

SUMMARY OF MY WORK ON OPTIMAL FORAGING THEORY

I have written about thirty papers on foraging. Many of these papers are unpublished, but most of the basic ideas are mentioned in published papers. I want to make my unpublished papers (and some drafts of published papers) available on my web page.

This is a summary of my papers on foraging, including published and unpublished papers. I hope that this summary will serve as a guide to anyone interested in Oaten's (1977) stochastic model for foraging in a patchy environment. It should also serve as an outline of work on Oaten's model for students who only want to know the basics of work on Oaten's model without reading many papers.

My papers on foraging can be divided into three groups. First, the majority of my papers are based on Oaten's model. These papers can be broken into three groups: (a) theoretical examples chosen for mathematical convenience, (b) empirical examples, and (c) examples used to study functional response.

Second is work on other problems; in particular, central-place foraging, sit-and-wait foraging and oviposition strategies of parasitoids. Not all of this work involves finding optimal strategies, but some of it does. This work is not as systematic or coherent as my work on Oaten's model, but it may show how ideas of optimization may shed light on problems where it had been used differently or had not been used at all.

Third is discussion and commentary. Even the constructive papers have comments to situate the work. In the case of foraging and functional response, the more-or-less incidental comments might be thought of as a critique of early attempts to use functional response to comment foraging behavior and predator-prey model stability. The main pieces of criticism are directed at the view of foraging theory offered by Stephens and Krebs (1986). I criticize this view because I think that it pays insufficient attention to foraging behavior itself on the one hand and to the ecological bases and consequences of foraging on the other hand. In my criticism I have tried to show how foraging theory has been done badly and how it can be done better.

All of my constructive work involves stochasticity--different foraging patches vary in quality and different excursions by foragers achieve different levels of success. I study particular examples and I treat them quantitatively, comparing different environmental conditions and different foraging strategies.

A. Work on Oaten's (1977) stochastic model of optimal foraging

Optimal foraging theory was begun by mathematically-inclined ecologists who were interested in niche breadth and food choice (MacArthur and Pianka 1966, Emlen 1966). Early work was intended to understand animal behavior or ecology, but within about ten years the emphasis was on strategies. Optimal foraging theory became part

of the emerging field of behavioral ecology in the sense Krebs and Davies (1978). The approach was to produce mathematical models of behavior, or of some consequence of behavior, to try to determine an optimal strategy, and to test qualitative predictions to see if animals use optimal strategies.

Pyke, Pulliam and Charnov (1977) published a very useful review of work on foraging during its first ten years. They listed four problems that foragers may be considered to face: (1) Which prey to take? (2) Which patches to search? (3) How long to remain in each patch? (4) How to move among patches? Most of the work on foraging is concentrated on problems (1) and (3). My work is concentrated on problem (3).

Charnov (1976) offered a solution to problem (3) under the assumption that the rate of finding prey within a patch decreases with time spent searching a patch. This decrease is due to the fact that the number of prey available to a predator in a patch is depressed by the presence of the predator, which may deplete the patch by removing prey, or may frighten prey into hiding or fleeing. Charnov assumed that there are a number of different patch types, distinguishable to the forager, and that the net energy gain in each patch type is a smooth function of time in a patch. The gain functions, which would be different for each patch type, are assumed to be differentiable, with (at least eventually) a negative second derivative. The solution to Charnov's optimization problem is a set of patch-leaving times for the different patch types. A forager should leave each patch when the net rate of energy gain (the derivative of the gain function) falls to a value equal to the highest possible long-term average net rate of energy gain in the environment. This result is Charnov's "marginal value theorem." Some predictions follow from Charnov's theorem: foragers should stay longer in better patches, foragers should spend longer in all patches if travel time between patches is longer, and all patches should be left when they reach the same quality.

Charnov's (1976) model was deterministic, but he claimed that his model could easily be made stochastic. Oaten (1977) disputed Charnov's claim, and he developed an explicitly stochastic model to prove that a stochastic model could produce very different results from a deterministic model. Oaten's model assumed that foragers search for prey distributed in superficially similar patches (all of the same "type" in the language of Charnov). Oaten's model assumes:

1. Prey are distributed in patches and the forager knows the distribution of the number of prey per patch.
2. The forager knows the joint distribution of the capture times, given the number of prey in a patch.
3. Prey are not replaced as they are captured.
4. The forager knows the time t that it takes to go from patch to patch. If travel time is variable, t is its mean.

5. The forager decides when to leave a patch based on its knowledge of 1, 2 and 4, and its experience in the patch.

6. Given a strategy we can calculate

$$R = E(G)/[E(S) + t]$$

where

R = the long-term average rate of finding prey,

E(G) = the expected number of prey caught in a random patch, using the strategy, and

E(S) = the expected length of time searching for prey in each patch.

7. The forager uses the strategy that maximizes R.

It is useful to make a distinction here. There are three ways to look at the Pyke, Pulliam and Charnov (1977) problem (3). First, there is the patch residence-time problem: How long should a forager stay in a patch? Second, Charnov (1976) gave a theorem which specified when to leave a patch, but the theorem did not say how a forager should decide when to leave a patch. Third, Oaten's model requires a strategy that specifies how the forager should decide when to leave a patch.

Unfortunately, Oaten's model is very general, and it is very difficult to determine the optimal patch-leaving rule. Optimal patch-leaving rules have been found in particular cases by a number of people, including Oaten himself (Oaten 1977, Stewart-Oaten 1982), Iwasa, Higashi and Yamamura (1981) and McNamara (1982). McNamara (1982) gave a theorem which can be used in some cases to find optimal patch-leaving rules. I have worked out quite a number of cases, finding the optimal patch-leaving rules, finding the rates that they achieve, and comparing their performance with the performance of other rules.

Theoretical examples of Oaten's model chosen for mathematical convenience

Here I will describe, paper by paper, a number of examples of Oaten's model that I have worked out. In general, I consider (a) a Problem, (b) an Approach, (c) some Technique, and (d) a Result and an Interpretation. Each example (each Problem) will be characterized by (i) a distribution of the number of prey in a patch and (ii) a pattern of search. In all cases I assume that whatever prey are in a patch will be randomly distributed there. My Approach is to find the optimal patch-leaving rule and the rate of finding prey that it achieves, and to compare this rate with that achieved by other possible patch-leaving rules, particularly one that I call the "naive" rule--one that ignores experience in a patch. The main Technique that I use is dynamic programming. This method relies on a calculation that I do using Bayes' rule. My Interpretation usually

involves the form (the simplicity) of optimal patch-leaving rules and the comparison of the performance of optimal and other patch-leaving rules.

Discrete-time examples of Oaten's model

Paper [1] = [3].

This is my first paper on Oaten's model. It uses a standard Bayesian setup. The distribution of the number of prey in each patch is treated as binomial, with parameters n and p . For the main example, $n = 20$, and p is assumed to be a random variable, being different for each patch. The parameter p is assumed to have a beta distribution. That is, prey distribution is assumed to be a beta mixture of binomials and search is assumed to be systematic. That is, a patch is assumed to consist of $n = 20$ bits, which are places, each of which contains either no prey or one prey. Thus, this example assumes discrete time and systematic search.

I find the optimal patch-leaving rule using dynamic programming. The method requires four steps, repeated: (1) Guess a value, C , for the maximum long-term average rate of finding prey R^* . (2) Find a rule that "tries to achieve" rate C [a rule that would be best if C were actually the maximum possible rate R^* .] (3) Find the rate $R(C)$ actually achieved by the rule found in (2). (4) Compare the rate achieved, $R(C)$ with the rate guessed, C . If they are the same, the best rule and the rate that it achieves have been found. If they are different, guess again. In fact, very few guesses are required to find the best rule. Dynamic programming, more specifically, backward recurrence, is used to find the rule in step (2).

The basic idea is that a forager who is searching a patch may leave the patch at any time, and from that time on achieve the long-term average rate, R^* . Therefore, the forager should only remain in a patch if it can do well enough for the rest of its stay in the patch. The rule that "tries to achieve" rate C is found by beginning one step before the end of a patch [when the end is reached, the forager leaves the patch]. For each possible number of prey found up to one step before the end, it is possible to calculate the probability of finding a prey in the last bit [this probability is found using Bayes' rule]. If this probability is greater than C , then the forager should remain. If this probability is less than C , then the forager should leave. Once it is determined whether to stay or leave for each possible number of prey found up to one step before the end, it is possible to decide whether to stay or leave for each possible number of prey found up to two steps before the end. Working in this way back to the beginning one can determine the optimal patch-leaving rule. The optimal patch-leaving rule is characterized by a set of points, (t,x) such that the forager should leave if only x prey have been found by time t . No matter how many prey have been found, all patches are left when they have been searched completely.

For the examples that I have considered, the patch-leaving points for the optimal rule fall almost on a line. That is, the optimal patch-leaving rule is a simple one, simple enough so that one might imagine that an animal might be able to use it. While I was

unaware of this when I wrote my paper, Waage (1979) offered a simple behavioral mechanism to model the patch-leaving decision for parasitic wasps. Waage's mechanism yields a rule quite similar to the optimal rule that I found in paper [1] using dynamic programming.

The optimal patch-leaving rule is not only simple, but it is efficient in the sense that it produces a higher rate of finding prey than a naive rule that ignores experience in a patch. In paper [1] I considered the best rule, the naive rule and two other rules, an omniscient rule and an instantaneous-rate rule. A forager using the omniscient rule--which uses more information than Oaten (1977) assumed was possible--could recognize patch type (they would know the value of the binomial parameter p) without searching the patch, and they would only search patches that were good enough (had high enough values of p). Omniscient foragers could do much better than foragers that had to use their experience in a patch to evaluate patch quality. Animals using the instantaneous-rate rule would leave each patch when the expected rate of finding prey there fell below some critical threshold. Such a rule was suggested by Pyke (1978) as a rule that would satisfy the marginal value theorem. It turns out, at least for my beta-binomial example, that the value of the optimal threshold for the instantaneous-rate rule is not equal to the rate achieved, which would be the case for the marginal value theorem, but the best instantaneous rate rule is almost as good as the best rule.

In paper [1] I compared the performance of the four conceivable patch-leaving rules using several sets of parameter values for the beta distribution (the prior or mixing distribution). I fixed the average number of prey per patch and the travel time and varied the variability of the prior distribution. The advantage of using information is greatest with the prior distribution is most variable. That is, the advantage of using information about patch quality is greatest when there is more information to use.

I also fixed the prior and varied travel time. The advantage of using information about patch quality is greatest when travel time is lowest.

This is the paper to read if one wants a fairly clear description of how to use dynamic programming to find the optimal patch-leaving rule for Oaten's model. The method is easier to understand in discrete time, as used in paper [1] = [3]. The published version of this paper [3] has a typo in the last expression of equation (7) that is not found in the draft version [1].

Paper [2]

The point of this paper was to use optimal foraging to investigate the functional response of a predator. I will discuss my papers on functional response separately, but here I will describe how I looked at the performance of an optimal forager. The model considered here is the same as that used in paper [1].

In paper [2] I give a table showing the beta-binomial distribution for various sets of parameters (assuming the same mean for all distributions). Then, for two sets of

parameters, I show the optimal patch-leaving rules (the same as in paper [1]) and also the distribution of the number of prey an optimal forager would find and the mean and variance of this number. Then, I show what I took at the time to be the functional response, that is, I calculated the average number of prey taken by an optimal forager from patches beginning with each of the possible numbers of prey. Finally, I found the distribution of the number of prey remaining in a patch after it has been searched by a foraging using the optimal patch-leaving rule, and I compared this distribution with the distribution that would be found if each patch had a random number of prey removed under the assumption that the average number would be the same as the average removed by an optimal forager. An optimal forager removes a larger proportion of the prey from rich patches (patches having more prey) than from poor patches, so I assumed that patches searched by an optimal forager would show less variability in the number of prey remaining than if prey are removed at random. However, the number of prey remaining in patches with random removal are more, not less, variable than the number remaining in patches visited by an optimal forager. That is, optimal foragers do not tend to homogenize prey distribution.

Paper [8]

Paper [8] uses the same beta-binomial model as used in papers [1] = [3] and [2], but it adds the treatment of another patch-leaving rule, the fixed giving-up-time rule. The idea of this rule is that if the forager searches for some fixed time without finding any prey, or if after finding some prey it goes for that fixed time without finding a prey, then it will leave the patch. The rate of finding prey achieved by a fixed giving-up-time rule was calculated by Murdoch and Oaten (1975), Breck (1978) and by Iwasa et al. (1981) for the examples that assumed random search and continuous time. My model, which assumed systematic search and discrete time, required a different model involving counting the number of ways to arrange successes and failure such that no run of failures before the last was as long as the fixed giving-up-time. (This method was also used in paper [26] = [27], which treated the optimal oviposition strategy of parasitoids assumed to be under threat of being attacked by hyperparasitoids using a fixed giving-up-time rule.)

In paper [8] I compared the fixed giving-up-time rule with the best rule and the naive rule. The performance of the giving-up-time rule is not as good as that of the best rule, but it is quite a bit better than that of the naive rule. For a given case (set of environmental parameters) the performance of the fixed giving-up-time rule is not very sensitive to the giving-up-time used. A similar result is seen for rules of the optimal type. For a given set of environmental parameters, I find the rules that "tries to achieve" different rates, C . For quite different rates, I find quite different rules, but if I plot the rate achieved by the rules, R_{ach} , against the guessed rate, C , I find that the rate achieved is insensitive to the rate guessed.

In another calculation I find the best patch-leaving rule for several different sets of environmental parameters and then I determine the rate achieved for one set of parameters by a patch-leaving rule that is best for a different set of parameters. It is

seen that a rule that is best for one set of parameters (for a slightly variable prey distribution, for example) is almost as good as the best for another quite different set of parameters (for a more variable prey distribution, for example). That is, rules of the right general form are robust in the sense that a rule that is best for one set of parameters is close to the best for a quite different set of parameters.

Thus, the results of papers [1], [2] and [8] show that the optimal patch-leaving rule has a simple form, that it is efficient, that its performance is not very sensitive to its exact form if its general form is correct, and that it is robust to moderate changes in prey distribution.

Paper [19]

Paper [19] treats another discrete-time example, a two-point mixture of binomials. Each patch has a binomially distributed number of prey, but the binomial parameter has one of two values, one high and the other low (but not zero). That is, patches are either "good" or "bad."

I compare the performance of three different patch-leaving rules: a best rule, the fixed giving-up-time rule, a naive rule that ignores patch quality and remains until each patch has been searched completely, and an "odds rule," in which a forager leaves a patch when, based on its experience in a patch, the the forager knows that the odds that the patch is good has fallen to some threshold value. The best odds rule is only slightly worse than the best rule, which is better than the best fixed giving-up-time rule, which, in turn, is better than the naive rule.

In one figure I plot the expected gain, $E(G)$, versus expected time in a patch, $E(S)$ for various candidate best rules (obtained by guessing values C , then finding the best rule that "tries to achieve" rate C) and candidate fixed giving-up-time rules (found by choosing particular values of the giving-up-time). This plot resembles the figures that are sometimes used to illustrate Charnov's (1976) "marginal value theorem."

There is one theoretical innovation. In this paper, as the candidate rules are found, the expected gains and expected time remaining in each patch are calculated, and the final values (when the algorithm has worked back to the origin) are used to calculate the rate, $R(C)$. Thus, step (4) in the iterative procedure used to find the best rule is eliminated.

Paper [22] = [24]

Paper [22] is basically a discussion paper, but it does give a discrete-time example, which is a version of the two-point mixture of binomials used in paper [19]. In this example, the bad patches have no prey, while good patches have a binomially-distributed number of prey. This example is given as a mathematically more tractable version of a distribution used in an experiment by Lima (1984). The optimal strategy is to leave a patch if the first few bits that are searched are empty, but if a prey

is found, to stay until the patch has been searched completely. The optimum number of empty patches to search before leaving is fixed value, but the rate of finding prey is not very sensitive to the value used.

Paper [33] = [36]

Paper [33] treats the exact distribution used by Lima (1984): patches are either empty or they contain a fixed number of prey. The best rule in this case is not as simple as that for the modified problem treated in paper [22].

Paper [33] gives quite simple code to find the optimal patch-leaving rule. The simplicity is due to two innovations. One is the elimination of step (4) as mentioned in the summary of paper [19]. The other innovation is to calculate the posterior probabilities, $r(t,x)$, directly, using Bayes' rule and the distribution of the number of prey per patch. Here $r(t,x)$ is the probability that the next bit will have a prey, given that exactly x prey have been found in the first t bits searched. In some earlier papers, such as [1], $p(t,c)$ was used for what is called $r(t,x)$ in [33]. In some ways, calculating $r(t,x)$ directly is more complicated than getting a simple formula based on a convenient assumption about prey distribution, but a direct calculation is conceptually simple, and it is much more general.

Continuous-time examples of Oaten's model

Continuous-time models are probably more realistic than discrete-time models, and the results are a bit prettier, but they are also more difficult to obtain. Continuous-time models may assume systematic or random search, but discrete-time models must use systematic search if the calculations are to be tractable. The reason for this is that in order to use backward recurrence, it is necessary to characterize succinctly the information obtained in patch up to a (late) decision point. If search is systematic (and whatever prey are discovered are discovered are found at random times--an assumption implicit is the binomial assumption), then the number of prey, x , found after spending time t in a patch (or after searching t bits), is sufficient for estimating patch quality and the capture times are irrelevant. For continuous time, random search is interpreted as systematic search at an exponentially decreasing rate. There is no such interpretation for random search and discrete time.

One technical idea is needed in order to use dynamic programming to find the optimal patch-leaving rule for continuous time that was not needed for discrete time, and another technical idea is needed for random search. For discrete time, the forager needs to decide whether or not to leave a patch at only a finite number of times, but for continuous time, the forager might leave at any time. To find the best patch-leaving rule we use the fact that for nice prey distributions, such as negative binomial, the patch leaving points, or rather, points at which the forager should leave or should never have been reached because the patch should have been left earlier, for a triangle set. We may break up points, (t,x) , into two sets, staying points, such that a forager should stay if x prey have been found by time t , and leaving points, such that a forager should leave,

or have left, if x prey have been found by time t . Then, for a negative binomial prey distribution, the leaving points are a triangle set, in the sense that if (t,x) is in the set, any point with the same t but smaller x , or any point with the same x but a smaller t , is in the leaving set; and if (t,x) is a staying point, than any point with the same t but larger x , or any point with the same x but smaller t , is in the staying set. Using the fact that the leaving points are a triangular set, we find the optimal patch-leaving rule by finding the highest value of x for which the forager should leave the patch. Then we use bifurcation to find the time for which the forager should leave the patch if it has found x prey. Then one does the same for $x - 1$ prey, and so on.

For random search there is, in a sense, no end to the search in a patch, so there is no natural place to begin and then work back from. So, the trick is to pick some long time and assume that the patch must be left by that time. Then the optimal patch-leaving rule and the rate that it achieves is calculated, assuming that the search must be truncated. Then, we ask what would happen if we used some longer long time, and we do the same calculations again. We repeat until the rule and the rate that it achieves does not change.

Paper [16]

This is the best paper that I have published on optimal foraging theory. In a sense, it is a reworking of the paper by Iwasa et al. (1981), which I think is the most important paper on Oaten's (1977) stochastic model of foraging in a patchy environment. In this paper I consider both systematic and random search, and for each pattern of search I find the optimal patch-leaving rule for three different distributions of the number of prey in a patch: (i) Poisson distribution, (ii) negative binomial distribution and (iii) a regular distribution in which each patch has the same number of prey.

Following Iwasa et al. (1981) I plot the instantaneous rate of finding prey, $r(t,x)$, as a function of time, t , for each of the combinations of search pattern and prey distribution. I also show what the optimal patch-leaving rule looks like for each of the six cases. [The figures in the type-script are correct, but some of them have been interchanged in the published version.] Unlike Iwasa et al. I did not consider a binomial distribution for the numbers of prey in patches. On the other hand, I did find the optimal patch-leaving rule for a negative binomial prey distribution.

Paper [16] contained an appendix that included formulas (some incorrect) needed to find the optimal patch-leaving rule for systematic search and a negative binomial prey distribution. No computer code was given in that paper, but computer code and corrections were given in a technical report [12]. In fact, many more examples and details related paper [16] are contained in a series of technical reports [12], [13], [14], [17], [18]. Paper [16] treats only one example of each of the prey distributions and it illustrates only one possible pattern of prey-finding experience.

Paper [4]

The first example of a continuous time model that I considered assumed systematic search and a two-point mixture of Poisson distributions. The point of the paper was to investigate functional response, but the example used was intended to be the basis of a psychology experiment. Joel Myerson was a researcher in the Psychology Department at the University of California Riverside. He was interested in Skinner box-type experiments. If animals are offered rewards on a random interval schedule, then "prey" appear according to a Poisson process. Reward rate is constant (for a time) on each response lever, but once the animal switches levers the new lever is governed by a constant, but possibly different, rate.

The experiment was never done, but the calculations were done; the optimal patch-leaving rule was found and the rate that it achieved was calculated. In fact, the results were not very impressive: optimal foraging was not much better than naive foraging. The reason was that in order to choose reasonable parameters a number of constraints had to be considered. (i) Poor patches should result in at least a few rewards. (ii) Good patches had to be enough better than poor patches so that the experimental animals should care about the difference. (iii) The total number of prey found in one experimental session should not be enough to satiate the animal. In fact, to satisfy constraints (i) and (iii) I chose values for the high and low rates that were not sufficiently different to matter much to the foragers.

Paper [12]

Paper [12] involves my favorite example, a negative binomial (gamma mixture of Poissons, or gamma-Poisson) prey distribution and systematic search. I think that this example is biologically realistic and mathematically pretty. This paper, and some other technical reports that I wrote around the same time ([14], [18]) elaborate the ideas found in my main published paper on continuous time [16], while other papers ([13], [17]) fill in some gaps or carry the ideas further. This paper, and many of the others, include formulas used to do the calculations, and computer code used to implement the calculations.

In paper [12] I give optimal patch-leaving rules for several parameter sets and calculate the expected gain, the expected time in a patch, and the rate of finding prey achieved by each. I compare the distribution of the number of prey in each patch with the distribution of the number of prey that would be found in each patch by optimal foragers. I also calculate the means and variances and the covariance of the numbers of prey found and the time spent searching each patch, and I use these to get an asymptotic expression for the variance in the long-term rate of finding prey. The conclusion is that the variance in the foraging rate is substantial. It is so large that it would be difficult for a statistician, let alone a bird or insect, to distinguish between the rate achieved by an optimal forager and a sub-optimal forager used a good alternative to optimal foraging.

I also plot the rate, R_{opt} , of finding prey achieved by foragers "trying to achieve"

rate C against rate C , and I plot the expected number of prey found against the expected time in a patch for various candidate best rules. This gives a figure like that often used to illustrate the marginal-value theorem, but with expected number of prey found replacing the gain function, and the expected time in a patch replacing time in a patch.

Paper [13]

This paper contains one of the most difficult calculations that I did in treating Oaten's model. I find the optimal fixed giving-up-time for systematic search and continuous time. The calculation uses some facts about the distribution of the sizes of gaps between uniform order statistics.

In this paper I show how to find the rate achieved by a fixed giving-up-time rule, I plot the rate of finding prey that such a rule achieves against the giving-up time for a negative binomial prey distribution. I also find the rate achieved by a similar rule that is even simpler than the fixed giving-up-time rule. I call this rule "Lima's rule," because it seems to be used by the downy woodpeckers studied in an experiment by Steve Lima (1984). A forager using "Lima's rule" would leave a patch if it spends some fixed time there without finding a prey, but if one prey is found before the fixed time, the forager remains in the patch until it has been searched completely. For the cases that I have considered, "Lima's rule" is not only simpler than the fixed giving-up-time rule, but it is better. The critical value for Lima's rule is substantially lower than the critical fixed giving-up time.

Paper [14]