

Collaborative Semantic Data Fusion with Dynamically Observable Decision Processes

Luke Burks

Graduate Research Assistant

Nisar Ahmed

Assistant Professor

International Conference on Information Fusion

July 3, 2019

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



Motivation: Leveraging Human Data for Robotic Tasks

Autonomy: great for tasks dangerous or inconvenient

Humans = Semantic Sensors

- Volunteering Information
- Answering Questions

If people can benefit from robots, why can't robots benefit from people?



Disaster Assistance



Airborne Search



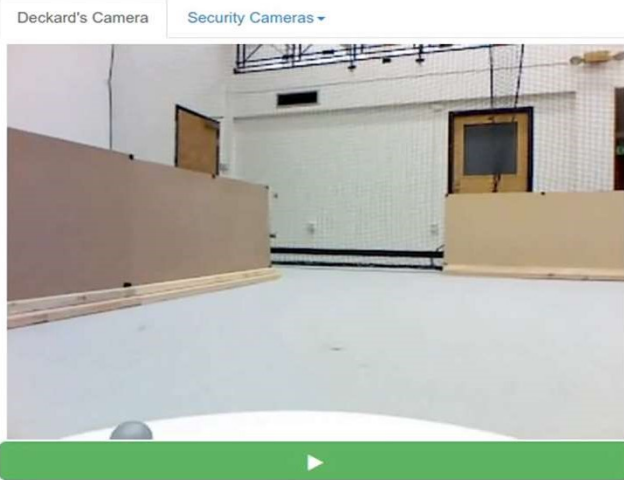
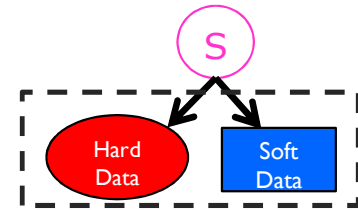
Military Robotics



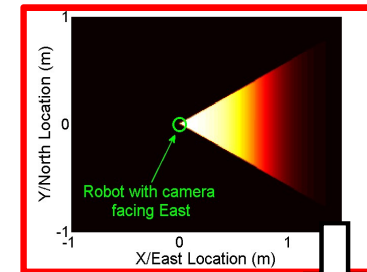
Space Exploration

Related Work: Semantic Sensing and Planning under Uncertainty

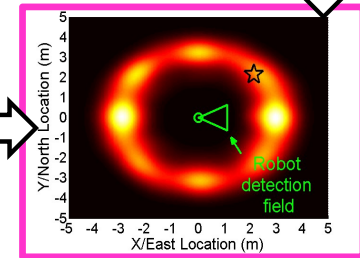
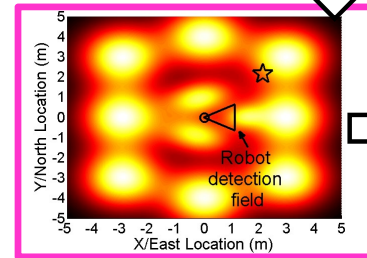
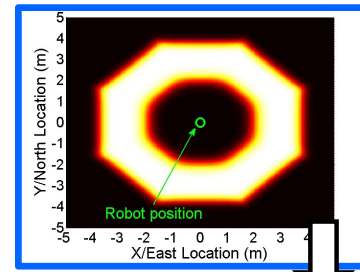
- Offline Continuous Partially Observable Markov Decision Process (CPOMDP) [Burks, Ahmed, FUSION 2018]
 - Requires known static models
- Adaptive Belief Trees [Kurniawati, Yadav, Robotics Research 2016]
 - Requires seed offline policy solution
- Simple Online Value Iteration (SOVI) [Shani, Brafman, Shimony, ECML 2005]
 - Restricted to slowly changing models
- Deep Learning [Lore, et al. ICCPS 2016]



UA Sensor Model for "Detection"



Sensor Model for "Nearby"



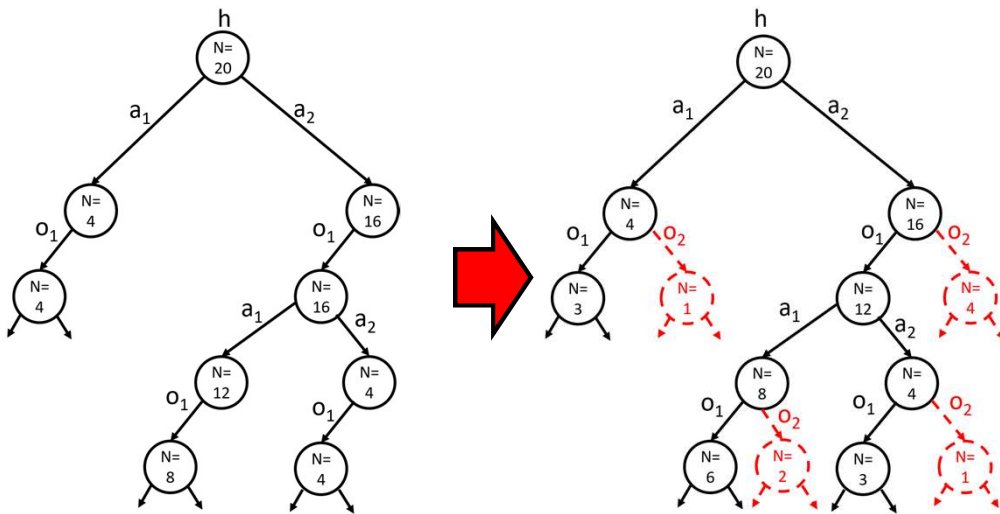
Research Vision

- Treat humans as taskable information providers for autonomous robots
- Integrate dynamic, ad-hoc models from human collaborators into tightly coupled optimal sensing and planning in unknown/dynamic environments



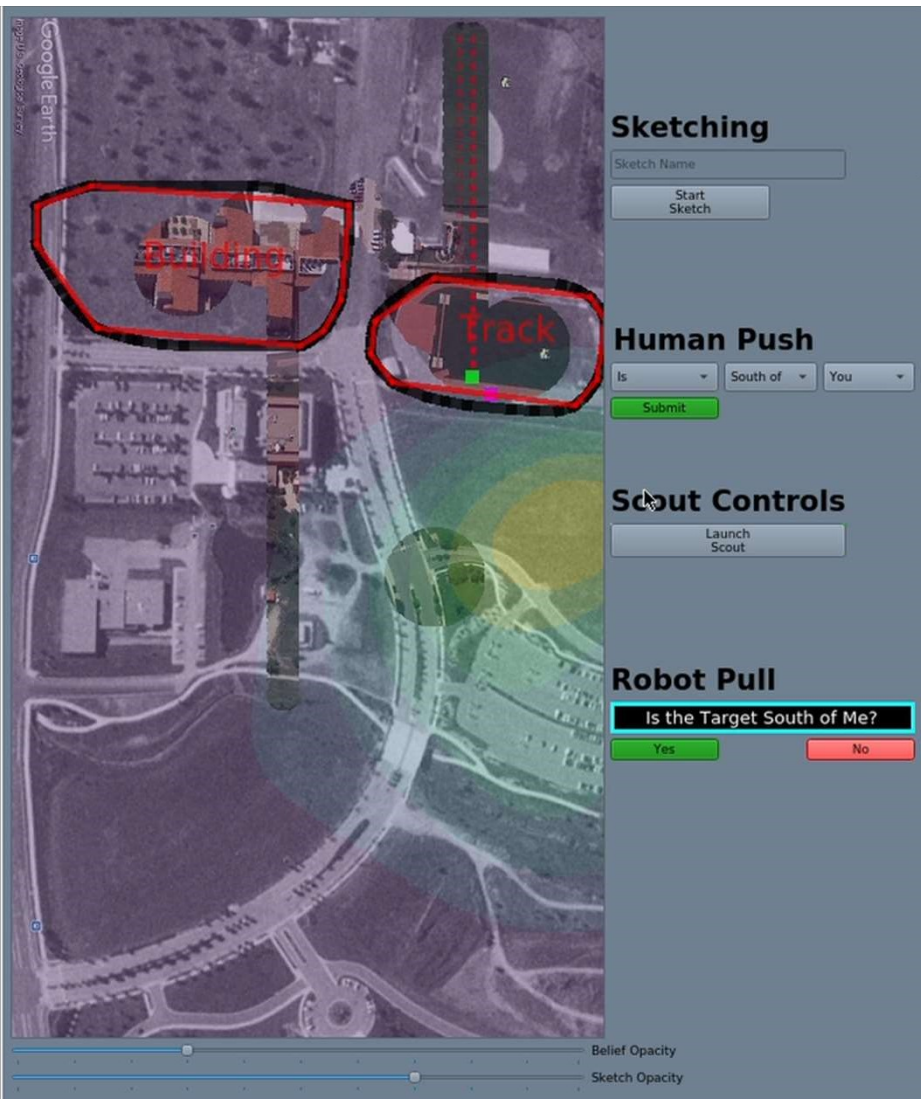
Contributions of this work

- Proposed Dynamically Observable Monte-Carlo Planning (DOMCP)
 - Ad-hoc observation models
 - Dynamically modifying optimal planning



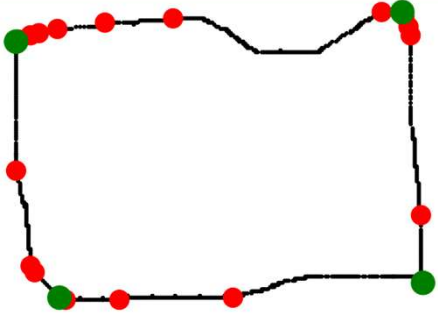
- Sketch based Human-Robot target search
 - Real-time semantic codebook building in unknown dynamic environments
 - Requesting and evaluating human semantic observations

Acquiring Models: Sketch Interface



- Starts with outdated map which updates when explored
- Create new observations with semantic labels and spatial extent
- Volunteer information for fusion
- Human can independently gather information
- Policy combines movement and questions as actions
 - “Move North and ask ‘Is the Target West of the Track’”

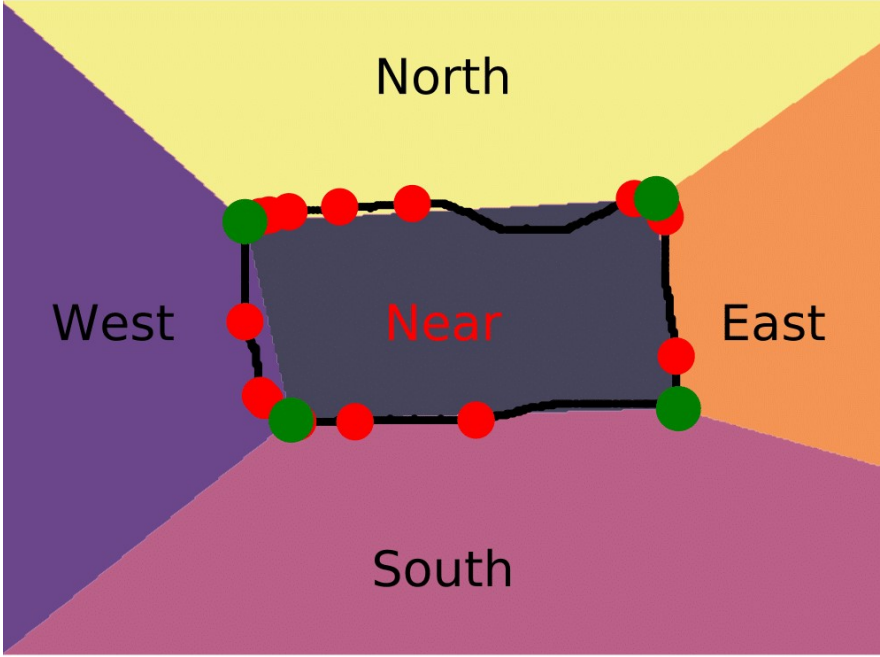
Extracting Models from Sketches



- 1. Build Convex Hull
- 2. Reduce by maximizing contributed angle

Softmax Model

$$p(o_k | s_k) = \frac{\exp(w_o^T s_k + b_o)}{\sum_c^{|\Omega|} \exp(w_c^T s_k + b_c)}$$



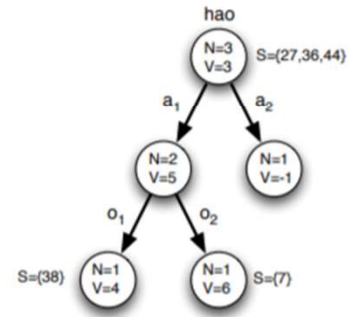
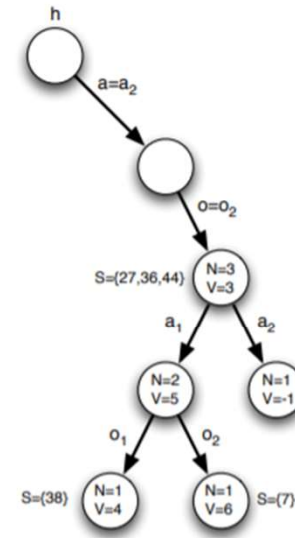
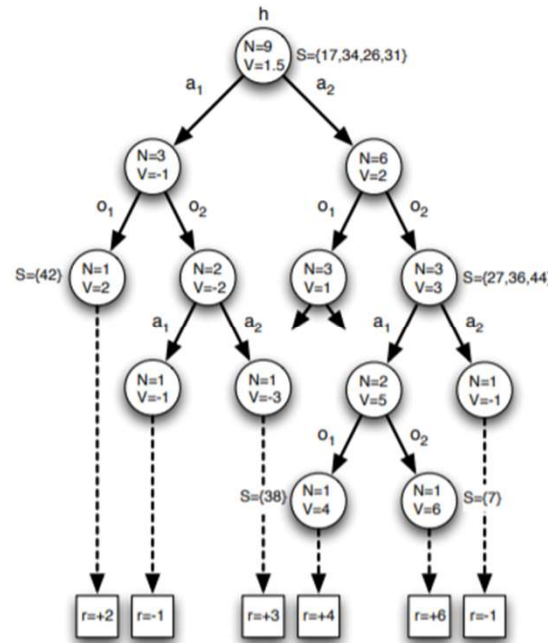
Online POMDPs: Partially Observable Monte Carlo Planning (POMCP)

[Silver, Veness, NIPS 2010]

Based on Monte-Carlo
Tree Search (MCTS)
planning algorithm

Optimizing for Minimum
“Time to Capture”

Uses a **generative model**
to build tree of histories



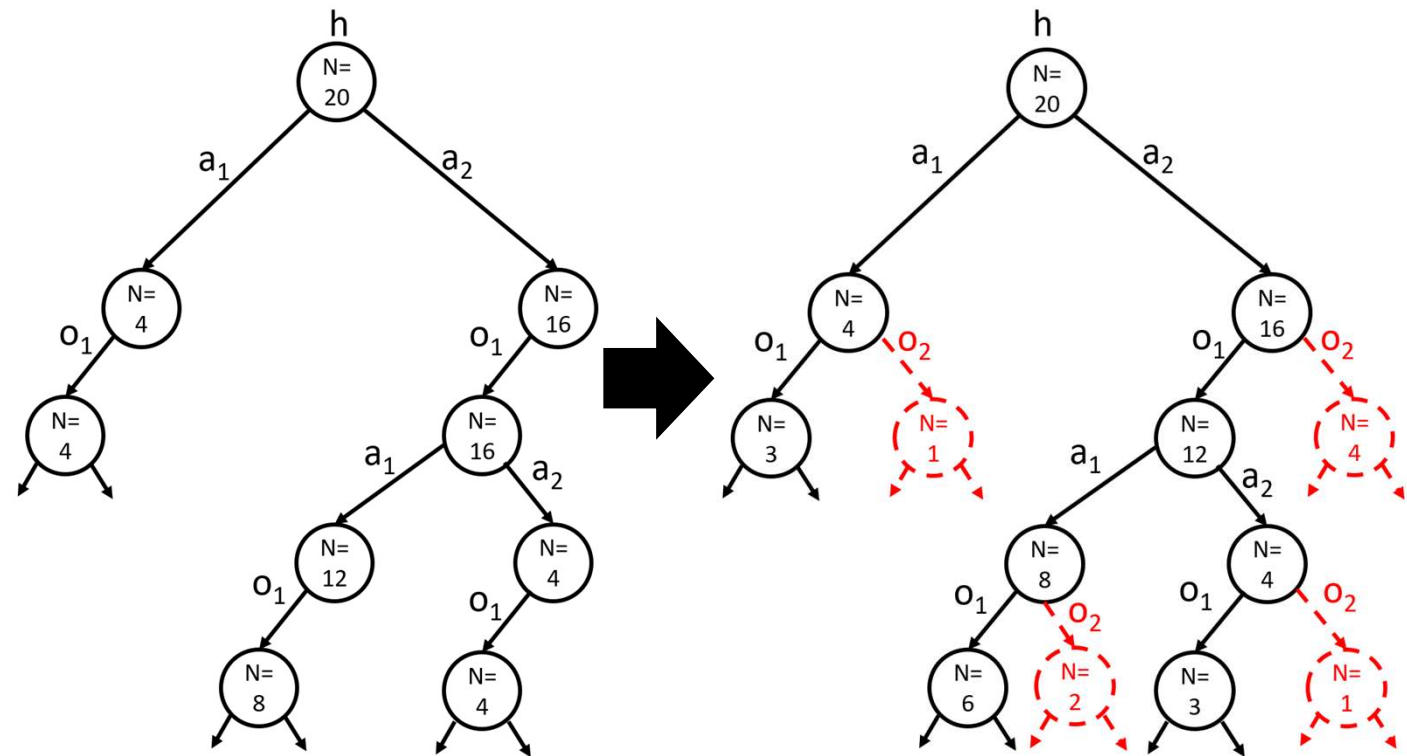
Originally assumed static models



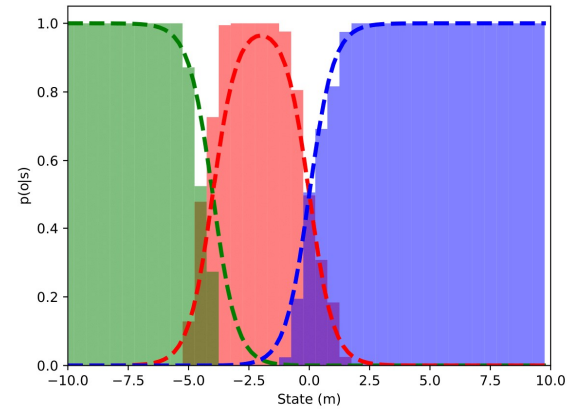
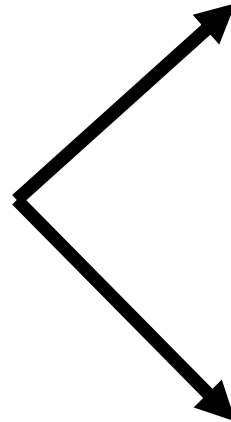
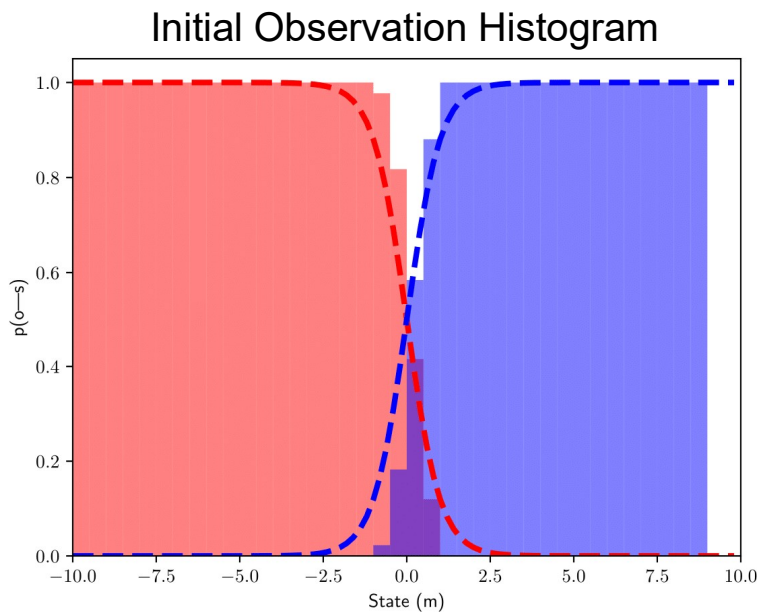
Minor modification accounts for
dynamic models

DOMCP: **Dynamically** Observable Monte-Carlo Planning

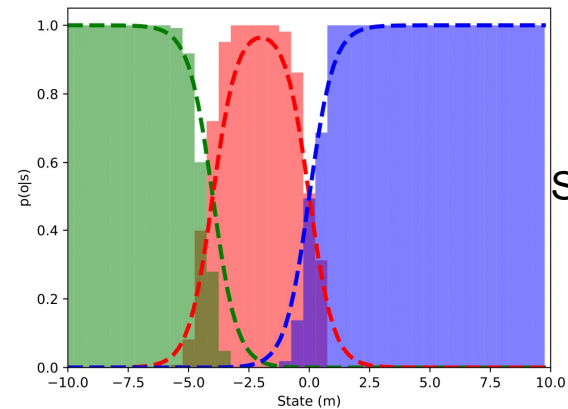
- Modify existing planning tree after model update
- Reallocates particles according to new observation likelihoods
- Retains prior information consistent with new model



DOMCP: **Dynamically** Observable Monte-Carlo Planning

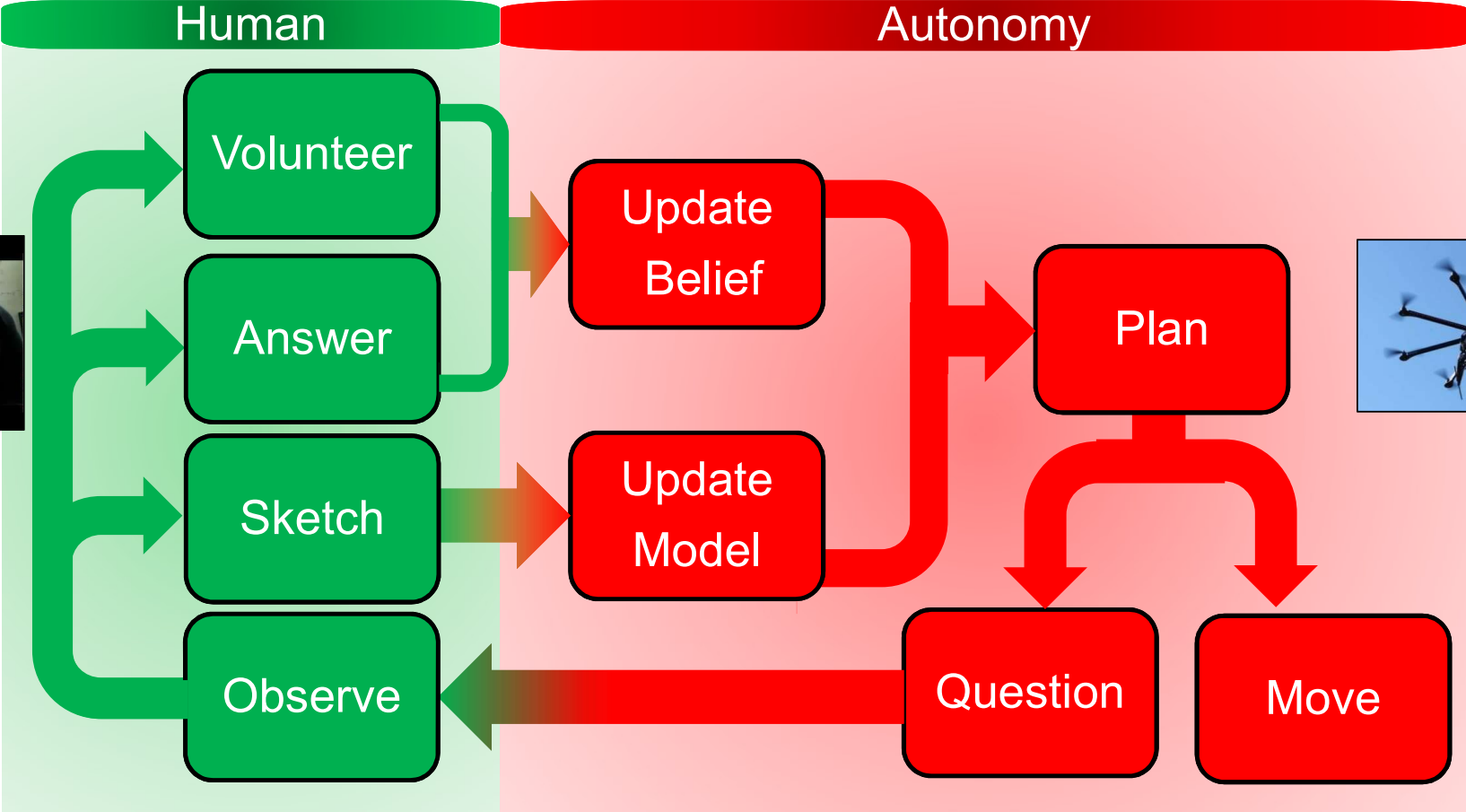


POMCP Replanning
Resimulates every state

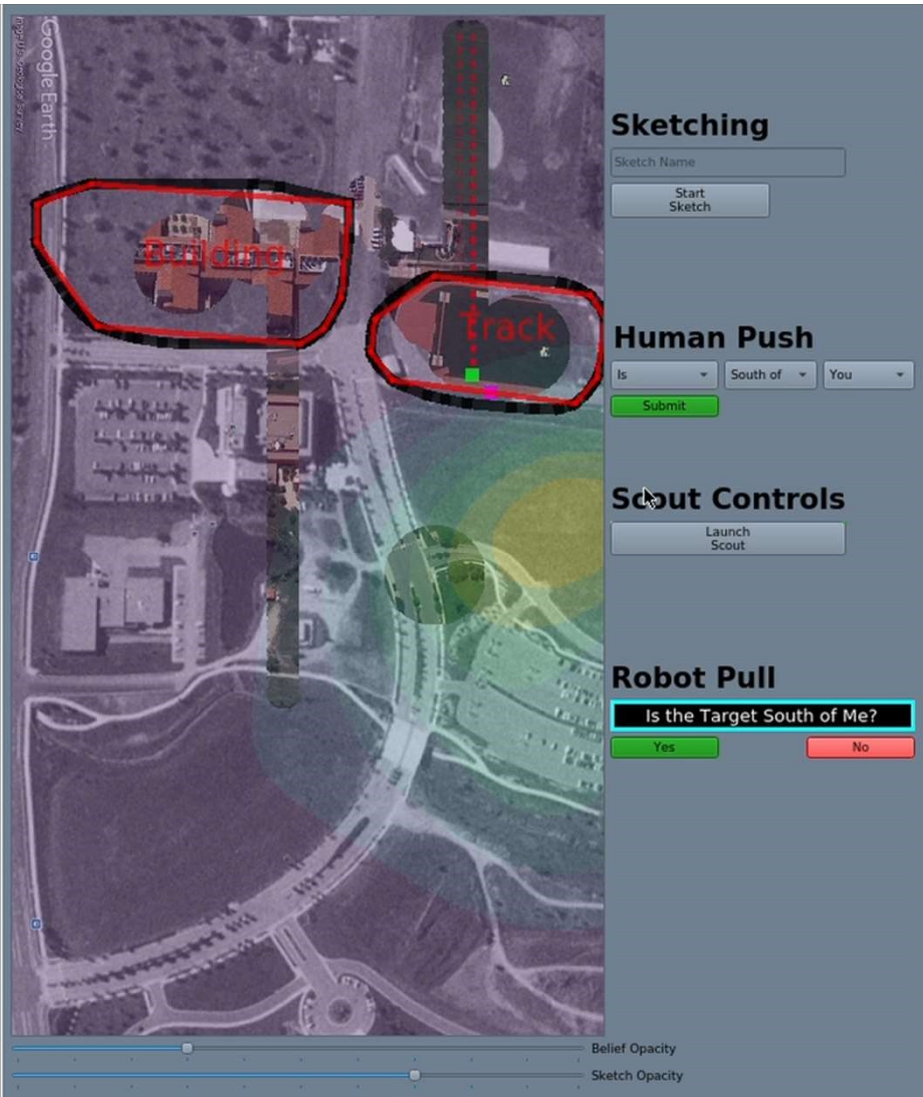


DOMCP Redistribution
Saves 73% of original states

Collaborating in Unknown Environments



Acquiring Models: Sketch Interface

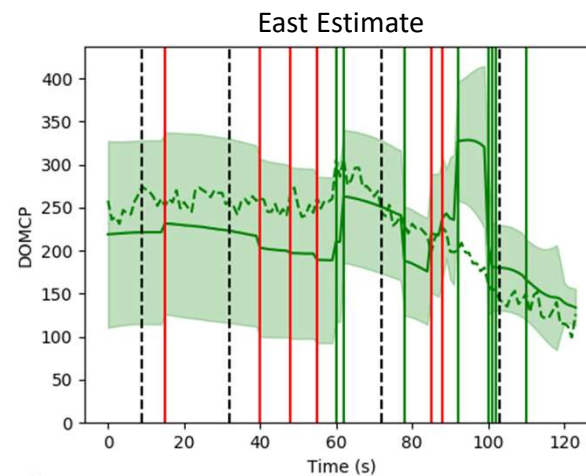


- Starts with outdated map which updates when explored
- Create new observations with semantic labels and spatial extent
- Volunteer information for fusion
- Human can independently gather information
- Policy combines movement and questions as actions
 - “Move North and ask ‘Is the Target West of the Track’”

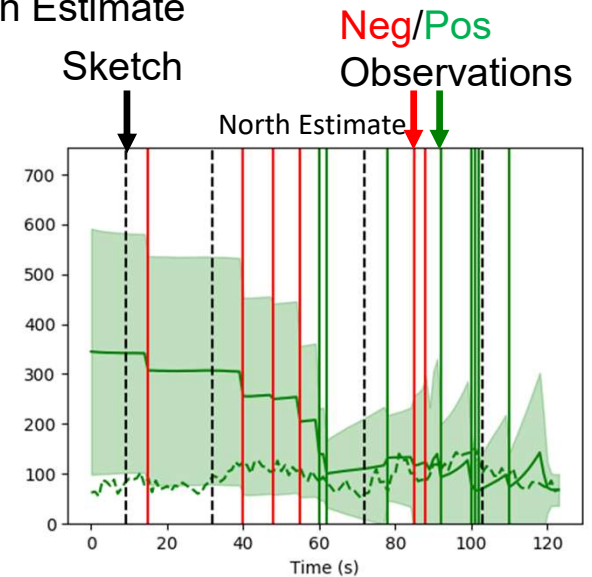
Simulations

- Three Approaches:
MAP, POMCP, DOMCP
- Sketches, volunteer observations, and target positions identical across runs
- N=5 for each method

Target Position Estimate

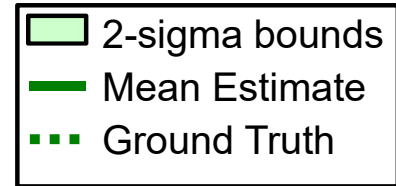


Sketch

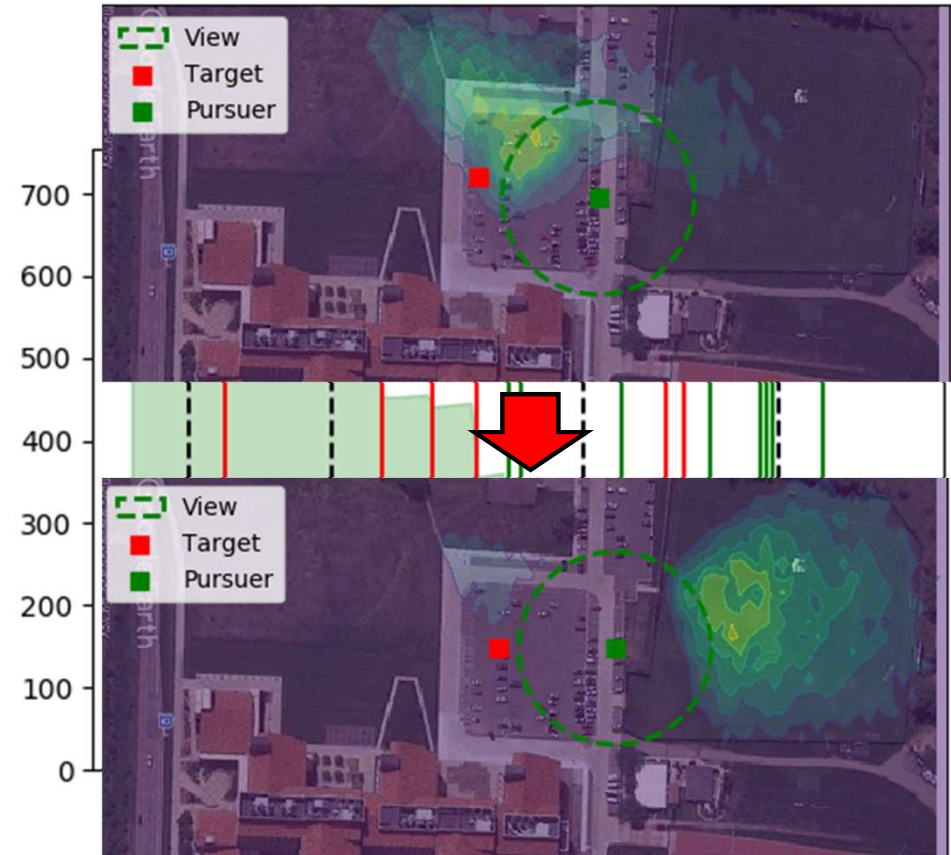
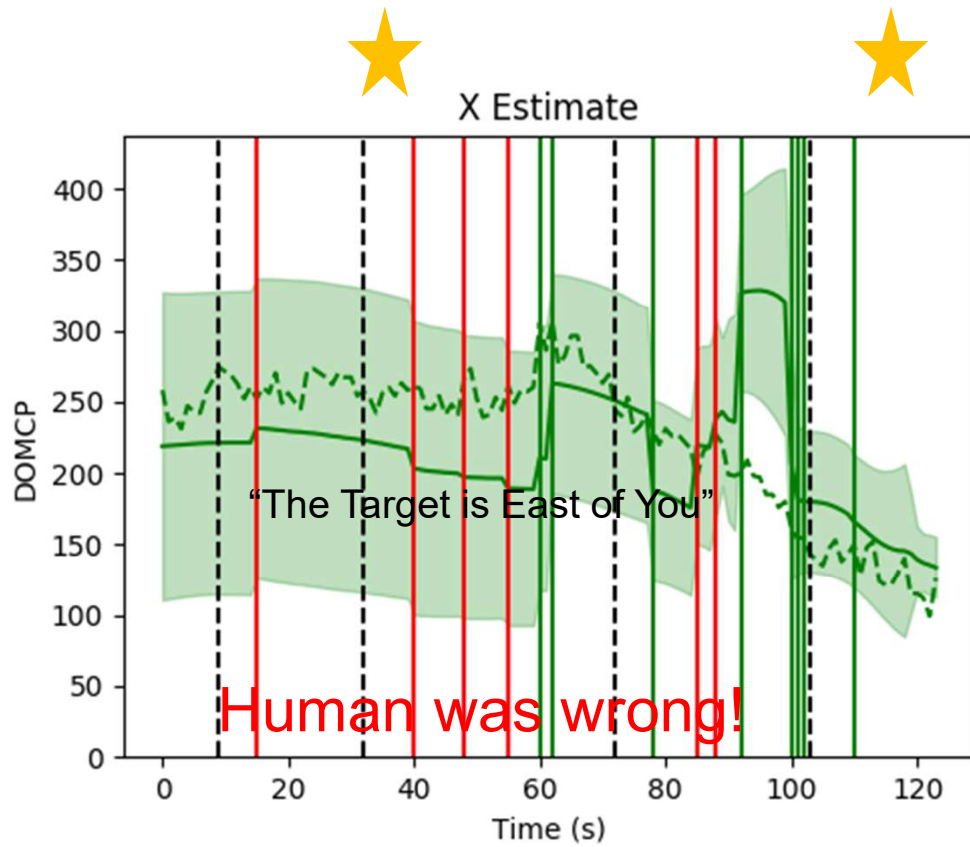


Target Search Interface Simulation Results

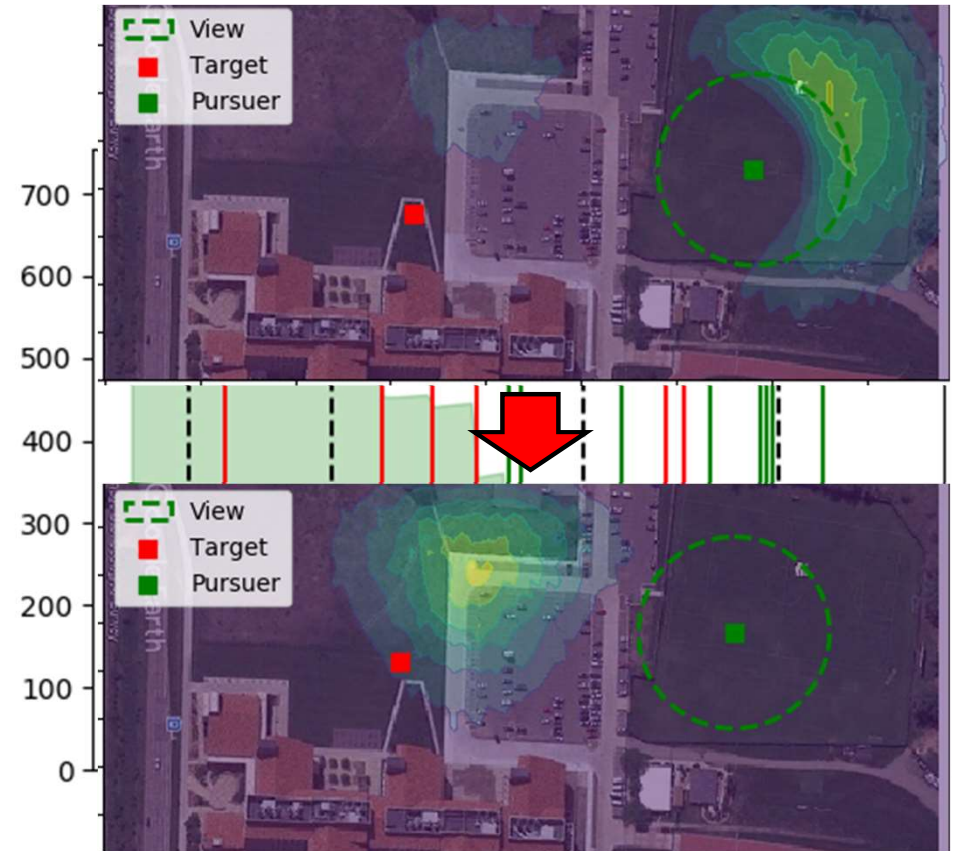
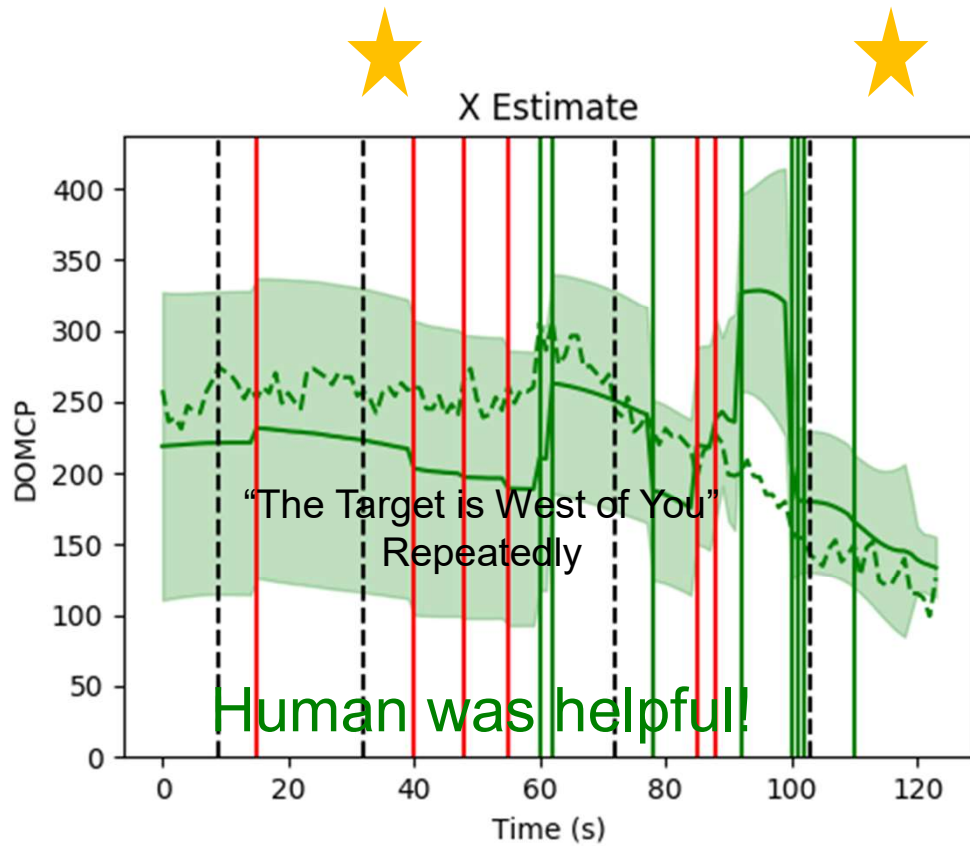
Method	Mean Time to Capture	Std Dev
MAP	126.2 s	± 37
POMCP	94.6 s	± 24
DOMCP	88.2 s	± 19



Erroneous Human Input



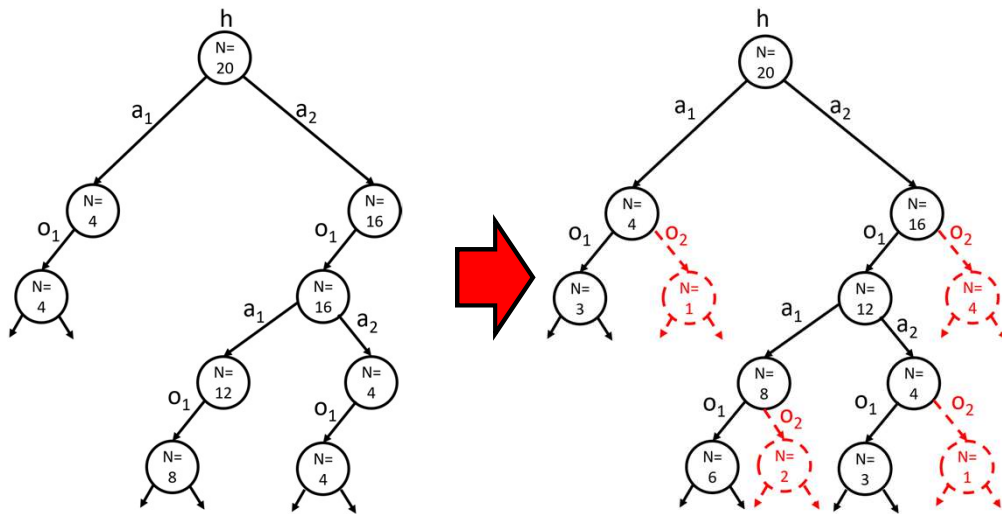
Correct Human Input



Conclusions

Presented Dynamically Observable Monte-Carlo Planning (DOMCP):

- Efficiently adapts prior planning to new observation models
- Retains relevant information from before model changes



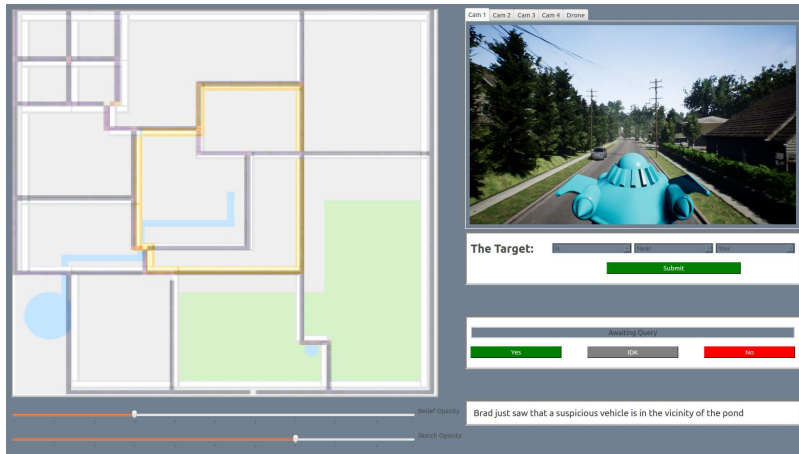
Online planning can effectively exploit available human sensors for target search

- Requesting and evaluating human semantic observations
- Adapting to hand sketched models

Ongoing Work

Further relax model assumptions

- Adapt policies to dynamic/unknown transition and reward models
- Incorporate velocity, contextual conditions, human states



Testing in realistic, scaled up environments

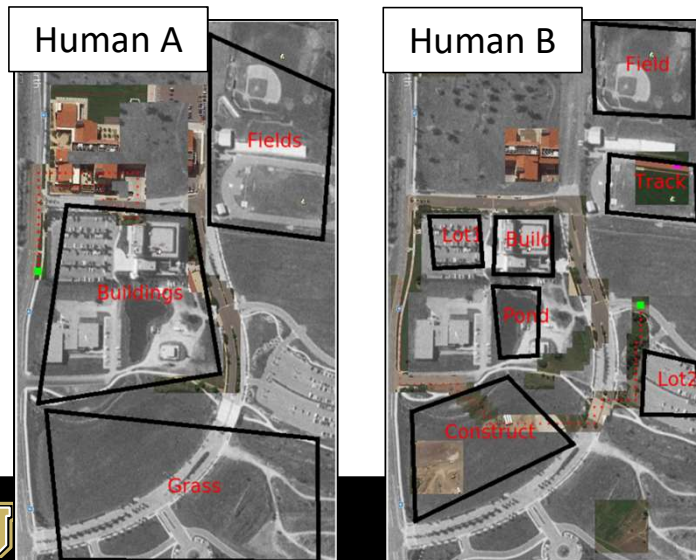
- Non-atomic, strategic movement actions
- Drone search in large open areas
- Fusion of additional sensor modalities such as object detection/recognition, depth sensors

Questions?

Backup Slides

Preliminary Results: Value of Human Sensor + POMDP Planning?

- VB-GM-POMCP vs. simple greedy MAP planning (i.e. go straight to largest pdf peak)?
- Human sensing quality (for spatial target location updates only):
 - no human: no sketch or linguistic location inputs provided
 - Human A: large/imprecise sketches, fewer location updates with some mistakes
 - Human B: small/precise sketches, more location updates with no mistakes
- Compare # steps taken by robot to find target (same data fed to both planners):



Preliminary Data: 1 test per condition

Number of Steps to Capture	MAP Planner	POMCP Planner
No Human	268	199
Human A	255	182
Human B	143	134



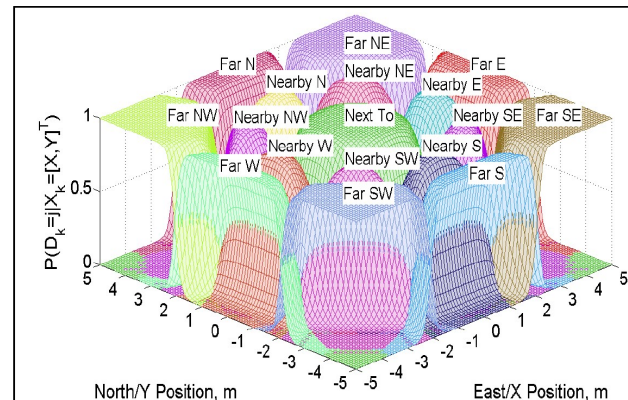
Semantic Observational Models [Sweet, Ahmed, ACC 2016]

A Useful Approach: Softmax Models

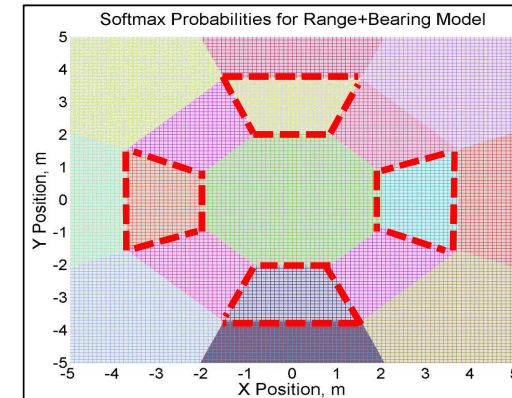
- Segment continuous state space into discrete classes
- Classes dominate spatial regions
- Generalizes to non-convex regions
- Sparse parameterization, easy to learn from data and embed constraints

$$p(o_k | s_k) = \frac{\exp(w_o^T s_k + b_o)}{\sum_c |\Omega| \exp(w_c^T s_k + b_c)}$$

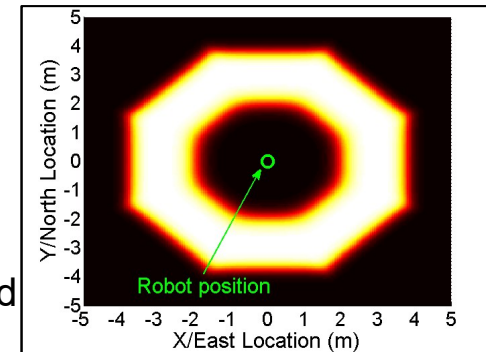
Likelihoods for all classes



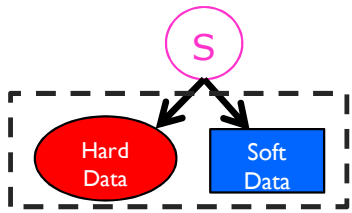
Dominant Regions



Compound "Near" Observation Likelihood



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]

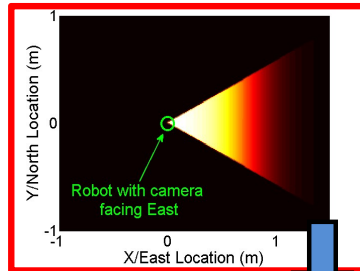
[Ahmed and Campbell, ICRA 2010; Ahmed, et al., T-RO 2013; Sweet and Ahmed, ACC 2016]

“Robot pull” (Active Sensing)

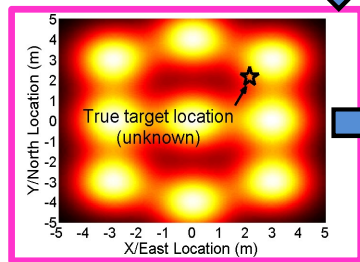
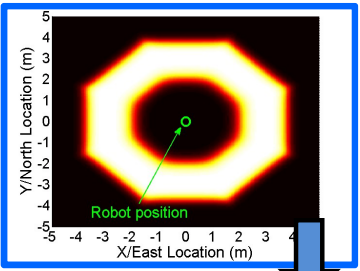
[Kaupp, et al., RAS 2010; Lore, Sweet et al., ICCPS 2015] (VOI, decoupled from planning)



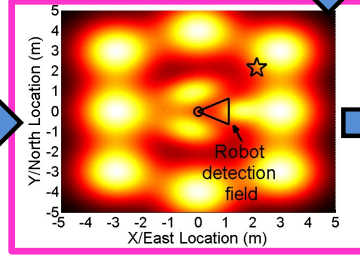
UA Sensor Model for “Detection”



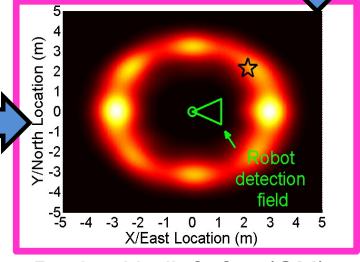
Sensor Model for “Nearby Trees”



$p(S)$
Initial target state prior (GM)



Revised belief after (GM)



Revised belief after (GM)

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

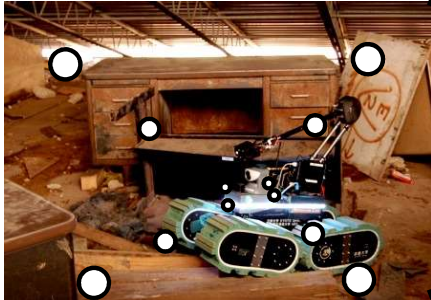
Target detected, range 500 m, bearing 36 deg

Is target turning around? What else do you see?

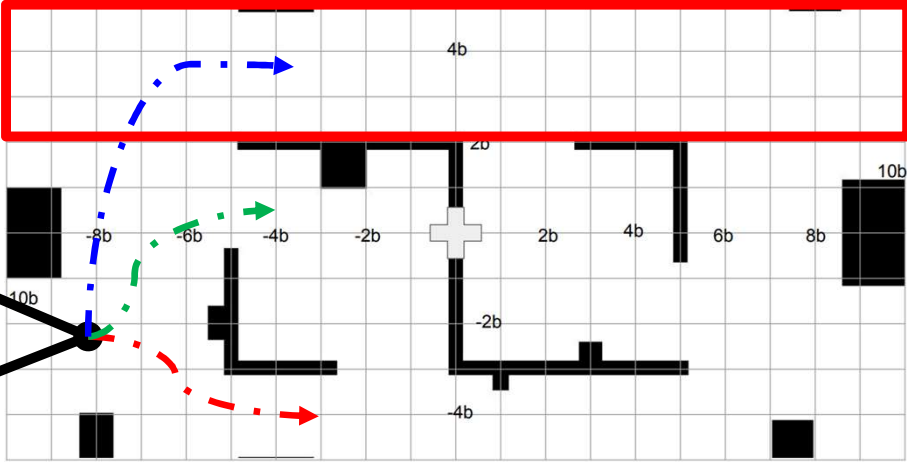


Research Vision

Could I climb that debris pile?



What room am I in?



Is that a person over there?

Is this area dangerous?

Questions a Human could help answer!

Active Semantic Sensing with Continuous State POMDPs (CPOMDPs)

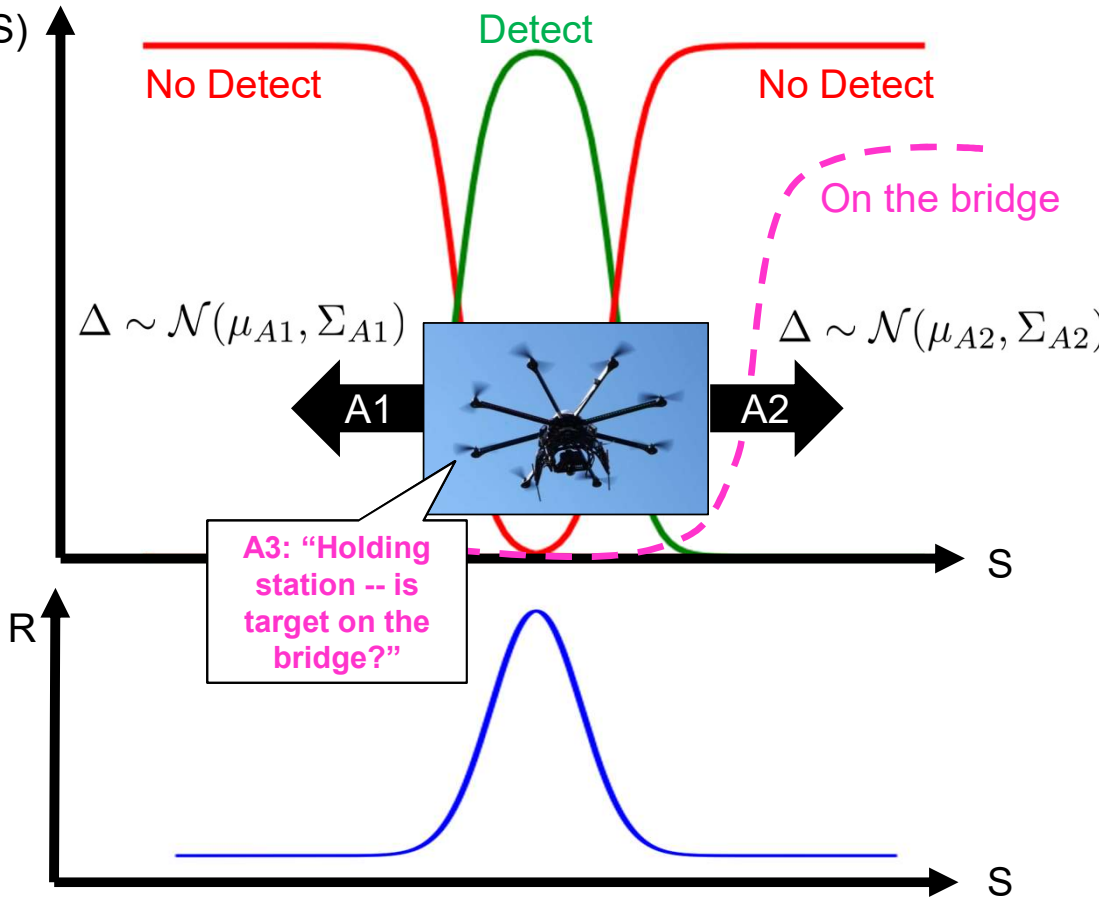
[Burks and Ahmed, CDC 2017]

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

Observations = Ω

Actions
 $= \{A1, A2, A3, \dots\}$
 (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:

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

CPOMDPs: Approximate Value Iteration in Belief Space

Alpha Element Backup for Point-Based Value Iteration

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

$$\alpha_n^i(s) = r_a(s) + \gamma \sum_o \arg \max_{\alpha_{a,o}^i} \langle \alpha_{a,o}^i, b \rangle$$

PBVI-type solution on discretized space x with α -vectors for policy π

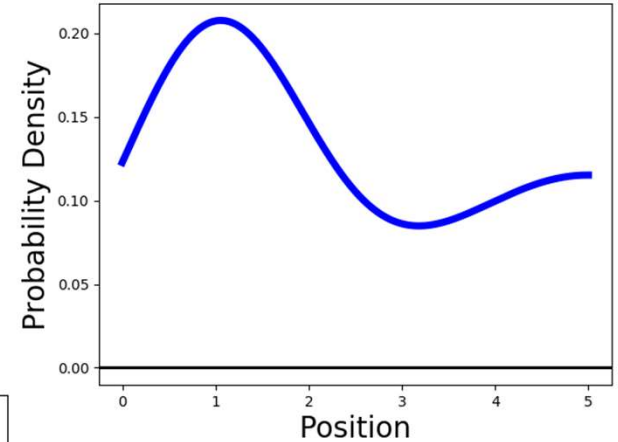
[Pineau, et al., JAIR 2006]

Gaussian Mixture (GM)

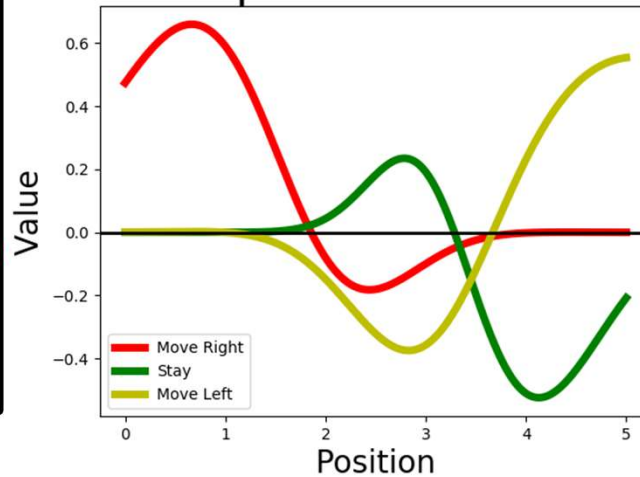
$$\sum_j w_j \phi(s|\mu_j, \Sigma_j)$$

Can represent arbitrary policy functions & pdfs

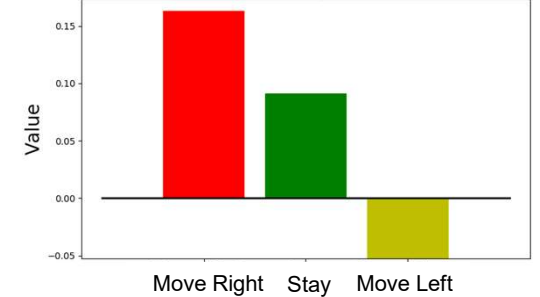
Belief



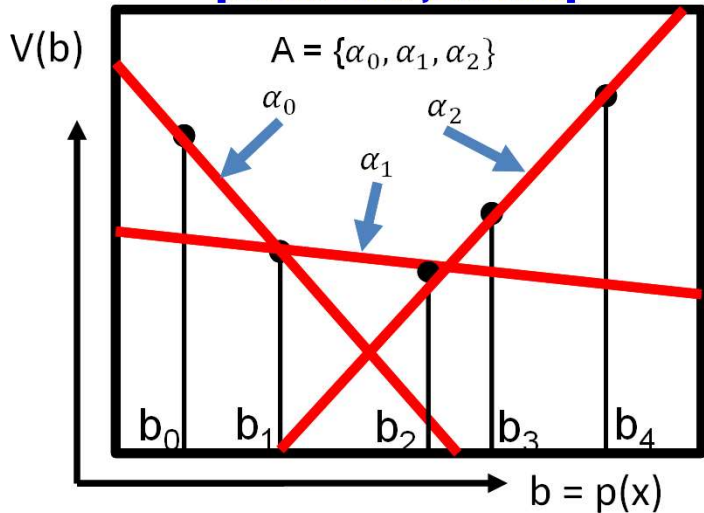
Alpha Functions



Dot Product of Alphas with Belief



Optimal
action
given
belief



The Value of Questions

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

Value of Human
Observation (o)

Expected Reward
After Observation

Expected Reward without
Observation

Most Valuable Questions

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

Robot Questions

Is Zhora inside the study?	<input type="button" value="YES"/>	<input type="button" value="NO"/>	<input type="button" value="?"/>
Is Zhora inside the library?	<input type="button" value="YES"/>	<input type="button" value="NO"/>	<input type="button" value="?"/>
Is Zhora right of the desk?	<input type="button" value="YES"/>	<input type="button" value="NO"/>	<input type="button" value="?"/>

Last question was:
Last answer was:

Research Vision

What kind of information do humans communicate?

- “Soft” semantic data

Humans = Semantic Sensors

- Volunteering Information
- Answering Questions

Robotic tasks can benefit from human information, if only they could ask the right questions!

Question: Where’s this conference room for your Comps talk?

