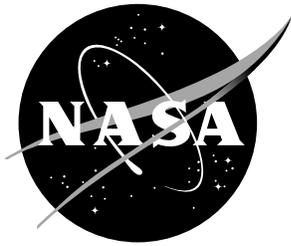# Information Foraging and Change Detection for Automated Science Exploration

(P. Michael Furlong)
(SGT Inc., Intelligent Robotics Group)

(Michael Dille)
(SGT Inc., Intelligent Robotics Group)

NASA STI Program...in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collection of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

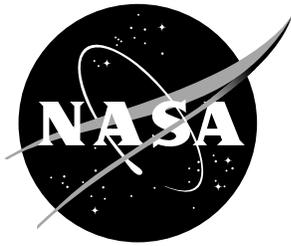- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI Program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

Information Foraging and Change Detection for Automated Science Exploration

(P. Michael Furlong)
(SGT Inc., Intelligent Robotics Group)

(Michael Dille)
(SGT Inc., Intelligent Robotics Group)

Abstract

This paper presents a new algorithm for autonomous on-line exploration in unknown environments. The objective is to free remote scientists from possibly-infeasible extensive preliminary site investigation prior to sending robotic agents. We simulate a common exploration task for an autonomous robot sampling the environment at various locations and compare performance against simpler control strategies. An extension is proposed and evaluated that further permits operation in the presence of environmental variability in which the robot encounters a change in the distribution underlying sampling targets. Experimental results indicate a strong improvement in performance across varied parameter choices for the scenario.

1   Introduction

Robots are our scientific surrogates as we peer into the depths of the ocean or into our solar neighbourhood. High-latency and low-bandwidth communication to these regions limits situational awareness and reaction times of the scientists controlling such robots. Therefore it is vital to increase the ability of robotic explorers to independently make in-mission decisions. A common exploration activity is remote sensing, in which a robot is tasked with collecting sensor data by sampling the environment at various locations. Many specialized sensors employed for activities such as biological collection and spectral mapping require long, energy-intensive sampling durations or the activation of single-use collection canisters. Constraints on mission length and payload capacity, coupled with limited remote operator awareness, necessitate some autonomy in sampling location selection for mission productivity and success.

Currently fielded robots either depend highly on operators for objectives or plan with considerable global knowledge. Operating in such conditions constrains them to rely on either remote human decision-making (requiring often impractical levels of situational awareness) or significant amounts of prior scouting, obviating the need to send a robotic agent. These limitations are mirrored in existing literature, which fails to provide principled reasoning about what to investigate *in situ* without such reliances.

This paper proposes an algorithm that addresses one example of such missions, in which objects or areas in the environment lie within some easily sensed class, and each class possesses some underlying data distribution (e.g. microbial colonization) that can only be sensed with expensive specialized sensors. The overall goal is to estimate the underlying distribution of each class with maximal accuracy.

For general applicability no global information, such as prior maps of sampling opportunities, is available. Sensing opportunities are assumed to arise nondeterministically (e.g. from classes present along a pre-determined trajectory or as currents draw objects past the robot), and the robot cannot return to objects it did not sample. Thus, the problem can be thought of as a stream of sensing opportunities providing varying reward (information about underlying class distributions), each requiring a decision to sample or move on.
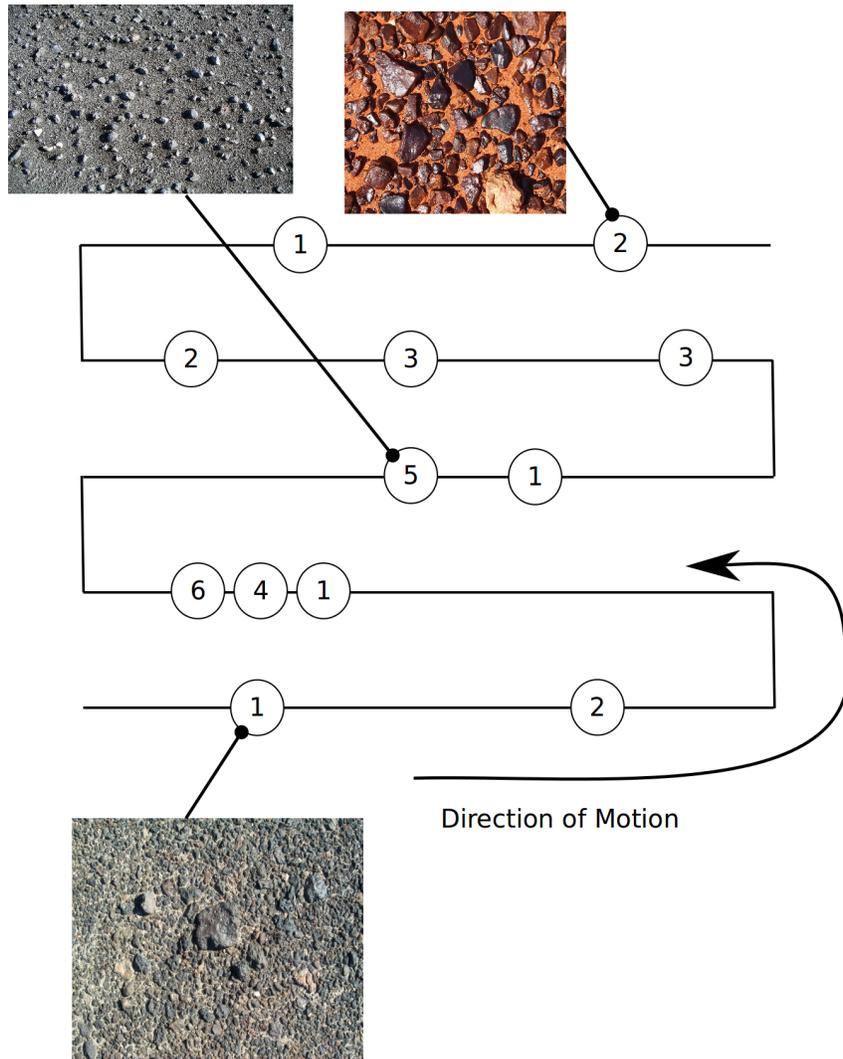
Figure 1. A cartoon of a path explored by a rover. The images represent different classes of desert pavements that may encountered by a rover as it follows a pre-determined path.

The proposed algorithm draws on techniques from optimal foraging theory and sequential experiment selection. Its use is motivated by observations of human and animal behavior, exemplified by geologists making decisions about investigating local phenomena without prior access to detailed maps, who are able to effectively choose between sampling materials in front of them or moving on to potentially more profitable sampling locations. These decisions may not be globally optimal, but they demonstrate an ability lacking in current exploration robots: to make decisions to stop and engage with the environment or to continue traveling in the hope of finding more informative sampling locations.

The algorithm is then extended to address another common situation in scientific exploration, namely environmental variability. As the robot traverses the environment, it may easily pass between larger regions in which the underlying distribution of classes differs. In the Atacama desert 100% of photosynthesis-promoting translucent rocks are colonized by microbes in semi-arid regions, but less than 50% of such rocks in semi-arid regions, and less than 1% in the hyperarid core [1, 2]. Detecting and reacting to such changes is relevant both for scientific interest and so sampling decisions will not be based on historically observed but now inaccurate class information. An extension is proposed incorporating an additional statistical test to detect a change in class distribution to notify operators, separate data segments, and reset empirical history that might otherwise misinform upcoming sampling decisions.

The remainder of this document begins with a brief survey of the relevant literature. Next, a detailed comparison of the proposed foraging algorithm and one based upon existing principles from the design of experiments literature. Finally, discussion of experimental results from a simulated exploration scenario indicates that under limitations on sample collection and overall mission time, the foraging algorithm presents a significant improvement for a realistic range of sampling costs.

2   Background

Automating experiment design and selection is not without precedent. Kristine Smith started the field of optimal experiment design in 1918. [3] Recently robots have been employed to conduct scientific exploration autonomously. [4, 5] Current robot scientists' reliance on global information prevents them from operating in truly unknown environments. Additionally, previous approaches in sequential decision making from statistics do not necessarily reflect the settings that autonomous robots encounter in the real world.

2.1   Multi-armed bandits

Sequential experiment selection, a type of active learning, is addressed in the multi-armed bandit (MAB) literature. This was introduced by Robbins [6] as a means of sequentially selecting which experiments to conduct with a limited budget. In Robbins' work, selecting experiments is modelled on determining the payouts of one-armed bandit machines – each machine representing a different experiment. The player has a fixed sampling budget and has to sequentially choose which machine to play, trading off exploiting expected rewards from well-studied arms against

exploring different arms, learning more accurately the payouts of those arms.

Lai and Robbins [7] use a value function in which uncertainty in arm rewards makes an arm more interesting. Recently decision rules like Thompson sampling [8] and Bayesian Optimal Control [9] have gained popularity. Other techniques addressing the exploration/exploitation problem use uncertainty as a reward metric. [10–12] In our setting, because the agent only needs to learn the distribution and not use it for anything, uncertainty is the only necessary reward.

Several factors distinguish the MAB setting from the problem explored in this paper. In MAB, the agent has access to any arm (analogous to a class in our setting) it chooses at any given time. The agent in our setting does not get to choose which of the classes it can investigate. Any previously seen classes are no longer available, and new classes arrive per a random model. Additionally, the standard MAB setting does not have switching costs, although there are some formulations which do include such costs. [13] In our setting, there is a cost incurred with every choice to continue exploring.

2.2   Optimal Foraging

Foraging is the problem encountered by animals seeking to maximize energy intake when operating in unknown environments. The central question of the problem is whether it is more valuable to continue extracting resources from the current location or to seek out resources in new locations. Charnov [14] introduced a technique for dealing with "patchy" environments, in which there are distinct regions that contain different classes of resources. The forager can extract value from these patches, with diminishing returns (modeling resources consumed), or choose to continue to wander randomly in the hopes of encountering more valuable locations.

The optimal time to leave a patch, according to Charnov's Marginal Value Theorem, is when the expected return from continuing to sample a patch is less than the expected return from searching the environment. In this formulation, the expected return from both the current patch and the environment are offset by the cost of extracting resources in this patch and the energy spent seeking a new patch.

Pirolli and Card [15] studied researchers attempting to acquire information. They modelled the rate of information gain and had their agent decide to leave a patch when the rate of information gain was lower than that of the environment. What differentiates their setting from ours is that their decision maker can choose which patch to sample, yet our exploring agent cannot.

Kolling *et al.* [16] studied humans engaged in a gambling task in which players have to consider the option they have before them and the opportunities the environment provides. Subjects were repeatedly presented with a choice of playing a gambling game or being randomly presented with a different game. Each game was a Bernoulli trial with some unknown probability of success. Kolling *et al.* identify possible neural substrates for foraging decisions in humans. The behaviour was near optimal, with some skewing of probabilities near 0 or 1.

2.3 Science Autonomy

Thompson and Wettergreen [17] maximize diversity of collected samples with mutual information sampling. This approach ensures diversity in the collected sample set, an act that reduces uncertainty in the input space of a function. However, this approach does not consider non-deterministic results of sampling.

Ferri *et al.* [18] address prospecting where an autonomous underwater vehicle (AUV) follows a predefined track and needs to decide when to deviate to sample anomalies. The AUV examines anomalies by searching in a spiral pattern, collecting data and characterizing the environment in that location. The AUV's sampling capacity is limited only by time. The decision to sample is based on a fixed threshold. While this may accurately encode subject matter experts' beliefs on what is interesting, it is fragile in the face of variable and unknown environments the AUV encounters. This exploration problem is an ideal application of the algorithm proposed in this paper.

Likewise, Girdhar *et al.* [19] present an approach to autonomous exploration wherein a robot investigates a scene when it encounters unexpected phenomena. Specifically, they use topic models to describe scenes and sample when scenes do not fit into the topic models they have constructed. In these works, the vehicle has no limit on its sampling capacity and is always collecting data. By slowing the vehicle down, more samples are collected in anomalous scenes. In this fashion this is very similar to work by Thompson *et al.* [20]

Girdhar *et al.* [21]build upon their anomaly detection techniques to develop a path planning method to maximize information gain of paths. In that respect, it belongs with the family of curiosity-driven algorithms pioneered by Sun *et al.* [22] Fundamental to these approaches is that explorers should spend time investigating regions of the world (or hypothesis space) where learned models are the least certain.

Previous work by the primary author with optimal foraging for science autonomy has considered robots with sampling budgets limited by a number of containers (**removed for anonymity**) and assumed knowledge of the number of sampling opportunities that would occur. While the limited sampling budget is realistic, foreknowledge of the transect is not. This paper improves upon the prior work by using productivity to reason about sampling choices and gives a constraint of time instead of an unknowable number of sampling opportunities.

3   Method

A simulated scenario is considered where a rover explores a path set for it by remote scientists. The exploration budget is 100 units of arbitrary time. While following this path the agent is repeatedly presented with one of K possible materials. The agent does not know how many different types of materials it may encounter during its travels. At every presentation the agent has a choice of sampling that material, represented by taking action $\xi_k \in \Xi$ and making an observation Z, or continuing along the path in the hopes of finding a more interesting sampling opportunity. The role of the agent is to determine $P(Z|\xi_k) = \theta_k^z (1 - \theta_k)^{1-z} \, \forall \xi_k \in \Xi$.

The experimental setup is a variation on Charnov's patchy foraging (see Section

2.2). In this case we assume a patch is exhausted by taking one sample. If the agent choses to continue searching it will be presented with a new material, drawn with probability $P(\xi_k)$. When patches are classes of random variables and the reward is information gained about the underlying distribution, the reward is less at every encounter, unless environmental conditions change the underlying distribution.

In this paper, we choose to model the different classes of materials as Bernoulli random variables, representing the common scientific exploration scenario of detecting the presence of a phenomenon of interest such as whether or not a material is colonized by microbes. We place a Beta prior on the parameter $\theta_k \sim Beta(\alpha_k, \beta_k)$ that determines the probability that a class of material is colonized, yielding a belief post-observation that $\mathbb{E}[\theta_k] = \alpha_k / (\alpha_k + \beta_k)$ where α_k is the number of times material k was observed being colonized ("success"), and β_k is the number of times material k was observed as not being colonized ("failure").

We anticipate that the agent will encounter a number $K = |\Xi| \leq \infty$ classes of random variables while exploring. In these experiments we set $K = 3$. However the agent is never informed of how many classes of materials exist in the environment.

3.1 Algorithms

Four algorithms for sampling decision-making are evaluated in these experiments. Three of these algorithms estimate the reward of action ξ_k by using Lindley's [23] value of an experiment, given in Equation 1. This reward represents the expected information gain over all possible observations that may result from choosing to take sampling action ξ_k.

$$R(\xi_k) = H(\theta_k|z_{k,1:t-1}, \xi_k) - \mathbb{E}_Z[H(\theta_k|z_{k,1:t}, \xi_k)], \tag{1}$$

where $z_{k,1:t}$ refers to the t observations that were collected for random variable ξ_k.

The first algorithm, **control 1**, will only choose to sample ξ_k if it is has the highest reward compared to any other $\xi_j, j \neq k$. This algorithm does not take into account the cost of moving to finding the next ξ_k, nor the rate at which they arrive. This algorithm corresponds to the simple greedy strategy of maximizing immediate reward.

The second algorithm, **control 2**, will choose to sample ξ_k if any other random variable, $\xi_j, j \neq k$, has been sampled more than ξ_k. Like **control 1**, this algorithm does not take into account the cost of traverse nor the cost of taking a sampling action. This algorithm attempts to distribute samples uniformly across all classes and provides valuable comparison as it has been previously shown to be a robustly successful strategy. [24]

The third algorithm, **foraging** (Algorithm 1), chooses to sample if the expected rate of reward of ξ_k is greater than or equal to the expected reward from continuing to explore the environment, that is whether greater *productivity* is to be had from loitering or from continuing on. This captures traversal and sampling costs (J in Algorithm 1). We place a Dirichlet prior on the occurrence of these random variables, estimating the probability of encountering class ξ_k as $\hat{P}(\xi_k) = n_k / \sum_{j=0}^{j=|K|} n_j$, where

6

n_k is the number of times ξ_k has been encountered. The distribution $\hat{P}(\xi_k)$ is used to compute the $prod_{continue}$ in Algorithm 1.

The fourth algorithm uses the same decision rule as Algorithm 1, but after it makes an observation it checks to see if the underlying distribution has changed, as in "DETECT_CHANGE" in Algorithm 2. It detects the change with a likelihood ratio test. It maintains two windows of observations for each ξ_k, one which is initially populated with *window_size* many observations, the other populated with the *window_size* most recent observations. The two windows represent hypotheses about the parameter θ_k. A third window of the *sample_size* most recent observations is used as the test population. We employ Wald's sequential probability ratio test [25] to determine if the observations in the second window represents a different distribution from the first. We select the threshold for detecting a change in the distribution, *change_threshold*, as specified in [25]. *window_size* is arbitrarily set to be 30, and *sample_size* to 5. If a distribution change is detected the current world model is cached, and the rover resets its sampling algorithm to an initial state.

Algorithm 1 Foraging Sampling Strategy

    function INIT_FORAGE_SAMPLING
        $\Xi \leftarrow \emptyset$
        $R(\cdot) \leftarrow \emptyset$
        $N. \leftarrow \emptyset$
    end function
    function FORAGE_SAMPLE(ξ_k)
        if $\xi_k \notin \Xi$ **then**
            $\Xi \leftarrow \Xi \cup \xi_k$
            $N_k \leftarrow 0$
            return sample
        end if
        $N_k \leftarrow N_k + 1$
        $prod_{sample} \leftarrow R(\xi_k)/J(sample)$
        $prod_{continue} \leftarrow \mathbb{E}_\Xi[R(\xi)]/(J(sample)+J(search))$
        if $prod_{sample} \geq prod_{continue}$ **then**
            return sample
        else
            return continue
        end if
    end function

3.2 Experiments

We conducted three experiments to demonstrate the effectiveness of our algorithm, varying the underlying distribution of each class, class arrival probability, and introducing a class distribution change during the experiment. The costs of sampling and searching were varied over $\{0.1, 0.2, 0.5, 0.75, 1.0, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ for experiments 1 and 2. In each experiment, for each setting of experiment parameters

and costs we ran 50 trials for each algorithm.

3.2.1 Experiment 1 - Underlying Distribution

In the first experiment the arrival probability is fixed with a constant uniform distribution. That is to say the probability that the next random variable to be presented to the agent is $P\left(\xi_k\right) = 1/3$. In this experiment we vary the underlying distribution of the random variables, $P\left(Z|\xi_k\right)$, which is the probability that the material represented by ξ_k is colonized.

Table 1. Experiment 1 Parameter Settings

Experiment	$P(Z|\xi_1)$	$P(Z|\xi_2)$	$P(Z|\xi_3)$
1.1	0.001	0.500	0.999
1.2	0.001	0.300	0.001
1.3	0.001	0.500	0.001
1.4	0.001	0.750	0.001
1.5	0.001	0.999	0.001

3.2.2 Experiment 2 - Arrival Probability

In the second experiment the probability of the different materials being colonized was held constant while the arrival probability is varied. The probabilities of being colonized are $P\left(Z|\xi_k\right) = \{0.001, 0.500, 0.999\}$. These values were because the expected rewards of the three random variables symmetrically span the range $\theta_k \in (0, 1)$. While the entropy of $P\left(Z|\xi_1\right)$ and $P\left(Z|\xi_2\right)$ are the same, $P\left(Z|\xi_2\right)$ has the maximum entropy possible for a Bernoulli distribution. This was done to determine if one random variable attracted the attention over the others because of either the expected value or the entropy of the underlying distribution.

Table 2. Experiment 2 Parameter Settings

Experiment	$P(\xi_1)$	$P(\xi_2)$	$P(\xi_3)$
2.1	0.6	0.3	0.1
2.2	0.3	0.1	0.6
2.3	0.1	0.3	0.6
2.4	0.1	0.6	0.3
2.5	0.3	0.6	0.1
2.6	0.6	0.1	0.3

3.2.3 Experiment 3 - Distribution Change

In this experiment we compare only the Foraging algorithm against the Foraging Algorithm with change detection. In the final experiment we fix the arrival probabilities $P\left(\xi_k\right) = \{0.4, 0.3, 0.3\}$, and fix the sampling and searching costs at 0.01 and 0.1, respectively. Halfway through this experiment the underlying distribution

is changed from $P\left(Z|\xi_k\right) = \{0.001, 0.500, 0.001\}$ to $P\left(Z|\xi_k\right) = \{0.001, 0.001, 0.300\}$. The change occurs when the agent gets halfway along its path. Because the foraging algorithm has a prior belief defined for all sampling outcomes we can use Kullback-Leibler (KL) divergence between true and estimated $P\left(Z|\xi_k\right)$ to measure algorithm performance.

4   Results

To determine the success of an algorithm we use Hoeffding's inequality defined by

$$P\left(|\theta_k - \hat{\theta}_k| > \gamma\right) \leq 2 \exp\left(-2\gamma^2 n_k\right) \tag{2}$$

to determine the error in estimating the parameter θ_i. Hoeffding's inequality was chose over the more standard KL divergence because there were some trials where agents observed either no successes or no failures for a given ξ_k, even though for all $\xi_k, \theta_k \in (0, 1)$. In this case the KL divergence is undefined, instead we computed the error by determining the setting of γ for the number of observations of ξ_i and a fixed probability of $P\left(|\theta_k - \hat{\theta}_k| > \gamma\right) \leq 0.05$.

For experiments 1 and 2 we present a 3D plot showing how the error in estimating the parameters θ_k is reduced by using the foraging algorithm over the control algorithm. In addition, we present 2D plots showing where either the foraging algorithm performs better than the control algorithm, or the control algorithm performs better than the foraging algorithm, or when their performance is indistinguishable.

We also report the effect size - the ratio of the mean to the standard deviation of the difference between the 50 paired trials. This is a variation of Cohen's d value [26]. Values greater than 1.3 are considered to be very large, above 0.8 to be significant, and below 0.5 to be insignificant. Tables 3 and 4 give the reduction in estimate error averaged over experiments 1.1-1.5 and 2.1-2.6, respectively.

4.1   Experiment 1 - Underlying Distribution

Table 3 shows the foraging algorithm reduces parameter estimate error relative to the two control algorithms. The effect size well exceeds the threshold for significance. Figure 2 demonstrates that the performance of the foraging algorithm generally performs at least as well the control algorithms, and often better. However we notice that when seaching costs are low the control algorithms can outperform the foraging algorithm.

Table 3. Experiment 1 Foraging vs Control Algorithms

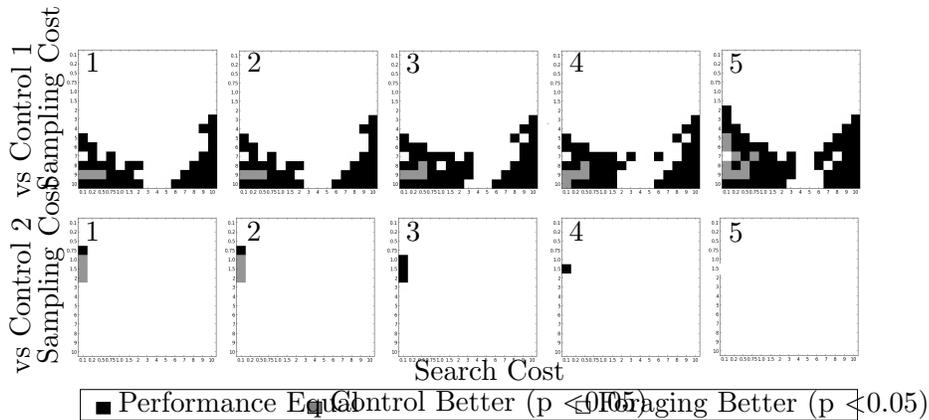| Metric	Control 1	Control 2
Ave. Reduced Error	0.178 ± 0.014	0.559 ± 0.001
Ave. Effect Size	$\mathbf{1.297 \pm 0.030}$	$\mathbf{1.818 \pm 0.007}$

Figure 2. The top row shows how the performance of the foraging algorithm compares to the first (greedy) control algorithm and the bottom row compared to the second (uniform) control algorithm. In the majority of settings for sampling and searching costs the foraging algorithm performs at least as good as the control algorithms and often (white regions) statistically significantly better.

4.1.1 Experiment 2 - Arrival Probability

Figures 3 and 4 demonstrate that, with the exception of when searching costs are small, the foraging algorithm performs at least as good and often statistically significantly better the control algorithms. As we can see in Table 4 the effect size of the error reduction reduction is very large by Cohen's d.

Table 4. Experiment 2 Foraging vs Control Algorithms

Metric	Control 1	Control 2
Ave. Reduced Error	0.202 ± 0.082	0.622 ± 0.015
Ave. Effect Size	$\mathbf{1.375} \pm 0.122$	$\mathbf{2.209} \pm 0.054$

4.1.2 Experiment 3 - Distribution Change

Figure 5 shows that the foraging algorithm with change detection performs substantially better than not using it. The leftmost bars show the performance of the foraging algorithm with and without change detection immediately before the change in the underlying distribution. By employing change detection ("After" in Figure 5) we see the error in estimating the underlying distribution is profoundly reduced. However it should be noted that if the number of opportunities to sample is not very large, then the agent will not be able to detect a change in an underlying distribution with any confidence. This must be considered when planning exploration missions.
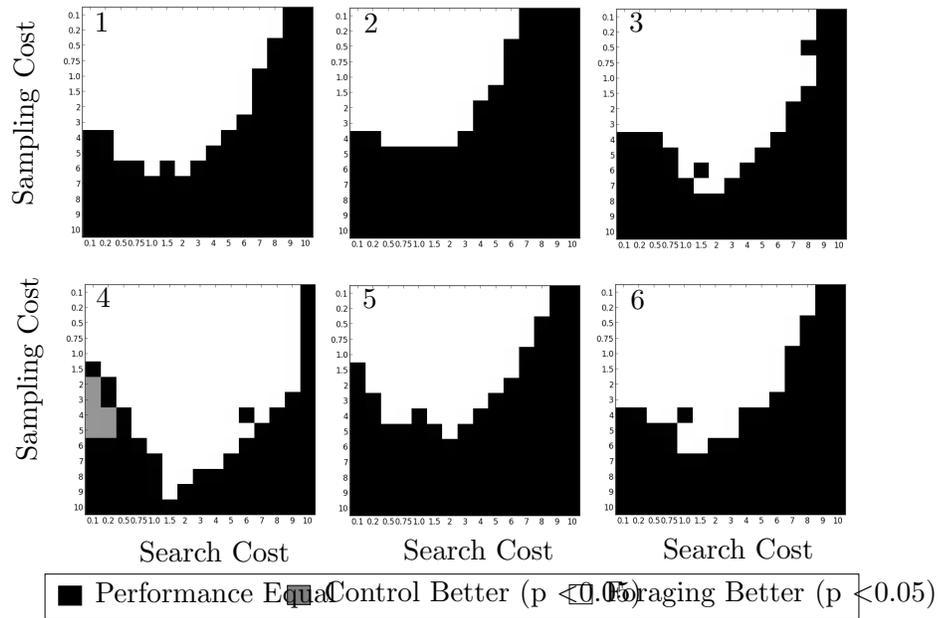
Figure 3. Compared to the first (greedy) control algorithm, the foraging algorithm generally does at least as well or better at estimating the underlying distributions, with statistical significance. However, for low search costs the control algorithm 1 does perform better than foraging.
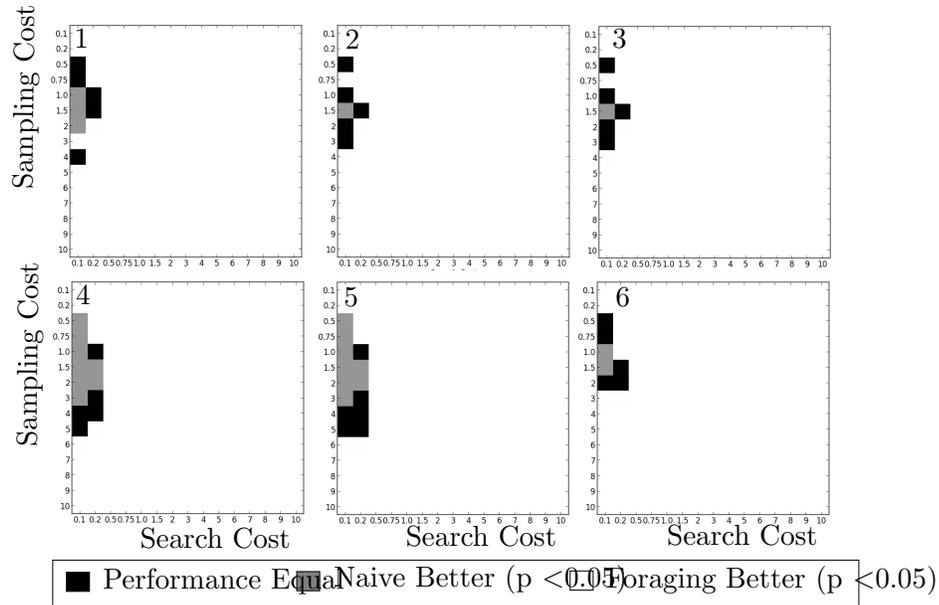

Figure 4. In the second experiment the foraging algorithm does profoundly better than the second (uniform) control algorithm. Statistical significcnace is achived in the majority of sample and search cost pairings, and across all experimental parameter settings.

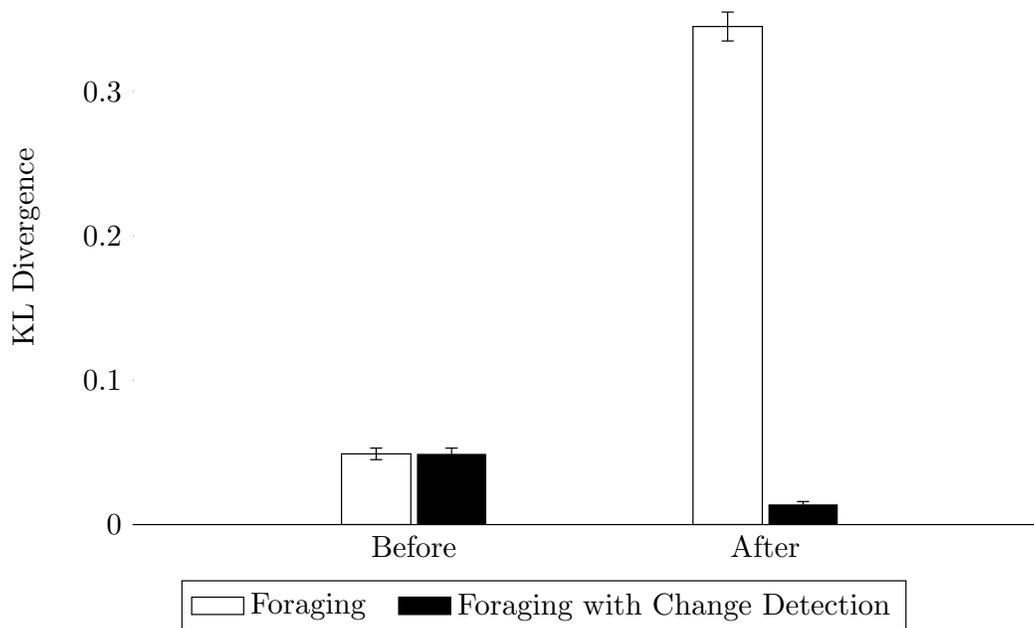KL Divergence when Distribution Changes, N=50

Figure 5. Both algorithms perform identically just before the underlying distribution change. At the end of the path the algorithm that detects changes performs substantially better than the one that doesn't. Error bars represent $1.96\times$ the standard error over the 50 trials.

5   Conclusions

We presented an algorithm that automatically samples while exploring while identifying changes in distributions underlying sampling targets. The foraging algorithm improves learning of unknown distributions in unknown environments. The change detection component identifies environmental changes that may be relevant to remote users, but it definitely improves the performance of the learning algorithm, making it more robust to unknown and variable environments.

We can draw three conclusions from these experiments. First, accounting for the costs of searching and sampling improves the performance of learning agents. Incorporating costs motivates helps discount possible future opportunities that may not arrive. Second, accounting for arrival probabilities of the random variables improves learning the underlying distributions. Again, it helps motivate not giving up on available opportunities. The reduced error achieved in the time budget speaks to the foraging algorithm's improved productivity over the control algorithms. Both these experiments show that low searching costs obviates accounting for environmental statistics. However, the cost of searching and sampling increases a foraging approach is favourable. When the arrival probabilities of classes deviate from uniform the reduction in error from the foraging algorithm is even more pronounced.

Third, change detection is a valuable component for learning in changing environments. We demonstrated a substantial reduction in final error after a change in the underlying distributions has occurred. By tracking the observations collected by exploring robots we can increase the performance of their learning mechanisms while identifying events of interest to remote scientists.

This work can be extended in several ways. First, employ the same change detection of the sample values to the arrival probabilities. This way the exploring agent can detect when the composition of the environment changes, which may be interesting to remote scientists. Second, model more complex underlying distributions. Third, integrate site selection with a path planner in order to determine costs of different sampling actions. Finally, account for possible misclassification of the identified random variables in a scene. These additions will make progress towards robust autonomous planetary exploration.

References

1. Warren-Rhodes, K. A.; Dungan, J. L.; Piatek, J.; Stubbs, K.; Gómez-Silva, B.; Chen, Y.; and McKay, C. P.: Ecology and spatial pattern of cyanobacterial community island patches in the Atacama Desert, Chile. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, vol. 112, no. G4, 2007.

2. Warren-Rhodes, K. A.; Rhodes, K. L.; Pointing, S. B.; Ewing, S. A.; Lacap, D. C.; Gómez-Silva, B.; Amundson, R.; Friedmann, E. I.; and McKay, C. P.: Hypolithic cyanobacteria, dry limit of photosynthesis, and microbial ecology in the hyperarid Atacama Desert. *Microbial Ecology*, vol. 52, no. 3, 2006, pp. 389–398.

3. Smith, K.: On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika*, 1918, pp. 1–85.

4. Wagner, M. D.; Apostolopoulos, D.; Shillcutt, K.; Shamah, B.; Simmons, R.; and Whittaker, W.: The science autonomy system of the Nomad robot. *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, IEEE, vol. 2, 2001, pp. 1742–1749.

5. King, R. D.; Whelan, K. E.; Jones, F. M.; Reiser, P. G.; Bryant, C. H.; Muggleton, S. H.; Kell, D. B.; and Oliver, S. G.: Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature*, vol. 427, no. 6971, 2004, pp. 247–252.

6. Robbins, H.: Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, vol. 58, no. 5, 1952, pp. 527–535.

7. Lai, T. L.; and Robbins, H.: Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, vol. 6, no. 1, 1985, pp. 4–22.

8. Thompson, W. R.: On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 1933, pp. 285–294.

9. Ortega, P. A.; and Braun, D. A.: A minimum relative entropy principle for learning and acting. *Journal of Artificial Intelligence Research*, 2010, pp. 475–511.

10. Burnetas, A. N.; and Katehakis, M. N.: Optimal adaptive policies for Markov decision processes. *Mathematics of Operations Research*, vol. 22, no. 1, 1997, pp. 222–255.

11. Auer, P.: Using confidence bounds for exploitation-exploration trade-offs. *The Journal of Machine Learning Research*, vol. 3, 2003, pp. 397–422.

12. Balcan, M.-F.; Beygelzimer, A.; and Langford, J.: Agnostic active learning. *Proceedings of the 23rd international conference on Machine learning*, ACM, 2006, pp. 65–72.

13. Jun, T.: A survey on the bandit problem with switching costs. *De Economist*, vol. 152, no. 4, 2004, pp. 513–541.

14. Charnov, E. L.: Optimal foraging, the marginal value theorem. *Theoretical population biology*, vol. 9, no. 2, 1976, pp. 129–136.

15. Pirolli, P.; and Card, S.: Information foraging. *Psychological review*, vol. 106, no. 4, 1999, p. 643.

16. Kolling, N.; Behrens, T. E.; Mars, R. B.; and Rushworth, M. F.: Neural mechanisms of foraging. *Science*, vol. 336, no. 6077, 2012, pp. 95–98.

17. Thompson, D. R.; and Wettergreen, D.: Intelligent maps for autonomous kilometer-scale science survey. *Proc. i-SAIRAS*, 2008.

18. Ferri, G.; Jakuba, M. V.; and Yoerger, D. R.: A novel trigger-based method for hydrothermal vents prospecting using an autonomous underwater robot. *Autonomous Robots*, vol. 29, no. 1, 2010, pp. 67–83.

19. Girdhar, Y.; Giguere, P.; and Dudek, G.: Autonomous Adaptive Underwater Exploration using Online Topic Modeling. *Experimental Robotics*, Springer, 2013, pp. 789–802.

20. Thompson, D. R.; Cabrol, N. A.; Furlong, M.; Hardgrove, C.; Low, B. K. H.; Moersch, J.; and Wettergreen, D.: Adaptive sensing of time series with application to remote exploration. *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, IEEE, 2013, pp. 3463–3468.

21. Girdhar, Y.; Whitney, D.; and Dudek, G.: Curiosity based exploration for learning terrain models. *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE, 2014, pp. 578–584.

22. Sun, Y.; Gomez, F.; and Schmidhuber, J.: Planning to be surprised: Optimal bayesian exploration in dynamic environments. *Artificial General Intelligence*, Springer, 2011, pp. 41–51.

23. Lindley, D. V.: On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 1956, pp. 986–1005.

24. Furlong, P. M.; and Wettergreen, D. S.: Budgeting Samples for Exploration in Unknown Environments. *i-SAIRAS*, 2014.

25. Wald, A.: Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, vol. 16, no. 2, 1945, pp. 117–186.

26. Cohen, J.: *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

Algorithm 2 Change Detection for Foraging
---
   function INIT_FORAGE_SAMPLING
      $\Xi \leftarrow \emptyset$
      $R\left(\cdot\right) \leftarrow \emptyset$
      $N. \leftarrow \emptyset$
      $window_{a,k} \leftarrow queue(\emptyset)$ **unless**$\neg empty(window_{a,k})$
      $window_{b,k} \leftarrow queue(\emptyset)$
      $sample_k \leftarrow queue(\emptyset)$ **unless**$\neg empty(sample_k)$
      $sample\_size \leftarrow 5$
      $window\_size \leftarrow 30$
   end function
   function DETECT_CHANGE$(k, z_{k,t})$
      if $size(window_{a,k}) < window\_size$ **then**
         $push(window_{a,k}, z_{k,t})$
      end if
      $push(window_{b,k}, z_{k,t})$
      if $size(window_{b,k}) > window\_size$ **then**
         $pop(window_{b,k})$
      end if
      $push(sample_k, z_{k,t})$
      if $size(sample_k) > sample\_size$ **then**
         $pop(sample_k)$
         $\theta_{a,k} \leftarrow sum(window_{a,k})/size(window_{a,k})$
         $\theta_{b,k} \leftarrow sum(window_{b,k})/size(window_{b,k})$
         $\Lambda \leftarrow \sum_{j=0}^{j=sample\_size} \log\left(\frac{P\left(sample_k(j)|\theta_{a,k}\right)}{P\left(sample_k(j)|\theta_{b,k}\right)}\right)$
         if $\Lambda > change\_threshold$ **then**
            $cache(window_{a,k})$ $\forall \xi_k \in \Xi$
            $window_{a,k} \leftarrow window_{b,k}$ $\forall \xi_k \in \Xi$
            init_forage_sampling()
         end if
      end if
   end function
---

1. REPORT DATE (DD-MM-YYYY) 01-08-2016	2. REPORT TYPE Technical Memorandum	3. DATES COVERED (From - To) 8/2015–10/2015

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Information Foraging and Change Detection for Automated Science Exploration	NNA14AA60C
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
P. Michael Furlong, Michael Dille	
	5e. TASK NUMBER Task 8
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Ames Research Center	8. PERFORMING ORGANIZATION REPORT NUMBER L–

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001	10. SPONSOR/MONITOR'S ACRONYM(S) NASA
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) NASA/TM–2016–219377

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified-Unlimited Subject Category Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES
An electronic version can be found at http://ntrs.nasa.gov.

14. ABSTRACT
This paper presents a new algorithm for autonomous on-line exploration in unknown environments. The objective is to free remote scientists from possibly-infeasible extensive preliminary site investigation prior to sending robotic agents. We simulate a common exploration task for an autonomous robot sampling the environment at various locations and compare performance against simpler control strategies. An extension is proposed and evaluated that further permits operation in the presence of environmental variability in which the robot encounters a change in the distribution underlying sampling targets. Experimental results indicate a strong improvement in performance across varied parameter choices for the scenario.

15. SUBJECT TERMS
change detection, foraging, active learning, planetary exploration, robotics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			STI Information Desk (help@sti.nasa.gov)
U	U	U	UU	24	19b. TELEPHONE NUMBER (Include area code) (757) 864-9658