Workshop on Knowledge Engineering for Planning and Scheduling, 2023

Autonomous Capability Assessment of Black-Box Sequential Decision-Making Systems

Pulkit Verma, Rushang Karia, Siddharth Srivastava

Arizona State University

Autonomous Agents and Intelligent Robots

Arizona State University

Personalized Assessment of SDM Systems

- Users would like to give AI systems multiple tasks.
  - How would users know what the AI systems can do?

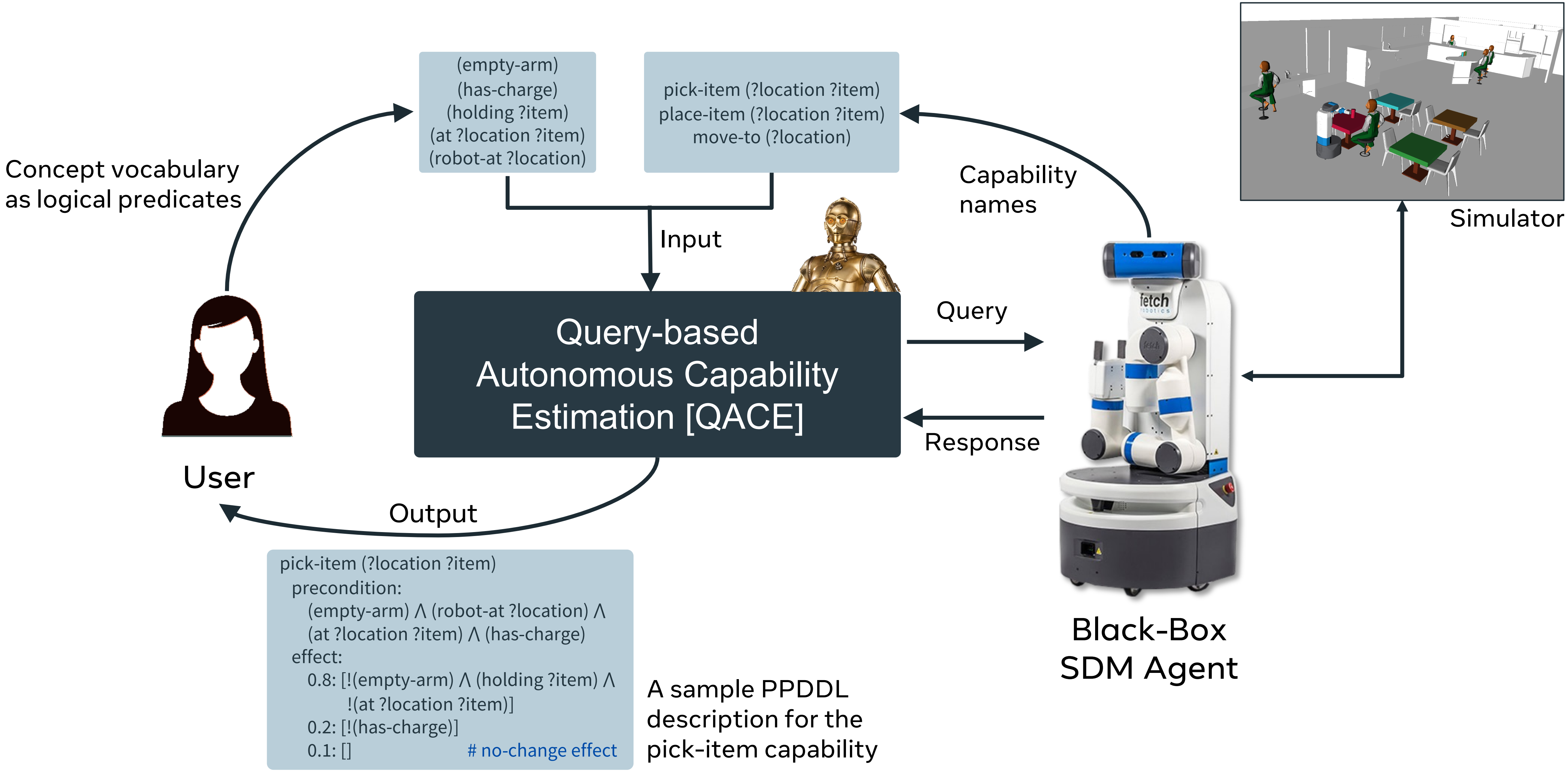
- AI systems should make it easy for their operators to learn how to use them safely.[†]

- The assessment should work with black-box AI systems.

[†]Srivastava S. *Unifying Principles and Metrics for Safe and Assistive AI.* In Proc. AAAI 2021.

Related Work

- Learning from passive observations: *Can learn incorrect models*.
  - Pasula et al. (JAIR 2007), Rodrigues et al. (ILP 2011), Mourão et al. (UAI 2012), Juba and Stern (2022), etc.
- Learning from sampled transitions: *Lower sample efficiency and correctness profiles*.
  - Ng et al. (IJCAI 2019), Chitnis et al. (AAAI 2021), etc.
- Autonomous Assessment for SDM systems: *Works only for deterministic settings*.
  - Verma et al. (AAAI 2021), Nayyar et al. (AAAI 2022), Verma et al. (KR 2022), etc.


(empty-arm)
(has-charge)
(holding ?item)
(at ?location ?item)
(robot-at ?location)
pick-item (?location ?item)
place-item (?location ?item)
move-to (?location)
Concept vocabulary as logical predicates
Capability names
Simulator
Input
Query-based Autonomous Capability Estimation [QACE]
Query
Response
User
Output
Black-Box SDM Agent
pick-item (?location ?item)
precondition:
(empty-arm) ∧ (robot-at ?location) ∧
(at ?location ?item) ∧ (has-charge)
effect:
0.8: [!(empty-arm) ∧ (holding ?item) ∧
!(at ?location ?item)]
0.2: [!(has-charge)]
0.1: [] # no-change effect
A sample PPDDL description for the pick-item capability

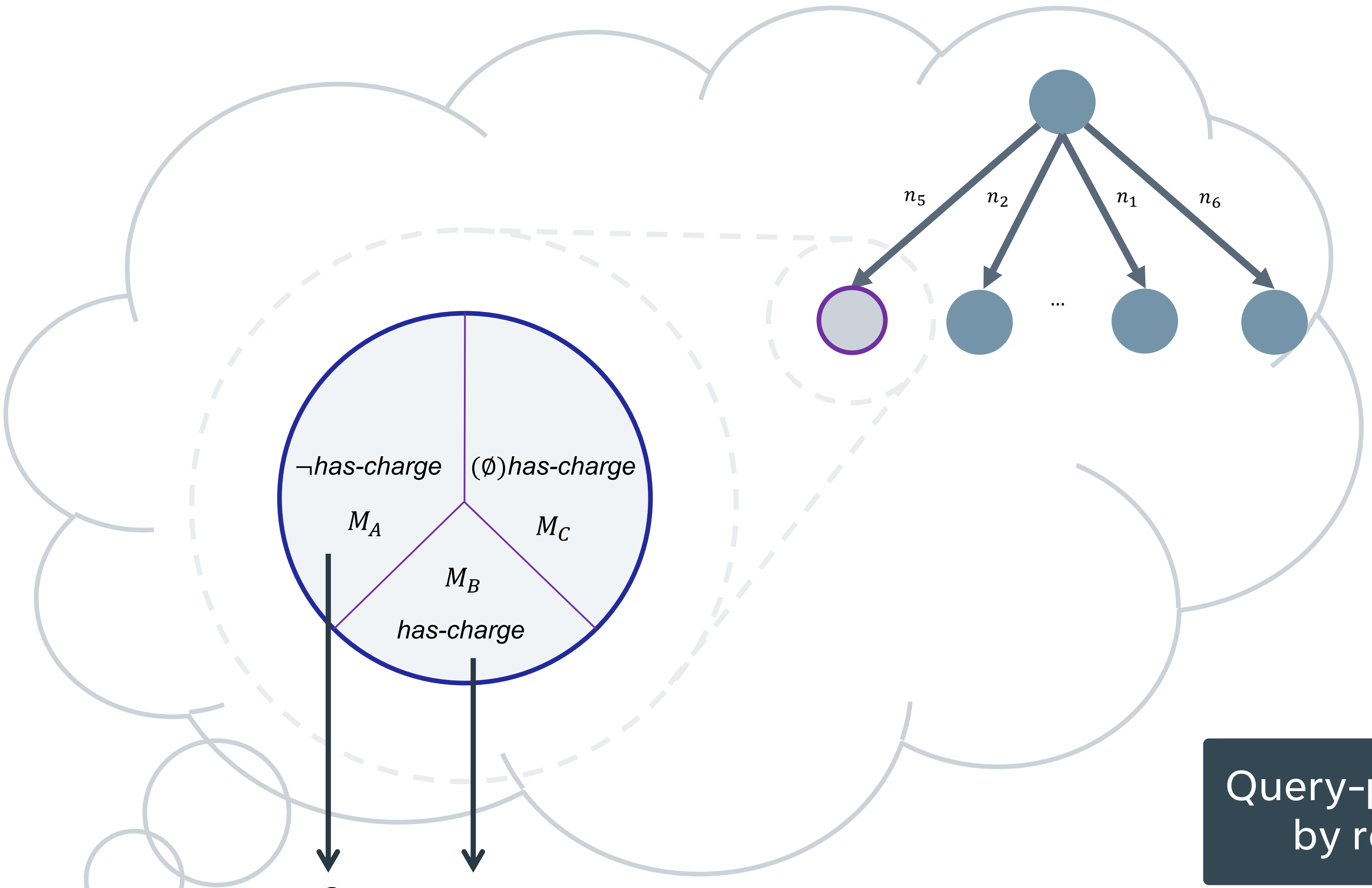
QACE works in 2 phases

Phase 1:
Learn a Non-Deterministic Model

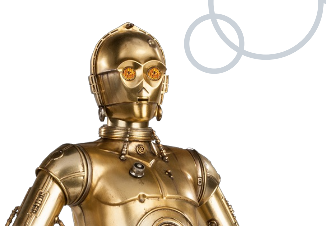
```
pick-item (?location ?item)
  precondition:
    (empty-arm) ∧ (robot-at ?location) ∧
    (at ?location ?item) ∧ (has-charge)
  effect:
    [!(empty-arm) ∧ (holding ?item) ∧
         !(at ?location ?item)]
    [!(has-charge)]
    []                  # no-change effect
```

Phase 2:
Convert Non-Deterministic Model
to Probabilistic Model

```
pick-item (?location ?item)
  precondition:
    (empty-arm) ∧ (robot-at ?location) ∧
    (at ?location ?item) ∧ (has-charge)
  effect:
    0.8: [!(empty-arm) ∧ (holding ?item) ∧
         !(at ?location ?item)]
    0.2: [!(has-charge)]
    0.1: []             # no-change effect
```



pick-item (?location ?item)
precondition:
n_1 (+/−/∅) (has-charge)
n_2 (+/−/∅) (robot-at ?location)
n_3 (+/−/∅) (at ?location ?item)
n_4 (+/−/∅) (empty-arm)
effect:
n_5 (+/−/∅) (has-charge)
n_6 (+/−/∅) (robot-at ?location)
n_7 (+/−/∅) (at ?location ?item)
n_8 (+/−/∅) (empty-arm)

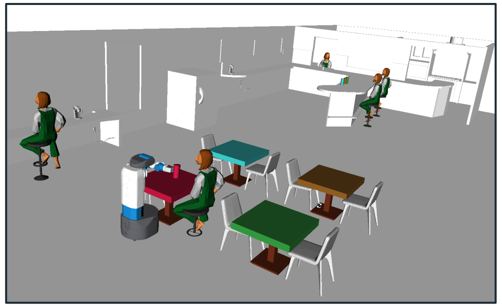


Generate a
distinguishing query:
Q such that $Q(M_A) \neq Q(M_B)$

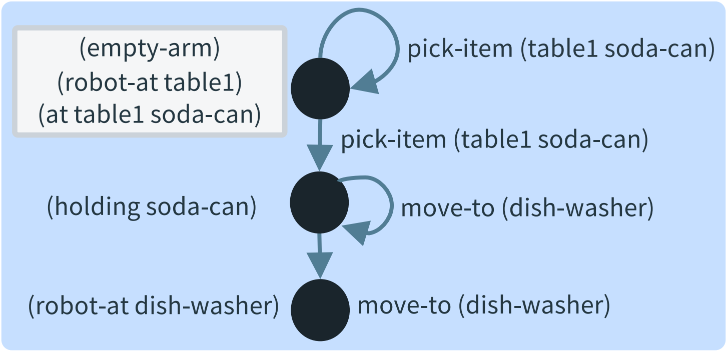
Query-policy generated automatically by reduction to FOND planning

Query Synthesis

	x	y	z	θ	φ	ψ
robot-base	1.0	-3.2	4.7	0.9	1.3	3.1
soda-can1	6.0	-2.8	3.5	8.3	6.7	9.2
⋮				⋮		
table4	-2.1	4.1	1.9	3.7	9.5	4.8



Simulator



A sample query policy

Reducing Query Synthesis to FOND Planning

Models differ in only one predicate in precondition or effect.

```
pick-item (?location ?item)
  precondition:
    (precondition)
  effect:
   (oneof
     (effect1)
     (effect2)
     !(has-charge)
   )
```

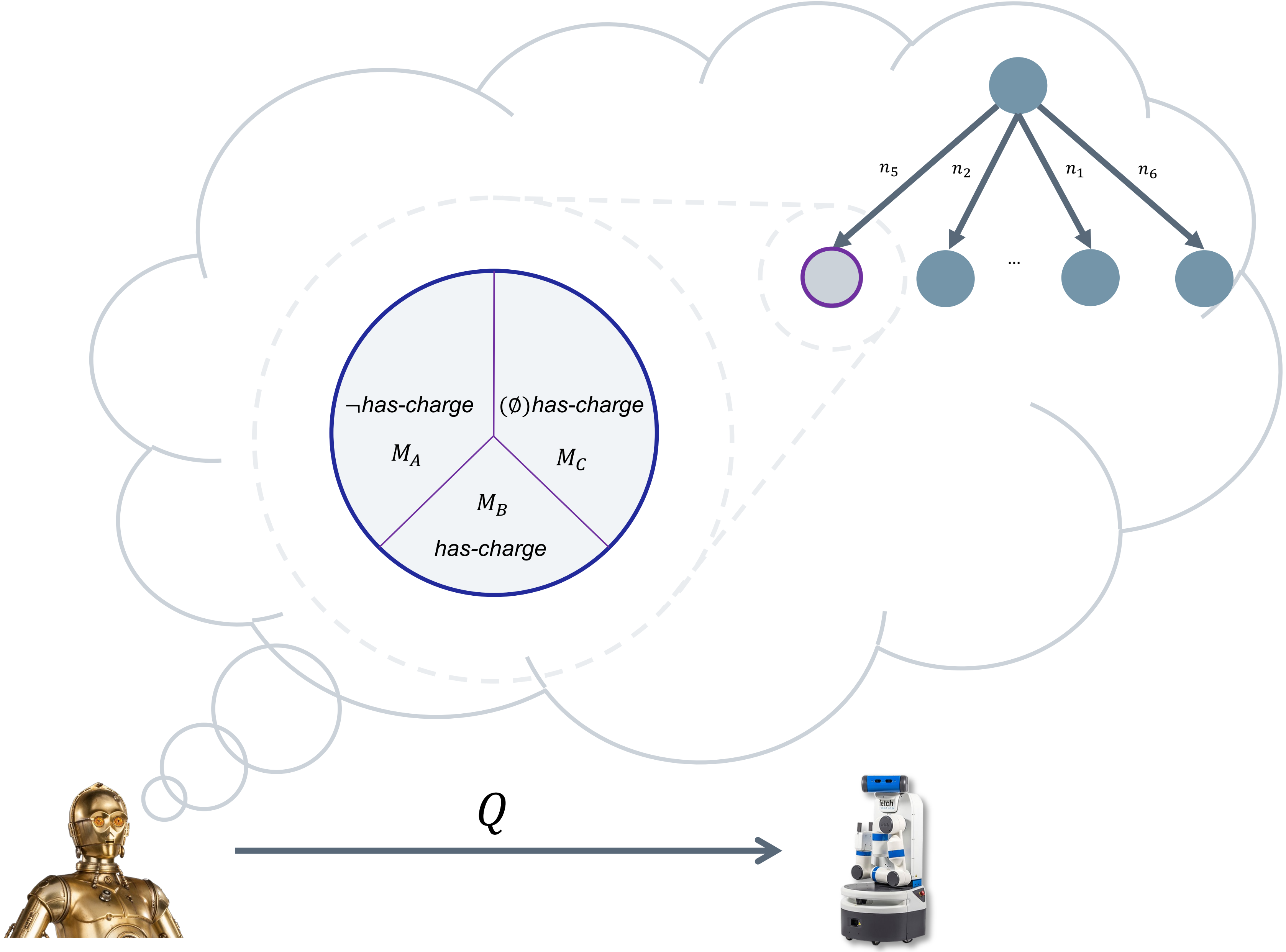
M_A

```
pick-item (?location ?item)
  precondition:
    (precondition)
  effect:
   (oneof
     (effect1)
     (effect2)
     (has-charge)
   )
```

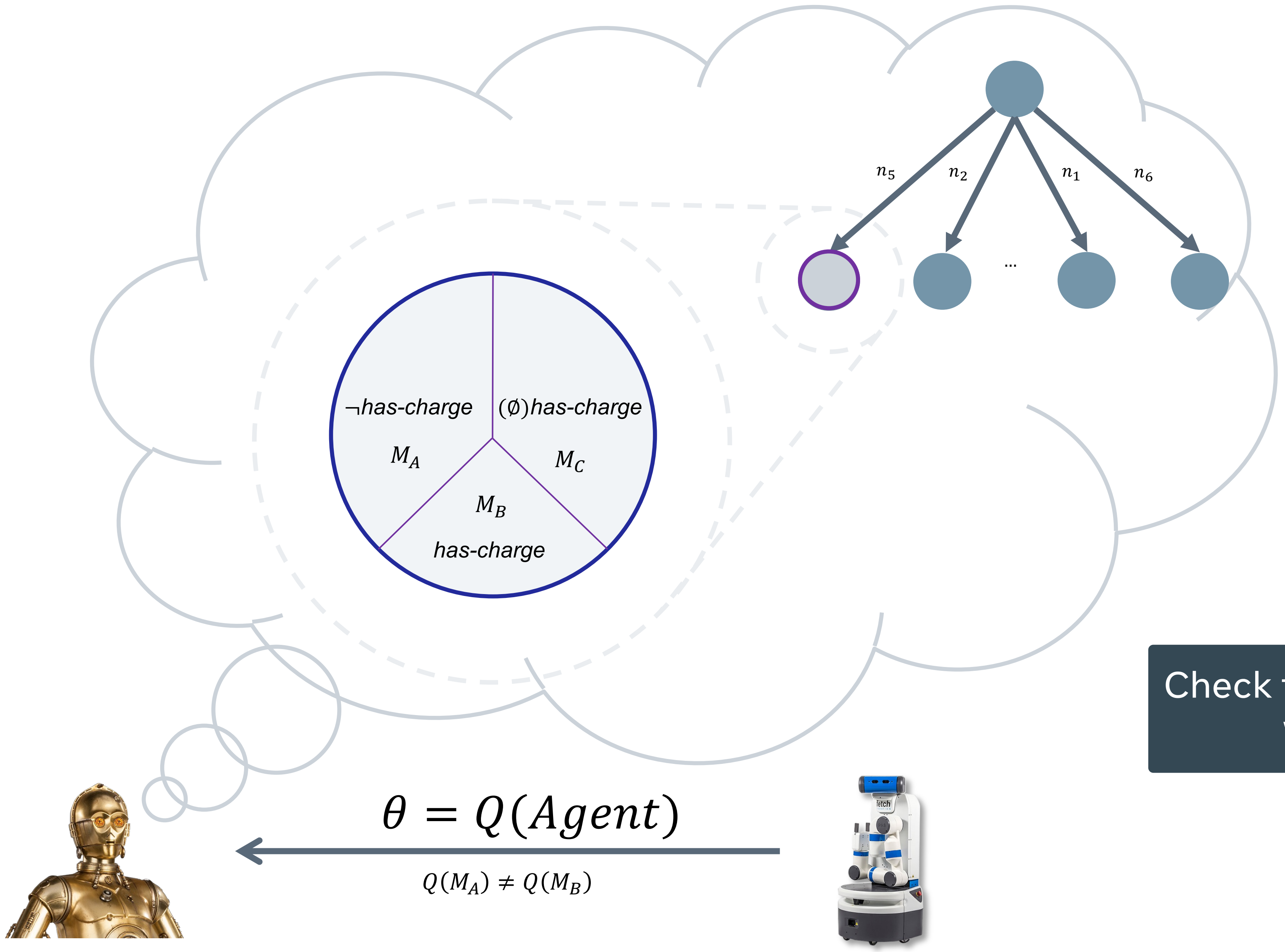
M_B

pick-item (?location ?item)
precondition:
$(\text{precondition})_{M_A} \lor (\text{precondition})_{M_B}$
effect:
$(\text{precondition})_{M_A} \land !\,(\text{precondition})_{M_B} \rightarrow (\text{goal})$
$!(\text{precondition})_{M_A} \land (\text{precondition})_{M_B} \rightarrow (\text{goal})$
$(\text{precondition})_{M_A} \land (\text{precondition})_{M_B} \rightarrow$
(one of
$((\text{effect1})_{M_A} \land (\text{effect1})_{M_B})$
$((\text{effect2})_{M_A} \land (\text{effect2})_{M_B})$
$(!\,(\text{has-charge})_{M_A} \land (\text{has-charge})_{M_B})$
)

Consolidated capability used to generate the FOND Planning Domain

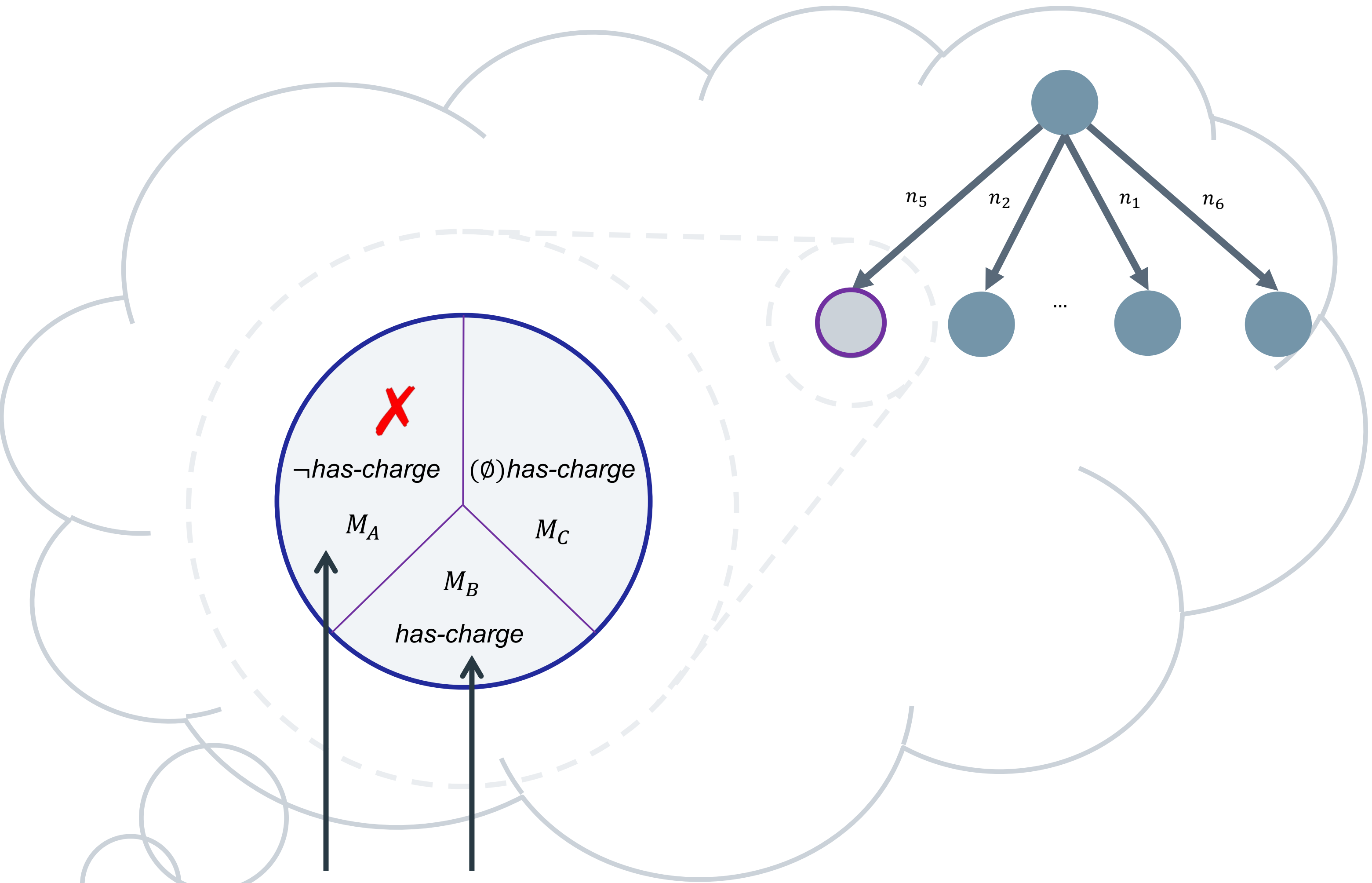


pick-item (?location ?item)
  precondition:
n_1 $(+/-/\emptyset)$ (has-charge)
n_2 $(+/-/\emptyset)$ (robot-at ?location)
n_3 $(+/-/\emptyset)$ (at ?location ?item)
n_4 $(+/-/\emptyset)$ (empty-arm)
  effect:
n_5 $(+/-/\emptyset)$ (has-charge)
n_6 $(+/-/\emptyset)$ (robot-at ?location)
n_7 $(+/-/\emptyset)$ (at ?location ?item)
n_8 $(+/-/\emptyset)$ (empty-arm)



pick-item (?location ?item)
  precondition:
n_1 (+/−/∅) (has-charge)
n_2 (+/−/∅) (robot-at ?location)
n_3 (+/−/∅) (at ?location ?item)
n_4 (+/−/∅) (empty-arm)
  effect:
n_5 (+/−/∅) (has-charge)
n_6 (+/−/∅) (robot-at ?location)
n_7 (+/−/∅) (at ?location ?item)
n_8 (+/−/∅) (empty-arm)

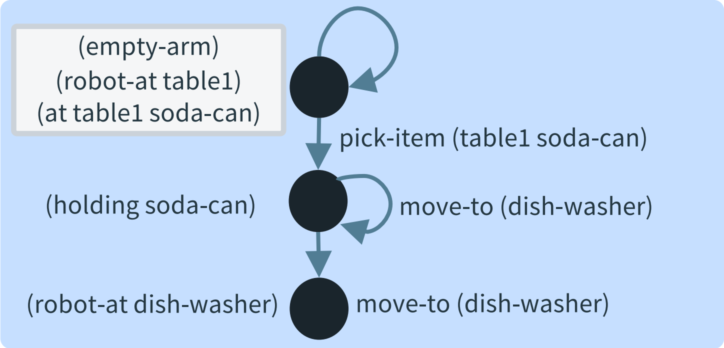
Check the consistency of refinements with the agent response



pick-item (?location ?item)
precondition:
n_1 (+/∅) (has-charge)
n_2 (+/−/∅) (robot-at ?location)
n_3 (+/−/∅) (at ?location ?item)
n_4 (+/−/∅) (empty-arm)
effect:
n_5 (+/−/∅) (has-charge)
n_6 (+/−/∅) (robot-at ?location)
n_7 (+/−/∅) (at ?location ?item)
n_8 (+/−/∅) (empty-arm)

Reject refinement(s) that are not consistent with the agent

FOND Model to Probabilistic Model: MLE




```
pick-item (?location ?item)
  precondition:
    (empty-arm) ∧ (robot-at ?location) ∧
    (at ?location ?item) ∧ (has-charge)
  effect:
    [!(empty-arm) ∧ (holding ?item) ∧
          !(at ?location ?item)]
    [!(has-charge)]
    []                          # no-change effect
```

1. Ask queries multiple times

2. Perform MLE on collected effects

```
pick-item (?location ?item)
  precondition:
    (empty-arm) ∧ (robot-at ?location) ∧
    (at ?location ?item) ∧ (has-charge)
  effect:
    0.8: [!(empty-arm) ∧ (holding ?item) ∧
          !(at ?location ?item)]
    0.2: [!(has-charge)]
    0.1: []                     # no-change effect
```

Formal Results

- QACE learns the models that are sound and complete wrt. the SDMA transition model.

Theorem 1. *Let $\mathcal{A}$ be a black-box SDMA with a ground truth transition model $\mathcal{T}'$ expressible in terms of predicates $\mathcal{P}$ and a set of capabilities $\mathcal{C}$. Let M^* be the non-deterministic model expressed in terms of predicates $\mathcal{P}^*$ and capabilities $\mathcal{C}$, and learned using the query-based autonomous capability estimation algorithm (Alg. 1) just before line 10. Let C_N be a set of capability names corresponding to capabilities $\mathcal{C}$. If $\mathcal{P}^* \subseteq \mathcal{P}$, then the model M^* is* sound *w.r.t. the SDMA transition model $\mathcal{T}'$. Additionally, if $\mathcal{P}^* = \mathcal{P}$, then the model M^* is* complete *w.r.t. the SDMA transition model $\mathcal{T}'$.*

- QACE learns models with accurate probabilities of probabilistic effects in the limit.

Theorem 2. *Let $\mathcal{A}$ be a black-box SDMA with a ground truth transition model $\mathcal{T}'$ expressible in terms of predicates $\mathcal{P}$ and a set of capabilities $\mathcal{C}$. Let M be the probabilistic model expressed in terms of predicates $\mathcal{P}^*$ and capabilities $\mathcal{C}$, and learned using the query-based autonomous capability estimation algorithm (Alg. 1). Let $\mathcal{P} = \mathcal{P}^*$ and M be generated using a sound and complete non-deterministic model M^* in line 11 of Alg. 1, and let all effects of each capability $c \in \mathcal{C}$ be identifiable. The model M is* correct *w.r.t. the model $\mathcal{T}'$ in the limit as η tends to ∞, where η is hyperparameter in query Q_{PS} used in Alg. 1.*

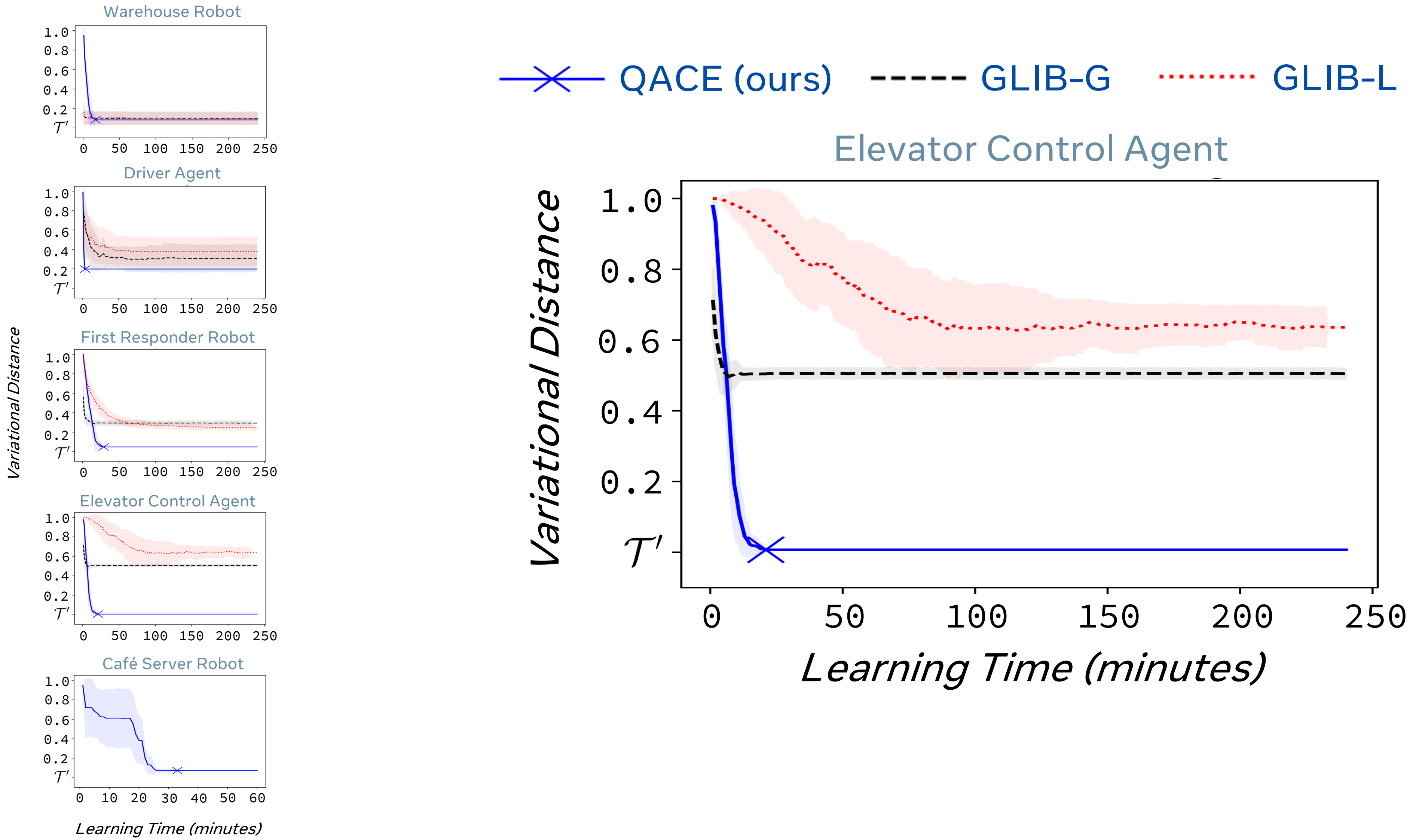
Empirical Evaluation

- 5 SDMAs: Café Server Robot (using OpenRave), Warehouse Robot, Driver Agent, First Responder Robot, Elevator Control Robot.

- Compared model accuracy in terms of variational distance with GLIB (Chitnis et al., AAAI 2021).

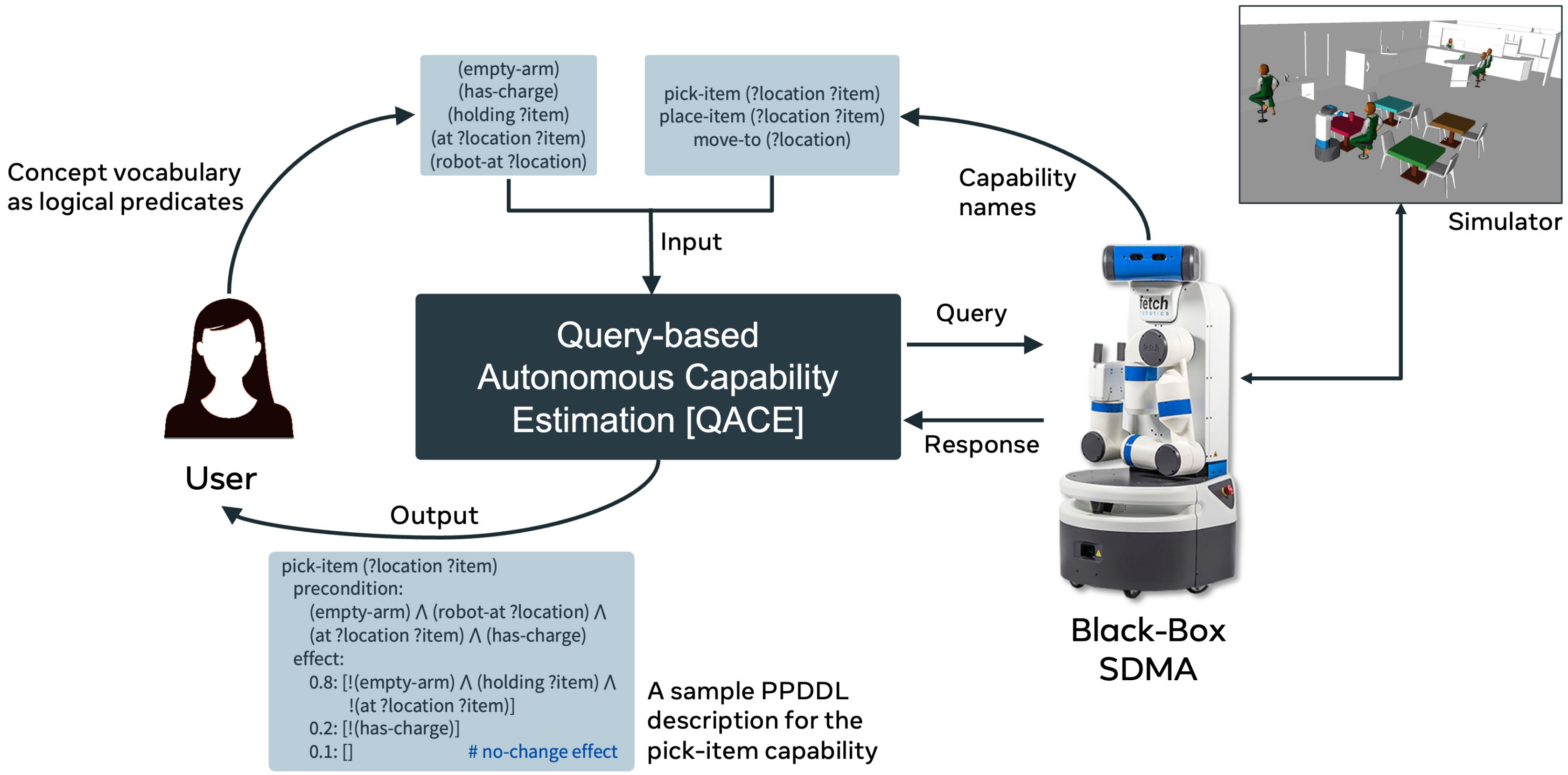
- Variational Distance = $\frac{1}{|D|}\sum_{d \in D} \mathbb{1}_{[s' \neq c_M(s)]}$, where
  - $d = \langle s, c, s' \rangle$
  - $c_M(s) =$ sample the transition using the capability in the model M.

QACE learns accurate probabilistic models faster



Autonomous Capability Assessment of Black-Box Sequential Decision-Making Systems

Pulkit Verma, Rushang Karia, Siddharth Srivastava



AAIR Autonomous Agents and Intelligent Robots

Arizona State University

verma.pulkit@asu.edu