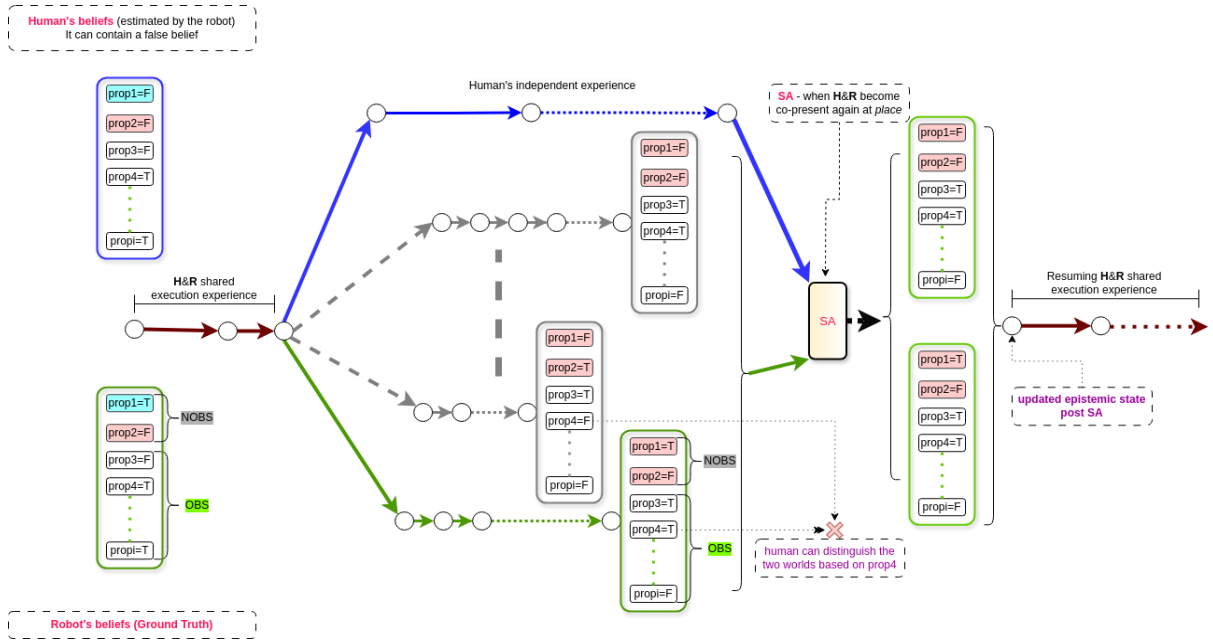
Our Planning Framework

(leverages tools developed for DEL-based epistemic planning)

Our Planning Approach (roughly…)

- An epistemic state (a set of tuples):

  $s = \{..., <RM,HM,HuMM>, ...\}$

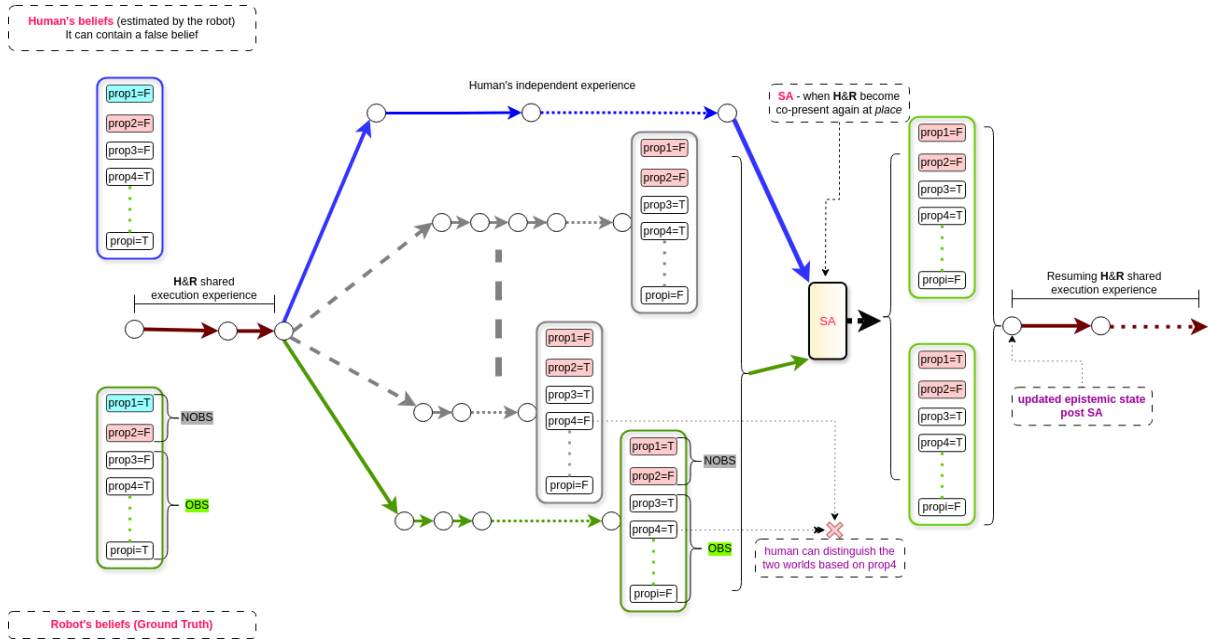


Our Planning Approach (roughly…)

- An epistemic state (a set of tuples):

  $s = \{..., <RM,HM,HuMM>, ...\}$

- **(ROBOT)** an epistemic action captures all possible events w.r.t. every $HuMM$ in s
  - One of these events is the **real R-action** that the robot performs (aligns with RM)

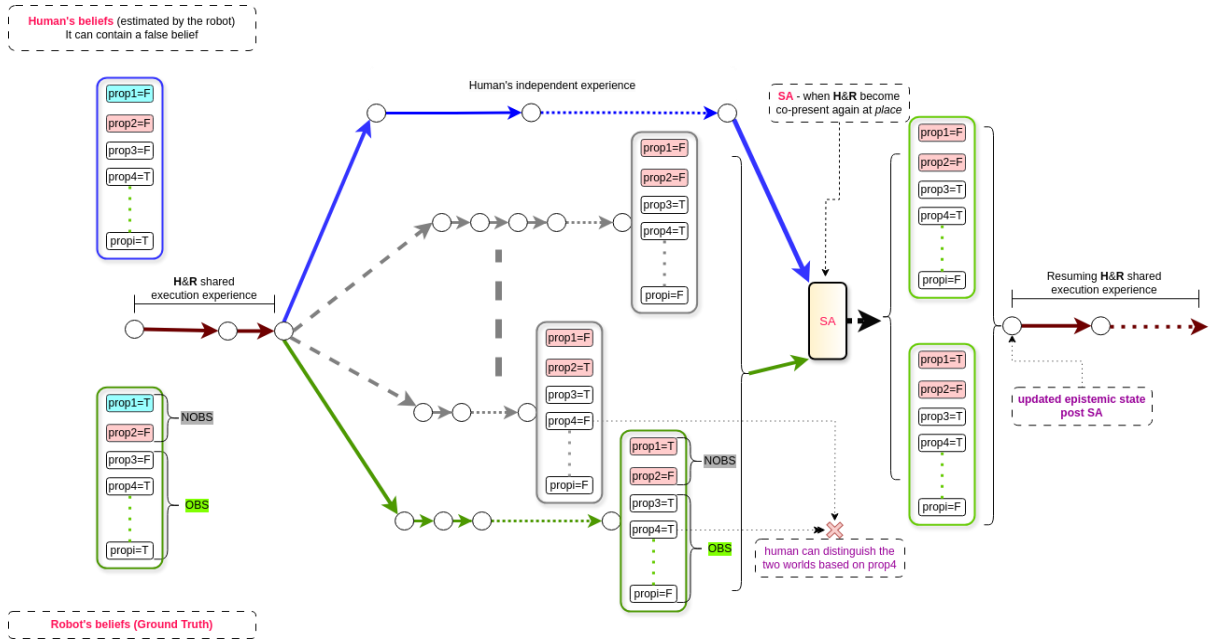


Our Planning Approach (roughly…)

- An epistemic state (a set of tuples):

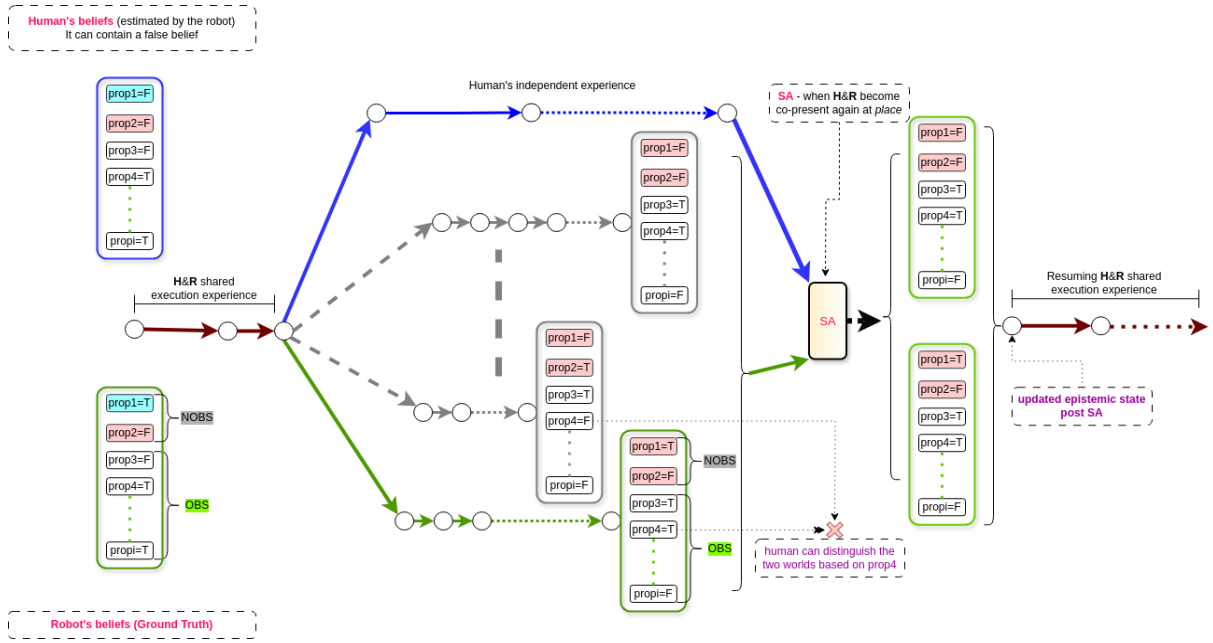
  s = {…, <RM,HM,HuMM>, …}

- **(ROBOT)** an epistemic action captures all possible events w.r.t. every HuMM
  - One of these events is the **real R-action** that the robot performs (aligns with RM)
  - The *indistinguishability* relation w.r.t. HUMAN is maintained as per co-present-HR



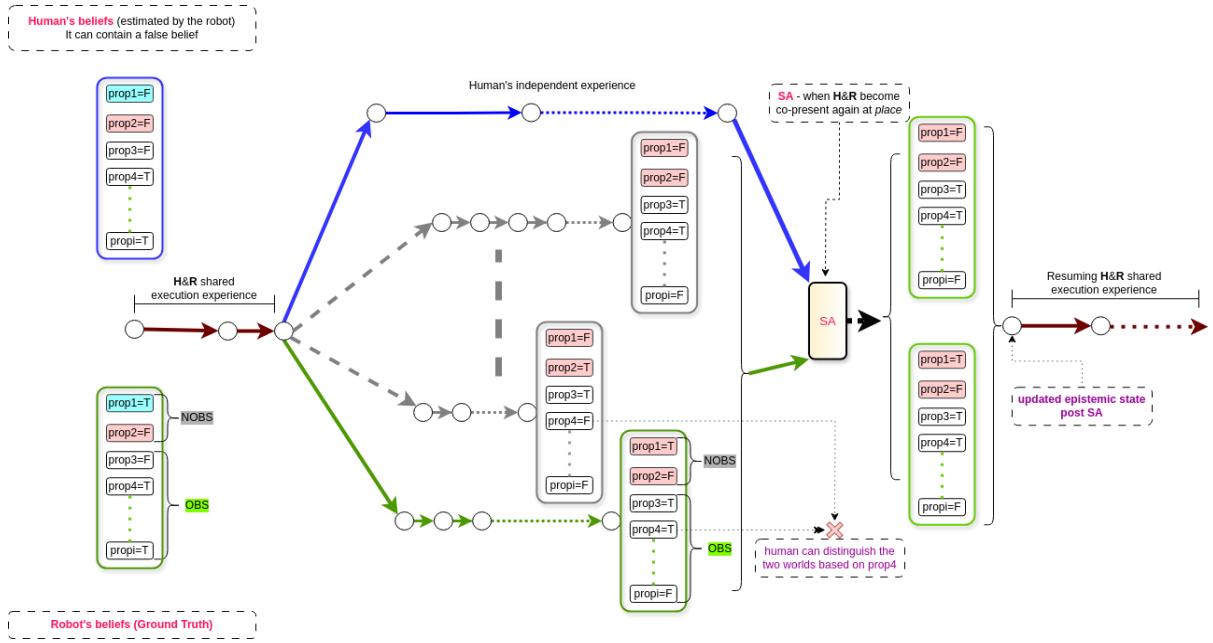
Our Planning Approach (roughly…)

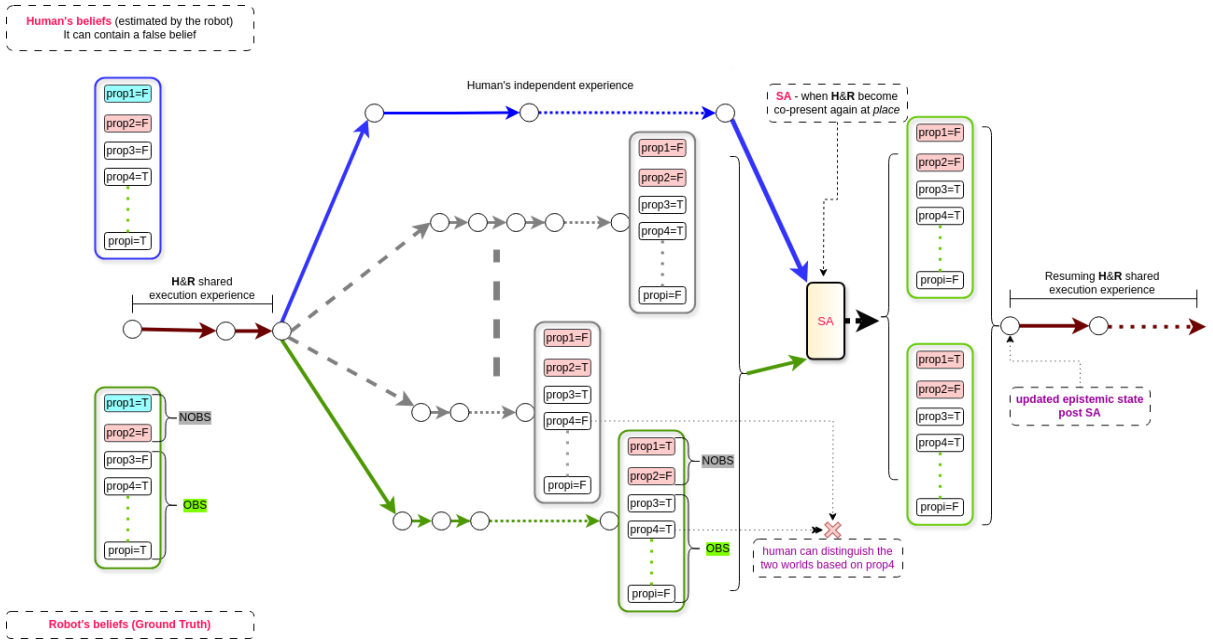
- **(HUMAN)** an epistemic action captures all possible events w.r.t. **HM**
  - The **real H-action** performed must be applicable w.r.t. each world (i.e., in **HM**) in the epistemic state



Our Planning Approach (roughly…)

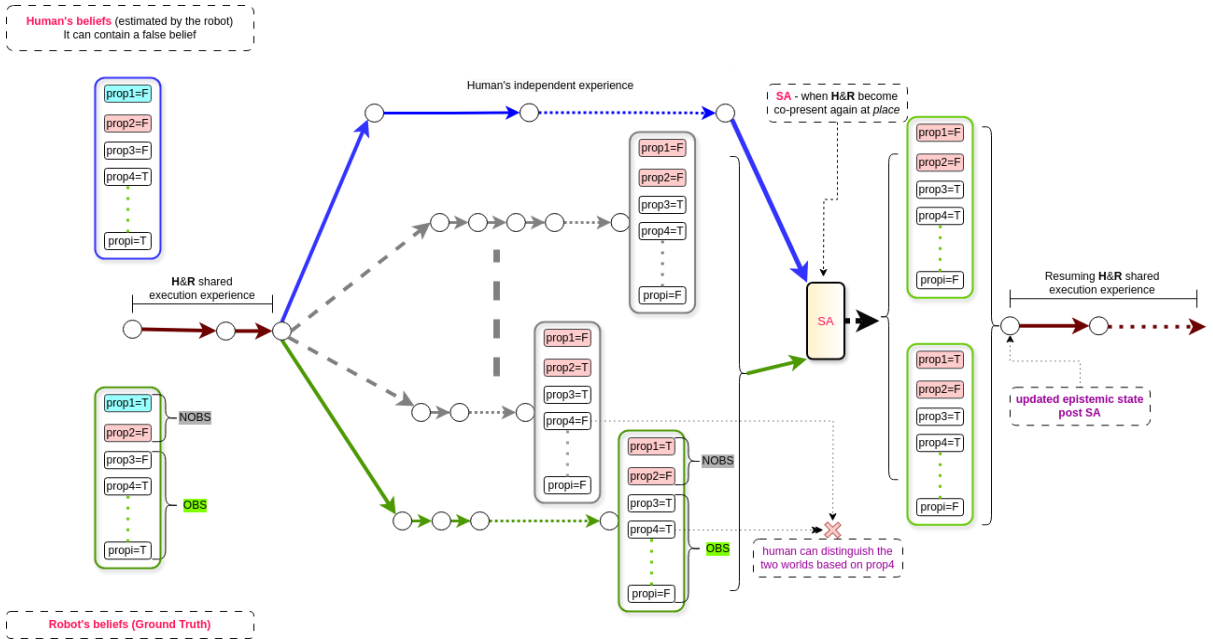
- **(HUMAN)** an epistemic action captures all possible events w.r.t. **HM**
  - The **real H-action** performed must be applicable w.r.t. each world (i.e., in **HM**) in the epistemic state
- We **apply** the epistemic action and generate the next epistemic state…
  - The size of this next state may grow (compared to the previous state)





Our Planning Approach (roughly…)

- Just after next state is generated:
  - We call the **situation assessment (SA)** subroutine that may shrink the overall possibilities

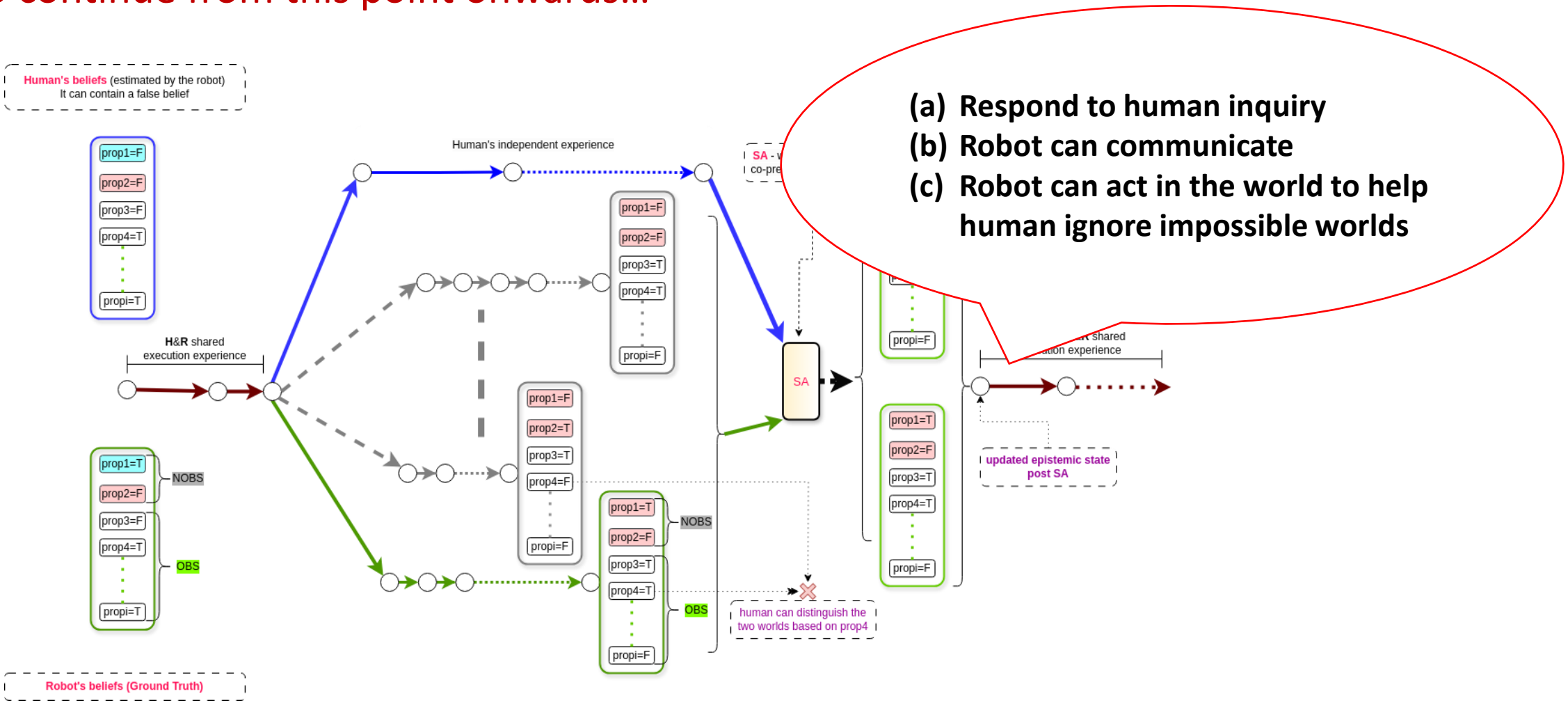


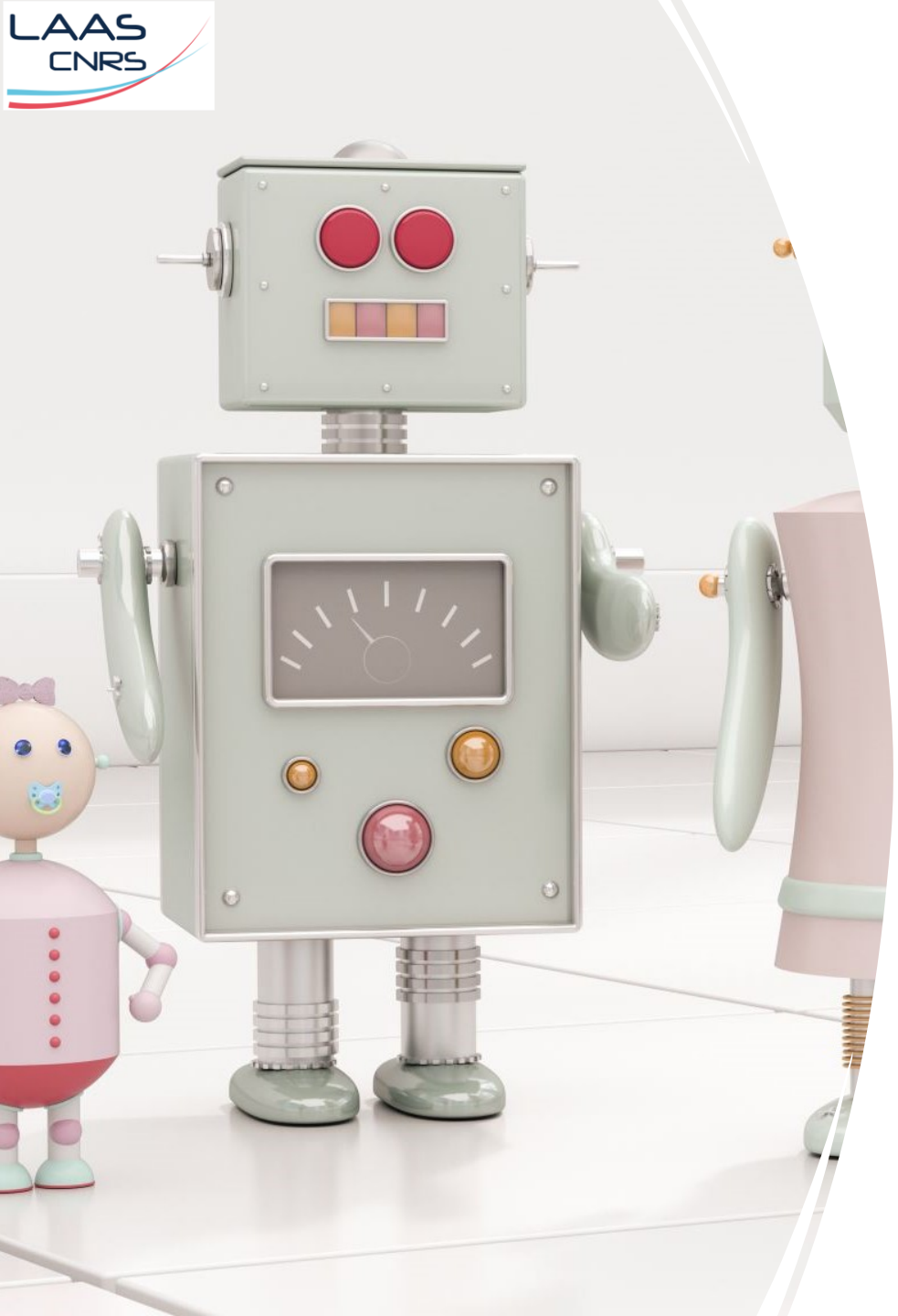
Our Planning Approach (roughly…)

- Just after next state is generated:
  - We call the **situation assessment (SA)** subroutine that may shrink the overall possibilities
  - Robot takes the perspective of human and ignores those possibilities that are impossible to be considered by humans

Human-Aware Epistemic Task Planner

To continue from this point onwards…





Human-Robot Communication

- About some state variable p
  - In an epistemic state s = {..., <RM,HM,HuMM>, ...}, **HUMAN** may not know the real value of p, but they always know that **ROBOT** knows the real world and real value of p
  - **HUMAN** can inquire about p (uncontrollable operator)
  - **ROBOT** can inform them the value of p if it optimizes the policy afterwards
  - **ROBOT** can also choose to act to implicitly communicate the value of p

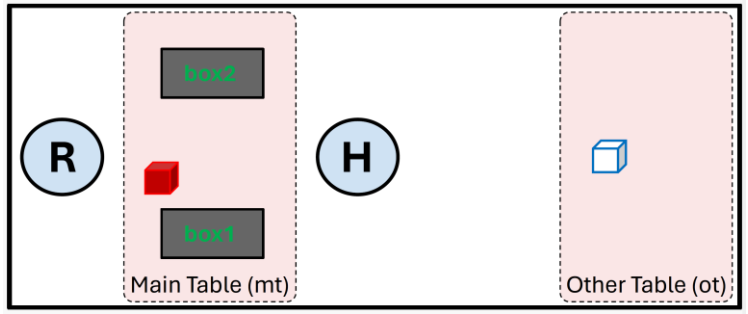
Implementation Details

- We build an **AND/OR** search based **planner** that performs breadth-first search
  - Its underlying idea is based on the HATP/EHDA planner (as described earlier)
- Our planner is implemented in **Python**

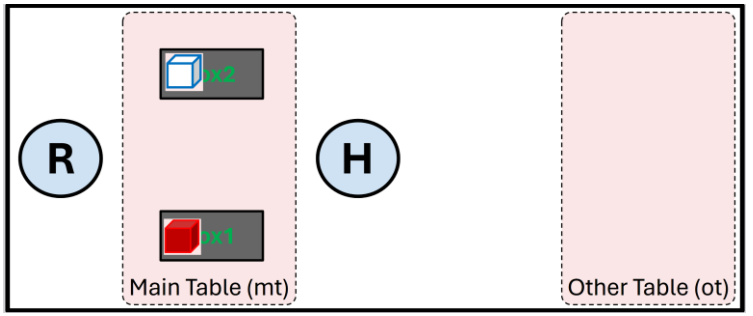
Empirical Evaluation

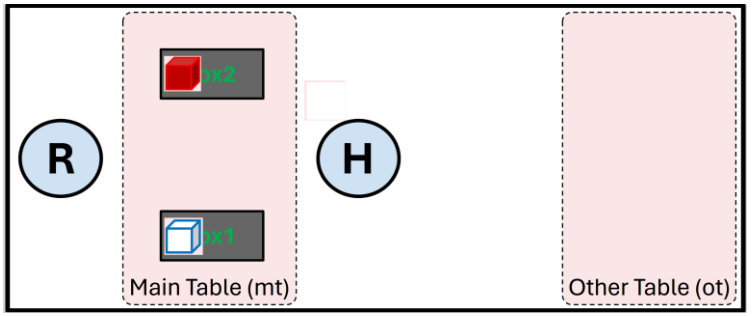
Cube Organization Scenario

- Two cubes: **cr (red)** and **cw** (white)
- **box1** and **box2** are placed on mt
  - They can be either ***transparent*** or ***opaque***.
- **Task** - organize the cubes in such a way that cubes from one table are placed in one box
- The choice of which box is flexible as long as each table's cubes end up in separate boxes



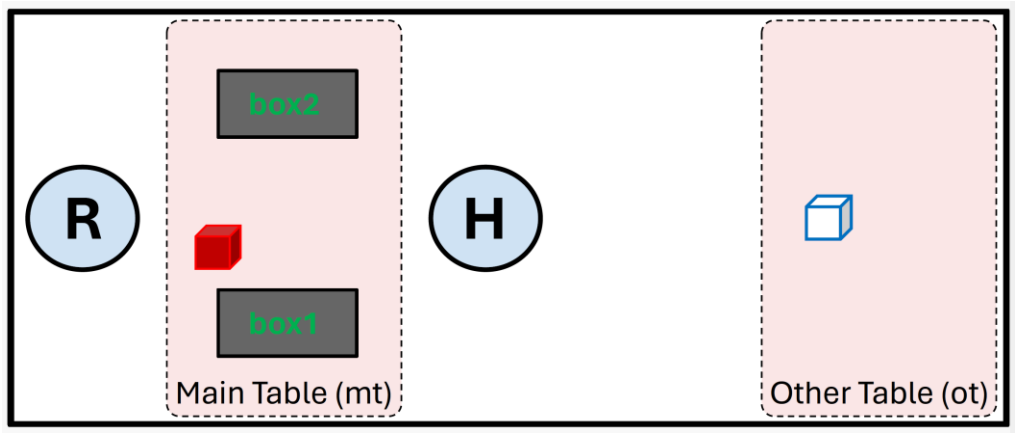
Initial state

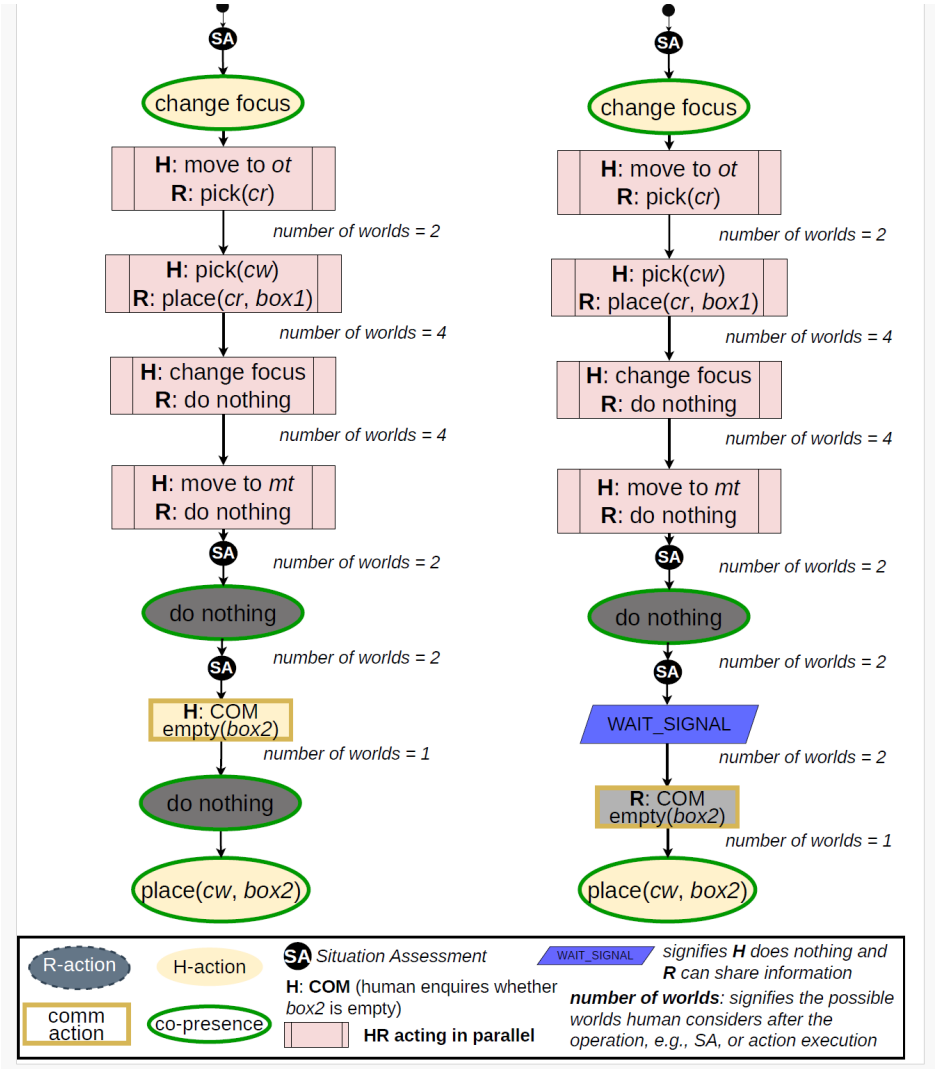




States in which the goal is satisfied

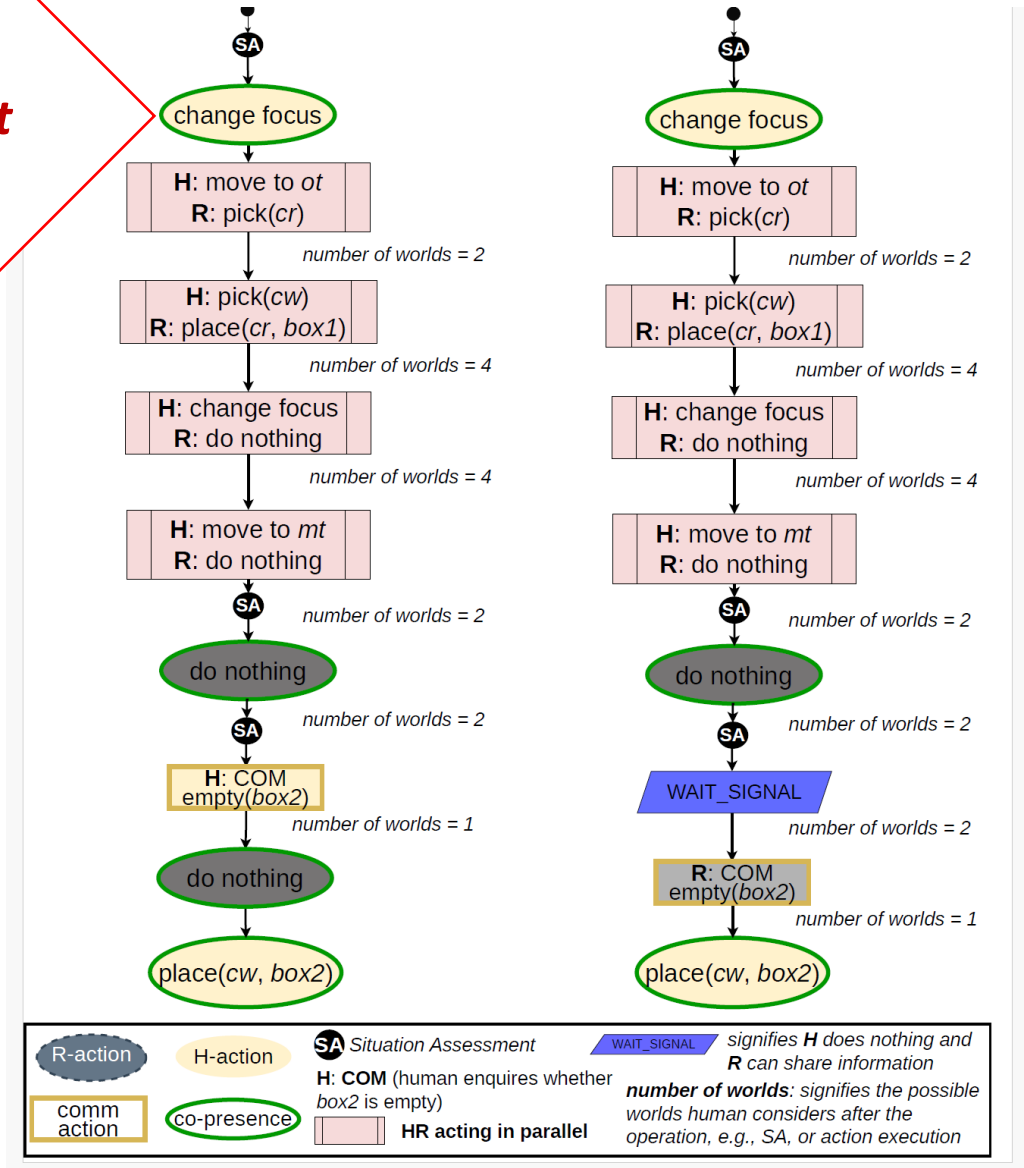
When the boxes are opaque



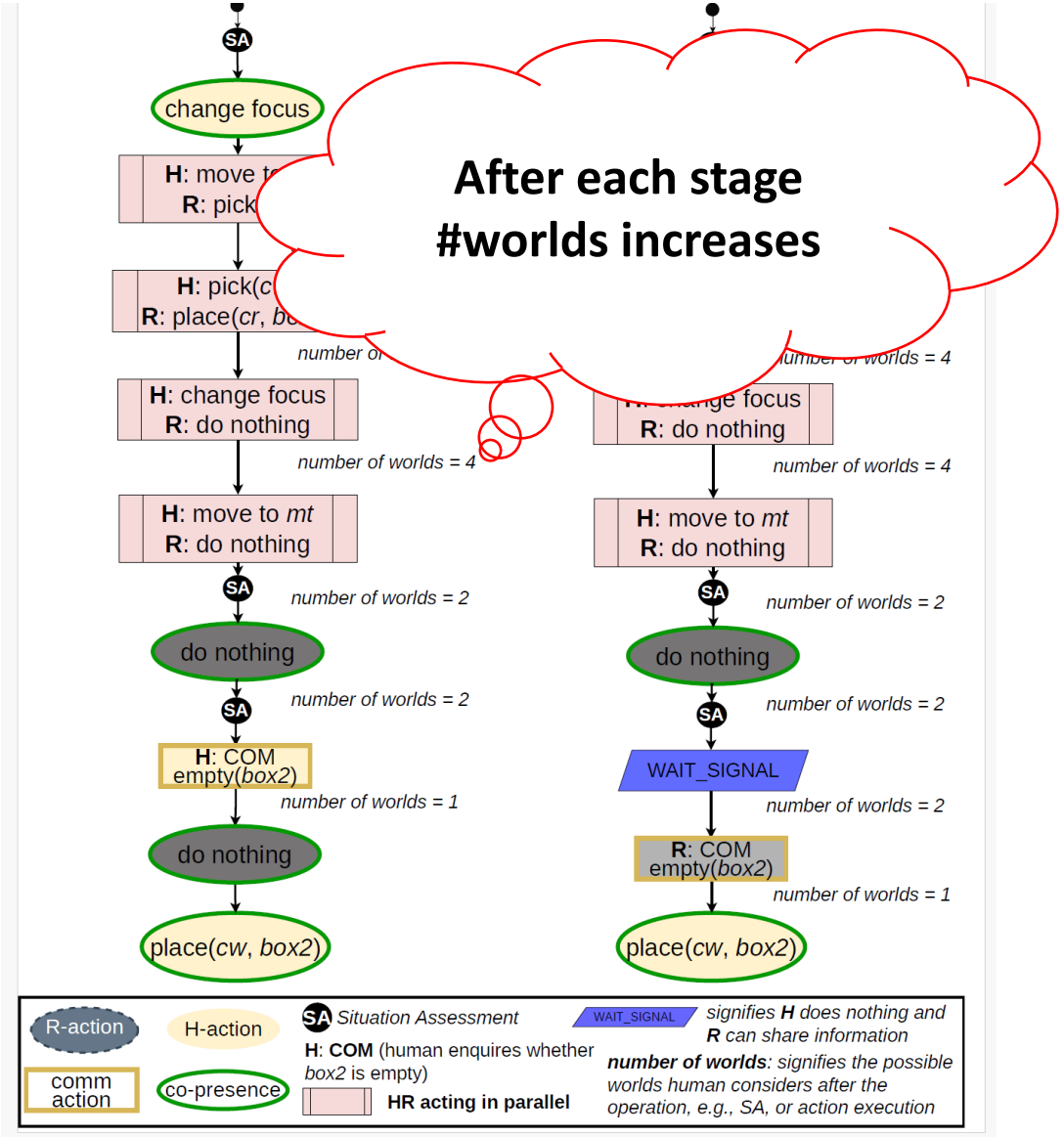


When the boxes are opaque

Human turns for *ot*

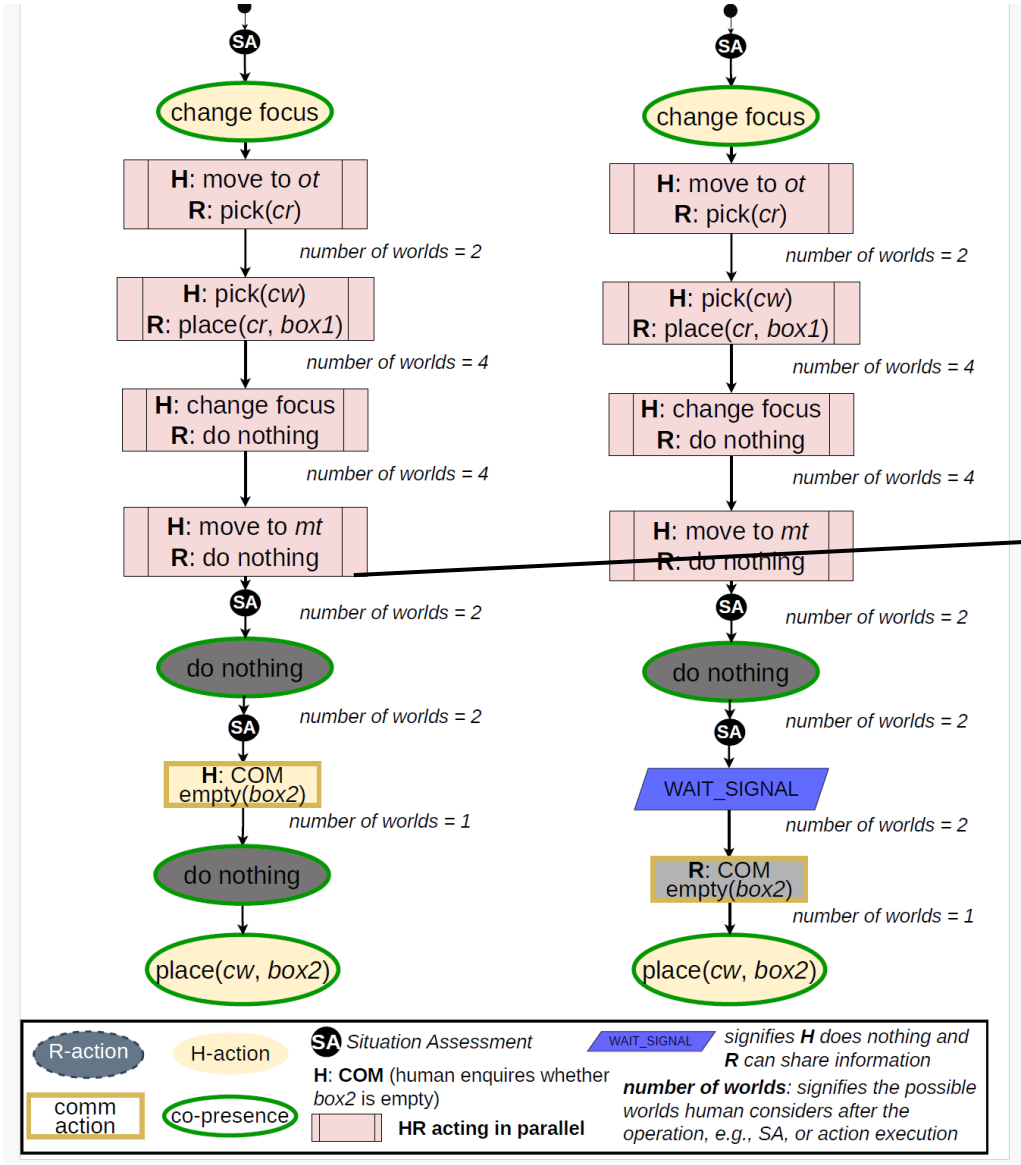


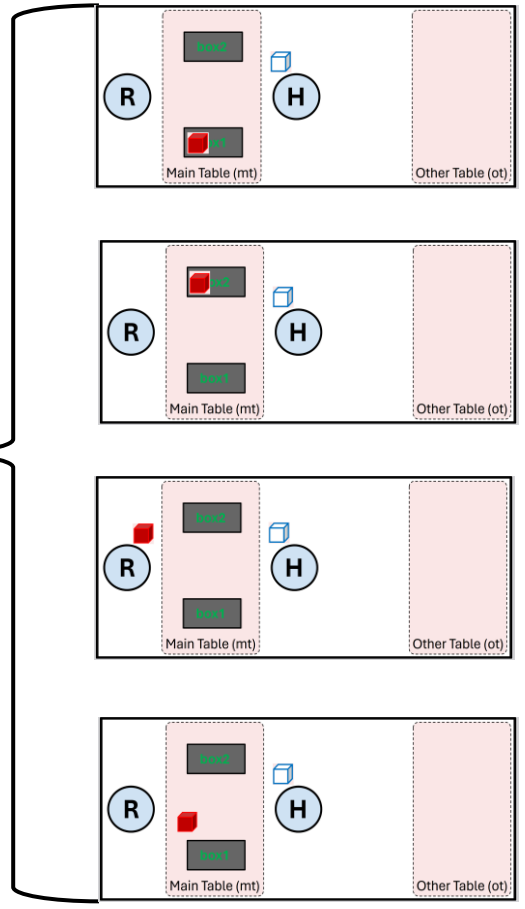
When the boxes are opaque



Just before Situation Assessment

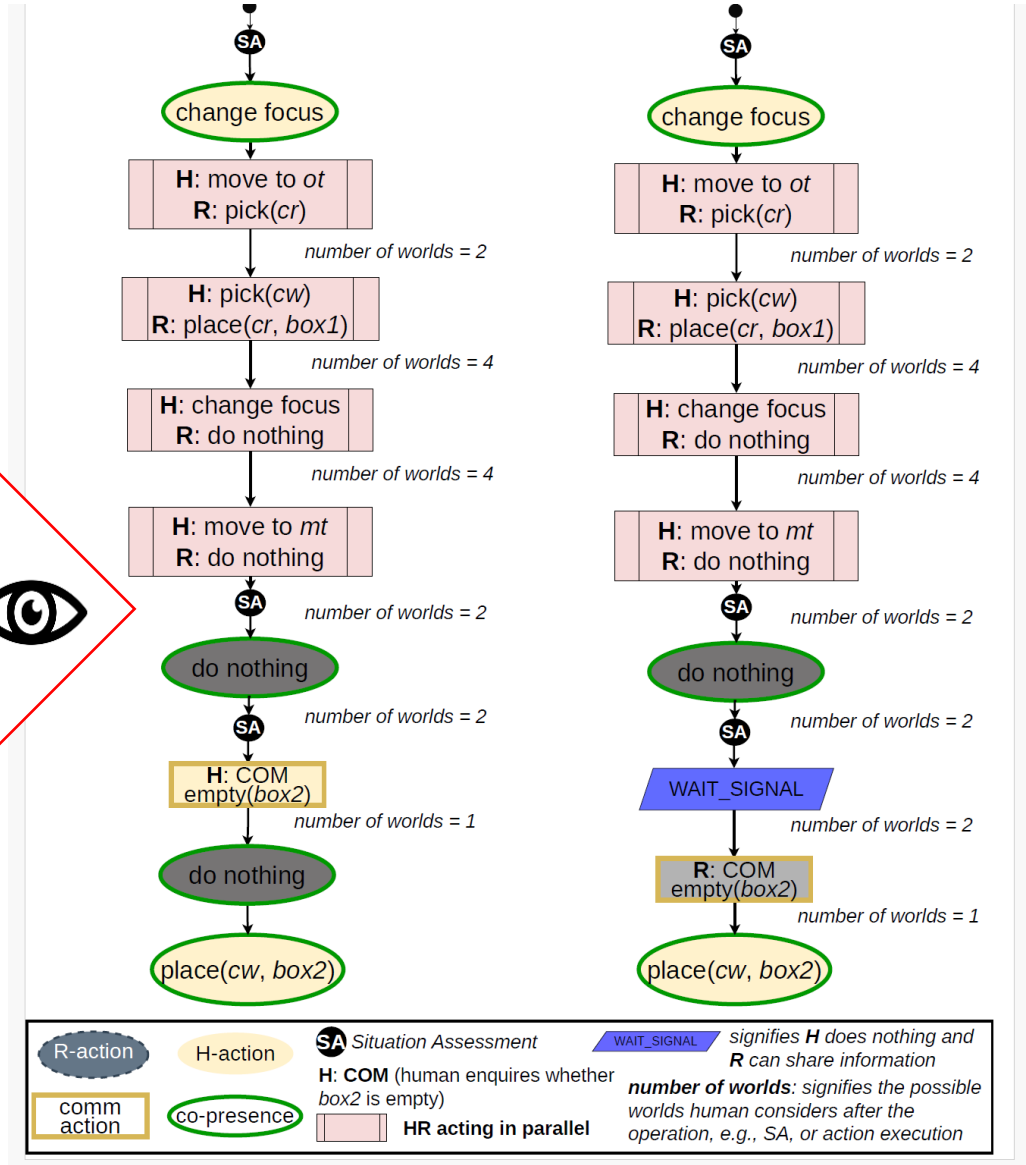
(all possible worlds from Human's perspective which the planner prepared)





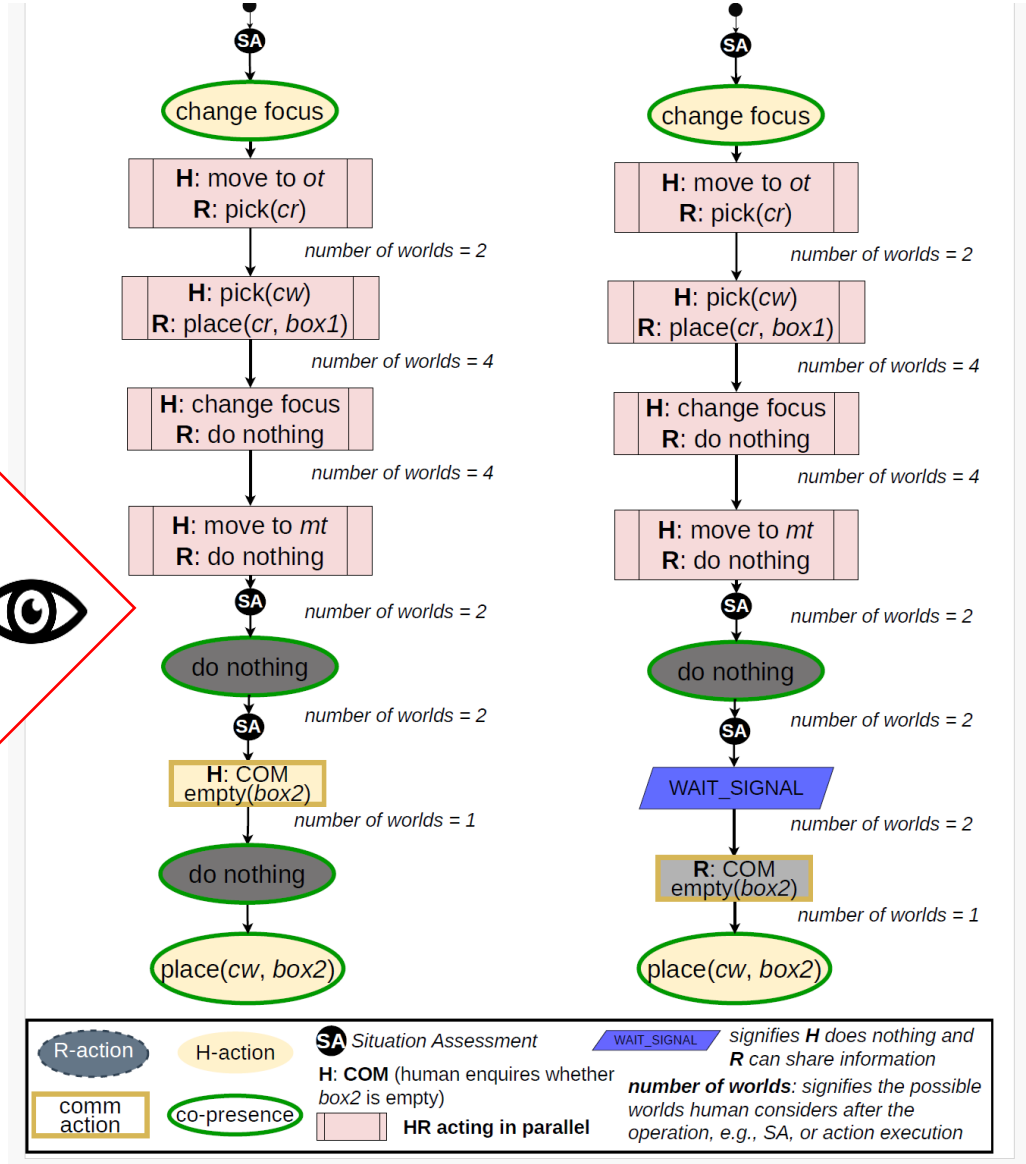
When the boxes are opaque

Human is back to the *mt*'s context, SA help ignore 2 worlds (out of 4)

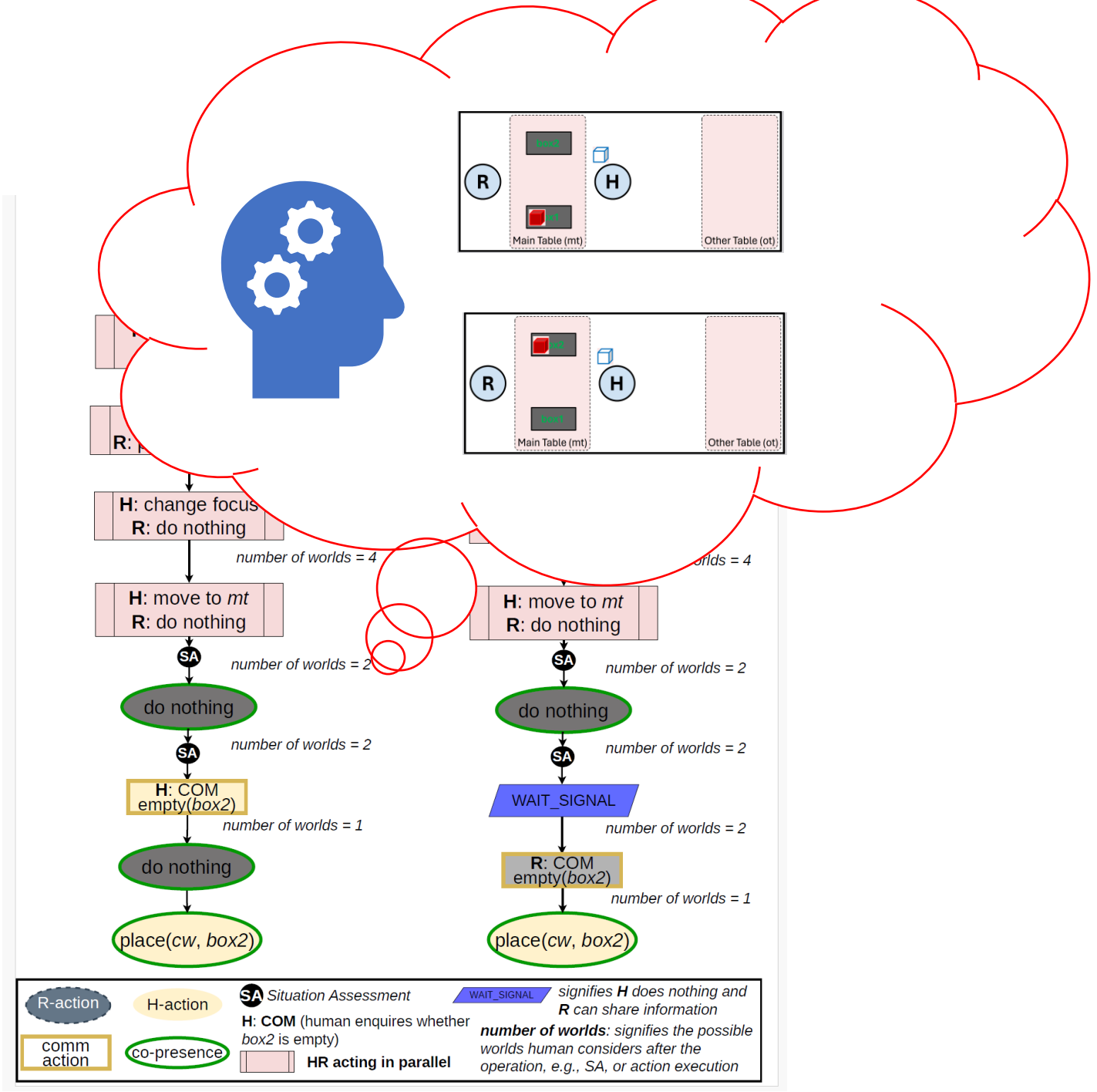


When the boxes are opaque

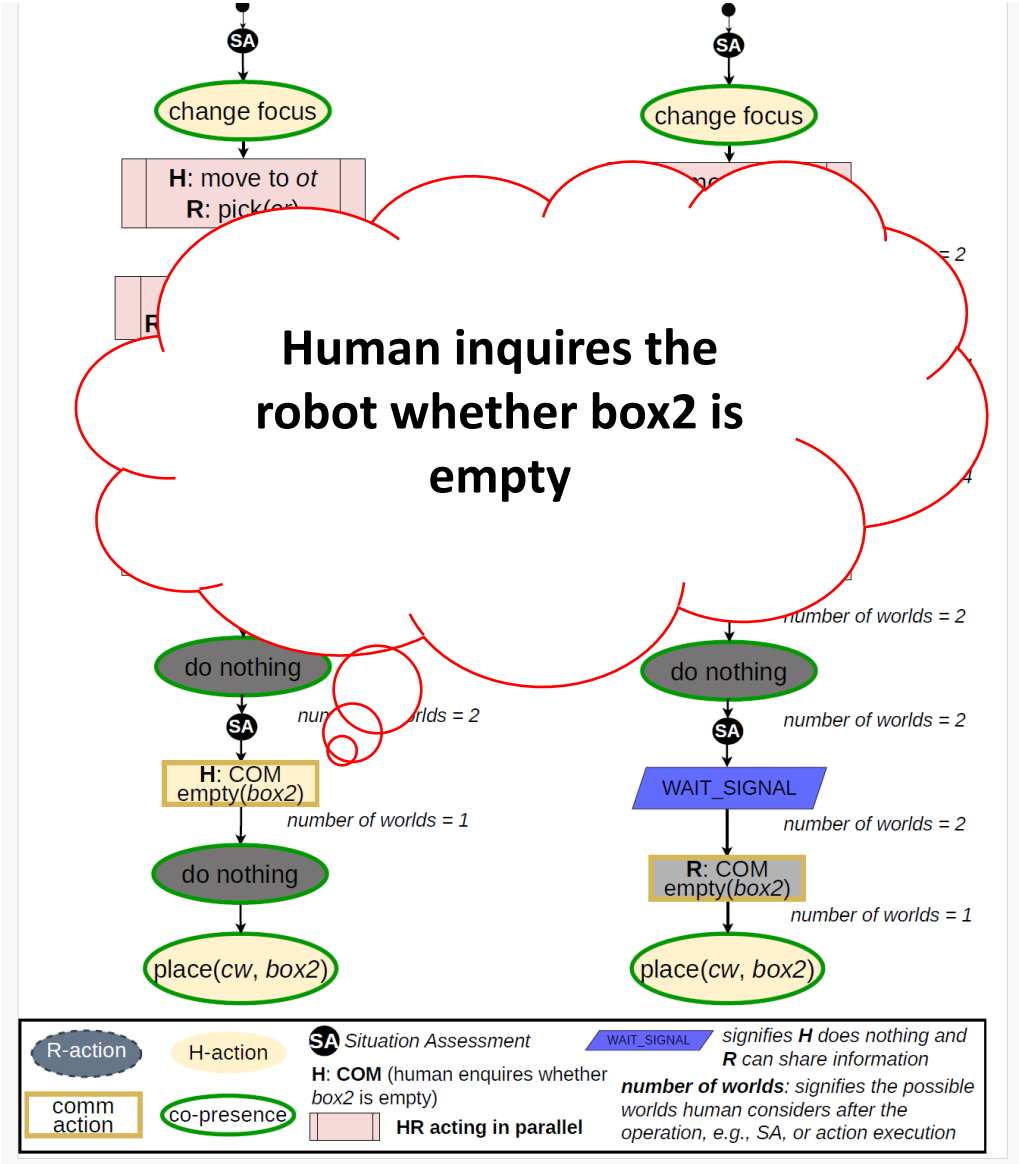
Possibilities remain are w.r.t. (inside cr box1) **and** (inside cr box2)



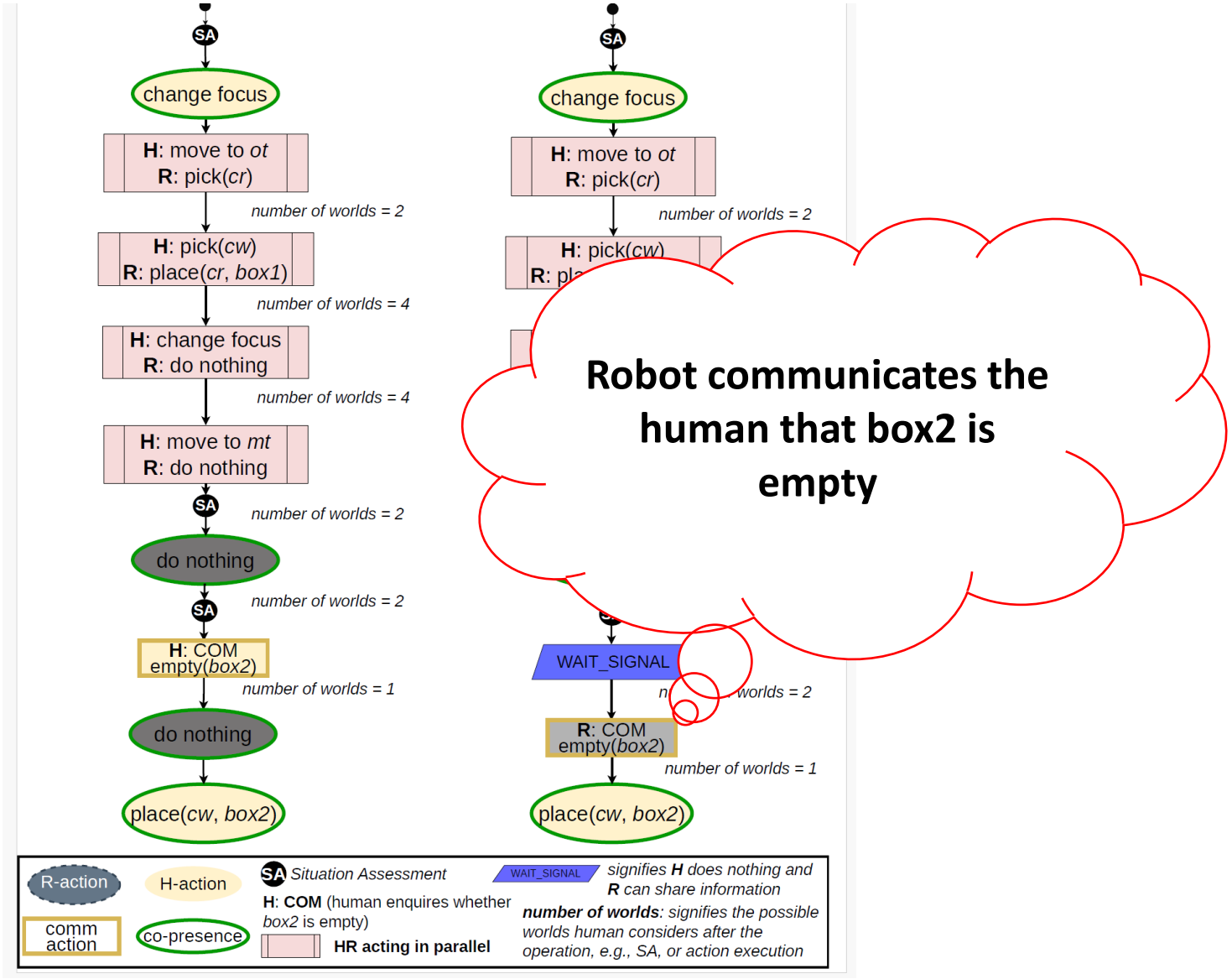
After SA


R
H
Main Table (mt)
Other Table (ot)
R
H
Main Table (mt)
Other Table (ot)
H: change focus
R: do nothing
number of worlds = 4
H: move to mt
R: do nothing
SA
number of worlds = 2
do nothing
SA
number of worlds = 2
H: COM
empty(box2)
number of worlds = 1
do nothing
place(cw, box2)
H: move to mt
R: do nothing
SA
number of worlds = 2
do nothing
SA
number of worlds = 2
WAIT_SIGNAL
number of worlds = 2
R: COM
empty(box2)
number of worlds = 1
place(cw, box2)
R-action
H-action
SA Situation Assessment
WAIT_SIGNAL signifies H does nothing and R can share information
H: COM (human enquires whether box2 is empty)
number of worlds: signifies the possible worlds human considers after the operation, e.g., SA, or action execution
comm action
co-presence
HR acting in parallel

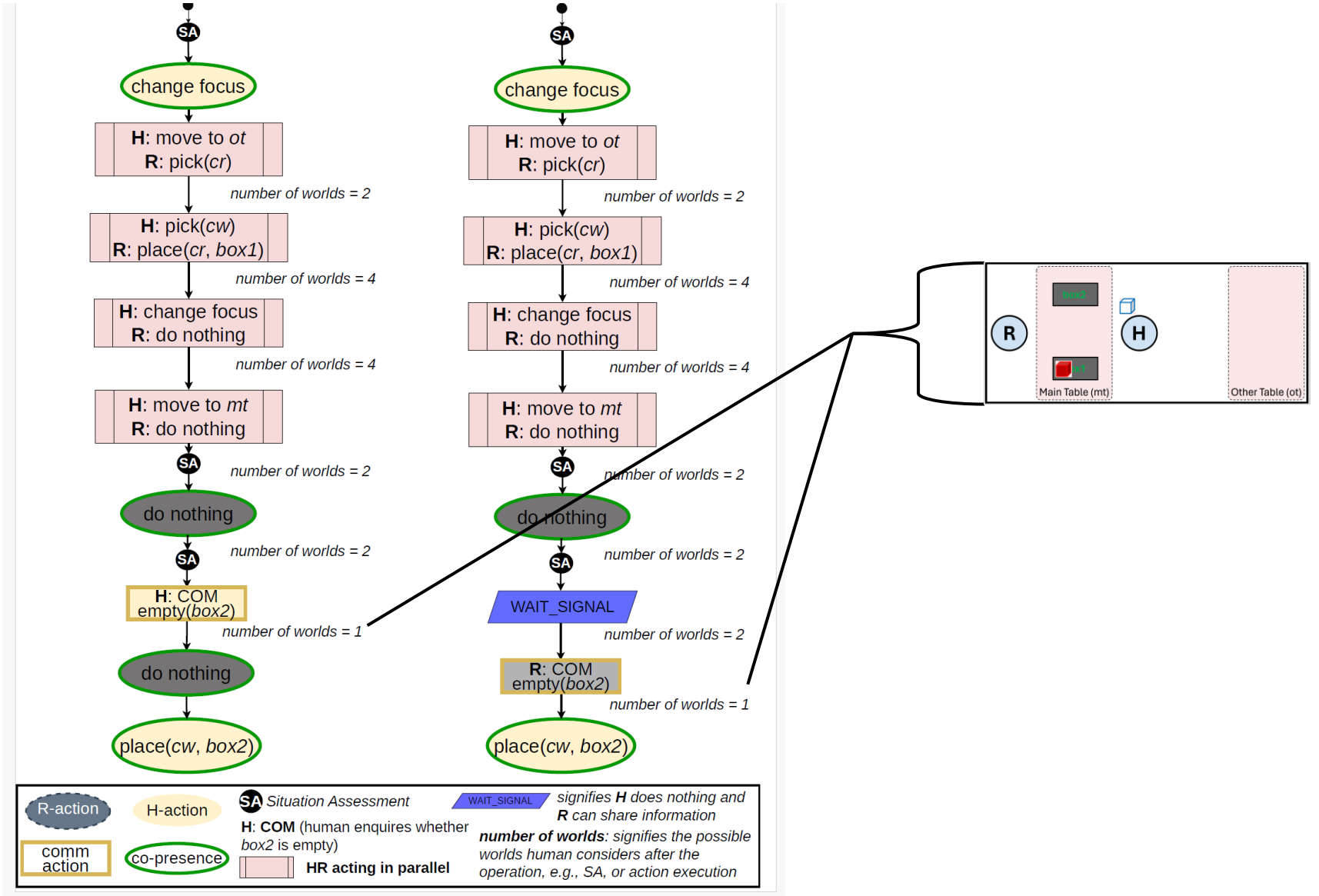
When the boxes are opaque: Human Inquires



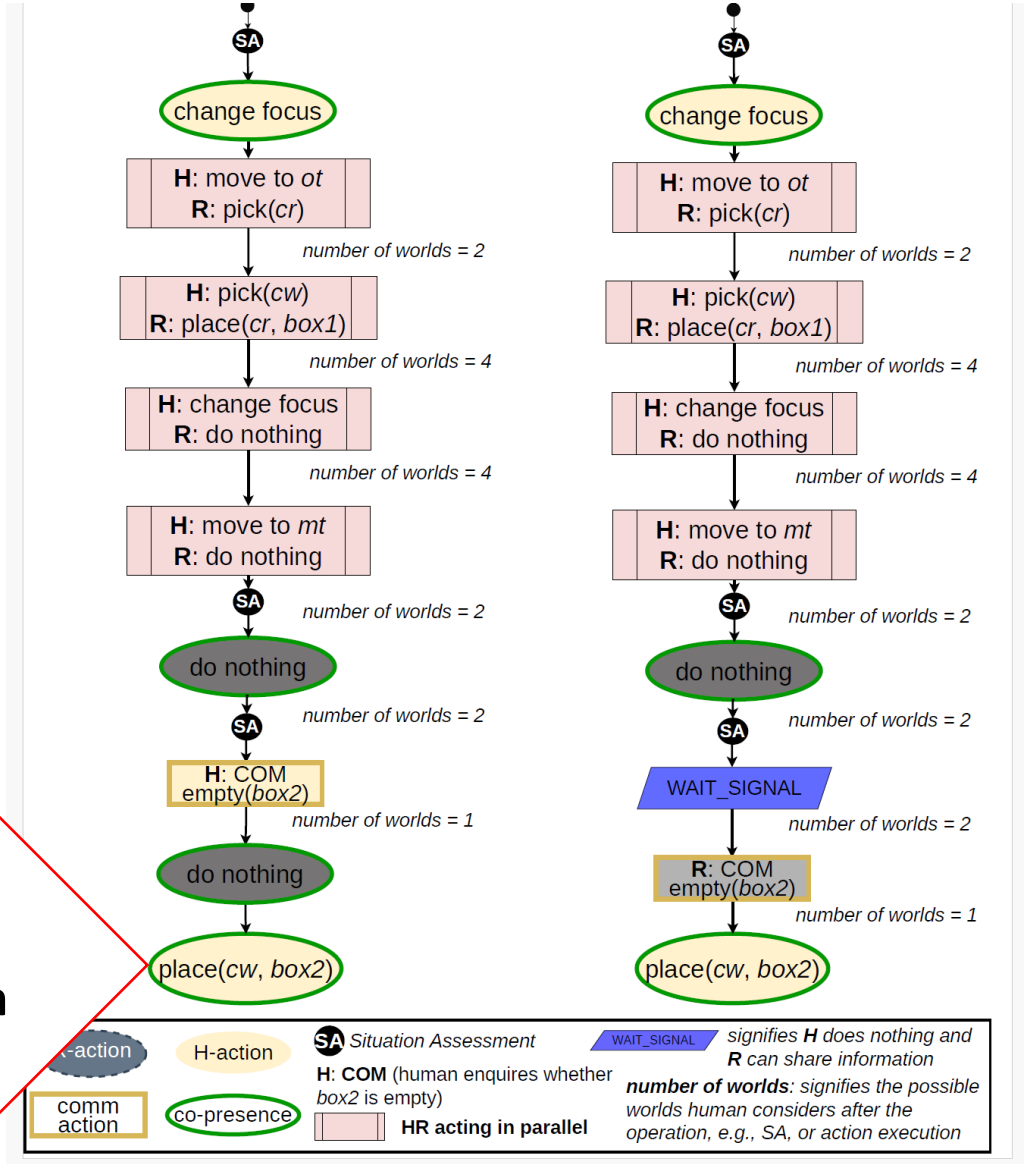
When the boxes are opaque: Robot Communicates



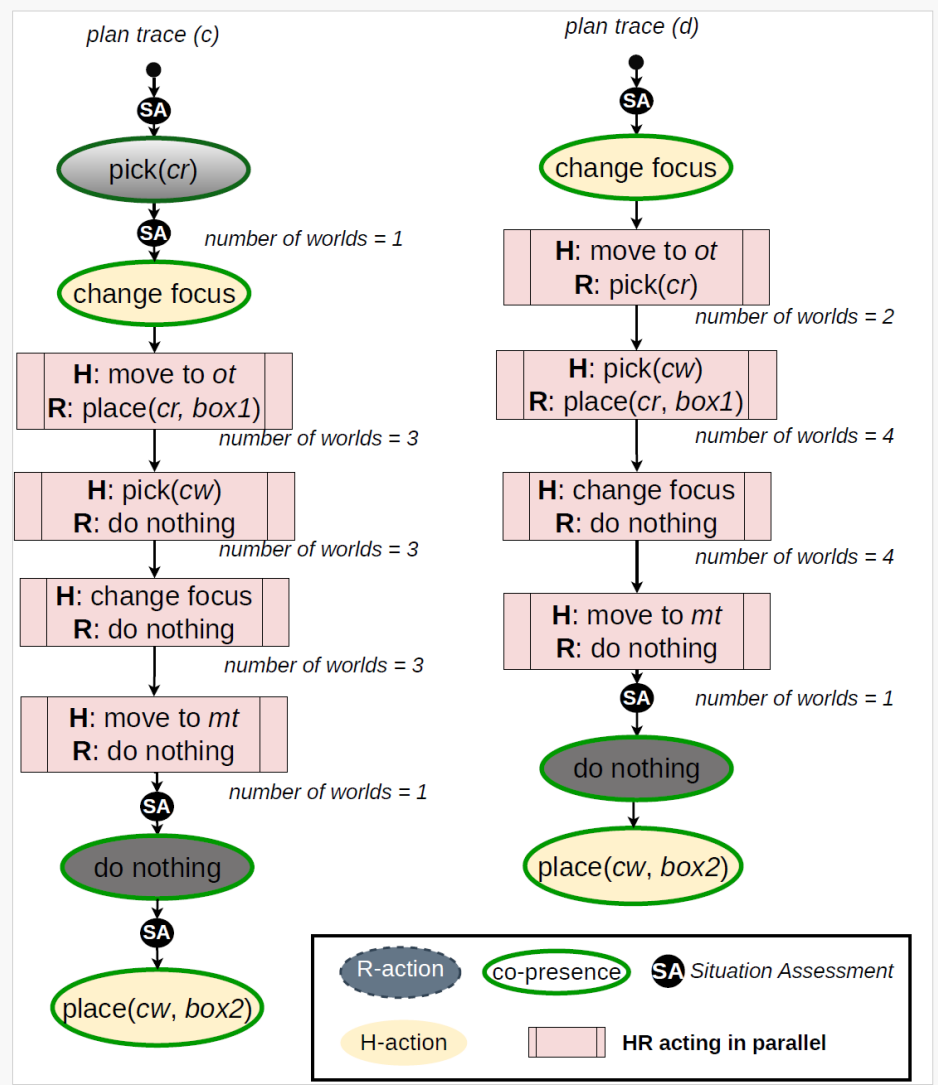
Human remains with the real world



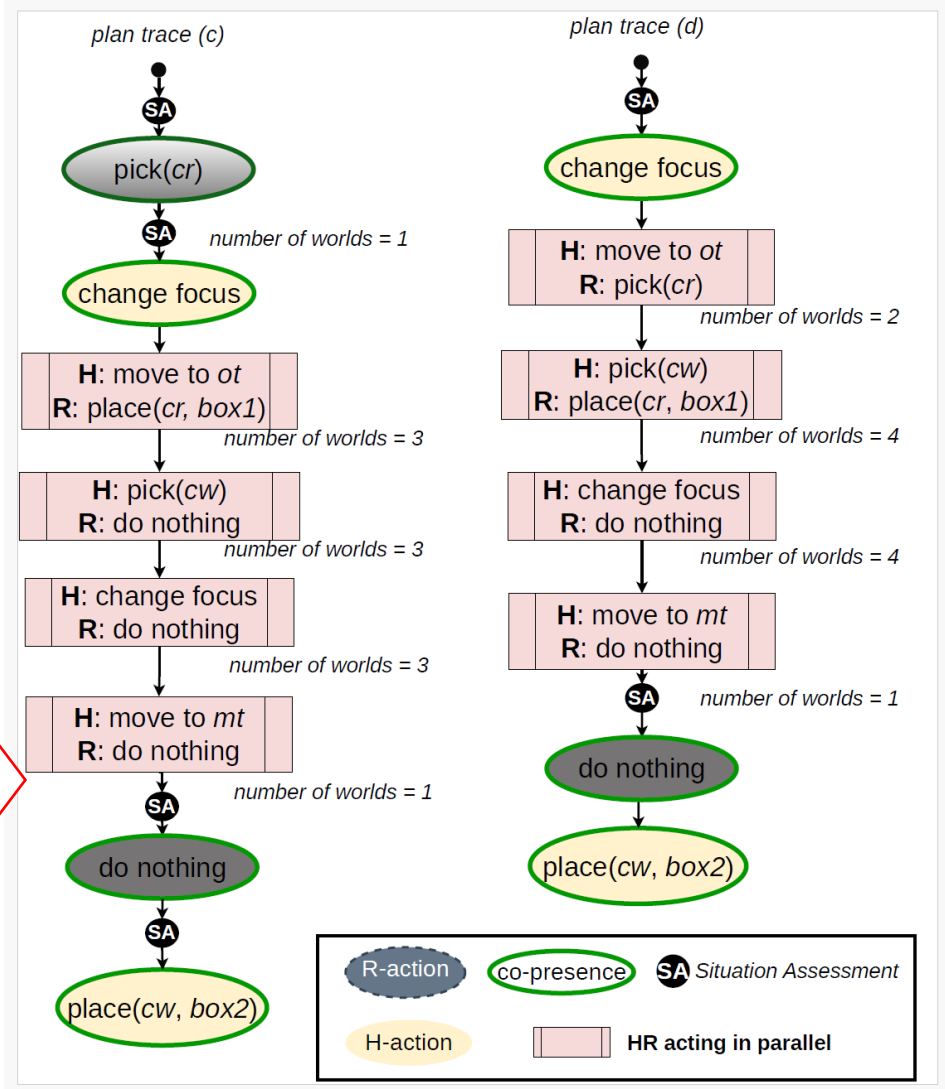
When the boxes are opaque



Human continues with the place action

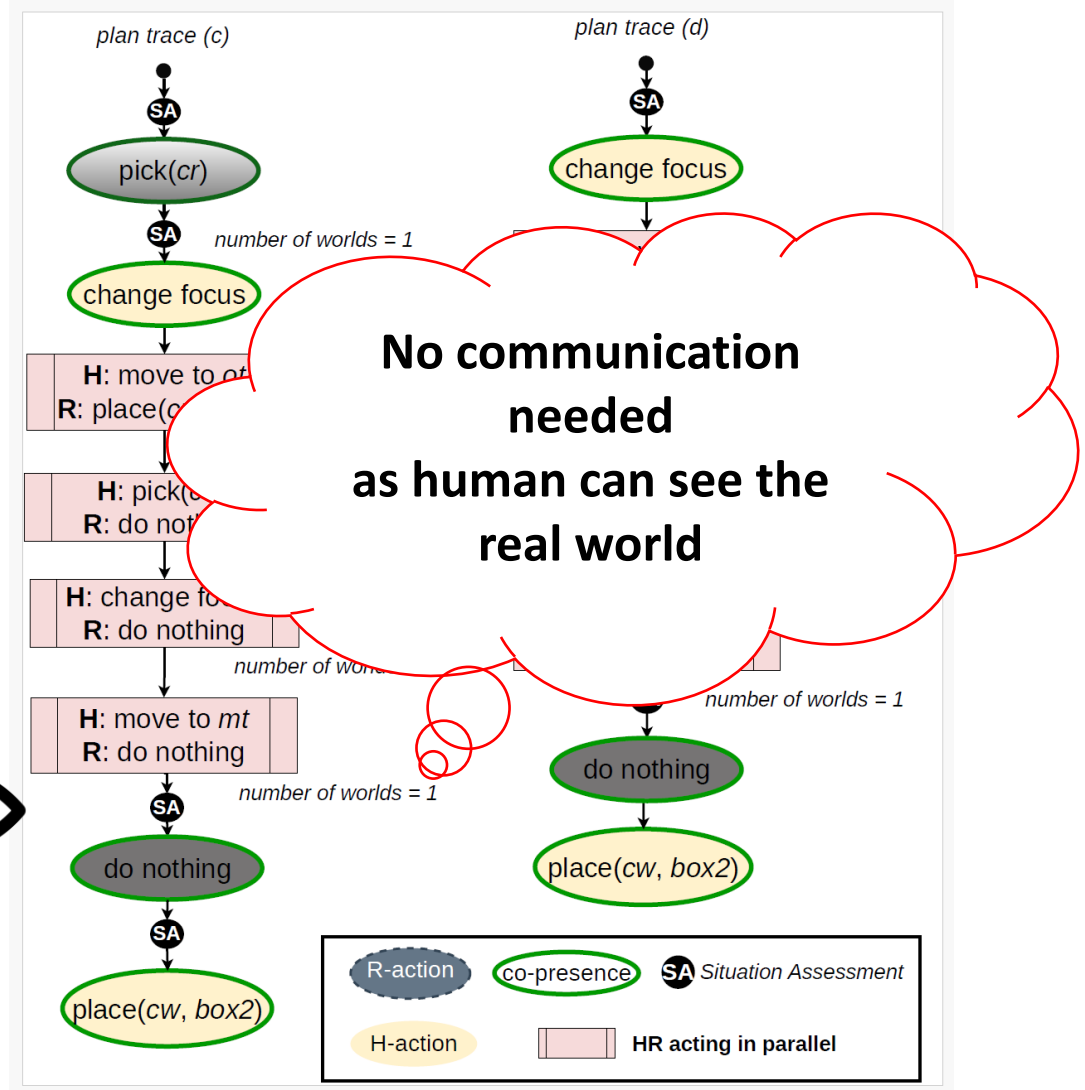


When the boxes are transparent

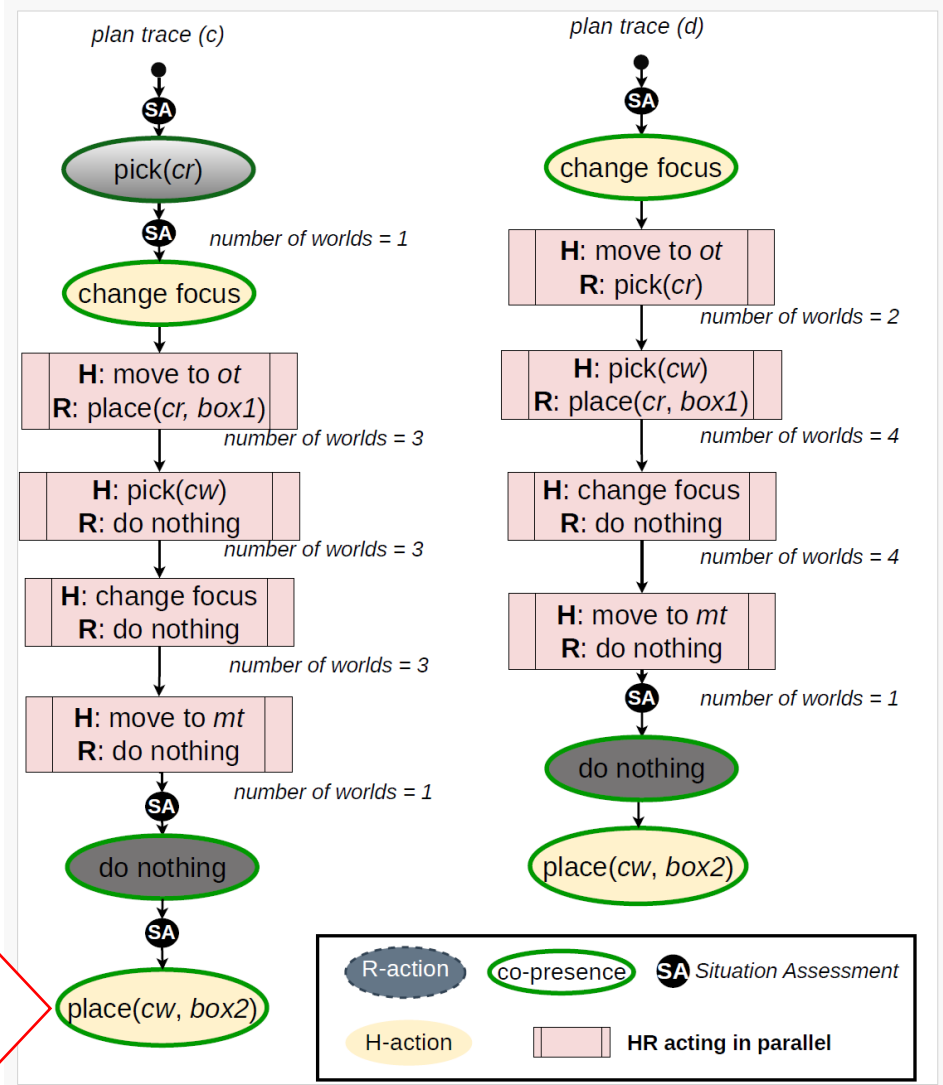


Human is back in the mt's context…

When the boxes are transparent



When the boxes are transparent



Human continues with the place action

When the boxes are transparent

Experiments: Quantitative Analysis

This is the preliminary data for our planner based on AND/OR search which conducts breadth-first search.

inst	K	*comm*	*#states*	$\|W\|$	*#leaves*	*time (ms)* $\times 10^5$
P1 (2,2,T)	2	N	218	4	3	0.008
P2 (2,2,O)	2	Y	236	4	3	0.014
P3 (3,2,T)	2	N	1643	7	6	5.906
P4 (3,2,O)	2	Y	2003	7	6	9.816
P5 (3,2,T)	4	N	4107	14	5	99.81
P6 (3,2,O)	4	Y	5607	14	5	125.3

Table 1: *The planner's performance is evaluated on different metrics.* inst *is instance description. Whether communication is employed –* comm. *The metrics include the total number of states explored (*#states*), the worst-case number of worlds (*$|W|$*) evaluated in a state, the number of traces (*#leaves*), and the execution* time *(measured in* 10^5 *ms).*

Experiments: Quantitative Analysis

This is the preliminary data for our planner based on AND/OR search which conducts breadth-first search.

inst	K	*comm*	*#states*	$\|W\|$	*#leaves*	*time (ms)* $\times 10^5$
P1 (2,2,T)	2	N	218	4	3	0.008
P2 (2,2,O)	2	Y	236	4	3	0.014
P3 (3,2,T)	2	N	1643	7	6	5.906
P4 (3,2,O)	2	Y	2003	7	6	9.816
P5 (3,2,T)	4	N	4107	14	5	99.81
P6 (3,2,O)	4	Y	5607	14	5	125.3

Table 1: *The planner's performance is evaluated on different metrics.* inst *is instance description. Whether communication is employed –* comm. *The metrics include the total number of states explored (*#states*), the worst-case number of worlds (*$|W|$*) evaluated in a state, the number of traces (*#leaves*), and the execution* time *(measured in* 10^5 *ms).*

Conclusion

- Our novel framework allows the robot to implement a ToM not only at execution time but also at planning time
- This is done, thanks to the following:
  - the use of epistemic reasoning,
  - the notion of shared experience, and
  - observable and non-observable facts, which allow anticipation of human situation assessment
- We showed the applicability and effectiveness of our human-aware planner.

Thank you