Decision Region Determination for Touch Based Localization

Shervin Javdani*, Yuxin Chen†, Amin Karbasi†, Andreas Krause†, J. Andrew Bagnell*, and Siddhartha Srinivasa*
*Carnegie Mellon University †ETH Zürich

I. Introduction

Many robotic systems deal with uncertainty by performing a sequence of information gathering actions. Unfortunately, it is not only NP-hard to solve optimally, but also NP-hard to approximate better than logarithmically [3].

Instead, many successful systems perform a shallow online search, optimizing for a specified metric on the resulting belief distributions. One often used metric, known as *information gain*, maximizes the expected decrease in Shannon entropy [2, 5, 1, 18, 6, 4, 12, 9, 15]. Shannon entropy measures the amount of uncertainty in a distribution. Thus, minimizing this is useful if the objective is to remove all uncertainty. In our prior work, we formulate a similar metric for manipulation, with the goal of finding the location of an object from an uncertainty distribution over pose. Notably, we prove our metric produces near-optimal sequences when selecting actions greedily, whereas optimizing Shannon entropy may not [13].

In many robotics problems, you don't need to reduce uncertainty completely - you simply need to reduce it enough to accomplish your task. Optimizing for a task-driven criteria directly enables the system to utilize fewer information gathering actions than reducing all uncertainty. For example, Hsiao [11] found that optimizing for probability of success outperformed optimizing for reducing Shannon entropy. In decision theory, value of information (VoI) [10] is a commonly used metric which attempts to capture this idea.

In our current and ongoing work, we formulate an objective for task-driven uncertainty reduction. We design metrics which are *adaptive submodular* [7], rendering the greedily policy with this metric near-optimal. We present the formalization of this idea for touch based localization, and an algorithm with near-optimality bounds published recently [14]. Additionally, we present our newly developed formulation, which is significantly faster to compute.

II. Formulation

As a running example, consider the task of pushing a button on a microwave, where the pose of the microwave is unknown. We suppose we have a prior discrete set of *hypotheses* representing uncertainty. Additionally, we have a set of button-pushing actions we can perform, which we call *decisions*. Each of these decisions will accomplish our task on a subset of the hypotheses. We call this subset a *decision region*. Note that a button pushing action may succeed for many object positions - thus, we do not need to reduce uncertainty completely, but just enough to accomplish our ask. See Fig. 1.

To gather information, we generate set of *tests* we can perform. In our button-pushing task, we use *guarded moves* [17], where the end effector moves along a path until contact is sensed. These tests operate on the space of hypotheses - that is, a test and the corresponding observation either keeps or removes each hypothesis. If the observation is consistent with that hypothesis, e.g. for a guarded move [17], we would have felt contact if the object was at the hypothesized location, we keep it. Otherwise, that hypothesis cannot be true, and as such is disproven and removed.

Our goal is to adaptively select a sequence of tests such that after performing those tests and receiving observations, all consistent hypotheses are encapsulated by one single decision region. We call this the *Decision Region Determination* (DRD) problem. See Fig. 2.

III. Contribution Overview

Which tests will help us make a decision? We have developed two metrics capturing this idea. For both, we prove that our metric is maximized iff all consistent hypotheses are encapsulated by one decision region. Additionally, both metrics are adaptive submodular, rendering the greedy algorithm near-optimal. They differ in the computation time of each metric, and the specific approximation bound they achieve.

Concretely, each approximation bound is of the form:

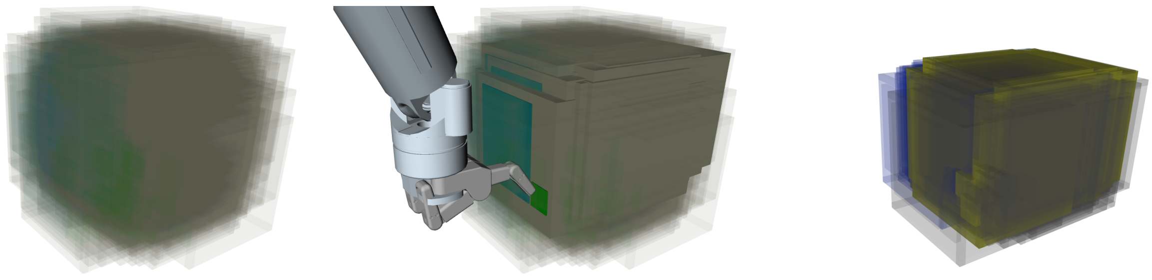
$$\mathcal{C}(\pi^G) \leq \alpha \mathcal{C}(\pi^*)$$

where π^G is the greedy policy with our metric, π^* the optimal policy, $\mathcal{C}(\pi)$ the cost we want to minimize, and α the algorithm-specific bound, which we will present below.

A. *Hyperedge Cutting (*HEC*)*

We know that if a single decision does not succeed for a set of hypotheses, some of those hypotheses must be disproven prior to making a decision. Our approach considers a particular hypergraph, which we call the *Splitting Hypergraph*, where hyperedges correspond to these sets. A hyperedge is removed by disproving any hypothesis that it connects. Our objective function is to maximize, in expectation, this hyperedge removal. We call this approach *Hyperedge Cutting* (HEC). Importantly, all hyperedges are removed iff all consistent hypotheses are encapsulated by one decision region.

We compute this edge removal by evaluating a particular polynomial, which exhibits additional structure. In particular,



(a) Hypotheses (b) A decision region (c) Two regions

Fig. 1: Touch based localization for pushing the button of a microwave. Given hypotheses over object location (a), decision actions are generated. The corresponding decision regions are computed by forward simulating to find hypotheses for which it would succeed (b). Decision regions will overlap. In (c), we see two regions (blue and grey) and their overlap (yellow).

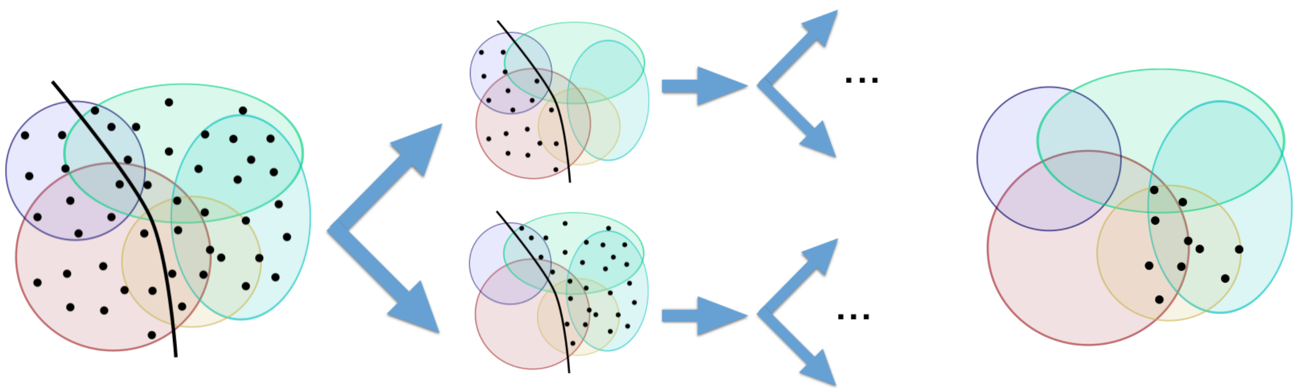


Fig. 2: An overview of the DRD setup. Hypotheses are shown as black dots, decision regions as colored circles, and tests as black lines. We start (left) with a prior set of hypotheses, and a set of decisions we might make. Each test can result in some observations (center), where we keep hypotheses consistent with that observation. We repeat this process until all remaining hypotheses are encapsulated by a decision region (right).

it can be exactly expressed as a sum of complete homogeneous symmetric polynomials, which can be computed efficiently. This insight enables to us utilize a more efficient algorithm.

For this algorithm, our approximation factor is $\alpha = k\ln(1/p_{\min}) + 1$, where k is a constant corresponding to the amount of overlap between regions, and $p_{\min}$ is the minimum prior probability of any hypothesis. Empirically, our results indicate that it outperforms similar algorithms for these tasks. See [14] for those results, and a specific definition of k.

Unfortunately, this metric can be slow to evaluate, even with our utilization of complete homogeneous symmetric polynomials. It is exponential in our constant k. Thus, when we have large overlap between regions, this constant becomes large and the algorithm intractable.

B. Noisy-Or

Suppose we have m decision regions. Instead of constructing one hypergraph over all decision regions, we will construct m objectives, one for each region. Each objective is maximized iff all remaining hypotheses are within the corresponding region. More specifically, each objective is an instance of EC2 [8], which handles the case of separated decision regions. The total objective function consists of using noisy-or formulation over each objective, such that the total objective is maximized if any objective is maximized (objective 1 is maximized, or objective 2, or…).

This algorithm is linear in the number of decision regions. Maximizing this has an approximation factor is $\alpha = 2m\ln(1/p_{\min}) + 1$, where the factor m is a result of taking the product of m EC2 instances.

However, we note that some instances of EC2 can be combined. We show that solving for which instances can be combined is equivalent to solving a graph-coloring problem, where each color corresponds to an instance of EC2. One can show that every graph can be colored with one more color than the maximum vertex degree using the greedy algorithm [16]. In our DRD problem, the maximum vertex degree is a constant $\widehat{k}$ corresponding to the overlap between regions. It is similar to k for HEC, though $k \leq \widehat{k}$. This formulation has an approximation guarantee $\alpha = 2\widehat{k}\ln(1/p_{\min}) + 1$

In practice, this algorithm is significantly faster than HEC, at the cost of only a small increase in query complexity. Furthermore, it can solve problems with high region overlap, which HEC cannot. Our preliminary results indicate that this algorithm is well suited for use in touch based localization.

References

[1] Frederic Bourgault, Alexei A. Makarenko, Stefan B. Williams, Ben Grocholsky, and Hugh F. Durrant-Whyte. Information based adaptive robotic exploration. In *IEEE/RSJ IROS*, 2002.

[2] Anthony R. Cassandra, Leslie Pack Kaelbling, and James A. Kurien. Acting under uncertainty: Discrete bayesian models for mobile-robot navigation. In *IEEE/RSJ IROS*, 1996.

[3] Venkatesan T. Chakaravarthy, Vinayaka Pandit, Sambuddha Roy Pranjal Awasthi, and Mukesh Mohania. Decision trees for entity identification: Approximation algorithms and hardness results. In *PODS*, 2007.

[4] Lawrence Erickson, Joseph Knuth, Jason M. O'Kane, and Steven M. LaValle. Probabilistic localization with a blind robot. In *IEEE ICRA*, 2008.

[5] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Active markov localization for mobile robots. *Robotics and Autonomous Systems*, 25:195–207, 1998.

[6] Jiaxin Fu, Siddhartha Srinivasa, Nancy Pollard, and Bart Nabbe. Planar batting under shape, pose, and impact uncertainty. In *IEEE ICRA*, 2007.

[7] Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *JAIR*, 42(1):427–486, 2011.

[8] Daniel Golovin, Andreas Krause, and Debajyoti Ray. Near-optimal bayesian active learning with noisy observations. In *NIPS*, 2010.

[9] Paul Hebert, Joel W. Burdick, Thomas Howard, Nicolas Hudson, and Jeremy Ma. Action inference: The next best touch. In *RSS Workshop on Mobile Manipulation*, 2012.

[10] Ronald A. Howard. Information value theory. In *IEEE Trans. Syst. Sci. Cybernetics*, 1966.

[11] Kaijen Hsiao. *Relatively Robust Grasping*. PhD thesis, Massachusetts Institute of Technology, 2009.

[12] Kaijen Hsiao, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. Robust belief-based execution of manipulation programs. In *WAFR*, 2008.

[13] Shervin Javdani, Matthew Klingensmith, J. Andrew (Drew) Bagnell, Nancy Pollard, and Siddhartha Srinivasa. Efficient touch based localization through submodularity. In *IEEE ICRA*, 2013.

[14] Shervin Javdani, Yuxin Chen, Amin Karbasi, Andreas Krause, J. Andrew (Drew) Bagnell, and Siddhartha Srinivasa. Near optimal bayesian active learning for decision making. In *AISTATS*, 2014.

[15] Andreas Krause and Carlos Guestrin. Near-optimal nonmyopic value of information in graphical models. In *UAI*, 2005.

[16] D. J. WELSH and M. B. POWELL. An upper bound for the chromatic number of a graph and its application to timetabling problems. *Computer Journal*, 10:85–86, 1967.

[17] Peter M. Will and David D. Grossman. An experimental system for computer controlled mechanical assembly. *IEEE Trans. Computers*, 24(9):879–888, 1975.

[18] Alice X. Zheng, Irina Rish, and Alina Beygelzimer. Efficient test selection in active diagnosis via entropy approximation. In *UAI*, 2005.