Closed-loop Bayesian Semantic Data Fusion for Collaborative Human-Autonomy Target Search

Luke Burks
Graduate Research Assistant

Nisar Ahmed
Assistant Professor

International Conference on Information Fusion

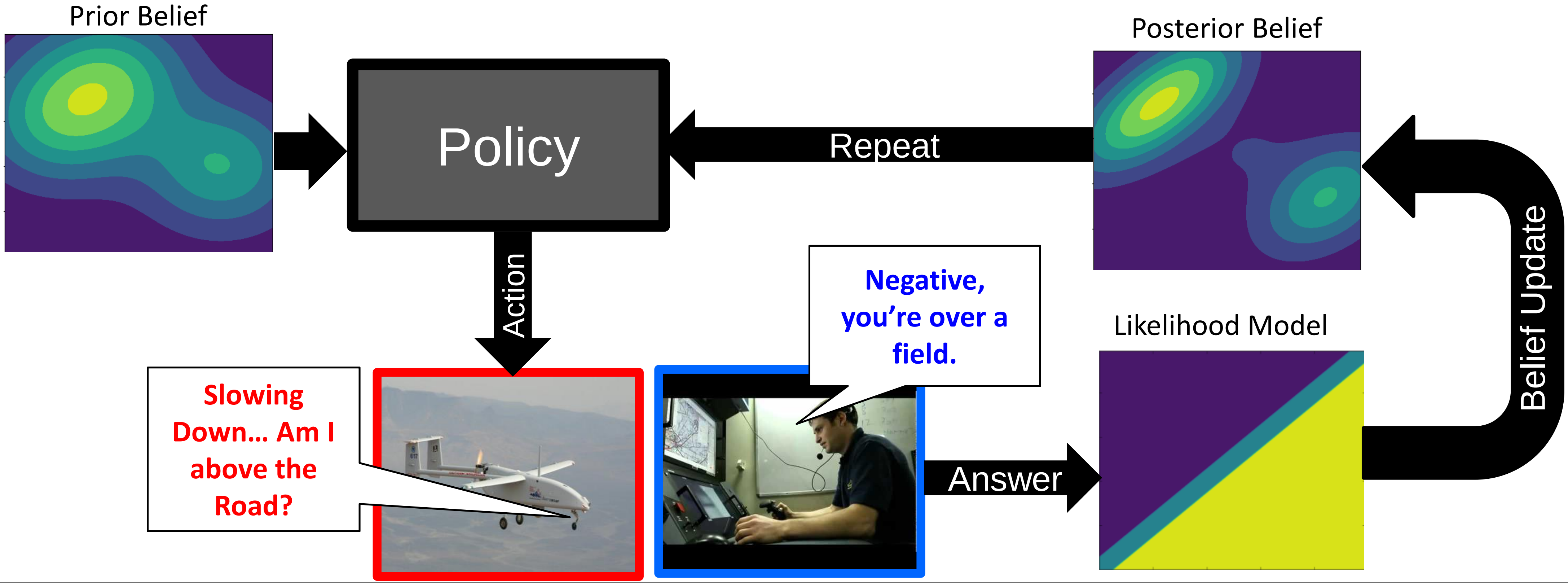
July 13, 2018

Cooperative Human-Robot Intelligence Laboratory
Ann and H.J. Smead Aerospace Engineering Sciences
University of Colorado at Boulder

COHRINT

Research Vision

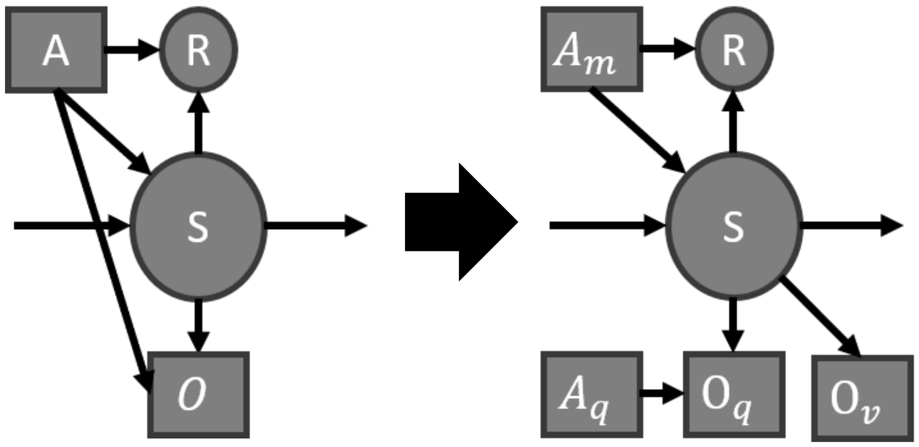
- Treat humans as taskable information providers for autonomous robots
- Formally integrate semantic human observations into tightly coupled optimal sensing and planning under uncertainty

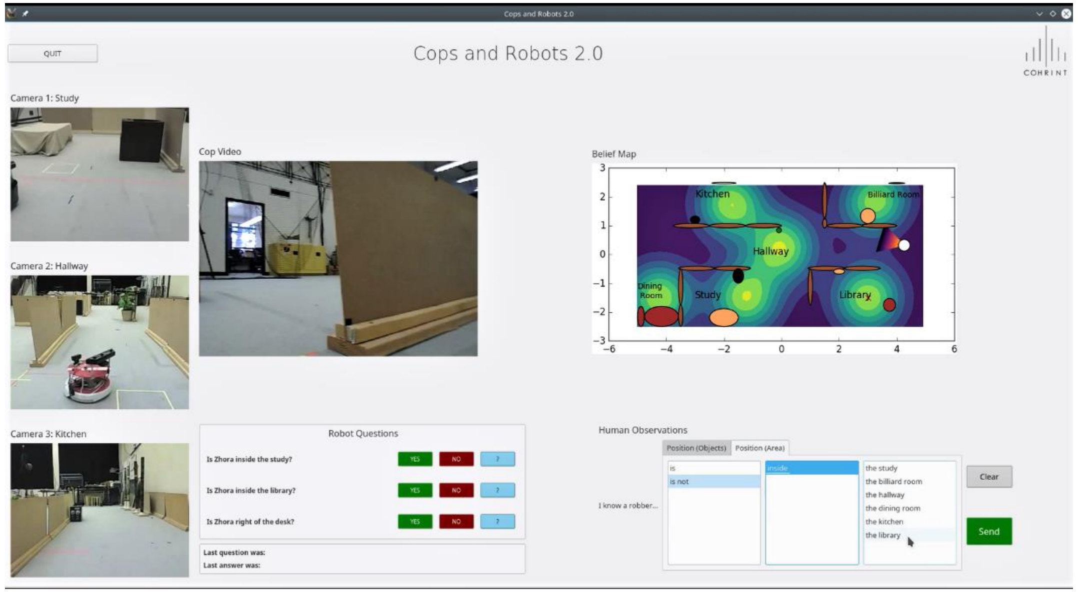


Cops and Robots Platform

Contributions of this work

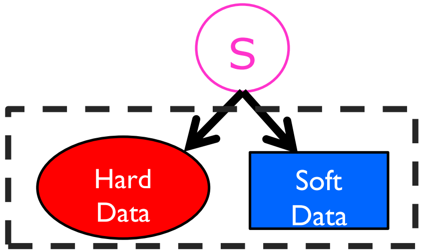
- Proposed Hierarchical Continuous State POMDPs (CPOMDPs)
  - Scale semantic planning to real-world sized problems
  - Combines movement and question planning





- Implemented on the Cops and Robots (CNR) platform
  - Physical robots with real human data
  - Requesting and evaluating human observations

Collaborative Bayesian Data Fusion for Target Localization



Large white truck is nearby the trees, heading North…

“Human push” (Supervisory Sensing)

[Kaupp, et al, FUSION 2005; Bourgault, et al, IROS 2008]

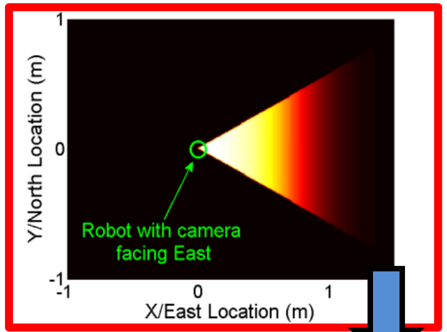
[Ahmed and Campbell, ICRA 2010; Ahmed, et al., T-RO 2013; Sweet and Ahmed, ACC 2016]

“Robot pull” (Active Sensing)

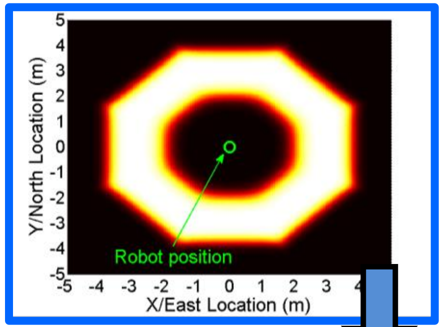
[Kaupp, et al., RAS 2010; Lore, Sweet et al., ICCPS 2015] (**VOI**, decoupled from planning)

Surveillance area

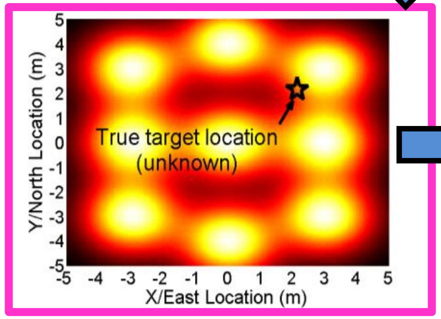
UA Sensor Model for “Detection”



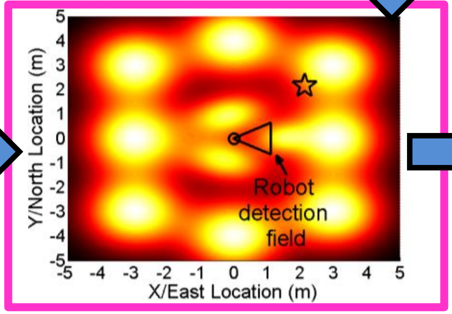
Sensor Model for “Nearby Trees”



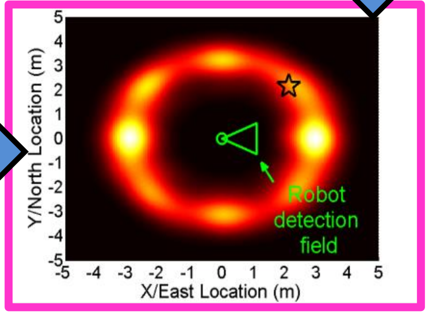
Target detected, range 500 m, bearing 36 deg



p(S)
Initial target state prior (GM)



Revised belief after (GM)

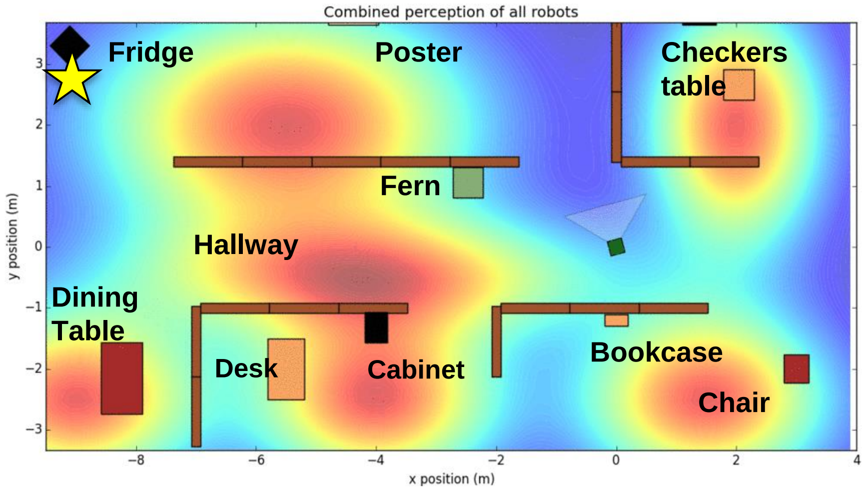


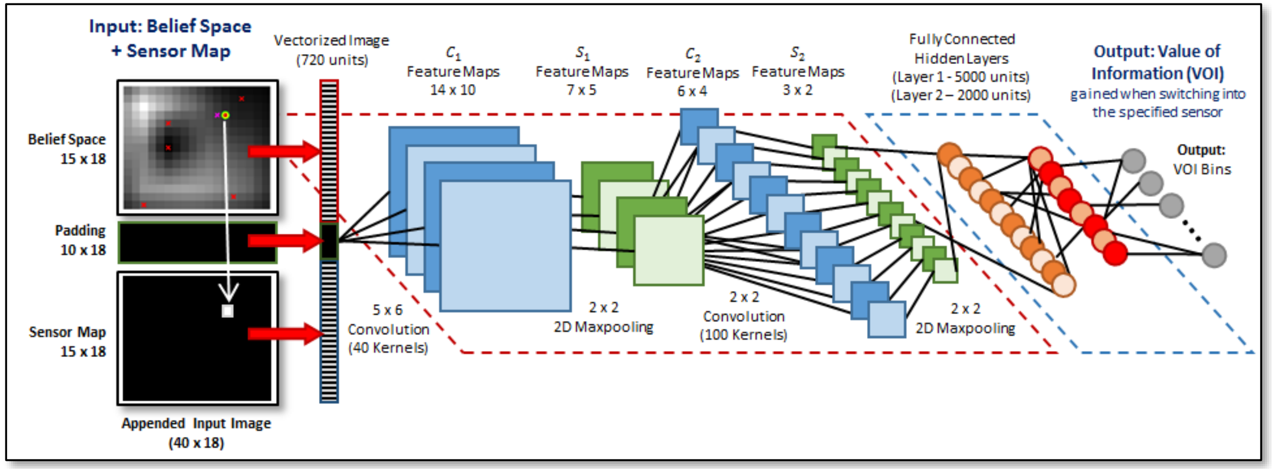
Revised belief after (GM)

Massive information gain → smarter autonomous robotic decision-making (ideally)

Related Work and Issues: Semantic Sensing and Planning

- Decoupled Planning and Control with VOI
  [Sweet, Ahmed, ACC 2016]
  – Doesn't account for information gathering actions
- Discretized State Space Partially Observable Markov Decision Process (POMDPs)
  [Kurniawati, Hsu, Lee, RSS 2008]
  – Difficult to scale to larger state spaces
- Deep Learning
  [Lore, et al., ICCPS 2016]
  – Requires large amounts of training data
- Online POMDP
  [Silver, Veness, NIPS 2010]
  – Brittle in scenarios with distant/sparse rewards
- CPOMDPs using Gaussian Mixtures
  [Porta, et al., IJCAI 2011]
  – Avoids issues above
  – Difficult to specify observation models with GMs, GM explosion





Active Semantic Sensing with Continuous State POMDPs (CPOMDPs)

[Burks and Ahmed, CDC 2017]

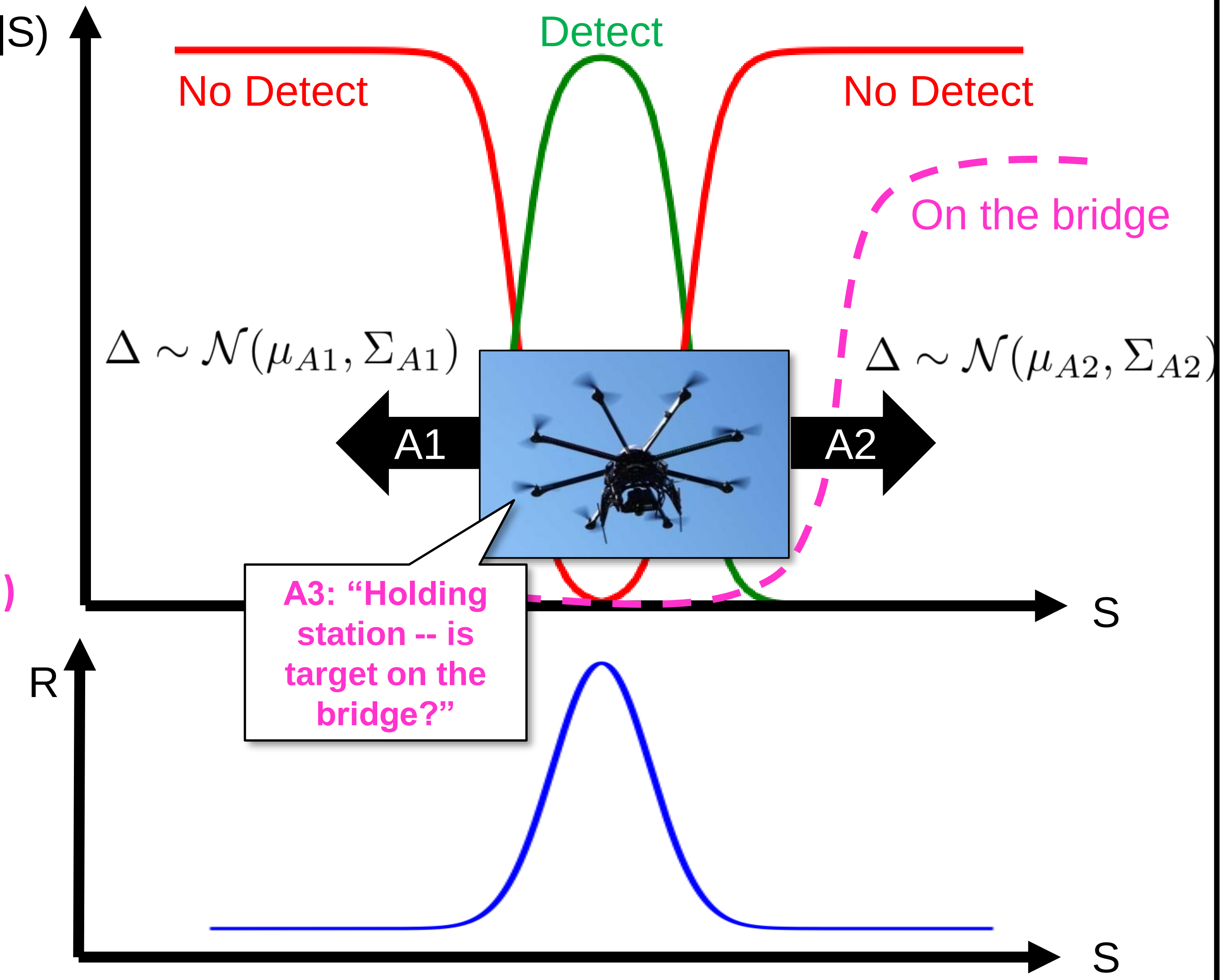
$$State = \mathbf{S}$$
$$= (-\infty, \infty)$$

$$Observations = \Omega$$

$$Actions$$
$$= \{A1, A2, A3, ...\}$$

(can include semantic queries to human sensor!)

Rewards based on proximity to target + action costs



POMDP solvers find policies π to map beliefs *b* to actions:

$$b = p(\mathbf{S}|\Omega, a)$$

$$\pi(b) \rightarrow a$$

Optimal policies maximize discounted total expected reward over time:

$$E[\sum_{t=0}^{\infty} \gamma^t r_t]$$

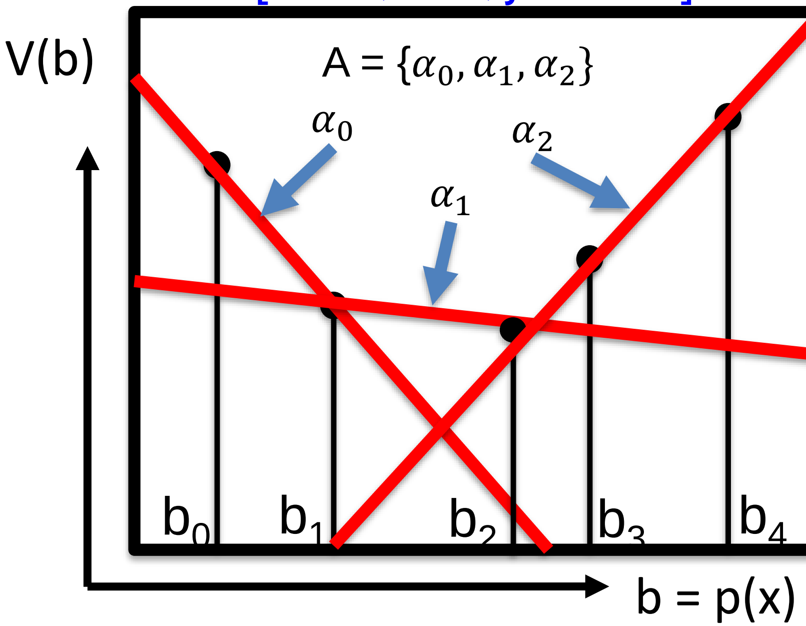
CPOMDPs: Approximate Value Iteration in Belief Space

Alpha Element Backup for Point-Based Value Iteration

$$\alpha^i_{a,o}(s) = \int_{s'} \alpha^i_{n-1}(s')p(o|s')p(s'|s,a)ds'$$

$$\alpha^i_n(s) = r_a(s) + \gamma \sum_o \arg\max_{\alpha^i_{a,o}}(< \alpha^i_{a,o}, b >)$$

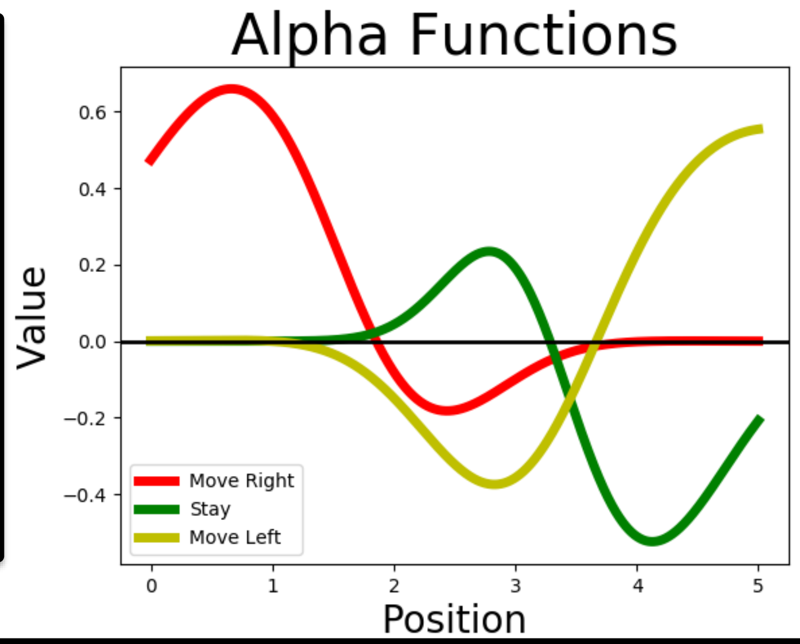
PBVI-type solution on discretized space x
with α-vectors for policy π
[Pineau, et al., JAIR 2006]

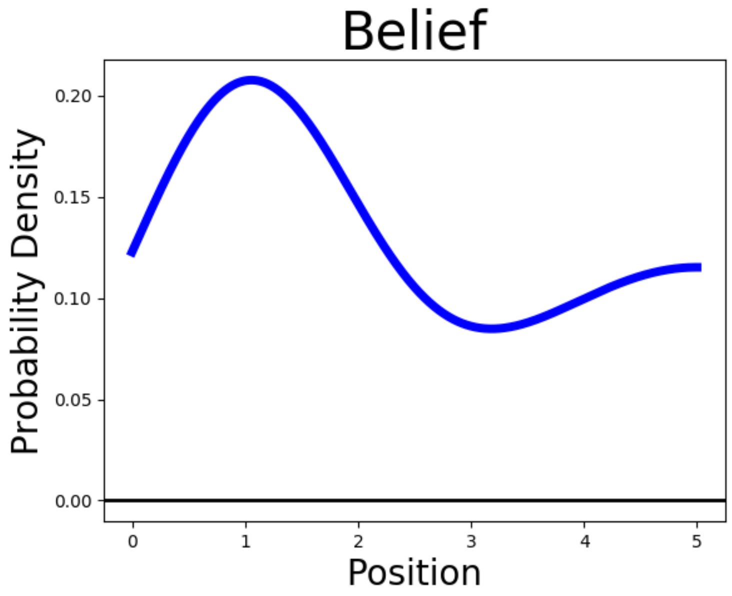


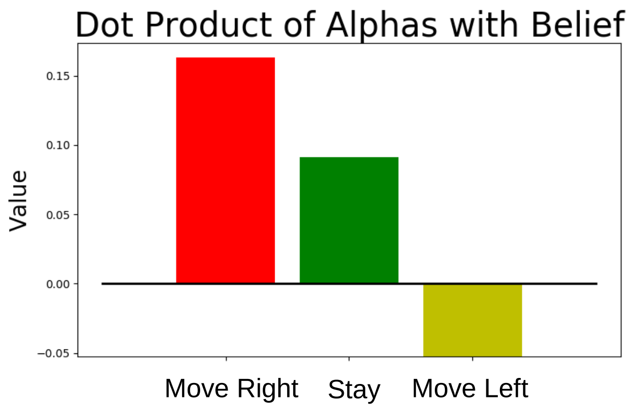
Gaussian Mixture (GM)

$$\sum_j^J w_j \phi(s|\mu_j, \Sigma_j)$$

Can represent arbitrary policy functions & pdfs







ptimal
tion
ven
lief

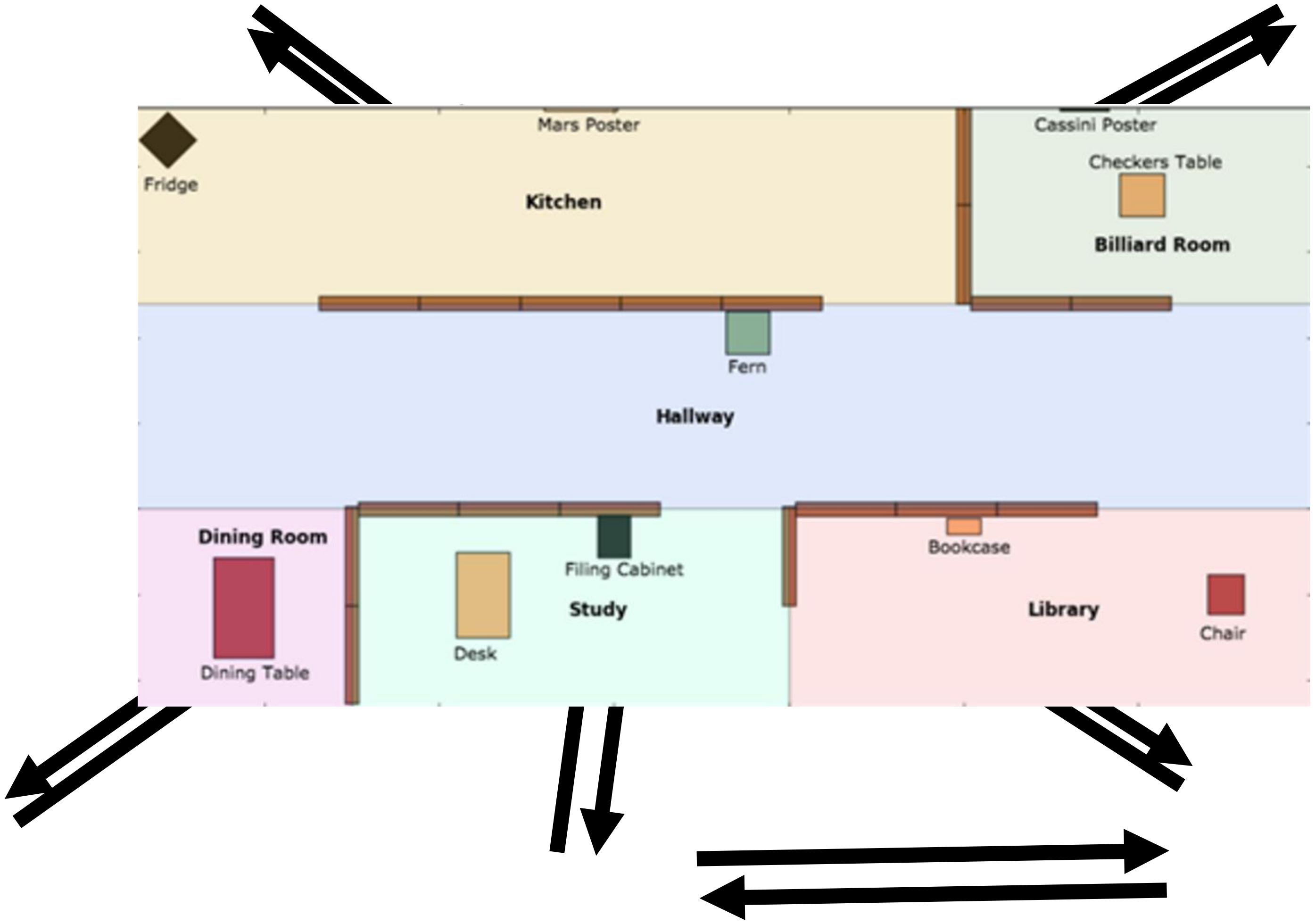
Hierarchical CPOMDPs

- Multiple transition modes (walls)
- Large observation space (many objects)
- Need for high precision beliefs

Less tractable policy solution

Hierarchical CPOMDPs

- Can exploit problem structure
- Rooms each become a bounded continuous space
- Set of rooms forms a higher level discrete state space
- **A Discrete POMDP that chooses among Continuous POMDPs**



The Value of Questions

$$VOI(o) = (\sum_{o_i \in O} p(o=o_i)[max_a \int p(s|o=o_i)R(a,s)ds]) - max_a \int p(s)R(a,s)ds$$

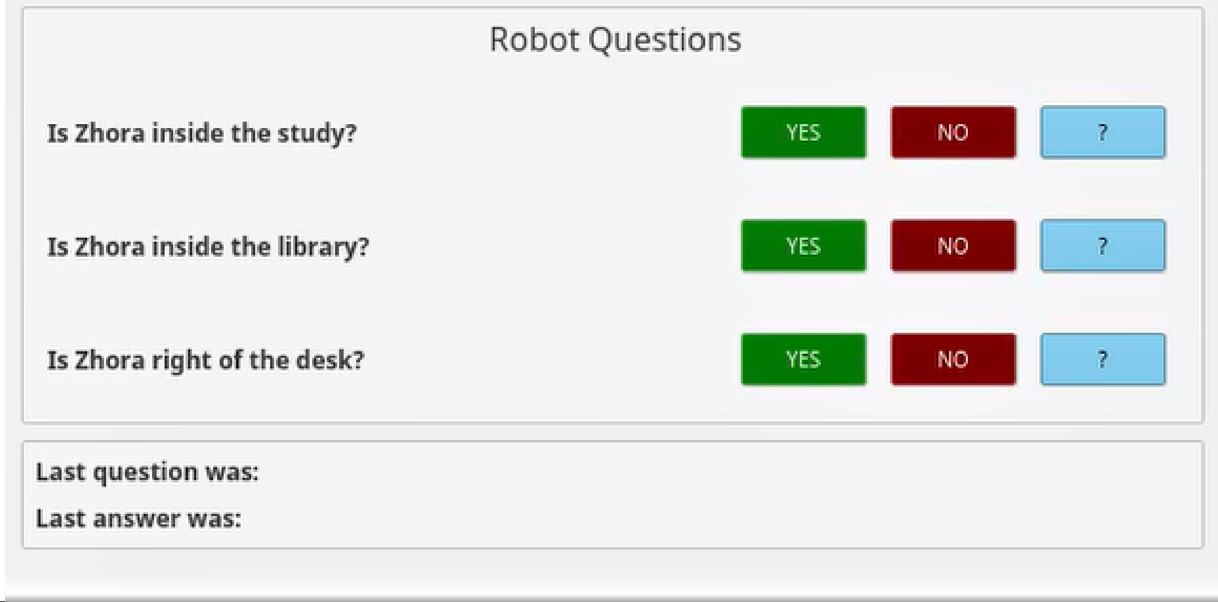
Value of Human Observation (o)

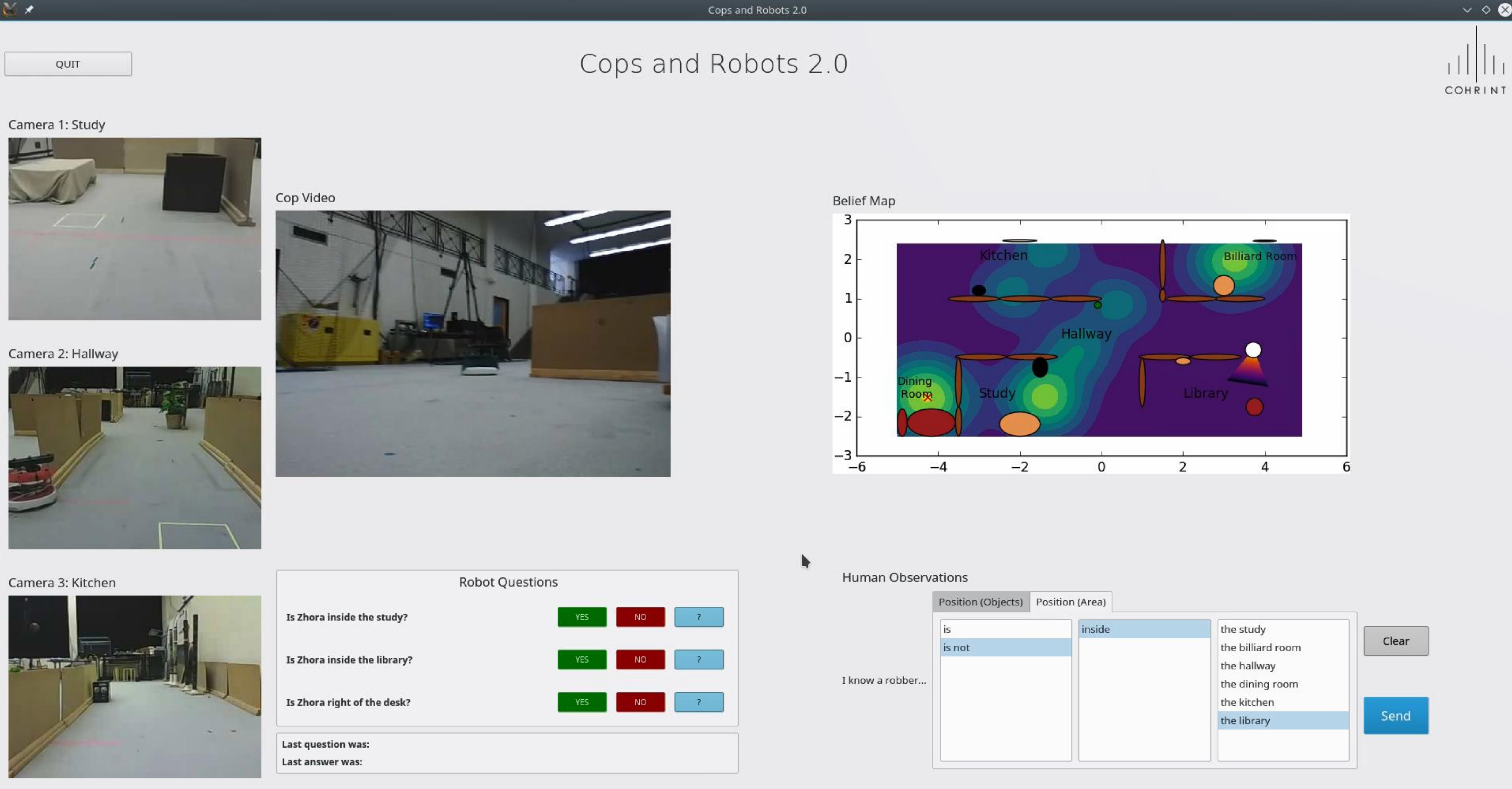
Expected Reward After Observation

Expected Reward without Observation

- POMDPs implicitly find VOI during policy solution
- Can extract 2nd, 3rd, …, Nth most valuable question for minimal additional computation

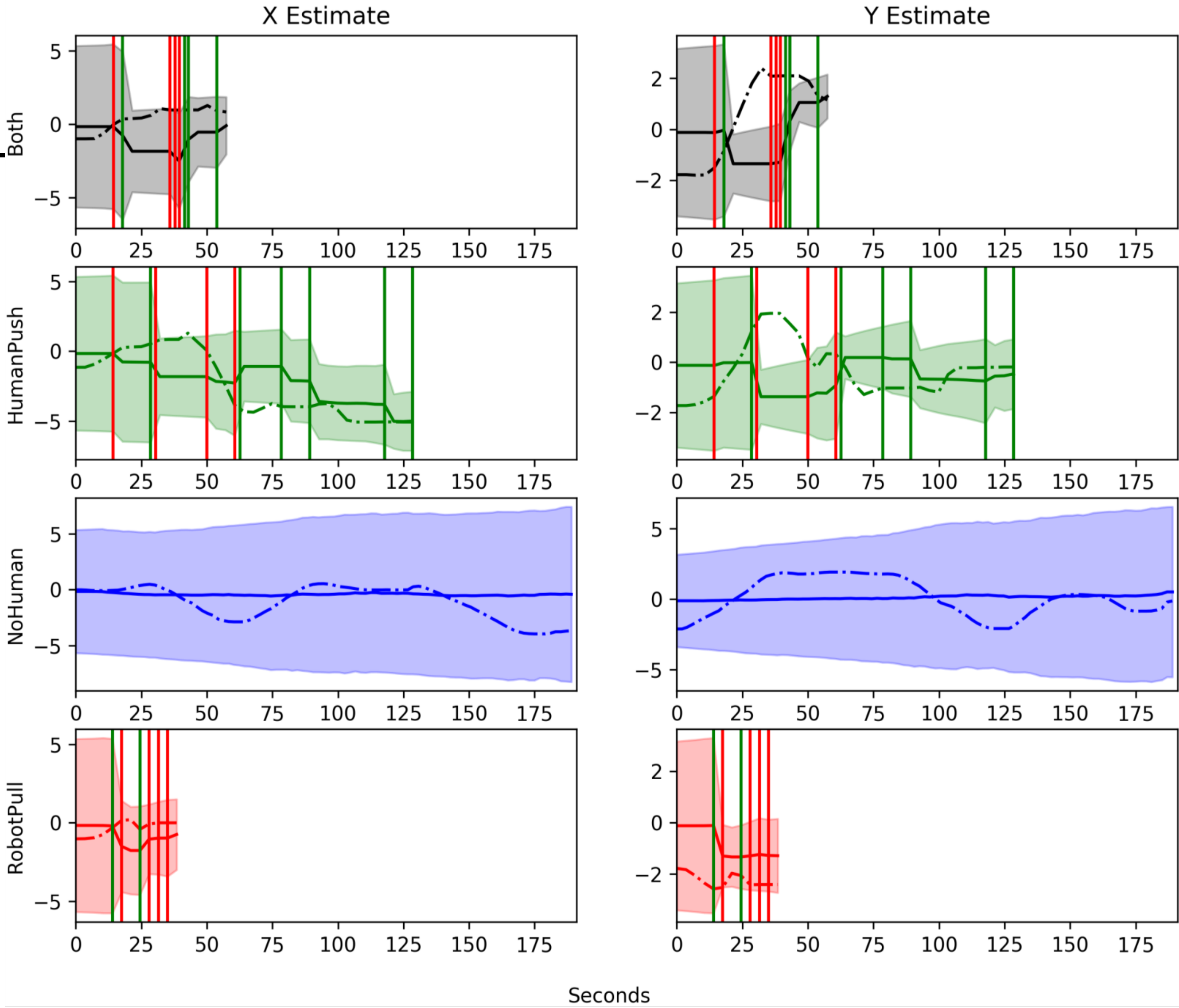
Most Valuable Questions



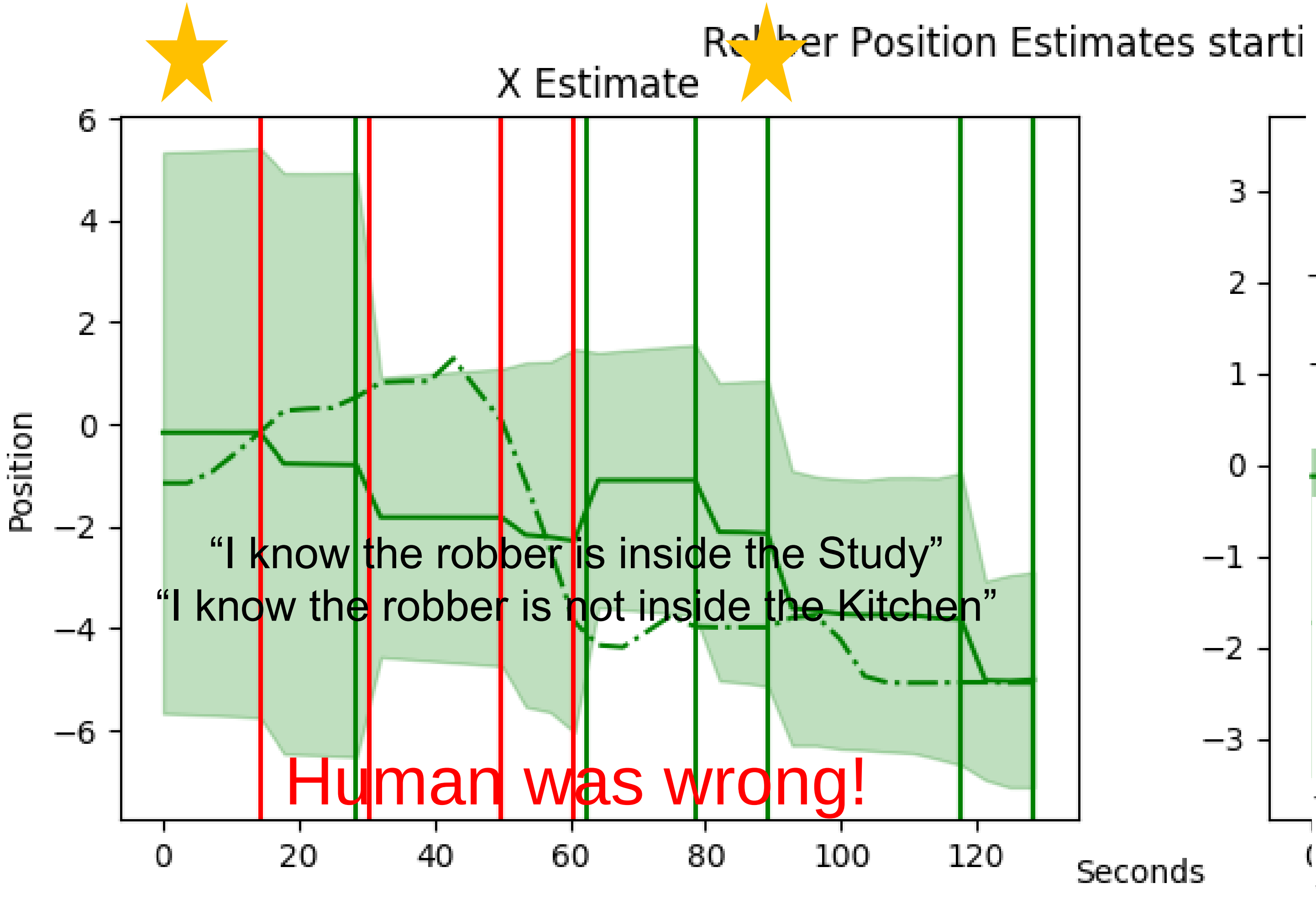

Cops and Robots 2.0
QUIT
Cops and Robots 2.0
COHRINT
Camera 1: Study
Camera 2: Hallway
Camera 3: Kitchen
Cop Video
Belief Map
Kitchen
Billiard Room
Hallway
Dining Room
Study
Library
3
2
1
0
−1
−2
−3
−6
−4
−2
0
2
4
6
Robot Questions
Is Zhora inside the study?
Is Zhora inside the library?
Is Zhora right of the desk?
YES
NO
?
Last question was:
Last answer was:
Human Observations
Position (Objects)
Position (Area)
I know a robber...
is
is not
inside
the study
the billiard room
the hallway
the dining room
the kitchen
the library
Clear
Send

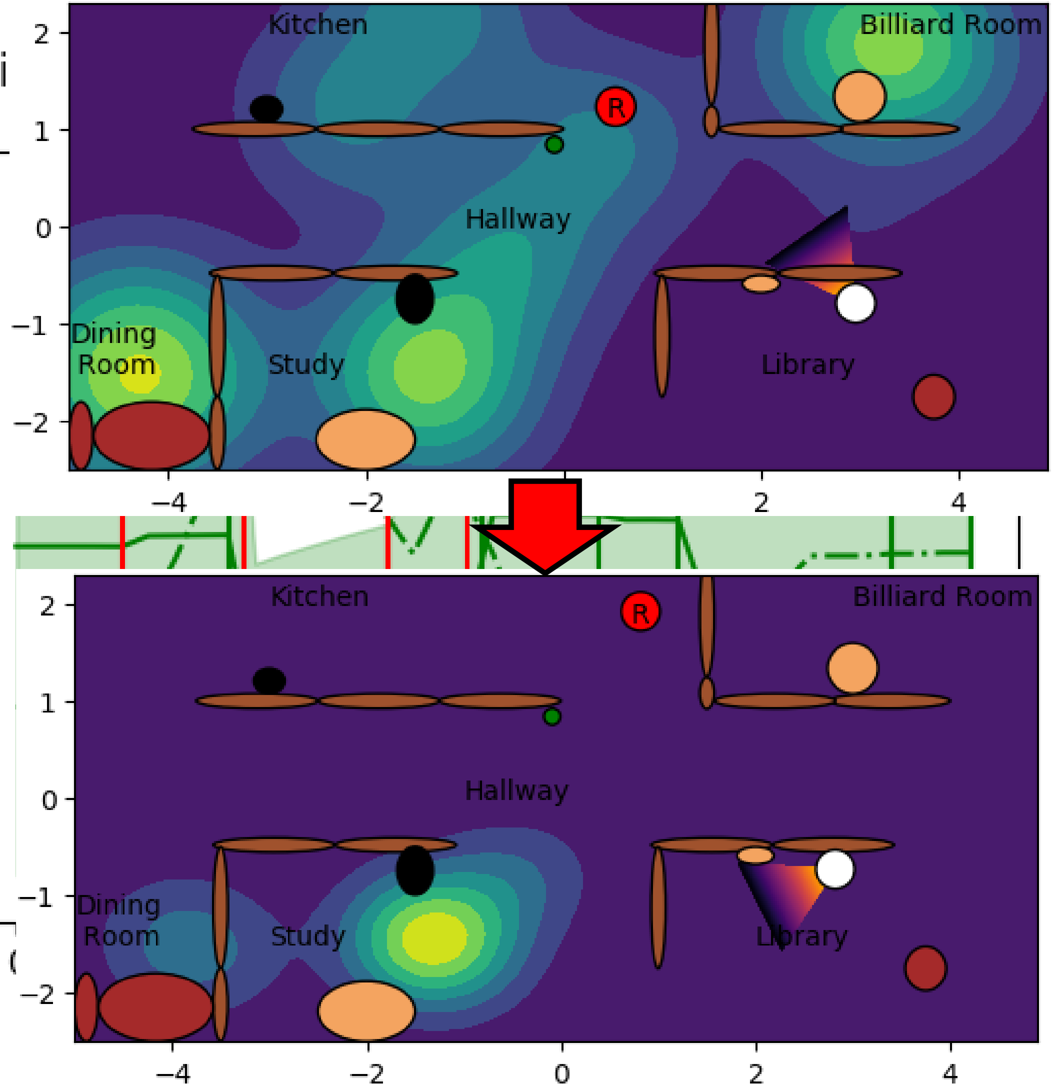
Hardware Simulations

- Two maps: Familiar and Unfamiliar to human subject
- Three Starting Configurations for Familiar, two for Unfamiliar
- Four Information Scenarios:
  - No Human Information
  - Only Human Push
  - Only Robot Pull
  - Both Push and Pull

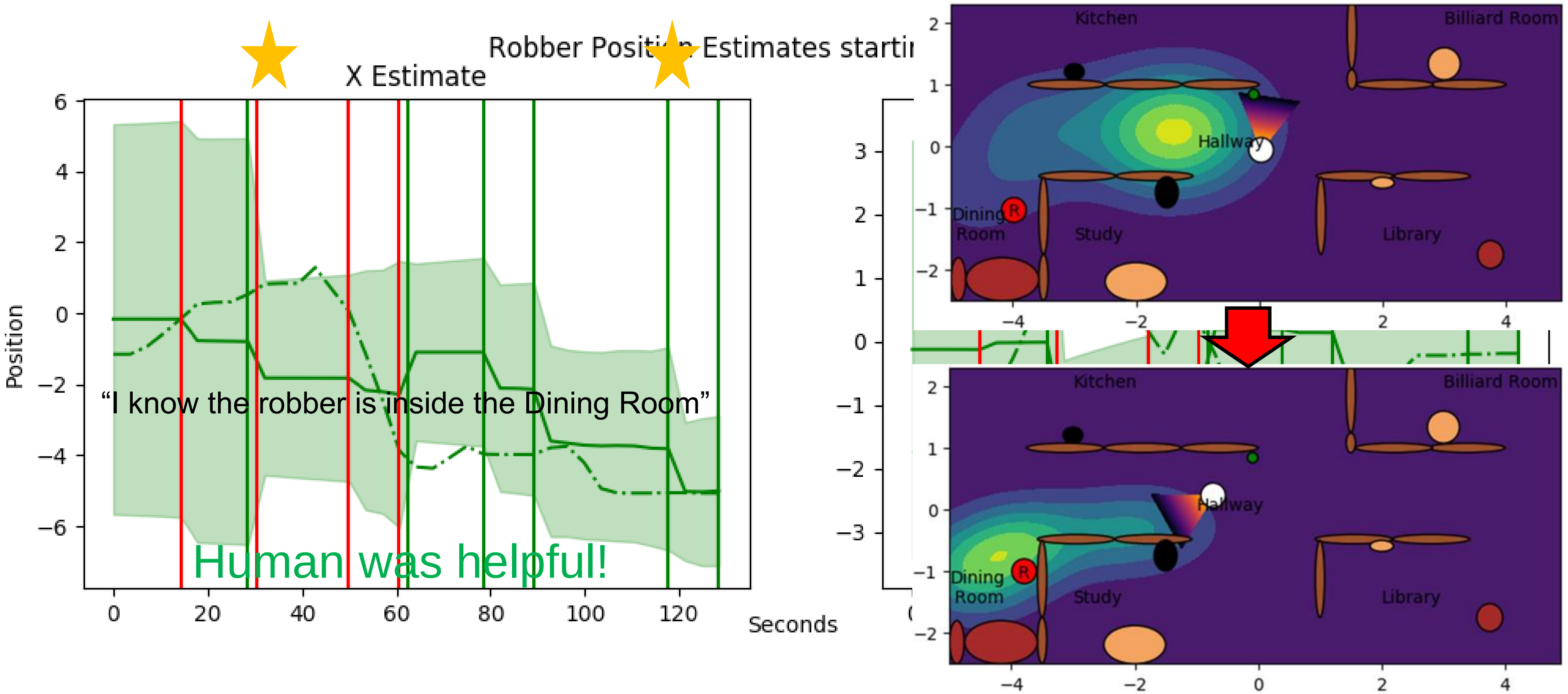


Misleading Human Input

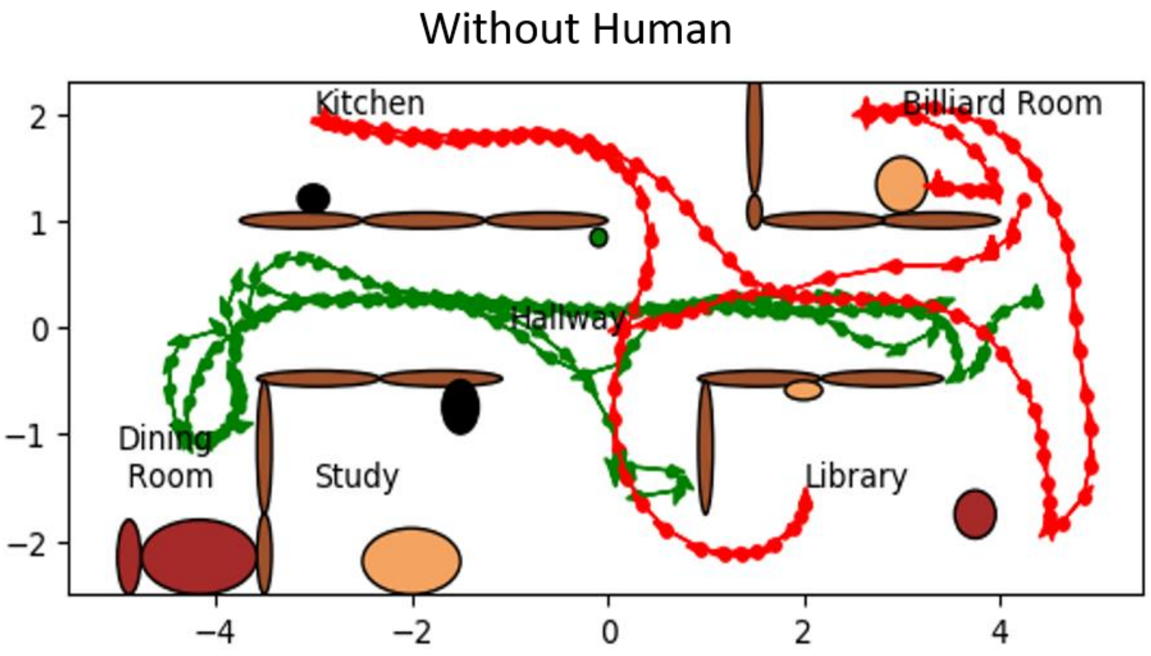




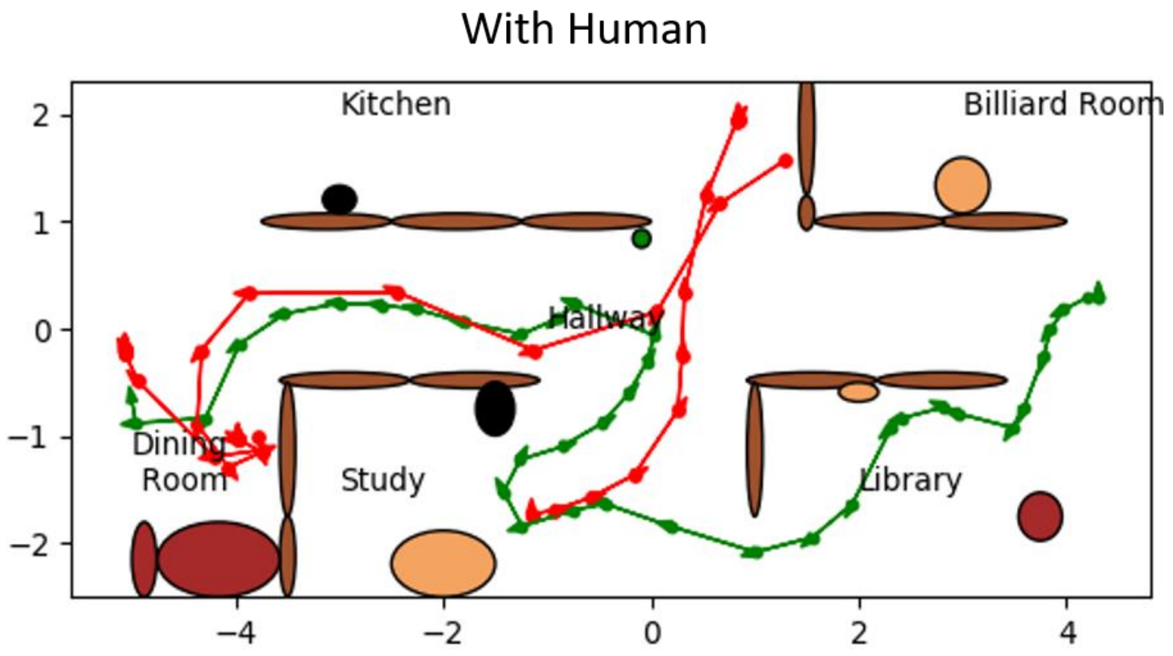
Helpful Human Input



Pursuit Behavior



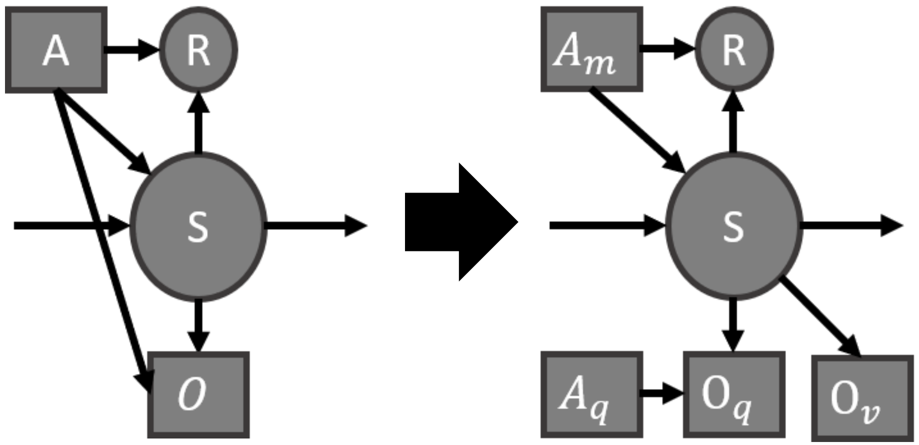
- Patrols Hallway, briefly surveys rooms
- Catches robber crossing hallway

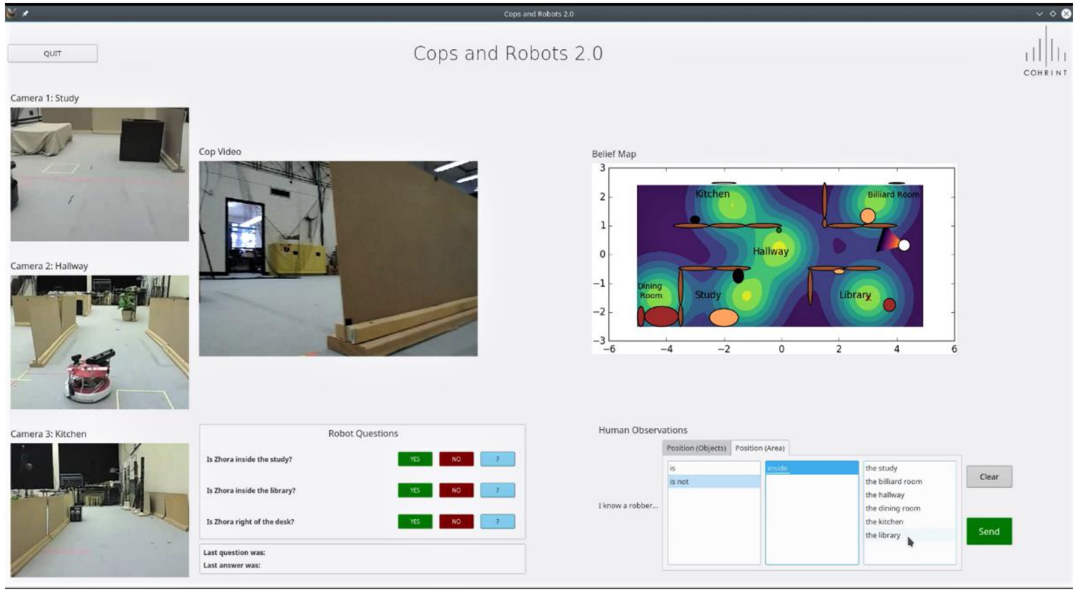


- Moves from room to room
- Pursues and corners robber

Conclusion

- Hierarchical CPOMDPs enhance collaborative tracking
  – Optimally exploits human information
  – Applies to realistic scale problems

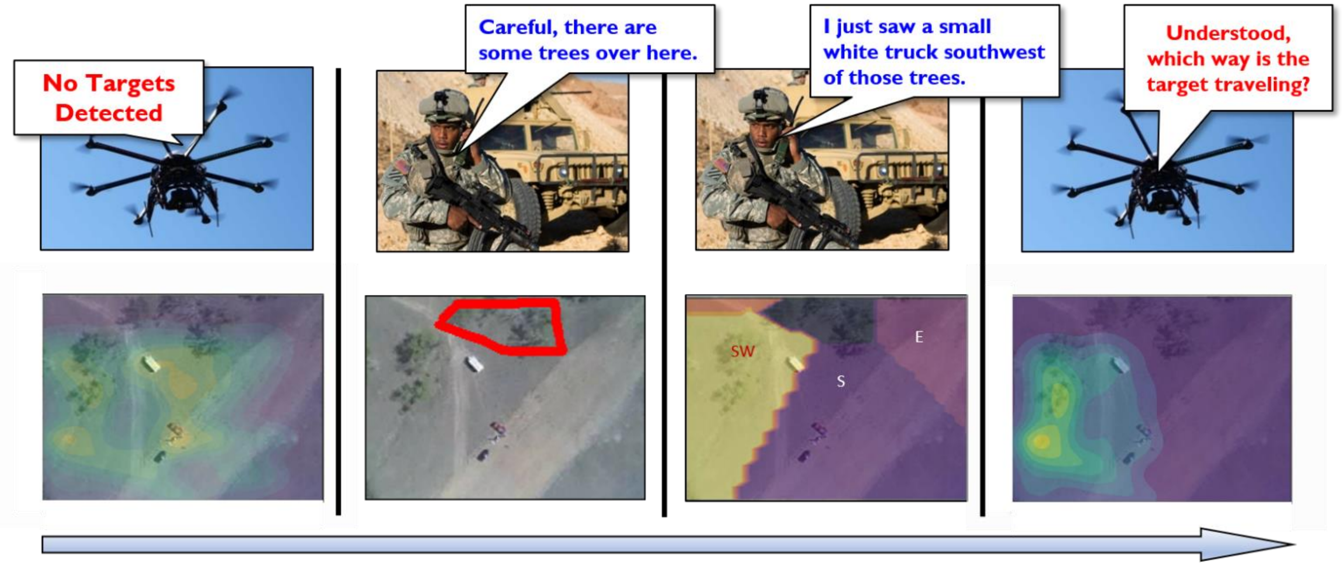


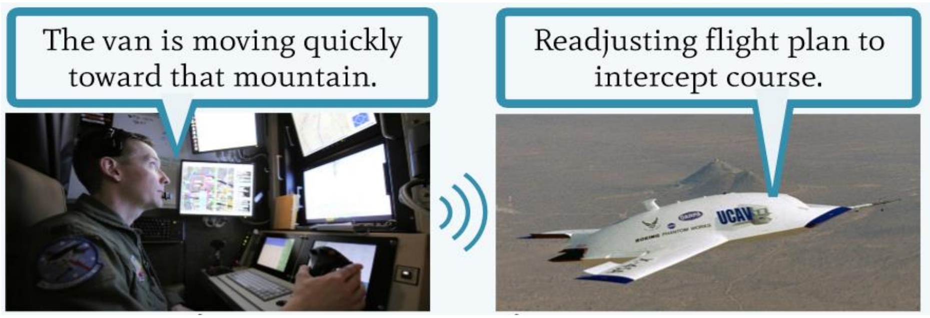


- Hardware implementation show advantage of **combined** human-robot sensing
  – Robust to occasional human errors
  – Human information significantly influences policy behavior
  – Generally consistent belief estimates

Open Directions

- Relax assumption of known structure
  - Use human to impose structure
  - Build likelihood models from sketches
  - Adapt planning dynamically





- Incorporating additional state information
  - Target Velocity
  - Contextual Conditions
  - Human States

Acknowledgements

COHRINT Team

- **Ian Loefgren:** MS Student
- **Luke Barbier:** BS Student
- **Jeremy Muesing:** MS Student
- **Jamison McGinley:** BS Student
- **Sousheel Vunnam:** BS Student
- **Nisar Ahmed:** Assistant Professor, Ph.D. Advisor

C-UAS
CENTER FOR UNMANNED AIRCRAFT SYSTEMS

A National Science Foundation Industry/University Cooperative Research Center (IUCRC)

Questions?