Flexible Semantic Human-Robot Sensing in Unknown Environments using Dynamic Information Gathering Policies

Luke Burks and Nisar Ahmed*

I. Introduction

Dynamic information gathering algorithms typically leverage well-defined models of target dynamics, robot platform motion, and sensor observations to solve challenging combined optimal control and estimation problems. While exact solutions are intractable, approximate algorithms for autonomous data fusion and decision making under uncertainty can be very brittle. This problem is exacerbated in unknown environments, where the demands of online localization, mapping, perception, and planning lead to even greater computing bottlenecks, uncertainties and risks.

To mitigate these issues, human operators and teammates can act as 'sensors' that contribute valuable information beyond the reach of autonomous robots. For instance, vehicle operators in search and tracking missions can provide 'soft data' to narrow down possible survivor locations using semantic natural language observations (e.g. 'Nothing is around the lake'; 'Something is moving towards the fence'), or provide estimates of physical quantities (e.g. masses/sizes or location of obstacles, distances from landmarks) to help autonomous vehicles better judge and understand search areas – thus improving online decision making. But how can autonomous reasoning *actively and opportunistically* engage human sensing in unknown environments?

This paper describes progress toward a framework for intelligent human-autonomy interaction that not only leverages combined robot-human sensing, but is also tightly integrated with dynamic platform decision making and planning. Our framework uses Bayesian data fusion to exploit human sensors and autonomous robotic sensor platforms in a 'plug and play' manner; this idea that has gained increased attention in various contexts over the last decade [1], [2], [3], [4], [5], [6], [7]. However, until now, these developments have only focused on structured and known environments, and have largely ignored coupling to planning/control problems. We combine our recent work on Bayesian semantic robot-human sensor data fusion in structured environments [8], [9], [10], [11] with concepts from optimal active sensing and online planning under uncertainty, in order to develop new methods for *interactive multi-level* human-robot sensing of dynamic states in unknown environments.

*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.

II. Proposed Technical Approach

For concreteness, we focus on dynamic target search and localization applications with human-robot teams, although the principles behind our approach are applicable to other problem domains. Previous work [11], [12], [13] focused on offline approximations of optimal collaborative human-robot information gathering policies in continuous state space target search problems – assuming given models for observations, rewards, and state transitions. In domains involving human interaction, these offline policies assumed a priori knowledge of environments and target behaviors in order to ground the meaning of semantic human sensor inputs (defined according to a structured natural language dictionary). These assumptions fail in unstructured and dynamic environments, as environment knowledge and target behaviors can change rapidly during policy execution.

We address these issues via online policy generation, so that a robot can plan under uncertainty using its most up to date understanding of the world. This allows the robot to tailor its actions both to improve its knowledge of the underlying models being used and achieve its mission objectives. Specifically, building on work in [11], [13], we employ online approximations for solving continuous state POMDPs to formulate joint *action-query policies*. The policies tell the robot how to respond to uncertainty in target location and other relevant planning variables. In this way, the robot simultaneously makes optimal decisions about how to sense and search the environment and about which semantic natural language questions (from a pre-defined dictionary) it should ask human sensors. This allows the robot to prioritize and 'pull' useful information from human sensors to help localize targets and perceive the unknown environment during the search mission. Additionally, the policy can direct the human to provide responses that refine the robot's probabilistic state transition, reward, and observation models. As in our previous work, human responses can be provided via structured natural language [10], [14] or as direct 'yes/no' answers to binary semantic queries [12], [13]. However, in unknown environments, the set of groundings and semantic references are likely to be highly limited in the initial phase of the mission. Thus, building on [9], human sensor data can be more conveniently provided via free-form semantic sketches on dynamically updated metric maps.

A. Illustrative Conceptual Example

Consider the scenario in Figure 1, in which a soldier and aerial robot team is out in a field searching for a moving target. While the extent of the space is known, and the

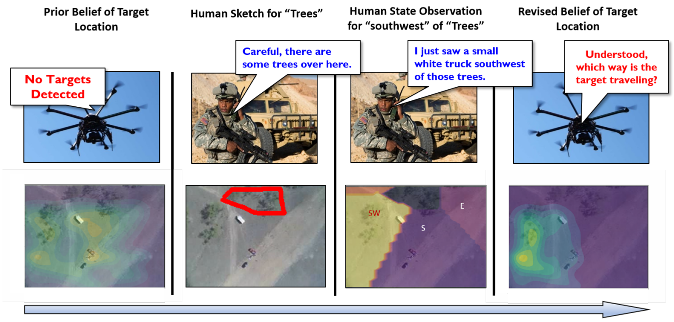

Fig. 1: Sequence of environment structure and target state belief updates following semantic human sensor inputs.

locations of the soldier and robot are known via GPS, the locations of obstacles and landmarks are not known a priori. The robot maintains a continuous belief $p(s)$ of the target's state s in the environment, modeled as a Gaussian Mixture (GM) pdf, with component weights, means, and covariances,

$$p(s) = \sum_{i=1}^{N} w_i \mathcal{N}_s(\mu_i, \Sigma_i).$$

The robot can use data from its onboard visual sensor to update and act on $p(s)$ in order to make decisions about how to localize and capture the target. Optimal search strategies can be found by approximating continuous POMDP policies via GM-based point-based value iteration (PBVI) [11], [15].

In traditional POMDP-based planning, observations change only an agent's belief about the state of the world (e.g. target location). However, observations from a human sensor need not be limited to state dependent observations; humans can also provide information about the structure of the environment around them. These structural observations change the models the robot uses for planning, thus changing the way the belief is changed. Structural semantic observations can be quite rich with information pertaining to state transition probabilities, observations, and reward functions (possibly all at once). For example, by specifying the location and name/type of an obstacle, human input can dynamically alter the transition function for a ground robot by blocking off a part of the search space that can't be safely entered. At the same time, the human input can create a new grounding reference for future semantic observations, as well as alter reward functions in the area of the object to avoid obstacle collisions. This allows the human to constrain and add structure to otherwise unknown and unstructured environments. More generally, human sensor data can be used to correct, add, or remove information about objects that are perceived and mapped by the robot itself.

In this example, the robot first receives semantic structural information about the environment (2D location and extent of a cluster of trees) from the soldier using a sketch interface, similar to the one in [9]. For simplicity, the sketch interface considered here (2nd column of Fig. 1) is adapted such that structural inputs from the human are treated as perfect observations, and the sketches are not gridded up over the search space. The trees are important to avoid for the flying robot, and thus can be encoded as obstacles in a configuration space map for motion planning by creating regions of high negative reward. But the trees also can serve as an anchor for semantic observations (and queries) in the future. Using the synthesis technique developed in [10], the sketch is converted into a softmax model $P(o = j|s)$ (3rd column of Fig. 1), with parameters w_j, b_j for semantic labels j,

$$P(o = j|s) = \frac{e^{w_j^T s + b_j}}{\sum_{c=1}^{m} e^{w_c^T s + b_c}}$$

This model encodes the likelihood for 'allowable' human observations, as determined by the type of object labeled by the human and the mission context. In this example, 'trees' must be pre-defined in a semantic dictionary, such that the robot understands that a human sensor can use this landmark to refer to the relative object positions via cardinal directions.

The robot then receives noisy semantic target state observation o from the soldier, indicating target position relative to the trees. Based on the previous structure observation, the robot translates "southwest of the trees" into a softmax label class. The target state pdf is then updated via Bayes' Rule, by approximating the product of the GM prior and softmax likelihood to get a GM posterior [8],

$$p(s|o) = P(o|s)p(s)/P(o) \approx \sum_{i=1}^{\tilde{N}} \tilde{w}_i \mathcal{N}_s(\tilde{\mu}_i, \tilde{\Sigma}_i)$$

To fully leverage the new human-provided structural and target state data, the robot must be able to adapt its continuous POMDP search policy top optimize decision making given its new environment model, semantic human sensor model, and target beliefs. As such, the policy should consider robot actions for planning to comprise both robot platform movements and queries to semantic human sensors. That is, the policy should actively suggest semantic observations for the human to provide about the target state and/or environment that aid robot motion planning, as well as consider robot motions that improve human-robot sensing (4th column of Fig. 1). This concept was recently explored for known search environments via offline PBVI-based CPOMDP policy approximations [13]; we are working to leverage insights from this offline approach (for speeding up GM-based PBVI with semantic softmax observations) for online policy approximation algorithms using Monte Carlo Tree Search methods. Partially Observable Monte Carlo Planning (POMCP) [16] is particularly suitable since it requires only a"black-box" model of the problem (to simulate dynamics, observations, and rewards) that can be dynamically updated with new structural information from the human sensor.

In the full paper, we will demonstrate a proof-of-concept dynamic policy implementation for a mock version of the Fig. 1 scenario. Open issues to be explored and discussed include: (i) strategies for framing human sensor queries (especially accounting for cognitive factors and instances where human input can potentially lead to more uncertainty/harm than good); (ii) rigorously accounting for ambiguous/incorrect semantic observations; and (iii) reasoning about how sketches should trigger CPOMDP model updates.

REFERENCES

[1] T. Kaupp, B. Douillard, F. Ramos, A. Makarenko, and B. Upcroft, "Shared Environment Representation for a Human-Robot Team Performing Information Fusion," *Journal of Field Robotics*, vol. 24, no. 11, pp. 911–942, 2007.

[2] 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 *International Conference on Intelligent Robots and Systems*, 2008, pp. 2342–2349.

[3] J. Frost, A. Harrison, S. Pulman, and P. Newman, "Probabilistic Approach to Modelling Spatial Language with Its Application to Sensor Models," in *Workshop on Computational Models of Spatial Language Interpretation (COSLI-2010) at Spatial Cognition 2010*, J. H. R. Ross and J. Kelleher, Ed., Portland, OR, 2010.

[4] B. Khaleghi, A. Khamis, and F. Karray, "Random finite set theoretic based soft/hard data fusion with application for target tracking," in *2010 Conf. on Multisensor Fusion and Integration for Intelligent Sys. (MFI 2010)*, Salt Lake City, 2010, pp. 50–55.

[5] A. Dani, M. McCourt, J. Curtis, and S. Mehta, "Information fusion in human-robot collaboration using neural network representation," in *2014 IEEE Int'l Conf. on Systems, Man and Cybernetics*. IEEE, 2014, pp. 2114–2120.

[6] S. Mehta, M. McCourt, E. Doucette, and J. Curtis, "A touch interface for soft data modeling in bayesian estimation," in *2014 IEEE Int'l Conf. on Systems, Man and Cybernetics*. IEEE, 2014, pp. 3732–3737.

[7] D. Yi, S. Choudhury, and S. Srinivasa, "Incorporating qualitative information into quantitative estimation via sequentially constrained hamiltonian monte carlo sampling," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 4648–4655.

[8] N. Ahmed, E. 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. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6301744

[9] 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*, ser. ICCPS '15. New York, NY, USA: ACM, 2015, pp. 80–89. [Online]. Available: http://doi.acm.org/10.1145/2735960.2735970

[10] N. Sweet and N. Ahmed, "Structured synthesis and compression of semantic human sensor models for Bayesian estimation," *Proceedings of the American Control Conference*, vol. 2016-July, no. 2, pp. 5479–5485, 2016.

[11] L. Burks and N. Ahmed, "Optimal continuous state pomdp planning with semantic observations," in *2017 IEEE Conference on Decision and Control*. IEEE, 2017, pp. 1509–1516.

[12] K. G. Lore, N. Sweet, K. Kumar, N. Ahmed, and S. Sarkar, "Deep value of information estimators for collaborative human-machine information gathering," in *Cyber-Physical Systems (ICCPS), 2016 ACM/IEEE 7th International Conference on*. IEEE, 2016, pp. 1–10.

[13] 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 *submitted to 2018 International Conference on Information Fusion (FUSION 2018)*, 2017.

[14] N. Sweet and N. Ahmed, "Toward natural language semantic sensing in dynamic state spaces," in *2016 RSS Workshop on Model Learning for Human-Robot Communication (MLHRC 2016)*, 2016.

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

[16] D. Silver and J. Veness, "Monte-Carlo Planning in Large POMDPs," *Advances in neural information processing systems.*, pp. 2164–2172, 2010. [Online]. Available: http://discovery.ucl.ac.uk/1347369/