2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Optimal "Off-Aiming": Stochastic Path Planning with One-Dimensional Features

Tom Temple and Emilio Frazzoli

Abstract— **This paper investigates a traditional navigational technique, known as "off-course navigation," "landfall intercept," "single line-of-position," and "aiming off," which has been extensively used by navigators on foot, ancient ships, pre-GPS aircraft, and modern submarines. Using this technique, the navigator deliberately aims to one side of their objective with the intention of following a line feature (e.g., a road, coastline, celestial bearing, or radio beacon) that is known to intersect the objective. Despite its extensive use, the question of "How much should one aim off?" has never been rigorously addressed.**

The main difficulty in quantifying the benefit of aiming off is that it entails optimal search as a sub-problem; how does one proceed once the line feature is reached? Recent scholarship has provided a strong heuristic policy for search on the real line. Given this policy, which we use as a black box, we are able pose the problem of "aiming off" as a straightforward optimization problem. This problem is relevant not only to path planning, e.g., in a GPS-denied environment, but also to search problems such as target acquisition.

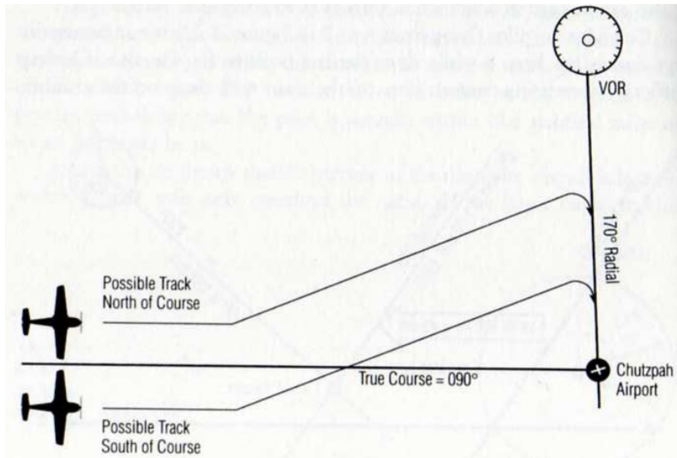

Fig. 1. from *The Proficient Pilot*[9]. Used with permission.

I. Introduction

In this paper we analyze a traditional navigation practice called "aiming off" in which a navigator deliberately aims at a point to one side his destination so that he can be more confident about which direction to search. Figure 1 is an example from *The Proficient Pilot*[9]. This is a sophisticated strategy because it considers the *value of information* about the direction to the destination. Unsurprisingly, then, path planning to mitigate uncertainty has been left to human experts—to quote a 1657 navigation text "[such things] are better learned by practice, than taught by pen."[5] In this paper we will attempt, by pen, to confirm this practice and determine the optimal amount of off-aim.

In addition to navigational applications, another practical application is in search for mobile targets. Suppose an Unmanned Aerial Vehicle (UAV) is being routed to take video of a suspicious mobile target on a road. Rather than move directly toward the mean of the target's position distribution it might be better for the UAV to aim off and acquire the target while following the road.

Previously this problem would have been difficult to tackle because it entails as a sub-problem the problem of optimal search. However recent work on the Cow Path Problem, [10] has provided an efficient means tackle search on the real line. Given a black-box algorithm for line-search, determining optimal heading off-aim is a tractable optimization problem.

By way of contrast, the previously accepted solution to the Cow Path Problem was a "doubling strategy."[1] Under such a strategy the navigator pursues a search pattern in which each time he changes direction he searches twice as far in the opposite direction. With such a search strategy it is never optimal to aim off.

A similar problem was posed in [8] which they called the Coastal Navigation Problem (CNP) in which a mobile robot attempted to robustly reach a goal location in the presence of features. The most general and exact form of their model encompasses our problem but is a continuous state Partially Observable Markov Decision Problem (POMDP) which is intractable. The approach taken was to use an augmented-state Markov Decision Problem (MDP) in which they tracked the mean and entropy of the agent location distribution. This approach is promising in that it predicts behaviors similar to aiming off, *e.g.*, wall following. However, to illicit these behaviours they had to apply an artificial reward model in which the robot was penalized for entropy when the distribution mean reached the goal. For the problem we will pose here, even with the modified reward structure, their approach would not aim off.

The main contribution of this paper is to demonstrate a methodology for determining an optimal control policy in a commonly arising navigation situation. We believe this is particularly interesting because it verifies a practice that human experts have known for centuries but could not be produced by any automated planner.

Another important contribution of this paper is a careful analysis of search of Gaussian distributions. In particular, we provide a Whittle index policy and compute the expected search distance of that policy as a function of the starting point and sensor radius.

978-1-4577-0079-8/11/$26.00 ©2011 AACC

II. Problem Specification and Assumptions

Assume that the objective is on a straight, featureless road. For the moment assume that the navigator can only detect the objective if he is at its location, Section V will discuss the effects of a non-zero sensor radius. Assume that the navigator starts at an initial position with no uncertainty and that the navigator–objective vector is perpendicular to the road and has length ℓ. We will assume that the navigator is a single integrator

$$\dot{x} = u + w,$$

with control $||u|| \leq v$ and noise $w \sim \mathcal{N}(0, I\sigma^2)$. This is a natural model for a navigator with a compass, for example.

The location of the navigator evolves according to a Wiener process with drift u. That is, at time t, the position is a Gaussian random variable with mean $\overline{x}(t) = \int_0^t u(\tau)d\tau$ and variance $t\sigma^2 I$. In the UAV search example, this would be equivalent to the target undergoing Brownian motion.

III. The Cow Path Problem

Once the navigator arrives at the road, he should pursue an optimal search strategy. Searching a one-dimensional road is an instance of the Cow Path Problem, which we briefly review.

In the Cow Path Problem (CPP), m agents are searching for a unique goal that lies on one of n rays diverging from a single origin (with $m < n$). The probability that the goal lies on path i at a distance less than z is given by $F_i(z)$. The agents know their positions with certainty, and each has a sensor that can detect the goal only if the agent is at the location of the goal; otherwise it gives no information. The goal is to find a routing policy for the cows that minimizes the expected time at which the goal is found.

In [10] the authors proved that Whittle's index heuristic can be applied to the CPP. This is significant because although optimal search is provably difficult, Whittle's heuristic has very strong empirical backing (see *e.g.*, [2], [4], [6]). Such a heuristic involves computing a function called a Whittle index γ_i^* for each path i and then pursuing the paths of highest index. For detailed requirements on the requirements of a Whittle index function refer to our references.

We here repeat the Whittle index for the CPP from [10]. For each path i, let $a_i \in \{0,1\}$ indicate whether there is a cow searching it and let z_i denote the maximum distance to which path i has been previously searched. For compactness, let $F(a,b) \equiv F(b) - F(a)$.

Policy III.1 (Whittle Index Policy for CPP[10]). *Pursue the m paths maximizing*

$$\gamma^*(a_i, z_i) = \inf_{z' > z_i} \frac{\mathbb{E}\left[c((a_i, z_i), z')\right]}{F_i(z_i, z')}. \tag{1}$$

The term $\mathbb{E}\left[c((a_i, z), z_i')\right]$ is the expected travel distance associated with extending the search from z_i to z_i' and returning if unsuccessful.

$$\mathbb{E}\left[c((a_i, z_i), z_i')\right] \equiv 2(1 - a_i)z_i + \int_{z_i}^{z_i'} \chi dF_i(\chi) + (1 - F_i(z, z_i'))2(z_i' - z_i) \tag{2}$$

A. Index Policy for Gaussian Search

We now describe the application of Policy III.1 to the case of a single agent searching a Gaussian distribution. Note that Equation 2 scales linearly under a scaling of distance. Since our error model ensures that the uncertainty will always be normally distributed we need only determine the search behavior in general by the search behavior for the standard normal. The actual search behavior can be recovered with appropriate transformation of units.

At the moment that the agent arrives at the road, define zero to be mean of the distance-to-objective distribution and define the unit of distance to be its standard deviation. With respect to this zero, let $\widetilde{x}$ denote the mean of the agent location distribution and z^+, z^- as the extent to which the agent has searched in the positive and negative directions, respectively. Let ϕ and Φ denote the density and cumulative of the standard normal, respectively.

We can simplify Policy III.1 by redefining the origin (of the CPP) to be the agent's location, which makes $a = 0$ for both directions. Rewriting Equation 2 and correcting signs in the negative direction,

$$\mathbb{E}\left[c^+\right] = \int_{z^+}^{z'} (\chi - \widetilde{x})\phi(\chi)d\chi + (1 - \Phi(z^+, z'))2(z' - \widetilde{x})$$
$$\mathbb{E}\left[c^-\right] = \int_{z^-}^{z'} (\chi - \widetilde{x})\phi(\chi)d\chi + (1 - \Phi(z', z^-))2(z' - \widetilde{x})$$

Making use of

$$\int_a^b \chi\phi(\chi)d\chi = \phi(a) - \phi(b)$$

we arrive at Policy III.2.

Policy III.2 (Index Policy for Gaussian Search). *The agent will move in the positive direction whenever*

$$\inf_{z' > z^+} \widetilde{x} - 2z' + \frac{\phi(z^+) - \phi(z') + 2(z' - \widetilde{x})}{\Phi(z^+, z')} \leq$$
$$\inf_{z' < z^-} 2z' - \widetilde{x} + \frac{\phi(z^-) - \phi(z') + 2(\widetilde{x} - z')}{\Phi(z', z^-)}$$

By forward-simulating Policy III.2, we determine the expected search time for $\mathcal{N}(0,1)$ as a function of starting location, shown in Figure 2. We will denote this function $\widetilde{d}_s(\widetilde{x})$ with domain and range in units of standard deviations. The expected search distance of a normal distribution with standard deviation σ is given by $\sigma\widetilde{d}_s(x/\sigma)$.

IV. Optimization

We assume that since the navigator is trying to minimize travel time he always applies maximum control $||u|| = v$ and attempts to travel in a straight line. Under this type of control, and given a search strategy, the problem is simply

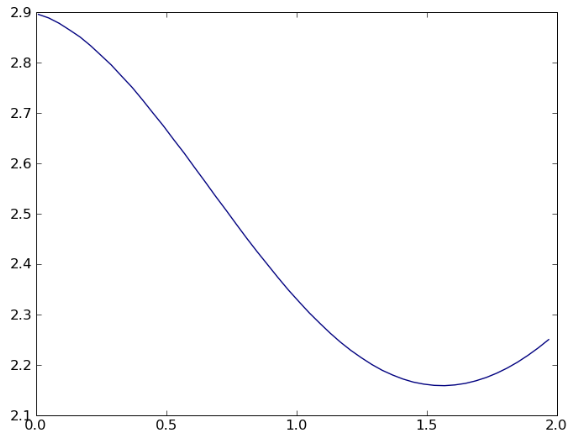

Fig. 2. Expected search distance (in standard deviations) as a function of starting location in standard deviations from the mean.

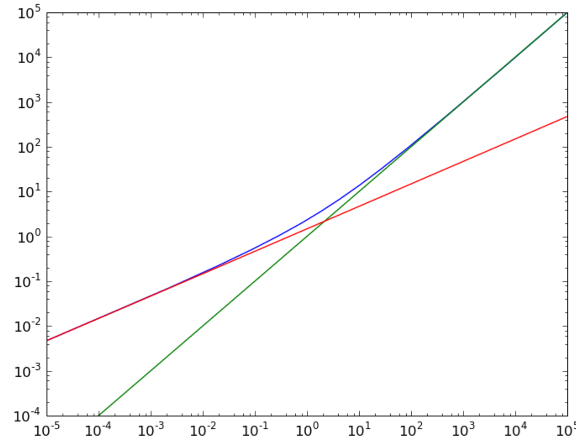

Fig. 3. Minimum expected search time as a function of initial distance from the objective, shown with linear and square-root asymptotes.

to select an initial heading. Given $\widetilde{d}_s$, we can now write the expected search time as a function of heading θ and the initial distance to the road ℓ.

$$d_T(\ell,\theta) \equiv \frac{1}{v}\left(\ell\sec(\theta) + \sigma_r \widetilde{d}_s\left(\frac{\ell\tan\theta}{\sigma_r}\right)\right) \tag{3}$$

The two terms are, respectively, the distance to the road and the search distance subject to a unit transformation. When the agent reaches the road the standard deviation of error is given by $\sigma_r \equiv \sqrt{\frac{\ell\sec(\theta)\sigma^2}{v}}$.

All that remains is to optimize Equation 3 over θ. The functions

$$d_T^*(\ell) \equiv \min_\theta d_T(\ell,\theta)$$
$$\theta^*(\ell) \equiv \arg\min_\theta d_T(\ell,\theta)$$

are shown in Figures 3 and 4.

A. Discussion

The interesting shape of these curves results from the fact that the difficulty of searching grows with the *square root* of the initial distance to the road, ℓ. At large values of ℓ, the amount of offset that corresponds to the minimum in Figure 2 grows at a rate proportional to $\sqrt{\ell}$ and so the optimal heading goes to zero. For the same reason, the search time becomes a vanishing component of the travel time.

When the navigator starts very close to the road, search time dominates and is proportional to $\sqrt{\ell}$. In this limit it is best to immediately move to the road. This can be seen by linarizing d_T about some particular θ_0 and examining the limit of $\ell \to 0$. In this limit we can approximate $\widetilde{d}_s(x) \approx a - bx$ for $a, b \geq 0$. Plugging this linearization into

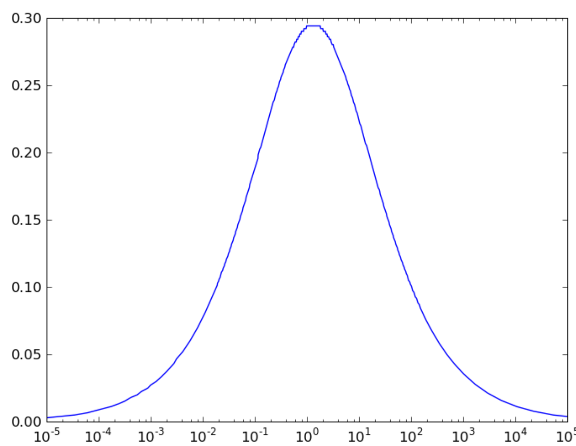

Fig. 4. Optimal heading as a function of initial distance from the objective.

Equation 3,

$$\begin{aligned}
&d_T(\ell,\theta_0+d\theta) \approx \ell\sec(\theta_0+d\theta)/v\ + \\
&\quad\sqrt{\ell\sec(\theta_0+d\theta)\sigma^2/v}(1/v)\left(a - b\frac{\ell\tan(\theta_0+d\theta)}{\sqrt{\ell\sec(\theta_0+d\theta)\sigma^2/v}}\right) \\
&= \ell\sec(\theta_0+d\theta)/v + \sqrt{\ell\sec(\theta_0+d\theta)\sigma^2/v}\ \times \\
&\qquad a - b\sqrt{\ell\sec(\theta_0+d\theta)v/\sigma^2}\sin(\theta_0+d\theta)) \\
&= (1/v)\Big[\ell\sec(\theta_0+d\theta)(1 - b\sin(\theta_0+d\theta))\ + \\
&\qquad a\sqrt{\ell\sec(\theta_0+d\theta)\sigma^2/v}\Big] \\
&\approx (\ell/v)(\sec(\theta_0) + d\theta\sec(\theta_0)\tan(\theta_0))\ \times \\
&\qquad (1 - b(\sin(\theta_0) + d\theta\cos(\theta_0))\ + \\
&\qquad a\sigma\sqrt{(\ell/v)\sec(\theta_0)}\,(1 + d\theta\tan(\theta_0)/2)
\end{aligned}$$

Small changes of θ decrease d_T proportionally to ℓ but increase it proportionally to $\sqrt{\ell}$. Therefore in the limit of $\ell \to 0$, $\theta^* \to 0$.

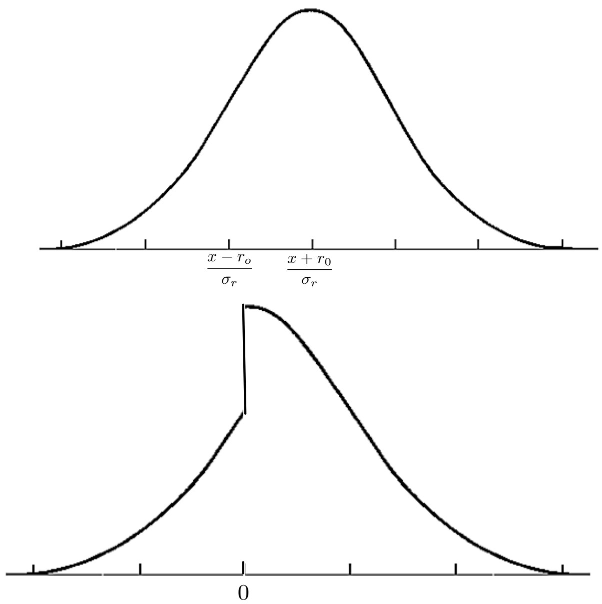

Fig. 5. Modification of the normal distribution to reflect non-zero sensor radius.

In between these two regimes we see that it is sometimes optimal to choose large heading offsets—over 15 degrees. Referring to Figure 2, this should not be surprising. There is a substantial benefit as we move the start point away from the mean. On the other hand, the extra distance only increases as $\sec(\theta)$ and the additional search time as $\sqrt{\sec(\theta)}$.

V. Sensor Radius

Suppose that the navigator can see a distance $r_s = \sigma_r \widetilde{r_s}$ in either direction along the road. We apply Policy III.1 to the distribution formed by removing the center $2r_s/\sigma_r$ about the agent's starting location $\widetilde{x} = x/\sigma_r$ from the normal distribution. This transformation is shown in Figure 5.

Let $\widetilde{d_s}'(\widetilde{x}, \widetilde{r_s})$ denote the expected search time of such a policy. To determine the actual search time we make the following correction. Let χ be the random variable denoting the distance to goal.

$$\begin{aligned} \widetilde{d_s}(\widetilde{x}, \widetilde{r_s}) &= (\Pr|\chi| \le \widetilde{r_s})\mathbb{E}\left[|\chi| | \chi \le \widetilde{r_s}\right] + \\ &\quad (\Pr|\chi| > \widetilde{r_s})(d_s'(\widetilde{x}, \widetilde{r_s}) + \widetilde{r_s}) \end{aligned}$$

The first term is the case in which the agent can immediately see the goal. Otherwise we simply add one sensor radius to the search time. Figure 6 shows $\widetilde{d_s}(\widetilde{x}, \widetilde{r_s})$ for sample values of $\widetilde{r_s}$ between 0 and 1.75. Above this value it no longer helps to aim off.

We adjust Equation 3 to now include the sensor radius

$$d_T(\ell, \theta) \equiv \ell \sec(\theta) + \sigma_r \widetilde{d_s}\left(\frac{\ell \tan\theta}{\sigma_r}, \frac{r_s}{\sigma_r}\right) \tag{4}$$

and optimize over θ.

Figure 7 shows the heading that minimizes Equation 4 for a variety of speeds and sensor radii. These parameters were chosen to bracket those of a person navigating on foot and the variances are scaled such that the asymptotic behaviors match.

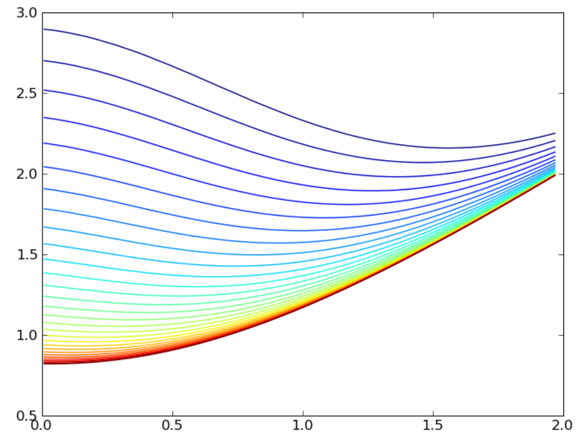

Fig. 6. Expected standardized search distance as a function of starting location for sensor radii between 0 (blue) and 1.75 (red) standard deviations.

As a function of starting distance ℓ, expected travel distance is qualitatively unchanged from Figure 3 and we do not show it. Primarily this is because the navigator must still move to the road and then to the actual location of the objective. In the $\ell \to 0$ limit, the performance improves by the ratio of $\mathbb{E}\left[|\mathcal{N}(0,1)|\right] \ : \ d_s(0)$ which is an improvement of about 45%. This improvement decreases along with the probability of immediately seeing the objective.

A. Discussion

Unsurprisingly, for large distances, Figure 7 matches Figure 4: the sensor radius is becoming increasingly negligible. The interesting behavior is the very steep rise in heading off-aim from zero up to the asymptote. Generally we see a rise from no offset to the asymptote in a doubling of distance. It seems that the once it is no longer "very likely" that the navigator will be able to see the objective immediately, off-aim very becomes attractive very quickly.

We see a maximum heading offset of 0.2 radians for the fast navigator with the most limited visibility. For these conditions it is advisable to aim off even over very short distances (a few tens of meters). On the other extreme the slower, more careful navigator with the best visibility has a maximum heading offset of 0.02 radians. In this case the navigator prefers the direct route for distances up to nearly two kilometers.

Orienteering books recommend between 2–3 and 10 degrees of off-aim (*e.g.*, [7],[3]). The surprising degree of non-linearity in this regime explains the lack of any standard 'rules of thumb' for aiming off in orienteering.

VI. Conclusion and Extensions

This problem was made easy by the fact that we were able to pre-compute a function that takes a measure over the road as an input and returns a mean search time. Given that,

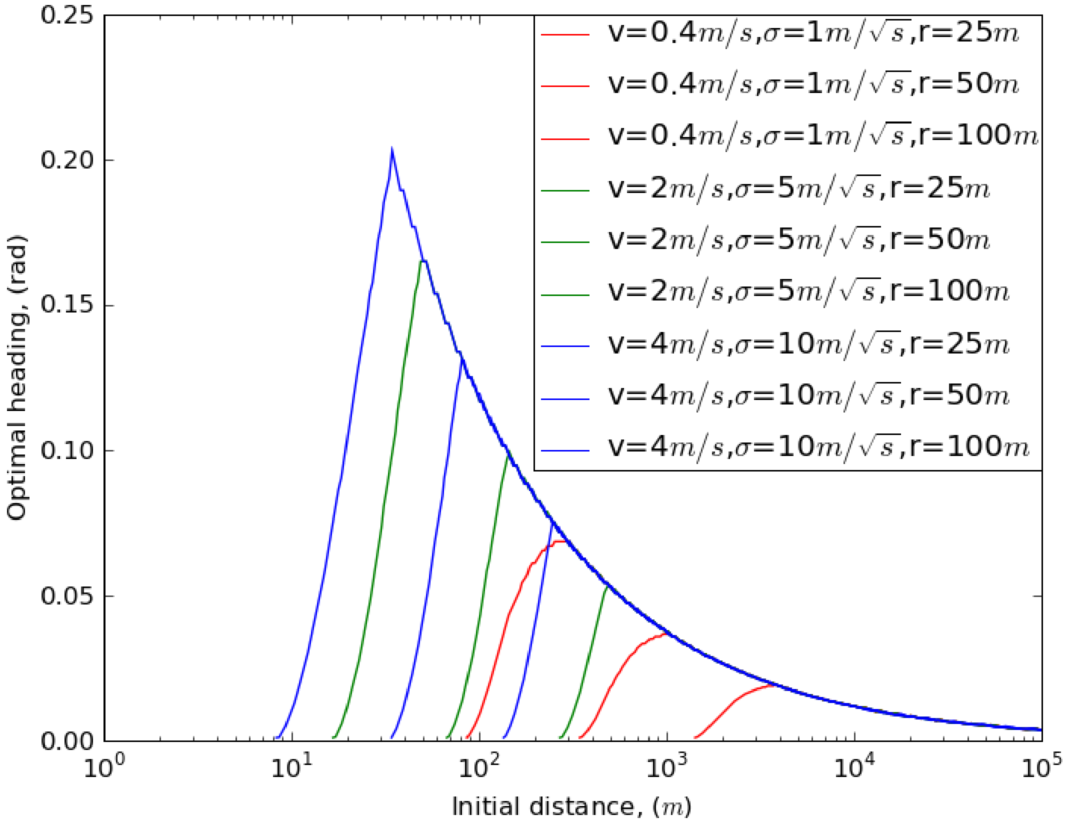

Fig. 7. Optimal heading as a function of starting distance from the road for various speeds and sensor radii. Within each color, the three curves show increasing sensor radii from left to right. These parameters were chosen to bracket those of a person navigating on foot and the variances are scaled such that the asymptotic behaviors match.

the path planning problem was just a matter of optimizing over trajectories. The approach here, then, can be straightforwardly applied to more difficult problems, for instance more complex noise or dynamics. The main difficulty in those cases would be determining the set of trajectories over which we should optimize. In any problem in which all "reasonable" trajectories can be efficiently enumerated, this approach will remain tractable.

Approaching this problem directly via dynamic programming would have been very difficult because the state space includes not only continuous variables, but also a *measure* over a continuous variable. What made this approach tractable was the fact that we evaluated a black-box policy to determine the value function for the tiny subset of states reachable from the initial conditions. In the terms of dynamic programming, what we are doing is using the value function from the Cow Path Policy to determine the cost-to-go for a searchable set of trajectories. On the surface this is trivial—given an optimal policy it is tractable simply to use it. More deeply, though, it represents an exploitation of hierarchy: We used the policy from a relatively simple, one-dimensional problem to tackle a seemingly difficult two-dimensional problem. Ongoing work extends this approach to dynamic search and surveillance problems on road networks.

Other work extends this approach to search problems on higher-dimensional manifolds. For example suppose a robotic arm is attempting place a component but only gets binary feedback about whether it is properly aligned. In this scenario it is likely that "aiming off," *i.e.*, following a longer trajectory than the deterministic optimum, will be preferable if it eases the process of alignment.

Acknowledgments

This research was done with support from the Michigan/AFRL Collaborative Center on Control Sciences, AFOSR grant no. FA 8650-07-2-3744. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the supporting organizations.

References

[1] S. Albers. Online algorithms: a survey. *Mathematical Programming*, 97(1):3–26, 2003.
[2] P. S. Ansell, K. D. Glazebrook, J. Niño-Mora, and M. O'Keeffe. Whittle's index policy for a multi-class queueing system with convex holding costs. *Mathematical Methods of Operations Research*, 57(1):21–39, 2003.
[3] U.S. Army. *Map Reading and Land Navigation*. Washington, D.C., fm 3-25.26 edition, July 2001.

[4] D. Bertsimas and J. Niño-Mora. Restless bandits, linear programming relaxations, and a primal-dual index heuristic. *Operations Research*, pages 80–90, 2000.
[5] John Davis. *Seaman's Secrets*. Gartrude Dawson, 8th edition, 1657.
[6] K. D. Glazebrook, H. M. Mitchell, and P. S. Ansell. Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research*, 165(1):267–284, August 2005.
[7] Don Paul. *The Green Beret's Compass Course*. Path Finder Publications (CA), 10th edition, January 2006.
[8] N. Roy, W. Burgard, D. Fox, and S. Thrun. Coastal navigation-mobile robot navigation with uncertainty indynamic environments. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 1, 1999.
[9] Barry Schiff. *Proficient Pilot, Volume 1*. Aviation Supplies & Academics, Inc., second edition edition, September 2001.
[10] T. Temple and E. Frazzoli. Whittle indexability of the cow path problem. *American Control Conference*, 2010.