NASA/TM–2016–

Productive Information Foraging

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

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

August 2016

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

NASA/TM–2016–

Productive Information Foraging

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

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

National Aeronautics and
Space Administration

Ames Research Center

August 2016

The use of trademarks or names of manufacturers in this report is for accurate reporting and does not constitute an offical endorsement, either expressed or implied, of such products or manufacturers by the National Aeronautics and Space Administration.

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Abstract

This paper presents a new algorithm for autonomous on-line exploration in unknown environments. The objective of the algorithm is to free robot scientists from extensive preliminary site investigation while still being able to collect meaningful data. We simulate a common form of exploration task for an autonomous robot involving sampling the environment at various locations and compare performance with a simpler existing algorithm that is also denied global information. The result of the experiment shows that the new algorithm has a statistically significant improvement in performance with a significant effect size for a range of costs for taking sampling actions.

1 Introduction

Human scientists have become accustomed to luxuries such as breathing, eating and drinking, and atmospheric and gravitational forces that do not vary significantly from "1." Consequently, robots have become 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 the situational awareness and reaction times of the scientists controlling such robots. For this reason it is important 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. The nature of many specialized sensors employed for activities such as biological collection and spectral mapping requires 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 common instance of such missions, in which objects or areas found in the environment lie within some respective class that is readily sensed, and each class possesses some underlying data distribution (e.g. spectral response or biochemical composition) that can only be sensed by activating the expensive specialized sensor. The overall goal is to estimate the underlying distribution of each class with maximal accuracy.

For scientific realism and general applicability, no global information such as a prior map 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 choose to sample. Thus, the problem can be thought of as a stream of sensing opportunities providing varying reward (information about the underlying distribution of a class), each requiring a decision to sample or move on.

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, in which they are able to effectively choose between sampling from materials in front of them or moving on to potentially more profitable sampling locations. While these decisions may not be globally optimal, they do demonstrate an ability that is 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 remainder of this document begins with a brief survey of the relevant literature. This is followed by a detailed comparison of the proposed foraging algorithm and one based upon existing principles from the so-called optimal 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 is not without precedent. Kristine Smith started the field of optimal experiment design in 1918. [1] It is only recently that robots have been employed to conduct scientific exploration autonomously. [2, 3] 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 The Secretary Problem

The secretary problem asks a decision maker to select the best candidate from sequentially presented candidates where it is not possible to return to rejected candidates. In the original setting, there is only one position for the candidate to fill, [4] and the optimal strategy is to reject the first $\frac{N}{e}$ candidates and then accept the first candidate who is ranked better than any of the previously seen candidates. Further, the decision maker was able to objectively score the candidates without cost. In our setting, we do not know the value until sampling, and sampling a class incurs a sampling cost.

2.2 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 [5] 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 *et al.* [6] use a value function that sums the mean and the standard deviation of rewards for an arm, in which uncertainty makes an arm more interesting. Other techniques addressing the exploration/exploitation problem use uncertainty in a reward metric. [7–9] In our setting, because the agent only needs to learn the distribution and not use it for anything, uncertainty is the only necessary reward.

There are a number of distinguishing factors between the MAB setting and the problem explored in this paper. First, in MAB, the agent has access to any arm it chooses at any given time. The arms in MAB are analogous to the classes in our setting. 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. [10] In our setting, there is a cost incurred with every choice to continue exploring, and it is a function of the arrival rates of the different classes.

2.3 Optimal Foraging

Foraging is the problem encountered by animals seeking to maximize the intake of energy when operating in an unknown environment. The central question to solving the problem is: Is it more valuable to continue extracting resources from the current location or to seek out resources in new locations? Charnov [11] introduced a technique for dealing with what he called "patchy" environments, in which there are localized 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 through the environment in the hopes of encountering a more valuable location.

The optimal time to leave the environment, according to Charnov's Marignal Value Theorem, is when the expected return from continuing to sample from a particular patch is less than the expected return from wandering in 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 as well as the energy spent seeking a new patch.

Pirolli and Card [12] introduced a model of 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 from which reservoirs to sample. Our exploring agent does not have that luxury.

Kolling *et al.* [13] studied how humans engage in a gambling task in which players have to consider the option they have before them and the opportunities the environment provides. In the described experiment, 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 decision making in humans. The behaviour was near optimal, with some skewing of probabilities at the extreme ends of the scale, i.e. $p \approx 0$ or $p \approx 1$.

2.4 Science Autonomy

Thompson and Wettergreen [14] maximize diversity of collected samples by using mutual information sampling. This approach ensures diversity in the collected sample set, an act that reduces uncertainty in the input space of a function. Neither mutual information nor maximum entropy sampling methods, when used with stationary Gaussian processes, take into account the dependent variable (the underlying class distribution in our setting) when selecting samples.

Bender *et al.* [15] make a modification to that work, instead using Gaussian processes to identified hypothesized distributions of life across the sea floor to direct exploratory actions. The prior maps were generated by vessels passing over the sea floor prior to the robot's exploration mission, not unlike Thompson and Wettergreen's use of satellite imagery. The advance of Bender *et al.* is the use of *in situ* measurements to update the Gaussian process being learned. Their rover can thus be said to be generating and testing hypotheses. However, they are severely limited by a budgeting size of six "gulpers" – devices for collecting seawater samples.

Ferri *et al.* [16] present an approach to prospecting where an autonomous underwater vehicle (AUV) follows a predefined track and needs to decide when to deviate to sample anomalies. The AUV in this work examines anomalies by engaging in a spiral search pattern, collecting data and characterizing the environment in that location. In this case, the rover is not limited in its sampling capacity. However the decision to sample is based on a pre-programmed threshold. While this may be an excellent way to encode subject matter experts' beliefs on what is interesting, it is fragile in the face of a changing environment and does not adapt to the actual environment the rover encounters. This exploration problem is an ideal application of the algorithm proposed in this paper.

Likewise, Girdhar *et al.* [17] present an approach to autonomous exploration wherein a robot investigates a scene when it encounters phenomena that do not reflect its current model of the world. Specifically, they use topic models to describe scenes and sample when they encounter scenes that 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 later work by Thompson *et al.* [18]

Additionally, Girdhar *et al.* build upon their anomaly detection techniques to develop a path planning method to maximize information gain of paths. [19] In that respect, it belongs with the family of curiosity-driven algorithms pioneered by Schmidhüber *et al.* [20] The fundamental concept behind these approaches is that an explorer should spend its time investigating regions of the world (or hypothesis space) where its models are the least certain.

Previous work by the primary author with optimal foraging for science autonomy has considered robots with limited sampling budgets [21] and assumed knowledge of the number of sampling opportunities that would occur. While the limited sampling budget is realistic, the foreknowledge of the transect is not necessarily so. 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.

As explained, real robots may not be able to predict the rewards they will earn from their actions and have to deal with unreliable arrival rates for sampling opportunities. These are concerns that are not modelled in typical sequential experiment selection algorithms such as the multi-armed bandit or secretary problems. This motivates the problem setting used in this paper, described in detail in the following section.

3 Method

The decision-making agents in this paper are tested on several randomly generated transects – one-dimensional paths representing a robot's trajectory. They are repeatedly presented with an object to sample and have to make the choice to either take a sample or continue travelling along the transect. As in Robbins' secretary problem, the agents are not permitted to backtrack to avail of a previous opportunity. Figure 1 gives an example of a simulated transect. Sampling opportunities of different types are scattered along the path that the robot travels.

The primary objective of the robot is to learn distributions behind different classes of objects, for example, the probability distribution governing the density of sub-surface microbial life in different classes of soil. Previous work has identified that texture information can successfully classify different types of soil material. [22] We imagine that the classes of objects in this research could correspond to those soil classes.

3.1 Experiment

The experiment presented in this paper is a modification of the experiments presented by Furlong and Wettergreen. [21,23] In that prior work, agents were equipped with limited sampling budgets. In this experiment, the agents have an unlimited capacity to take samples, but the time to take the sample is non-zero, and there is an overall limit on the duration of the mission. The sampling cost and the overall mission time are given in units of arbitrary time.

As in the previous experiments, agents are not permitted to backtrack to previously seen objects. Disallowing backtracking drives the robot to the end of the transect, as maximizing coverage is an important part of exploration. Additionally, making decisions between a current opportunity, a hypothetical future, and any number of previously seen but unsampled opportunities is considered a much more complex problem and outside the scope of this paper.

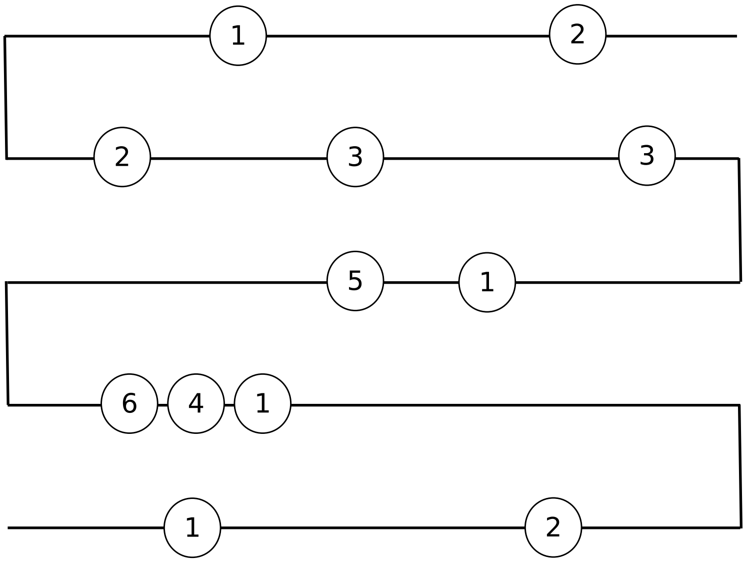

Figure 1. A cartoon illustrating a transect a robot might be encountering and how sampling opportunities of different types may be distributed along it. In the field, the robot would start at one end of the path and follow it to the other end. This is one path that the rover may follow across the terrain resulting in it encountering different types of sampling opportunities.

In this experiment, there are six different classes of objects the agent may encounter. They each have their own arrival rate, and their appearance along the transect is generated with a Poisson process. In this paper, the arrival rates of the different sampling opportunities do not change over the course of the experiment. While this is almost certainly not the case for long range desert traversals targeted by prior work, it is a reasonable approximation for shorter-range traverses. A total transect length of 1000 units is used, and sampling cost is varied from 0.001 to 150.

It is the objective of the robot to learn the true underlying distributions given in Table 1. The algorithms' performance is scored as the L1 difference between the true and learned distribution, p_k and $\hat{p}_k$, respectively. Limits on the integral are placed keep the integration time reasonably small. The integral of the L1 distance, summed over the K known classes, was chosen instead of the more typical KL-divergence to permit a finite error measure in the case of a class never being observed.

$$score = \sum_{k=1}^{K} \int_{\mu_k - 4\sigma_k}^{\mu_k + 4\sigma_k} |p_k(x) - \hat{p}_k(x)| \, dx \tag{1}$$

3.2 Algorithms

Two algorithms are compared on the simulated transect described above.

Class	Mean (arbitrary units)	Standard Deviation (arbitrary units)	Arrival Rate (arbitrary time)
1	0	1	1
2	10	0.1	0.8
3	0	5	0.9
4	2	4	1.1
5	-2	4	0.05
6	0	0.1	1.1

Table 1. The classes the robot is investigating all have values derived from Gaussian random variables with means and standard deviation given. Different instances of those classes are encountered in accordance to a Poisson process with the rates specified in the above table. All units are arbitrary, and all time quantities – arrival rate, mission time, and sampling cost – can be scaled to the order of the mission at hand.

3.2.1 Uniform Sampling

The Uniform Sampling algorithm attempts to distribute the number of samples it can collect evenly between the different classes of objects present on the transect. This is chosen because it was a robustly successful algorithm, as seen in prior work. [21, 23]

The Uniform Sampling algorithm does not consider the time remaining in the transect, nor the time to complete sampling. In this setting, the algorithm chooses to sample a class either if it does not have the most samples of all the encountered classes or if all classes have been sampled an equal number of times.

3.2.2 Foraging

The proposed foraging algorithm is an attempt to maximize the productivity of the learning agent along the transect. We attempt to maximize the amount of information learned per unit time.

The reward for sampling a class is an analog for surprise as defined by Koch *et al.* [24] Koch looked at the change in the distribution that resulted in a Bayesian update. Because this work uses an empirical non-parametric kernel density estimation for each class's distribution, we compare $\log\left(\frac{\hat{p}(x|D\cup\{x\})}{\hat{p}(x|D)}\right)$. To be compatable with optimal foraging algorithms, specifically the Marginal Value Theorem of Charnov, [11] the reward function must have diminishing returns. In the case of information update, the Bayes Factor will eventually converge to approximately 1, likewise our estimated empirical Bayes factor. We take the log of this approximation such that it converges to zero as more samples are collected. Figure 2 demonstrates the diminishing rewards of sampling a distribution using our reward function.

The innovation in this work is valuing actions in terms of productivity. Previous work in foraging maintained the variables of interest in the same units, energy consumed or spent in searching for and extracting resources. To enable the foraging

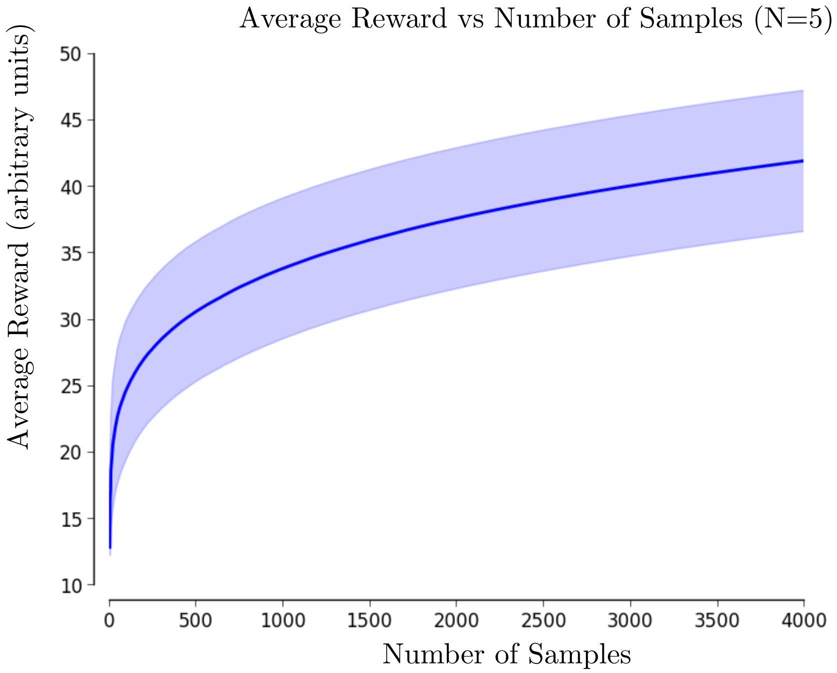

Figure 2. The reward function plotted is the cumulative reward for sampling from a uniform distribution over $[0, 1]$. The cumulative reward is averaged over five trials of 4000 samples. Increasing the number of samples from a distribution decreases the information gained per sample. The reward can be viewed as the reduction in Shannon surprise of an instantiation of the random variable as a result of adding that value to the learned distribution. The returns of the reward function diminish with more samples from the random variable. The diminishing returns are necessary to use the Marginal Value Theorem formulation from Charnov. [11]

agent to compare actions, we measure productivity as the average surprise experienced from a sample per unit of time spent to acquire that sample. In keeping with foraging work of Charnov [11] and Pirolli and Card, [12] the agent decides to move on when the productivity of the current sample is less than the expected productivity of exploring the environment. Namely, the decision rule is

$$\frac{\mathbb{E}_T\left[surprise_t\left(k\right)\right]}{t_{cost}} \geq \frac{\mathbb{E}_K\left[\mathbb{E}_T\left[surprise_t\left(k\right)\right]\right]}{\mathbb{E}_K\left[t_{interarrival(k)}\right] + t_{cost}}, \tag{2}$$

where $\mathbb{E}_T\left[\cdot\right]$ is the empirical expected value over the history of samples the agent has taken, $\mathbb{E}_K\left[\cdot\right]$ is the empirical expected value over the different classes. $surprise_t(k)$ is the surprise due to the t-th sampling of a class k. t_{cost} is the sampling cost in time, and $t_{interarrival(k)}$ is the average inter-arrival time for a class as the rover has encountered them.

4 Results

Figure 3 summarizes a comparison of the performance of the uniform and foraging algorithms, measured as total error in estimating underlying class distributions as specified in Equation 1, across a wide range of sampling cost. When the sampling cost is small relative to the duration of the transect, the performance of the foraging algorithm is significantly better. For a portion at the higher end of the spectrum, the uniform algorithm outperforms the foraging algorithm. However, at the extreme end, the algorithms' performance converges.

At a sampling cost of 100 time units, at the far end of this plot, only ten sampling actions can be taken during a transect that lasts for 1000 time units. In this case, one would expect performance to be limited regardless of decision rule, as it is difficult to learn much about six different random variables with only ten samples.

4.1 Analysis

For each of the twenty-one trials, the agents were scored on the data they collected when given a fixed sampling cost. For each sampling cost the performance was tested with a Bayesian paired t-test. [25] The paired t-test returns an average difference between the paired trials as well as a 95% credible interval around that difference. We accept that the difference is non-zero when the credible interval does not contain 0.

This test also gives an effect size. The effect size is the ratio of the mean to the standard deviation of the difference between the paired trials. This is a variation of Cohen's d value. [26] With this number, we consider a value greater than 1.3 to be very large, above 0.8 to be significant, and below 0.5 to be insignificant. Table 2 gives the results of the Bayesian paired t-test at different values of sampling cost.

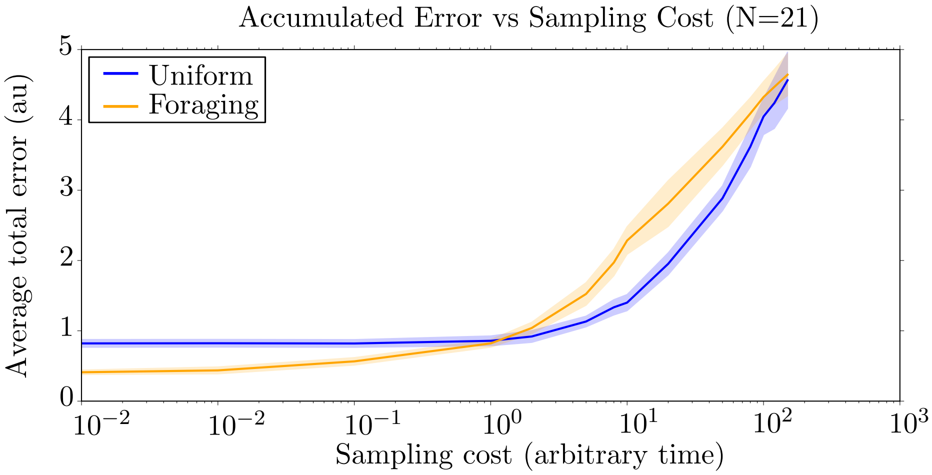

Figure 3. For sampling costs that are small, relative to the duration of the transect, the foraging algorithm presents a significant improvement over the uniform algorithm. The shaded region covers 1.96× the standard error from the 21 trials, approximating a frequentist 95% confidence interval, to indicate the variablility of the algorithms' performance across trials.

Cost	Reduction in Error	Credible Interval	Effect Size
0.001	**0.40**	**[0.34,0.46]**	**3.5**
0.01	**0.38**	**[0.32,0.45]**	**2.8**
0.1	**0.24**	**[0.17,0.31]**	**1.8**
1.0	0.04	[-0.041,0.11]	0.24
10.0	-0.88	[-1.1,-0.62]	1.6
100.0	-0.27	[-0.53,-0.006]	0.48

Table 2. Selected datapoints along the graph in Figure 3 along with the associated credible intervals of the difference and the effect size. Bold rows are cases where the foraging algorithm provides a statistically signficant improvement over the uniform sampling algorithm.

5 Conclusion

In this paper we present a new algorithm created by combining sequential experiment selection and models of optimal foraging with an information theoretic reward function. For certain regimes of operation, the new algorithm is significantly better the control algorithm based on optimal experiment design alone. Additionally, this work continues the process of introducing sequential selection to the field of science autonomy.

From the experiment presented in this paper, we can conclude three things. First, for small sampling costs relative to the mission duration, the foraging algorithm produces about a 50% reduction in accumulated error. Further, the effect size is substantial, and our Bayesian paired t-test gives us 95% confidence that the increase in performance is non-zero.

Second, when the sampling cost is large relative to the mission duration, uniform sampling is as good as or better than foraging. This makes sense, as the first samples one collects are the most informative about a distribution. Distributing samples across the different classes of objects increases the overall rate of information gained, ensuring the greatest short-term reduction in error.

Thirdly, for even larger sampling costs, foraging again becomes a competitive algorithm. This convergence in performance is mainly due to the very limited ability to sample and consequent poor performance of both algorithms.

This work does not address perception and requires some system to parse scenes to identify the classes available to the robot. This is necessary for fielding this algorithm on a robot. The algorithm does not account for variations in the problem setting such as more than one type of sensor with differing sampling costs, class arrival rates that change over time, and class distributions that change over time. These need to be addressed in the future to make a more plausible robot scientist. It is the desire of the authors to make the agent responsible not just for collecting data, but to generate and test hypotheses about the environment.

Ongoing work includes evaluation of these algorithms in a real field setting using terrain classification and with parameter changes such as varied arrival rates and non-stationary class distributions.

References

1. 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.

2. 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.

3. 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.

4. Ferguson, T. S.: Who solved the secretary problem? *Statistical science*, 1989, pp. 282–289.

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

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

7. 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.

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

9. 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.

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

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

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

13. 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.

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

15. Bender, A.; Williams, S. B.; and Pizarro, O.: Autonomous exploration of large-scale benthic environments. *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, IEEE, 2013, pp. 390–396.

16. 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.

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

18. 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.

19. 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.

20. 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.

21. Furlong, P. M.; and Wettergreen, D. S.: Budgeting Samples for Exploration in Unknown Environments. *i-SAIRAS*, 2014.
22. Dunlop, H.: Automatic rock detection and classification in natural scenes. Ph.D. Thesis, Carnegie Mellon University, 2006.
23. Furlong, P. M.; and Wettergreen, D.: Sequential Allocation of Sampling Budgets in Unknown Environments. *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE, 2014.
24. Itti, L.; and Baldi, P.: Bayesian surprise attracts human attention. *Vision research*, vol. 49, no. 10, 2009, pp. 1295–1306.
25. Bååth, R.: Bayesian First Aid: A Package that Implements Bayesian Alternatives to the Classical *.test Functions in R. *UseR! 2014 - the International R User Conference*, 2014.
26. Cohen, J.: *Statistical power analysis for the behavioral sciences*. Academic press, 2013.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*
01-08-2016

2. REPORT TYPE
Technical Memorandum

3. DATES COVERED *(From - To)*
07/2015–09/2015

4. TITLE AND SUBTITLE
Productive Information Foraging

5a. CONTRACT NUMBER
NNA14AA60C

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
P.Michael Furlong, Michael Dille

5d. PROJECT NUMBER

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–

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 of the algorithm is to free robot scientists from extensive preliminary site investigation while still being able to collect meaningful data. We simulate a common form of exploration task for an autonomous robot involving sampling the environment at various locations and compare performance with a simpler existing algorithm that is also denied global information. The result of the experiment shows that the new algorithm has a statistically significant improvement in performance with a significant effect size for a range of costs for taking sampling actions.

15. SUBJECT TERMS
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`)
					19b. TELEPHONE NUMBER ***(Include area code)***
U	U	U	UU	20	(757) 864-9658

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18