Active Semantic Sensing and Planning for Human-Robot Collaboration in Uncertain Environments

Luke Burks, Jamison McGinley, Sousheel Vunnam, Holden Kjerland-Nicoletti, and Nisar Ahmed*

***Abstract*— Autonomous robots can benefit greatly from human-provided semantic characterizations of uncertain task environments and states. However, the development of integrated strategies which let robots model, communicate, and act on such ‘soft data’ remains challenging. We present a framework for active semantic sensing and planning in human-robot teams which addresses these gaps by formally combining the benefits of online sampling-based POMDP policies, multi-modal semantic interaction, and Bayesian data fusion. This approach lets humans opportunistically impose model structure and dynamically extend the range of semantic soft data in uncertain environments, which otherwise yield little information to a lone robot. It also lets robots actively query humans for new semantic data which update understanding and beliefs of unknown environments for improved online planning. Dynamic target search simulations show that active collaborative semantic sensing leads to significant improvements in time and belief state estimates required for interception versus conventional planning, which relies on robotic sensing only.**

I. INTRODUCTION

Autonomous robotic vehicles will greatly extend human capabilities in domains such as space exploration [1], disaster response [2], environmental monitoring [3], infrastructure inspection [4], and defense [5]. Yet, the uncertain dynamic nature of these settings, coupled with vehicle size-weight-power-compute constraints, often necessitates some form of human oversight to cope with the brittleness of autonomy [6]. This has created interest in new forms of human-robot interaction that can efficiently leverage human reasoning to enhance robotic reasoning abilities. Probabilistic techniques based on semantic language-based human-robot communication have gained considerable attention for information fusion [7]–[12] and task planning [13]–[19]. However, existing approaches only allow robots to reason about limited kinds of uncertainties, i.e. as long as task environments and conditions are known a priori, or do not change in unforeseen ways. This in turn limits the flexibility and utility of semantic communication for adapting to new or unknown situations.

This work examines how robots operating in uncertain environments can use semantic communication with humans to solve the combined issues of model augmentation, multi-level information fusion, and replanning under uncertainty in an online manner. Such problems practically arise, for instance, with time-sensitive dynamic target search in areas

*Smead Aerospace Engineering Sciences Department, 429 UCB, University of Colorado Boulder, Boulder CO 80309, USA. E-mail: [luke.burks;nisar.ahmed]@colorado.edu. This work is funded by the Center for Unmanned Aircraft Systems (C-UAS), a National Science Foundation Industry/University Cooperative Research Center (I/UCRC) under NSF Award No. CNS-1650468 along with significant contributions from C-UAS industry members.

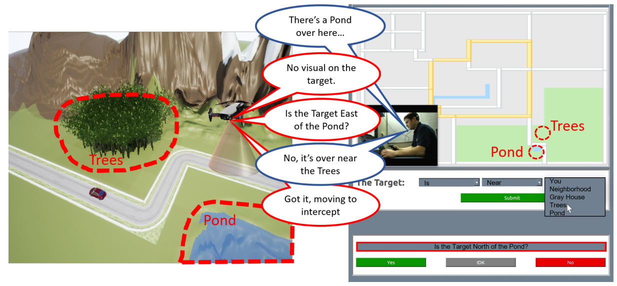

Fig. 1: Aerial robot searching for ground target with human aiding via semantic labeling and referencing of environment.

with outdated/poor prior maps, which lead to uncertain robot and target motion models as well as uncertain human input models. We present a novel framework for integrated active semantic sensing and planning in human-robot teams that formally combines aspects of online partially observable Markov decision process (POMDP) planning, sketch and language-based semantic human-robot interfaces, and model-based Bayesian data fusion.

As shown in Fig. 1, our approach features three key technical innovations. Firstly, humans act as ‘ad hoc sensors’ that push multi-level semantic data to robots via structured language, enabling robots to update task models and beliefs with information outside their nominal sensing range. Secondly, humans can use real-time sketch interfaces to update semantic language dictionaries grounded in uncertain environments, thus dynamically extending the range and flexibility of their observations. Finally, robots actively query humans for specific semantic data at multiple task model levels to improve online performance, while also non-myopically planning to act with imperfect human sensors. These features effectively enable online ‘reprogramming’ of uncertain POMDPs together with human-robot sensor fusion to support online replanning in complex environments.

II. MOTIVATING PROBLEMS AND BACKGROUND

We focus on dynamic target search problems in which a robot must intercept a moving target as quickly as possible through an uncertain environment. Figure 2 shows two versions this problem for an unmanned ground vehicle (UGV) and an unmanned aerial system (UAS), which will be used as examples. Major uncertainties arise in (but are not limited to): vehicle motion and transition models (due to unknown terrain or wind fields that are difficult for robots to sense); target models and states (if targets display different dynamic behaviors); and models of language-based semantic ‘human sensor’ observations (e.g. Fig. 1), where ‘soft data’ can be expressed with variable consistency/accuracy [9], [20].

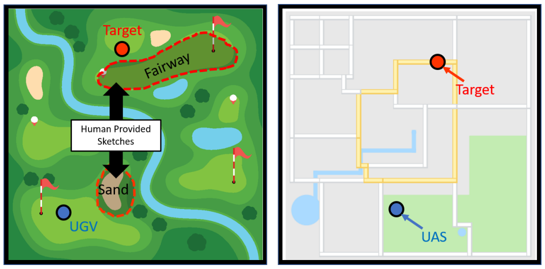

Fig. 2: Left: ‘Golf Course Problem’: UGV navigates terrain to intercept ground target; Right: ‘Road Network Problem’: UAS tracks a ground target can deviate from road.

In each case, the environment includes map regions with differing probabilistic state transition models that affect robot and target motion dynamics. The robot carries a visual sensor, which allows target proximity with a high probability and subject to false alarms, without informing about the environment. Finally, the robot can communicate with a remote human, who can view the robot’s telemetry data, sensing data, and map information with negligible time delay and generate soft data, but issues no commands to the robot. The robot plans its own movements to intercept the target using available models and observations, assuming an initial prior target state pdf. We next consider how the robot can use semantic soft data to update knowledge of the uncertain environment and target states as part of an optimal autonomous planning process.

A. Language-based and Sketch-based Semantic Soft Data

The human primarily acts as an auxiliary (and imperfect) semantic information source that can communicate with the robot at any time via either one of two interfaces. The first interface allows the human to compose linguistic statements that are parsed and interpreted as target state observations. As shown in Fig. 1, these are modeled in the structured form ‘Target *is/is not Desc Ref*’, where the *Desc* and *Ref* elements are taken from a defined semantic codebook, and the *is/is not* field toggles between positive and negative data [21]. In prior work [11], [12], [20], [22], [23], probabilistic likelihoods were developed for all possible statements in a given codebook to support recursive Bayesian fusion of linguistic human and robot sensor data. Since these works considered only *fully known* environments, all relevant semantic references could be enumerated in advance. However, the environments here are not fully known in advance, so the codebook is at best only partially defined at mission start.

This leads to the second interface: since new map information cannot be obtained from the robot’s sensor to augment the codebook, the human may instead do this by providing labeled 2D free-form sketches, which each depict a spatially constrained region on a 2D map display (as in Fig. 1 with ‘Pond’ and ‘Trees’, or in Fig. 2 with ‘Sand’ and ‘Fairway’). Building on [24], the codebook is automatically augmented with the labels of new landmarks/references, so that the corresponding spatial sketch data can also be used to generate suitable soft data likelihood models for the linguistic statement interface (see Fig. 1). However, unlike in [24], human sketches here may also provide direct information about probabilistic state transition models, to constrain how the robot and target may traverse certain map areas. While similar 2D sketch interfaces have also been developed for robotics applications [13], [15], [25], [26], this is the first time (to our knowledge) they have been used in the context of multi-level data fusion for planning under uncertainty.

B. Optimal Planning and Sensing under Uncertainty

The robot must autonomously generate its own plans for minimum time target interception, whether or not any semantic soft data are provided. The presence of model uncertainties, sensing errors, and process noise makes optimal planning quite challenging. One family of decision-making algorithms that accounts for these combined uncertainties are partially observable Markov decision processes (POMDPs). While POMDPs are impractical to solve exactly for all but the most simple problems [27], a variety of powerful approximations can exploit various features of particular problem formulations. A POMDP is formally specified as a 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \Omega, \mathcal{O}, \gamma)$, where the goal is to find a policy π which maps from a Bayesian posterior distribution, i.e. the belief $b = p(s)$ over the set of states $\mathcal{S}$, to a discrete action $a \in \mathcal{A}$. The transition model $\mathcal{T}$ is a discrete time probabilistic mapping from one state to the next given an action, $p(s'|s,a)$, after which the robot is rewarded according to $R(s,a)$. During policy execution, the robot receives observations $o \in \Omega$ according to observation likelihood $\mathcal{O} = p(o|s)$. For infinite horizon planning problems with discount factor $\gamma \in [0,1)$, the optimal policy $\pi[b(s)] \rightarrow a$ maximizes the expected future discounted reward: $\mathbb{E}[\sum_{k=0}^{\infty} \gamma^k R(s_k, a_k)]$.

The primary challenge arising from casting our motivating problems as POMDPs lies in the ability of the human to modify the models $\mathcal{T}$ and $\mathcal{O}$ online in an unmodeled ad-hoc fashion via the sketch interface. These modifications can happen rapidly, and might change large swathes of each model with a single sketch. Bayes-Adaptive POMDPs [28] allow POMDPs to learn the parameters of $\mathcal{T}$ and $\mathcal{O}$, but require gradual changes to their parameters and static $\mathcal{T}$ and $\mathcal{O}$. In the motivating problems, $\mathcal{A}$ and $\mathcal{O}$ change unpredictably with new sketches. This issue renders “full-width” offline point-based POMDP planners [29], [30] inapplicable, as they require models of how $\mathcal{T}, \mathcal{O}$ and Ω change.

Online POMDP approaches [31], which eschew the process of pre-solving the policy prior to execution in favor of interleaving steps of policy execution and search, have successfully been used to address large observation spaces [32], continuous state spaces [33], and more recently, dynamic ad-hoc models [24]. These algorithms make use of a ’black-box’ generative model, requiring only the ‘current POMDP problem’ at time of execution, making them good candidates for solving problems with dynamic model uncertainty.

Cast as a POMDP, the motivating problem easily incorporates information gathering actions. Unlike the Value-of-Information [8] and Human Observation Providers POMDP [9] frameworks, which relied on static environments and fixed codebooks regarding state information, our problem requires the ability to dynamically encode previously unknown environments and leverages multiple types of human input.

C. Formal Problem Statement

A mobile robot with continuous state s_r attempts to localize and intercept a target with continuous state s_t in some search environment. The joint state space is $[s_r, s_t]^T = s \in \mathcal{S} = \mathcal{R}^N$. The human sensor has full knowledge of s_r at all times as well the belief $b(s_t) = p(s_t|o_h^{1:k}, o_r^{1:k}, a^{1:k-1})$, which is the updated Bayesian posterior pdf given all observations made by the robot $o_h^{1:k}$ and the human $o_h^{1:k}$ through time k. Both s_r and $b(s_t)$ are displayed over a terrain map, which dictates transition model $p(s'|s,a)$ at each point space according to hazards and open areas. Action space $\mathcal{A}$ is a combination of movement action made by the robot which affect its state $\mathcal{A}_m$ and query actions which pull information from the human $\mathcal{A}_q$. The robot initially holds a uniform $p(s'|s,a)$, $\forall s$, which the human can update by sketching and labeling areas of interest on the map. Each sketch creates an enlarged action space $\mathcal{A}'$ such that $\mathcal{A} \subset \mathcal{A}'$, where each new action is a possible query to the human regarding the relative location of s_t with respect to the new sketch.

III. Methodology

A. Human Querying and Data Fusion

The human can act as either: a passive semantic sensor, which volunteers information without request whenever possible; or as an active semantic sensor, which can be queried by the robot to provide information on request. Here we assume the human will respond to queries based on a static a priori known responsivness value (ξ), such that, for all human observations o_h at all times, $p(o_h \neq None|s) = \xi$. This leads to an additional observation $o_h = None \in \Omega_h$. We also assume the human has a static a priori known accuracy η, such that, for all human observations o_h at all times, $\bar{p}(o_h = j|s) = p(o_h = j|s) \cdot \eta$, where probability is redistributed to the "inaccurate" observations,

$$\bar{p}(o_h = (k \neq j)|s) = \quad (1)$$
$$\frac{1}{|\Omega|-1} p(o_h = j|s) \cdot \eta \cdot p(o_h = (k \neq j)|s)$$

The notation $\bar{p}(o|s)$ and $\bar{p}(s'|s,a)$ denote models used during online execution, in contrast to the nominal distributions $p(o|s)$ and $p(s'|s,a)$. This parameterization of accuracy ignores similarity between any two observations. While this simplifies implementation and allows comparative testing of accuracy levels, other models may yield more realistic results; e.g. linear softmax model parameters [23], [24], [34], can be used to determine the likelihood of mistaking semantic labels given s.

We focus here on humans as active sensors. This requires an explicit dependency on actions to be included in the observation model, as well as additional actions which trigger this dependency. Just as POMDP observations can be typically modeled as $o \sim p(o|s), o \in \Omega$, human responses to robotic queries $a \in \mathcal{A}_q$ can be modeled as $o_h \sim p(o_h|s, a_q)$, $a_q \in \mathcal{A}_q$, $o_h \in \Omega_h$. These additional query

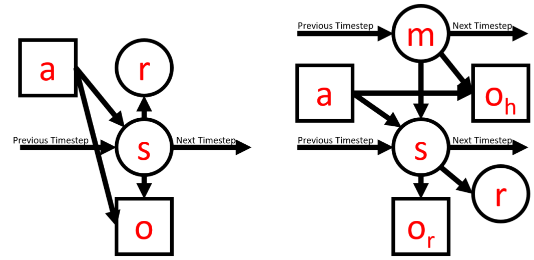

Fig. 3: Left: Graphical model for Golf Course POMDP; Right: Graphical model for the Road Network POMDP.

actions can be introduced into $\mathcal{A}$ in one of two ways. Casting them as exclusive options, where either movement or a query can be pursued but not both, minimally expands the action space to the sum of queries and movements, $\mathcal{A} = \mathcal{A}_m + \mathcal{A}_q$. But this can be limiting in situations which benefit from rapid information gathering about models and states together. Instead, we cast the queries as inclusive actions, in which every time step permits any combination of movement and query, $\mathcal{A} = \mathcal{A}_m \times \mathcal{A}_q$. This can drastically increase the size of the action space, but allows rapid information gathering.

The robot's on-board sensor produces a single categorical observation per time step o_r , which must be fused with o_h. We assume the robot's sensor is independent of the human observations given the state such that,

$$p(s|o_r, o_h) = \frac{p(s)p(o_h, o_r|s)}{p(o_r, o_h)} \propto p(s)p(o_r|s)p(o_h|s) \quad (2)$$

Fig. 3 summarizes the probabilistic dependencies for the POMDPs describing the Fig. 2 search problems. While all observations are state dependent, some now depend on actions chosen by the policy. Also, query actions have no effect on the state, and are pure information gathering actions. In these new POMDPs, a combined action might be "Move to North, and Ask human if target South of Lake", which may return a combined observation of "Target is Far from me, and human says target is not South of Lake". So in this case $a = \{North, South/Lake\}$, and $o = \{Far, No\}$.

B. Online Planning and Model Revision

POMDP policy approximations typically assume a fixed problem during policy search and execution. This limits their ability to find policies in poorly modeled but learnable environments. A human sensor can address this shortcoming: in addition to answering queries about target states, the human can help edit the POMDP by providing model-related information while requiring minimal expert knowledge. Uncertain state transition models $\mathcal{T}$ due to partial terrain information can be updated with soft information during execution. The spatial location of alterations to $\mathcal{T}$ can also be labeled as semantic references for new soft data options. However, this newly available information creates its own issues.

Most POMDP solvers make two key assumptions on transition models. First is that the model is temporally static

$$p_{k+1}(s'|s,a) = p_k(s'|s,a), \forall k \quad (3)$$
$$\hat{p}_{k+1}(s'|s,a) = \hat{p}_k(s'|s,a).$$

where $\hat{p}_k(s'|s,a)$ is the internal transition function in use by the robot at time k, in contrast to the true underlying (unknown) model $p_k(s'|s,a)$. The second is that the model being used by the solver is identical to the model being used during the execution of the policy, $\hat{p}_k(s'|s,a) = p_k(s'|s,a)$.

The first assumption is particularly important for infinite horizon offline solvers, where Bellman-backup steps generate approximations of the value function agnostic of whichever timestep a reward may be achieved on. Finite horizon offline and online solvers can cope with this limitation [31], [35], but require that even non-stationary transition models be known for all times prior to generating a policy.

Here we propose a sketch-based method for the human to modify the POMDP during execution. Sketches made by the human are labeled as landmarks within the map contained within $\mathcal{S}$. In addition to growing the possible query action space by a set amount for each sketch, each labeled spatial sketch area gives the human the opportunity to describe terrain features or likely motion parameters in that area.

For $\mathcal{T}$, we assume the underlying model isn't changing, while the robot's understanding of it might, i.e.

$$\begin{aligned} p_{k+1}(s'|s,a) &= p_k(s'|s,a) \\ \hat{p}_{k+1}(s'|s,a) &\stackrel{?}{=} \hat{p}_k(s'|s,a) \end{aligned} \tag{4}$$

However, we make no assumption on the nature of this understanding change nor the timing. For the observation model however, both the true and internal model change when given a new sketch. All of these model changes imply that the robot must solve a different but related POMDP after each human sketch. With this in mind, we next consider viable approximations for solving such POMDPs.

The Partially Observable Monte Carlo Planning (POMCP) algorithm proposed in [31] is particularly promising due to its use of generative 'forward' models and online any-time characteristics. While the original implementation of POMCP uses an unweighted particle filter for belief updates, the authors of [36] note that, for problems with even moderately large $\mathcal{A}$ or $\mathcal{O}$, a bootstrap particle filter allows for more consistent belief updates without domain specific particle reinvigoration. This weighted particle filter approach coincidentally also provides a solution to the dynamic modeling problem. Each belief update is carried out using the most up-to-date model, while changes to the model only affect future timesteps. As model alterations require solving a different POMDP after each sketch, this approach allows each planning phase to be treated as the start of a brand new POMDP solution. Our procedure for carrying out POMCP planning while handling dynamic model updates from a human is detailed in Algorithm 1. Certain aspects of this procedure, such as our use of a discretized grid to store transition model modifiers (α) or the semantic observations included with each sketch (described in more detail in the next section), readily adapt to more general problems.

Algorithm 1 Planning with Human Model Updates

B_k = Particle_Set(Size = N)
$\alpha(s)$ = Discrete_Grid()
repeat
    $[a_m, a_q]$ = POMCP(B) (Ref. [31])
    $s \sim p_k(s'|s, a_m)$ (unknown state)
    $o_r \sim \bar{p}_k(o_r|s)$ (robot sensor observation)
    $o_h \sim \bar{p}_k(o_h|s, a_q)$ (human query answer)
    B_{k+1} = Bootstrap_Filter(B_k, a_m, a_q, o_r, o_h)
    if New Human Sketch (W) **then**
        V = W.Value, (human assigned state modifier)
        L = W.Label, (human assigned label)
        for $s \in W$ **do**
            $\alpha(s)$ = V
        end for
        $\Omega_{k+1} = \Omega_k \cup (L \times [Near, E, W, N, S])$
    end if
until Scenario End

C. Multi-Level Active Information Gathering

To accommodate stochastic switching dynamics, it is generally convenient to include mode states m as shown in Fig. 3 (right) for the Road Network problem, which dictate alterations to the target state transition models,

$$\bar{s} = [s, m]^T, m' \sim p(m'|m) \;\; s' \sim p(s'|s, a, m'). \tag{5}$$

This additional discrete state variable can be easily handled by POMCP and other Monte Carlo tree search approximations, which rely on generative black box simulators and support edits to $\mathcal{T}$ and $\mathcal{O}$. Offline approximations based on switching-mode POMDPs [37] can also accommodate hybrid dynamics, but require stationary models. Such hybrid model extensions also open the door to active semantic queries and specific human observations o_h pertaining to m, which can greatly enhance the Bayesian belief for m.

With this modification to the generative model, the probability of different state transition models being in effect can be explicitly represented via particle approximation as

$$P(m = x) = \frac{1}{N} \sum_{n=0}^{N} \mathbf{1}(m_n = x), \tag{6}$$

where $\mathbf{1}(m_n = x)$ is an indicator function applied to the mode of particle (n). This allows queries to be constructed in the same way as those presented in Section III-a, but referring only to the mode segment of the state vector $\bar{s}$. For instance, in Section IV-b we consider a problem where the modes governing the transition probabilities represent whether or not a target is on a road. Thus, actions can be taken such as "Ask the human if the target has gone off-road", which can improve the robot's estimated belief of the target's speed. In this way, the model changes brought about by queries are anticipated by the solver, and don't imply the need to solve a different POMDP. For human responses modeled as $p(o_h|m)$, the belief update equation becomes

$$p(s|o_r, o_h) = \sum_m p(s, m|o_r, o_h) \tag{7}$$

$$= \sum_m p(s|o_r, o_h, m) p(m|o_r, o_h) = \sum_m \frac{p(r|s)p(m|s)p(h|m)}{p(h)p(r)}$$

With the mode treated an additional state variable as in Equation 5, this equation reduces to Equation 2. The bootstrap particle filter used in Algorithm 1 approximates this belief update through use of weighted particles.

IV. Dynamic Target Search Applications

We present the results of two simulated scenarios depicted in Fig. 2 in which the robot attempts to localize and intercept a moving ground target. In the first 'Golf Course Problem', an unmanned ground vehicle (UGV) pursues the target in an unmapped area with unknown terrain features for which the state transition dynamics are unknown a priori. In the second scenario, an unmanned aerial system (UAS) pursues a ground based target in a large road network; the target moves according to a two-layer HMM as shown in Fig. 3 (right), where m indicates whether the target is road-bound.

A. Golf Course Problem: Search in Unknown Environment

The golf course shown in Fig. 2 is the underlying environment. The course contains numerous regions which either slow or speed movement through them by ground vehicles. We model these differences through the use of a modifier variable (α) applied to the nominal transition model. The UGV is initialized with a uniform value of (α) across the state space. The UGV's human teammate, for cases wherein human assistance is provided, is able to modify the UGV's understood value of (α) through sketches. Here we assume the human's provided modifiers are perfectly accurate even if their observations about their sketches aren't necessarily. Expected transitions for the UGV $E[s']$, with $\Delta(a)$ being the expected move resulting from an action, are given as $E[s'] = s + \alpha\Delta(a)$, $\alpha = [\mathit{Water} : 0.05, \mathit{Sand} : 0.5, \mathit{Fairway} : 1.5, \mathit{Green} : 2, \mathit{otherwise} : 1]$. The UGV has a nominal speed of 5 m/s in any of the 4 cardinal directions, while the target nominally moves at 2 m/s using a nearly constant velocity (NCV) kinematic state model [38].

Observations made by the robot's on-board proximity sensor have three possible values, either indicating that the target is "Far", "Near", or "Captured", all with true positive rates of 98%. The robot's field of view, and thus the "Near" observation, extend out to 15m while distances beyond are tagged "Far". Capture is declared at a distance of 7.5m.

In the "Human" test case, a simulated human was used to provide sketches and linguistic observations using these. Sketches were drawn from a list of 15 pre-drawn sketches, focusing on areas of interest such as water, sand, and greens. For each of the 100 simulated runs, the human provided a sketch at the beginning of the run, and added an additional sketch every 20 time steps up until the end of the run at 100 time steps. Each time step corresponded to 2 seconds, which was the planning time allocated to the POMCP solver. Using the observation model described in Section III-a, the simulated humans were assumed to have an accuracy and responsiveness of $\eta = \xi$ =95%. The queries made by the robot were constructed in binary fashion similarly to those in [23], [24]: questions a_q took the form of queries about the target position relative to cardinal directions for each sketched map landmark, e.g. "Is the target east of Fairway 1", with potential replies $o_h \in [Yes, No, None]$.

Golf Course Problem Results

Method	Mean TTC	Standard Deviation
Non-human	34.16	±23.97
Non-human Informed	31.22	±22.04
Human	25.46	±15.78

TABLE I: Golf Problem Comparison of simulation TTC

Road Network Problem Results

Method	Mean TTC	Standard Deviation
Non-human	39.43	±23.74
Human	30.45	±23.28

TABLE II: Road Network Comparison of TTC

The UGV gains a reward of 100 for capturing the target, which requires being within 7.5m, and incurs a small cost of 1 for asking questions of the human. This reward function encourages fast target capture. Since POMDP policies maximize the cumulative reward utility resulting from $R(s, a)$ and thus seek to achieve a real-world goal encoded by the resulting utility, we report the results of our simulations in terms of our goal metric, Time to Capture (TTC), which is the first timestep the UGV is within 7.5 meters of the target.

Three deployment cases are evaluated to determine the effect of introducing human sketches and observations. In the first "Non-human" case, the UGV is given a uniform α modifier over the map. The "Human" case gives the UGV a human, capable of providing both sketches and observations in response to requests from the UGV. The final case, "Non-human Informed", gives the UGV the true α values across the state space. The TTC results are shown in Table I.

In order to examine the effects of differing levels of human accuracy η and responsiveness ξ, each variable was tested with select values ranging from 30% to 99%. The number of questions asked as a percentage of all actions, as well as the TTC metric for each test, are shown in Figure 5. η and ξ here are known and static for the robot. Additional research, possibly extending the HOP-POMDP framework [9] or [39] could allow the robot to learn these online.

B. Road Network Problem: Search with Switching Dynamics

In this scenario, the target moves through the search area in one of two modes, m_0 = *on-road* and m_1 = *off-road*. An on-road target has a chance to transition to off-road, and will follow an NCV model until it encounters another road. Thus the target dynamics in the continuous state space s are conditionally constrained with respect to the mode m. When available, the human is given a binary query asking "Is the target on the road", which yields response $o_h \in [Yes, No, None]$ in accordance with η and ξ. The reward, robot observation models, and human responsiveness and accuracy statistics are otherwise identical to the 'Golf Course' scenario presented above. Transitions are modeled with a uniform $\alpha = 1$ across the space.

The road-network scenario was tested in cases with and without a human collaborator answering questions pertaining to the target mode. As in the 'Golf Course' problem, the true simulated human's accuracy and responsivity are $\eta = \xi =$ 95%. The TTC results are displayed in Table II.

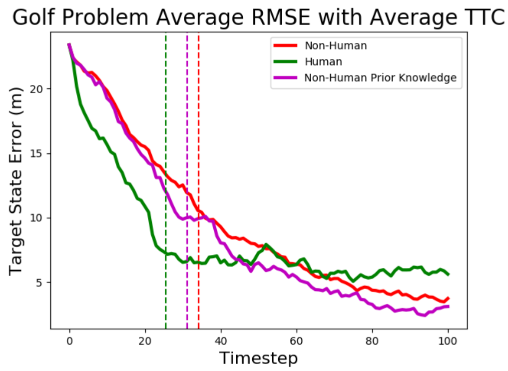

Fig. 4: Average Root Mean Squared Error for Golf Course problem (vertical lines: average TTC per method).

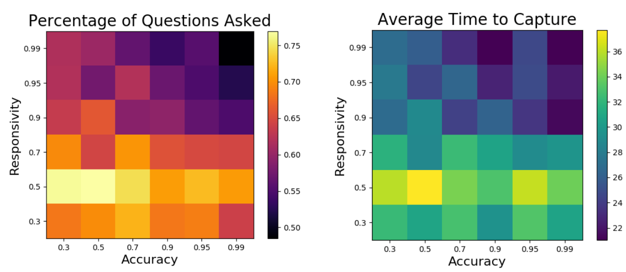

Fig. 5: Given humans with various levels of accuracy and responsivity, Left: The number of questions asked by the robot as a percentage of all actions; Right: The average TTC

C. Discussion of Results

In the Golf Course problem, the human-robot team was able to capture the target faster on average than the robot alone,($p < 0.05$). A major factor in their success was the ability of the simulated human to rapidly improve the robot's estimate of the target's location, as shown in Figure 4. A better estimate leads to more effective querying and planning, with in turn produce better estimates and faster capture times.

When testing differing η and ξ values in Fig. 5, the most accurate and responsive humans were asked the fewest questions and achieved the best average TTC. Similarly, the simulated humans with lower η and ξ received more questions. The robots knowledge of η and ξ indicates it adjusts the frequency of questions to obtain a similar quality of target state estimate regardless of human. Furthermore, these tests support our claim that a human-robot team generally outperforms a lone robot, despite introducing increased complexity. In 89% of these tests, the TTC was lower than that achieved in the Non-human case from Table I, and the combinations of η and ξ which did not surpass the 'Non-human' case still carried similar performance. This suggests the 'plug-and-play' nature of our framework allows it to make the best of a poorly suited human as well as take advantage of a highly useful one – and in the worst case see comparable performance with a more complex problem.

In the Road Network problem, the human was able to positively impact the robot's search even without the ability to provide semantic data about the target's continuous state. As indicated in Table II, the option to query a human about the current mode of the target, i.e. whether or not it was bound to the road-based dynamics model, resulted in significantly lower TTC, ($p < 0.05$). This is demonstrated by Fig. 6, where the robot's estimate of the target's mode is clearly superior when the human is available. The robot is able to periodically constrain its estimate of m by querying the human, and maintain a better understanding of the target's future movements. Thus, the use of a human to measure an indirect variable such as the target's mode of travel allows the robot to more effectively carry out its task.

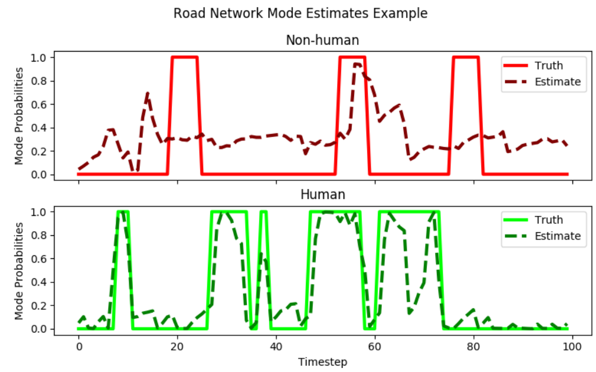

Fig. 6: Example of robot's ability to estimate whether target is bound to road network, with and without human input.

V. Conclusion

We proposed and demonstrated a novel approach to multi-level active semantic sensing and planning in human-robot teams in unknown dynamic environments. Our solution extends online POMDP planning frameworks to incorporate semantic soft data from a human sensor about the location of a tracked target as well as relevant terrain information. Such information is propagated from human to robot through the use of a sketch based, natural language interface. This approach provides a novel formulation of a POMDP with a "human-in-the-loop" active sensing model, as well as innovations to the use of soft-data fusion as applied to higher level modal information. We demonstrated our approach on two example problems. The first showcased the improvements a queryable human can bring to target search problems, while the second demonstrated that the incorporation of human information even with regard to higher level mode observations can improve a robot's target search ability.

Moving forward, research will examine the effect of inaccurate or incomplete sketches on the part of the human. Pairing the robot with a visual object detector has potential to allow feedback to ensure that the robot's perception of the object or region is consistent with a sketch. Additionally, a 3D physics driven simulated search environment for large scale robotics testing is being developed in parallel with our interface to extend and test the ideas contained in this work using real humans in realistic search tasks. Hardware implementations with a UAS and live human are also ongoing.

REFERENCES

[1] S. McGuire, P. M. Furlong, C. Heckman, S. Julier, D. Szafir, and N. Ahmed, "Failure is not an option: Policy learning for adaptive recovery in space operations," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1639–1646, 2018.

[2] J. Delmerico, S. Mintchev, A. Giusti, B. Gromov, K. Melo, T. Horvat, C. Cadena, M. Hutter, A. Ijspeert, D. Floreano *et al.*, "The current state and future outlook of rescue robotics," *Journal of Field Robotics*.

[3] M. Dunbabin and L. Marques, "Robots for environmental monitoring: Significant advancements and applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 24–39, 2012.

[4] M. Hutter, R. Diethelm, S. Bachmann, P. Fankhauser, C. Gehring, V. Tsounis, A. Lauber, F. Guenther, M. Bjelonic, L. Isler *et al.*, "Towards a generic solution for inspection of industrial sites," in *Field and Service Robotics*. Springer, 2018, pp. 575–589.

[5] R. A. David and P. Nielsen, "Defense science board summer study on autonomy," Defense Science Board Washington United States, Tech. Rep., 2016.

[6] J. M. Bradshaw, R. R. Hoffman, D. D. Woods, and M. Johnson, "The seven deadly myths of" autonomous systems"," *IEEE Intelligent Systems*, vol. 28, no. 3, pp. 54–61, 2013.

[7] F. Bourgault, A. Chokshi, J. Wang, D. Shah, J. Schoenberg, R. Iyer, F. Cedano, and M. Campbell, "Scalable bayesian human-robot cooperation in mobile sensor networks," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008, pp. 2342–2349.

[8] T. Kaupp, A. Makarenko, and H. Durrant-Whyte, "Human–robot communication for collaborative decision makingA probabilistic approach," *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 444–456, 2010.

[9] S. Rosenthal, M. Veloso, and A. K. Dey, "Learning accuracy and availability of humans who help mobile robots," in *Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[10] A. N. Bishop and B. Ristic, "Fusion of spatially referring natural language statements with random set theoretic likelihoods," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 2, pp. 932–944, 2013.

[11] N. R. Ahmed, E. M. Sample, and M. Campbell, "Bayesian multicategorical soft data fusion for human–robot collaboration," *IEEE Transactions on Robotics*, vol. 29, no. 1, pp. 189–206, 2013.

[12] N. Sweet and N. Ahmed, "Structured synthesis and compression of semantic human sensor models for Bayesian estimation," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 5479–5485.

[13] M. Skubic, C. Bailey, and G. Chronis, "A sketch interface for mobile robots," in *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, vol. 1. IEEE, 2003, pp. 919–924.

[14] S. Tellex, T. Kollar, S. Dickerson, M. R. Walter, A. G. Banerjee, S. Teller, and N. Roy, "Understanding natural language commands for robotic navigation and mobile manipulation," in *Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[15] D. Shah, J. Schneider, and M. Campbell, "A sketch interface for robust and natural robot control," *Proceedings of the IEEE*, vol. 100, no. 3, pp. 604–622, 2012.

[16] C. Matuszek, E. Herbst, L. Zettlemoyer, and D. Fox, "Learning to parse natural language commands to a robot control system," in *Experimental Robotics*. Springer, 2013, pp. 403–415.

[17] T. M. Howard, S. Tellex, and N. Roy, "A natural language planner interface for mobile manipulators," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 6652–6659.

[18] F. Duvallet, M. R. Walter, T. Howard, S. Hemachandra, J. Oh, S. Teller, N. Roy, and A. Stentz, "Inferring maps and behaviors from natural language instructions," in *Experimental Robotics*. Springer, 2016, pp. 373–388.

[19] R. Paul, J. Arkin, N. Roy, and T. M Howard, "Efficient grounding of abstract spatial concepts for natural language interaction with robot manipulators," 2016.

[20] E. Sample, N. Ahmed, and M. Campbell, "An experimental evaluation of bayesian soft human sensor fusion in robotic systems," in *AIAA Guidance, Navigation, and Control Conference*, 2012, p. 4542.

[21] W. Koch, "On exploiting "negative" sensor evidence for target tracking and sensor data fusion," *Information Fusion*, vol. 8, no. 1, pp. 28–39, 2007.

[22] K. G. Lore, N. Sweet, K. Kumar, N. Ahmed, and S. Sarkar, "Deep value of information estimators for collaborative human-machine information gathering," in *Proceedings of the 7th International Conference on Cyber-Physical Systems (ICCPS 2016)*. IEEE Press, 2016, pp. 3–12.

[23] L. Burks, I. Loefgren, L. Barbier, J. Muesing, J. McGinley, S. Vunnam, and N. Ahmed, "Closed-loop Bayesian semantic data fusion for collaborative human-autonomy target search," in *2018 21st International Conference on Information Fusion*. IEEE, 2018, pp. 2262–2269.

[24] L. Burks and N. Ahmed, "Collaborative semantic data fusion with dynamically observable decision processes," in *2019 22nd International Conference on Information Fusion*. IEEE, 2019.

[25] F. Boniardi, B. Behzadian, W. Burgard, and G. D. Tipaldi, "Robot navigation in hand-drawn sketched maps," in *2015 European conference on mobile robots (ECMR)*. IEEE, 2015, pp. 1–6.

[26] N. Ahmed, M. Campbell, D. Casbeer, Y. Cao, and D. Kingston, "Fully bayesian learning and spatial reasoning with flexible human sensor networks," in *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. ACM, 2015, pp. 80–89.

[27] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[28] S. Ross, B. Chaib-draa, and J. Pineau, "Bayes-adaptive pomdps," in *Advances in neural information processing systems*, 2008, pp. 1225–1232.

[29] J. Pineau, G. Gordon, S. Thrun *et al.*, "Point-based value iteration: An anytime algorithm for POMDPs," in *2003 International Joint Conference on Artificial Intelligence*, vol. 3, 2003, pp. 1025–1032.

[30] M. T. Spaan and N. Vlassis, "Perseus: Randomized point-based value iteration for POMDPs," *Journal of Artificial Intelligence Research*, vol. 24, pp. 195–220, 2005.

[31] D. Silver and J. Veness, "Monte-Carlo planning in large POMDPs," in *Advances in neural information processing systems*, 2010, pp. 2164–2172.

[32] Z. N. Sunberg and M. J. Kochenderfer, "Online algorithms for pomdps with continuous state, action, and observation spaces," in *Twenty-Eighth International Conference on Automated Planning and Scheduling*, 2018.

[33] A. Goldhoorn, A. Garrell, R. Alquézar, and A. Sanfeliu, "Continuous real time POMCP to find-and-follow people by a humanoid service robot," in *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2014, pp. 741–747.

[34] N. R. Ahmed, "Data-free/data-sparse softmax parameter estimation with structured class geometries," *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1408–1412, Sep. 2018.

[35] A. Somani, N. Ye, D. Hsu, and W. S. Lee, "DESPOT: Online POMDP planning with regularization," in *Advances in neural information processing systems*, 2013, pp. 1772–1780.

[36] M. Egorov, Z. N. Sunberg, E. Balaban, T. A. Wheeler, J. K. Gupta, and M. J. Kochenderfer, "POMDPs.jl: A framework for sequential decision making under uncertainty," *Journal of Machine Learning Research*, vol. 18, no. 26, pp. 1–5, 2017. [Online]. Available: http://jmlr.org/papers/v18/16-300.html

[37] E. Brunskill, L. P. Kaelbling, T. Lozano-Pérez, and N. Roy, "Planning in partially-observable switching-mode continuous domains," *Annals of Mathematics and Artificial Intelligence*, vol. 58, no. 3-4, pp. 185–216, 2010.

[38] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2004.

[39] J. Muesing, L. Burks, M. Iuzzolino, D. Szafir, and N. R. Ahmed, "Fully bayesian human-machine data fusion for robust dynamic target surveillance and characterization," in *AIAA Scitech 2019 Forum*, 2019, p. 2208.