Information Foraging in Information Access Environments

Peter Pirolli
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
pirolli@parc.xerox.com

Stuart Card
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
card@parc.xerox.com

ABSTRACT

Information foraging theory is an approach to the analysis of human activities involving information access technologies. The theory derives from optimal foraging theory in biology and anthropology, which analyzes the adaptive value of food-foraging strategies. Information foraging theory analyzes trade-offs in the value of information gained against the costs of performing activity in human-computer interaction tasks. The theory is illustrated by application to information-seeking tasks involving a Scatter/Gather interface, which presents users with a navigable, automatically computed, overview of the contents of a document collection arranged as a cluster hierarchy.

KEYWORDS: Information foraging theory, information access.

INTRODUCTION

Recent years have witnessed an explosive increase in public interest in information access and communication technologies. Along with this burgeoning excitement, the rapid growth of electronically available information sources, such as those available on the Internet, has further exacerbated the need for effective and efficient tools for information workers and consumers. For researchers and developers in human-computer interaction, this increases the need for models and analysis techniques that allow us to determine the value added by particular information access, manipulation, and presentation techniques, and to reveal design elements that may yield further enhancements.

To motivate previous designs of interactive information systems [6], we appealed to mechanisms of cognitive science [5], and to general principles of information science [17]. We have argued that in an information-rich world, the real design problem to be solved is not so much how to collect more information, but rather, how to optimize the user's time, and we have deployed these principles in an attempt to increase relevant information gained per unit time expended. But for task analysis, design exploitation, and evaluation of information systems, a more developed theory is needed.

In this paper, we lay out the framework for an approach we call *information foraging theory.* This approach considers the adaptiveness of human-system designs in the context of the information ecologies in which tasks are performed. Typically, this involves understanding the variations in activity afforded by some space of human-system design parameters, and understanding how these variations trade-off the value of information gained against the costs of performing the activity. While complementary with information processing approaches to computer interaction theory, such as those in the GOMS family [5, 14], information foraging theory emphasizes a larger time-scale of behavior, the cost structure of external information-bearing environments, and human adaptation.

Consider the time-scales of activity outlined by Newell [13]. The sorts of information-seeking and sensemaking activities of interest to us span from the middle of the *cognitive band* of activity (~100 ms - 10 s), across the *rational band* (minutes to hours), and perhaps into the *social band* (days to months). Typically, information-processing models of cognition have addressed behavior at the cognitive band, and elementary cognitive mechanisms and processes (e.g., such as those summarized in the Model Human Processor, [5]) play a large part in shaping observed behavior at that grain size. As the time scale of activity increases, "there will be a shift towards characterizing a system...without regard to the way in which the internal processing accomplishes the linking of action to goals" (Newell [13], p 150). Such explanations assume that behavior is governed by *rational* principles and largely shaped by the constraints and affordances of the task environment. Rather than assuming classical normative rationality, one may assume that the rationale for behavior is its adaptive fit to its external ecology [4]. This is the essence of an *ecological stance* (Neisser, as cited in Bechtel [4]) towards cognition. Whereas information-processing models, such as GOMS, provide mechanistic accounts of *how* cognition operates, ecological models address *why* it operates that way, given the ecological context in which it occurs. This kind of integrated explanatory framework has been promoted by Marr [12] and, more recently, Anderson [1, 2] in cognitive science.

INFORMATION FORAGING TASKS

Information foraging refers to activities associated with assessing, seeking, and handling information sources. Such search will be adaptive to the extent that it makes optimal use of knowledge about expected information value and expected costs of accessing and extracting the relevant information. We use the term “foraging” both to conjure up the metaphor of organisms browsing for sustenance and to indicate a connection to the more technical *optimal foraging theory* found in biology and anthropology [21, 22].

Animals adapt their behavior and their structure through evolution to survive and reproduce to their circumstance. Essentially animals adapt, among other reasons, to increase their rate of energy intake. To do this they evolve different methods: a wolf hunts (“forages”) for prey, but a spider builds a web and allows the prey to come to it. Humans seeking information also adopt different strategies, sometimes with striking parallels to those of animal foragers. The wolf-prey strategy bears some resemblance to classic information retrieval, and the spider-web strategy is like information filtering. Human hunter-foragers have been observed to hunt in groups when the variance of finding food is high. They accept a lower expected mean in order to minimize the probability of several days without food. Similarly, we have observed, in the field, professional market analysts who had developed an ethic of cross-referring information, essentially information-foraging in groups, so as to reduce the probability of missing important literature.

Optimal foraging theory is a theory that has been developed within biology for understanding the opportunities and forces of adaptation. We believe elements of this theory can help in understanding existing human adaptations for gaining and making sense out of information. It can also help in task analysis for understanding how to create new interactive information system designs.

Optimality models in general include the following three major components.

- *Decision assumptions* that identify which of the problems faced by an agent are to be analyzed. This might involve decisions about whether to pursue a given type of information item upon encountering it, the time to spend processing a collection of information items, the choice of moves to make in navigation, the choice of strategy under uncertainty, or degree of resource sharing.

- *Currency assumptions*, which identify how choices are to be evaluated. The general assumption in ecological analyses is that some feature x will exist over other features, if x satisfies some existence criteria. Existence criteria have two parts (a) a *currency*, and (b) a *choice principle*. Typically, optimal foraging models in anthropology and biology assume energy as a currency. Information foraging theory will assume information value as currency. Choice principles include maximization, minimization, and stability.

- *Constraint assumptions*, which limit and define the relationships among decision and currency variables. These will include constraints that arise out of the task structure, interface technology, and the abilities and knowledge of a user population.

We assume that information foraging is usually a task that is embedded in the context of some other task and its value and cost structure is consequently defined in relation to the embedding task and often changes dynamically over time [3, 18]. The *value* of information [16] and the *relevance* of specific sources (Schamber, et al., 1990) are not intrinsic properties of information-bearing representations (e.g., documents) but can only be assessed in relation to the embedding task environment.

Usually, the embedding task is some ill-structured problem for which additional knowledge is needed in order to better define goals, available courses of action, heuristics, and so on [15, 20]. Such tasks might include such things as choosing a good graduate school, developing a financial plan for retirement, developing a successful business strategy, or writing an acceptable scientific paper. The structure of processing and the ultimate solution are, in large part, a reflection of the particular knowledge used to structure the problem space. Consequently, the value of the external information may often reside in the improvements to the outcomes of the embedding task.

The use of optimality models should not be taken as a hypothesis that human behavior is classically rational, with perfect information and infinite computational resources. A more successful hypothesis about humans is that they exhibit *bounded rationality* or make choices based on *satisficing* [19]. However, satisficing can often be characterized as localized optimization (e.g., hill-climbing) with resource bounds and imperfect information as included constraints [23]. Optimality models do *not* imply that animals or information foragers will necessarily develop so as to embrace the simple optimum. Rather, they describe the possibilities of a niche, a possible advantageous adaptation if not blocked by other forces (for example, the consequences of another adaptation). For us, these models help fill in what Anderson [1] calls the Rational Level theory of information access.

OVERVIEW OF THE EXAMPLES

We present several examples of foraging analyses to illustrate some of the range of problems and insights that may be addressed. Our coverage has to be limited, so we use three relatively concrete and detailed analyses from a particular system.

In the context of a an analysis of the Scatter/Gather document browser [9] we introduce two simple models, the *information patch model*, and the *information diet model*, borrowed rather directly from optimal foraging theory. These "conventional" models derive from Holling’s disc equation (Stephens & Krebs, 1986), which states that rate of currency intake, R, is the ratio of net amount of currency gained (energy in the case of biological systems; information value in our case), U, divided by the total amount of time spent searching, T_s, and exploiting, T_h,

$$R = \frac{U}{Ts + Th}. \quad (1)$$

The net gain, U, is the overall currency intake (the gross amount foraged), U_f, less the currency expended in foraging, C_f or $U = U_f - C_f$. It can be shown that broadly applicable stochastic assumptions approximate Holling's disc equation.

If we assume that information workers or consumers encounter information items as a linear function of time, the total number of items encountered while searching can be represented as λT_S, where λ is the rate of encounter with items per unit time. The average rate of currency intake will be $\bar{u} = \frac{U_f}{\lambda T_S}$ and the average cost of handling items will be $\bar{h} = \frac{T_h}{\lambda T_S}$, and so $U_f = \bar{u}\lambda T_S$ and $T_h = \bar{h}\lambda T_s$. If s is the search cost per unit time then the total cost of search will be sT_S. Substituting these into Equation 1, we have

$$\begin{aligned} R &= \frac{\bar{u}\lambda T_S - sT_S}{T_S + \bar{h}\lambda T_S} \\ &= \frac{\lambda\bar{u} - s}{1 + \lambda\bar{h}} \end{aligned} \quad (2)$$

The information patch model and the information diet model are formulated as variants of Equation 2. We discuss the analytic optimal solutions to these models in the context of the illustrations.

We also develop a more comprehensive dynamic model that incorporates the information patch model and information diet model as subcomponents. Using dynamic programming we illustrate how one may determine the optimal human-system strategies using dynamic programming techniques.

EXAMPLE 1: FORAGING IN INFORMATION PATCHES

A user's encounters with valuable or relevant information will typically have a clumpy structure over space and time. Information items are often grouped into collections such as libraries, databases, and wire services. The biological analogy is that an organism's ecology may have a variety of food patches of differing characteristics and the organism must decide how to best allocate its foraging time. Models of this situation are called *patch models* [7, 10, 22]. We discuss an information patch model in the context of the Scatter/Gather text database browser.

Scatter/Gather Interface

Figure 1 presents a typical view of the Scatter/Gather interface.[1] The emphasis in this browsing paradigm is to present users with a kind of automatically computed overview of the contents of a document collection, and to provide a method for navigating through this summary at different levels of granularity. This is achieved by organizing the collection into a *cluster hierarchy*. Conceptually, a collection may be clustered, for instance, into B = 10 groups of related documents. Each cluster is represented by a separate area as in Figure 1. For each cluster, the user is presented with typical words that occur in the text contents of documents in a cluster, as well as the titles of the three most prototypical documents. The user may *gather* some subset of those clusters, by pointing and selecting buttons above each cluster, and then ask the system to *scatter* that subcollection into another B subgroups, by selecting a Scatter/Gather button at the top of the display in Figure 1. The clustering is based on a form of inter-document similarity computation based on representations of text contents. Scatter/Gather browsing and clustering employs methods that can occur in constant interaction-time (Cutting, Karger, & Pedersen, 1993).

Patch Analysis

We can view Scatter/Gather clusters as *information patches*. Foraging in a cluster-patch corresponds to selecting a cluster, displaying the document titles belonging to the cluster in a scrollable window, scanning/scrolling through the listed titles, and for each title deciding if it is relevant or not to the query at hand. If the title is judged relevant, then it is handled by selecting, cutting, and pasting it to a query record window. Relevant and irrelevant documents will be randomly intermixed in the display window.

This simple loop of activity can be characterized by a cumulative gain function $g_i(t)$ that indicates how much information value is acquired over time t in cluster-patches of type i. In our empirical studies, we used specific task instructions that indicated that the information value was simply the number of relevant documents collected. The proportion of relevant documents in a cluster is the *precision*, P, of that cluster

$$P = \frac{N_R}{N_T}, \quad (3)$$

where N_R is the number of relevant documents and N_T is the total number of documents in the cluster. The rate of encounter, λ_P with relevant items while scanning/scrolling through a list is

$$\lambda_P = \frac{P}{t_s}, \quad (4)$$

where t_S is the time it takes to scan and judge the relevance of a title. If the time to handle a relevant item were $t_h = 0$, then we would have

$$g_i(t) = t\lambda_P. \quad (5)$$

However, when handling costs are non-negligible,

[1]This interface was developed by Marti Hearst at Xerox PARC.

Figure 1. The Scatter/Gather interface for navigation through large document collections

$$g_i(t) = t\left(\frac{\lambda_p}{1+\lambda_p t_h}\right) = \frac{N_R t}{N_T t_s + N_R t_h} \tag{6}$$

In this case, $g_i(t)$ is just a linear function of t. This is illustrated qualitatively by function $g_A(t)$ in Figure 2. Next we illustrate how a foraging model reveals the improvements that would be made by introducing an interface method that displayed cluster titles ordered by the similarity of document contents to a user's query rather than in the unordered presentation captured by Equation 6. This method is currently unavailable in Scatter/Gather, but certainly possible and advisable given our analysis.[2] This foraging model illustrates explicitly how the overall rate of information gain would change, as well as how the users' foraging time in each cluster would be reduced.

If the document titles in a cluster were ordered by their probability of relevance to a query, then the gain function would shift qualitatively to a diminishing returns curve such as $g_B(t)$ in Figure 2. That is, there would be higher number of relevant documents encountered while scanning the initial parts of the display of document titles, but this encounter rate would diminish as one continues scanning. Simple patch-foraging models (Stephens & Krebs, 1986) identify the overall maximum rate of information and the optimal time at which to stop foraging in a patch. In these traditional patch models, each patch type i with gain function $g_i(t_i)$, is additionally characterized by the rate at which such patch-gain functions will be encountered during prolonged foraging, λ_i. For instance, over a sufficiently long run of work with the Scatter/Gather interface one might be able to classify encountered clusters by their expected gain functions, g_i, and to additionally associate with each cluster type i the rate, λ_i, at which one expects to encounter clusters with gain functions g_i.

The optimization problem is to determine the optimal vector of patch residence times (t_1, t_2, ..., t_n) that maximizes the currency intake rate R.

$$R(t_1, t_2, \ldots, t_n) = \frac{\sum_{i=1}^{n} \lambda_i g_i(t_i)}{1 + \sum_{i=1}^{n} \lambda_i t_i}. \tag{7}$$

Charnov's marginal value theorem [7] states that the long-term rate of currency intake is maximized by choosing patch residence times so that the marginal value (instantaneous rate) of the currency gain at the time of leaving each patch equals the long-term average rate across all patches in the ecology.[3] More technically, the full vector of rate maximizing t_i's, $(\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_n)$, must fulfill the condition specified by

[2]The current Scatter/Gather does have a similarity search operator that works over the entire collection but cannot be restricted to selected clusters.

[3]We are assuming that it is long-term average net gain that is being maximized. See (Hames, 1992)Hames [10] for other optimization policies.

$$
\begin{aligned}
g'_1(\hat{t}_1) &= R(\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_n) \\
g'_2(\hat{t}_2) &= R(\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_n) \\
&\vdots \\
g'_n(\hat{t}_n) &= R(\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_n)
\end{aligned}
\tag{8}
$$

Charnov [7] shows that this theorem implies that residence times decrease as overall rate of gains from the information ecology improve and the optimal patch residence time occurs when the instantaneous gain rate equals the overall rate of gain from all patches that are foraged.

Figure 2 shows the graphical representation of Charnov's marginal value theorem that appears in many discussions of optimal foraging theory. Figure 2 captures the basic relations for the situation in which there is just one kind of patch-gain function. Travel time between patches (the inverse of patch encounter rate) is plotted on the abscissa, starting at the origin and moving to the left. To determine the overall maximum rate of gain R_B, one draws a line tangent to the gain function $g_B(t)$ and passing through $1/\lambda$ to the left of the origin. The point of tangency also provides the optimal maximum foraging time $\hat{t}_B$. Reasoning graphically with Figure 2, one should see that if one further improves the method of ordering of relevant documents presented to a user, then one changes the gain function $g_B(t)$ to $g_C(t)$, and that (keeping the tangent line anchored at $1/\lambda$) the rate of return improves to R_C, and the optimal foraging time in a cluster-patch decreases to $\hat{t}_C$. Given sufficient information about the relevant quantities and functions, one could make precise predictions about the impact of interface changes on optimal user performance. However, using the graphical reasoning just illustrated one may also make useful qualitative predictions.

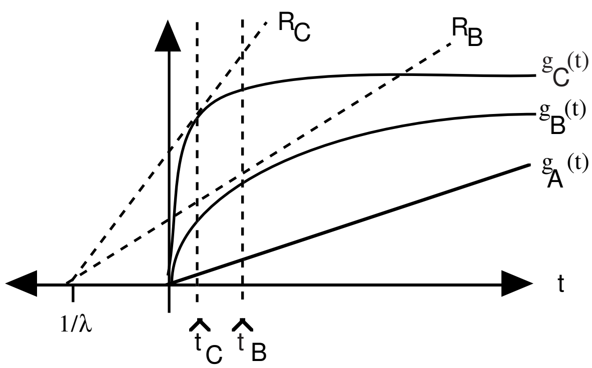

Figure 2. Optimal foraging in information cluster-patches. $g_A(t)$ is the current information-gain function, and $g_B(t)$ and $g_C(t)$ illustrate the effects of ordering by relevancy.

We suspect that in most tasks the rate of gaining information relevant to some embedding task from external collections diminishes (or is finite) with time. This occurs, for instance, in situations in which there is some redundancy among relevant documents. Consequently, the patch model analysis ought to have broad applicability.

EXAMPLE 2: INFORMATION DIET

The problem of selecting which clusters to pursue can be viewed as analogous to the problem faced by a predator who must select the optimal mix of prey to pursue upon encounter. Models of such choices are often called *prey selection* or *diet selection* models in optimal foraging theory. In our case, we have a problem of determining the *optimal information diet*.

Discussion of the analytic formulation and solution of prey selection models may be found in Stephen and Krebs [22], but here we focus on three general results: (a) ranking by profitability, (b) the zero-one rule, and (c) the independence of inclusion from encounter rate. Assume that the items being selected into the diet can be categorized as to their *profitability*, which is the expected net gain divided by the amount of time it takes to handle the item. In our case, the cluster profitabilities will be proportional to λ_p. To select the optimal diet, the items should be ranked by their profitability. Starting with the most profitable items, include lower-ranked items in the diet only if their profitability is greater than or equal to the average rate of gain of all the higher ranked items. Including items with profitability lower than the average rate of gain of the all the higher ranked items can only decrease the average. This is an instance of the *principle of lost opportunity*: by handling lower-ranked items in the diet one loses the opportunity to go after higher-ranked items. The zero-one result simply states that the optimal diet will be one in which items of a given profitability level are chosen all-or-none. The independence of inclusion from encounter rate result states that the inclusion of lower-ranked items in a diet is solely dependent on their profitability and not upon the rate at which they are encountered. To use an everyday analogy: if reading junk mail is categorized as a too-low profitability item (because there are better things to pursue), then the decision to ignore junk mail when it is encountered should be made regardless of the amount or rate of junk mail received.

EXAMPLE 3: A DYNAMIC FORAGING MODEL

Information patch models and information diet models are examples of *static* optimization models that do not take into account changes in human-system state that occur over time. Furthermore, the analysis above focused on localized subcomponents of the overall interaction with Scatter/Gather. It ignored the fact that users navigate through the cluster hierarchy presumably making choices about the best subclusters of documents to pursue. We now illustrate a more complete analysis characterized as a *dynamic* optimization model.

Tasks

For our model development we analyzed a version of Scatter/Gather that operates on the NIST Tipster corpus. This contains over one million documents collected from the Wall Street Journal, the AP newswire, Department of Energy technical abstracts, the Federal Register, and computer articles published by Ziff-Davis. This is a test corpus that is extensively used by the information retrieval community. One reason for using Tipster is that there are sets of information retrieval tasks (queries) that have been defined with associated listings of relevant and non-relevant Tipster documents, as judged by experts. If not an objective measure of task success, this at least provides us with a common standard to compare performance against.

Subjects, Measurements, and Procedures

To get measures of basic costs of using the Scatter/Gather interface we asked four subjects to retrieve as many relevant documents as possible for as many TREC queries as possible in one hour. Subjects were Xerox PARC researchers (including the two authors) whose experience with Scatter/Gather-Tipster were 96 hrs, 20 hrs, 10 hrs, and 2 hrs The purpose of this study was to obtain measurements of time costs associated with various Scatter/Gather browsing operations. Video tapes of subjects were parsed into the relevant events and time-coded to the nearest second. In addition, we conducted separate experiments on system response times. The relevant time costs are presented in Table 1.

Table 1: Empirical costs of Scatter/Gather use.

Average time on a query	$T = 720$ s
Scanning a cluster and judging relevance	$t_{sc} = 3$ s
Adding a cluster to gather list	$t_{gc} = 5$ s
Scanning a document title and judging relevance	$t_s = 1$ s
Selecting, cutting, and pasting title to record window	$t_h = 5$ s
System time to scatter new clusters	$t_{cl} = 23$ s
System time to display titles in a cluster	$t_d = 20$ s

Dynamic Programming

Dynamic programming has also been used to develop optimal foraging models [11]. Assuming a state variable X, one may define an evaluation function $\phi(x)$ that specifies the value of a state at some task deadline time T. In our task, this is the number of relevant documents collected by the end of the task. A state-change operator, $\delta i(x_t) = x_{t+1}$, that produces a new state x_{t+1} from a current state x_t, can be additionally characterized by the time, $C_i(t)$, it takes to perform the operator. The final fitness value F(x,T)—in our case number of documents one can expect to collect by the end of the task—for any state x at any time $0 \leq t \leq T$, can be defined recursively as

$$F(x,t) = \begin{cases} \phi(x), \text{ if } t = T \\ \max_i [F(\delta_i(x), t + C_i(x)], \text{ if } t < T \end{cases} \quad (9)$$

That is, assuming that the optimal choices (or *optimal policy*) will be followed from states at $t + 1$, make the optimal choice at time t.

Dynamic Programming Model for Scatter/Gather

For Scatter/Gather, we developed a multidimensional, approximate representation of the state variable X. We define a set of accessor functions that return the relevant components of a state variable x:

$$\begin{Bmatrix} N(x) = \text{No. documents in current cluster,} \\ R(x) = \text{No. relevant in current cluster,} \\ f_D(x) = \text{No. foraged in current cluster,} \\ f_C(x) = \text{No. subclusters foraged,} \\ Q(x) = \text{No. documents collected so far,} \\ H(x) = \text{Clustering history.} \end{Bmatrix} \quad (10)$$

where "clustering history" represents the path by which the user has navigated through the cluster hierarchy by iteratively gathering then scattering clusters.

A reasonably average initial state x_0 would be

$$\begin{Bmatrix} N(x_0) = 10^6 \\ R(x_0) = 10^3 \\ f_D(x_0) = 0 \\ f_C(x_0) = 0 \\ Q(x_0) = 0 \end{Bmatrix} \quad (11)$$

Next, we define a set of operators, $\delta_k(x)$, $k = 0, 1, ..., K$, that specify the possible changes in human-system interaction state, and an associated set of time-cost functions, $C_k(x)$, that specify how long it takes to perform the operations.

The set of activity involved in displaying a cluster, scanning and judging the displayed titles, and collecting the relevant ones can be captured by an operator $\delta_0(x)$. When applied to a state x_n, it yields a new state x_{n+1}

$$x_{n+1} = \delta_0(x_n). \quad (12)$$

At a given time, t, with $T - t$ time remaining until the deadline is hit, the number of relevant documents collected from the displayed cluster titles will be either (a) all the relevant cluster titles, if there is sufficient remaining time to go through the whole cluster, or (b) as many relevant documents as can be collected in the remaining time, as characterized by Equation 6:

$$Q(x_{n+1}) = \begin{cases} Q(x_n) + R(x_n), \text{ if } C_Q(x_n) < (T - t) \\ Q(x_n) + g_A(T - t), \text{ otherwise} \end{cases} \quad (13)$$

The associated time-cost function is

State-Space Analysis

Figure 3. Expected gains and optimal operator choices in Scatter/Gather navigation.

$$C_Q(x_n) = \min\begin{Bmatrix} N(x_n)t_s + R(x_n)t_h \\ T - t \end{Bmatrix} \qquad (14)$$

Unless the time cost of the operator takes the human-system interaction past the deadline time, T, interaction returns to the cluster that is superordinate in the cluster hierarchy, and the number of foraged documents, remaining relevant documents, and foraged subclusters are updated appropriately, as is the history.

It remains to specify the operators for navigating through the Scatter/Gather cluster hierarchy. At each cluster in the hierarchy, there is a set of 10 subclusters that may be gathered and then rescattered. Of the $2^9 - 2$ possible gatherings at each node in the cluster hierarchy, the analysis of information diet suggests that the optimum can be found by considering only nine possibilities. The clusters can be ranked by their profitabilities λ_p and we can consider gatherings of the top 1, top 2, ... top 9 clusters to find the optimal "cluster diet" (in the following, for practical reasons, we only considered the top 3).

Operators for choosing the highest-precision i = 1, 2, ...9 clusters ca be defined by operators δ_i,

$$y_{n+1} = \delta_i(y_n). \qquad (15)$$

If a cluster is divided into B subclusters (e.g., B = 10, in our case), then the number of documents collected together by gathering i subclusters together is, in the average case,

$$N(y_{n+1}) = \frac{N(y_n)i}{B}. \qquad (16)$$

The number of relevant documents that will be found in the newly gathered cluster will be

$$R(y_{n+1}) = D_i(y_n) - f_D(y_n), \qquad (17)$$

where $D_i(y_n)$ is the Cumulative Cluster Relevancy Distribution for a gathering of i subclusters, and $f_D(y_n)$ is the number of relevant documents that have already been foraged through. $D_i(y_n)$ assumes that the relevant documents in the gathered clusters are distributed exponentially when the clusters are ranked from highest to lowest in terms of their precisions, λ_p, bounded by the maximum number of documents that can be allotted to a cluster. We used the exponential exp(-.8 * cluster rank) in our calculations.

The cost of the gathering operation is

$$C_i(y_n) = Bt_{sc} + it_{gc} + t_{cl} \qquad (18)$$

Figure 3 presents a summary of the expected value and the optimal operator-choice policies for states in the state-space of Scatter/Gather navigation. Each point in Figure 3 represents a state reachable from the initial state specified in (11) by the iterative application of the Scatter/Gather operators specified above: display-scan-select relevant documents from a cluster (δ_0), scatter-gather one (δ_1), two (δ_2), or three clusters (δ_3). For instance, the leftmost point in Figures 3a-3c plots the initial state and four of the state-points to right of that point were generated by application of the four operators (each state-point in Figures 3a-3c is similarly associated with four points to the right, up to the deadline time).

Figure 3a presents the number of documents that are optimally expected to yet be gained before the task deadline arrives, $E^*[g_i(T - t)]$, for each Scatter/Gather state. Each state is plotted as a box, whose degree of fill is proportional to the precision, λ_p, of clusters at that state. $E^*[g_i(T - t)]$ decreases as available time decreases and as the precision decreases, as suggested by Equation 6.

Figure 3b presents, $Q(x)$, the number of documents already collected at each state of Scatter/Gather, as a function of time. State precisions are plotted as in Figure 3a. It takes about half the task time to begin collecting documents. Figure 3c plots (ignoring precisions) the sum of $Q(x)$ (Figure 3b) and $E^*[g_i(T - t)]$ (Figure 3a), which is the fitness, $F(x, t)$ of each state x, as defined in Equation 9.

Figure 3d shows the optimal operator-choice policy for each state encountered up to the decision to begin to collect documents, as a function of state precision and time. Circles represent states from which optimal operators are δ_1, δ_2, and δ_3 for scatter-gathering clusters, with circle size proportional to the optimal number of clusters selected. Dots represent states for which the optimal operator choice is to begin displaying, scanning, and selecting documents (δ_0). Scatter-gathering clusters is optimal at the beginning of the task when time available and cluster precision is low and displaying-selecting is optimal later or at higher precisions. The optimal number selected clusters tends to increase with overall state precision, and precision increases with time.

CONCLUSION

Stephen's and Krebs begin the preface of their book on *Foraging Theory*, thus:

> This book analyzes feeding behavior in the way an engineer might study a new piece of machinery. An engineer might ask, among other questions, about the machine's purpose: is it for measuring time, wind speed, or income tax? This is a worthwhile question for the engineer, because machines are built with a purpose in mind, and any description of a machine should refer to its purpose. . . .
>
> Biologists also ask questions about purpose, about what things are for in contrast to the engineer, the biologist thinks of design or purpose as the product of natural selection, rather than as the product of a conscious creator. (Stephens & Krebs, [22], p. ix]

In this paper, we have tried to make a start at reversing the above analogy—exploiting theory developed in the service of behavioral ecology to analyze information ecologies and the design of interactive information systems. It is hoped that this line of research will give us a new set of tools to aid in the design of these systems.

REFERENCES

1. Anderson, J. R. (1990). *The adaptive character of thought* . Hillsdale, NJ: Lawrence Erlbaum Associates.

2. Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

3. Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*, 407-424.

4. Bechtel, W. (1985). Realism, instrumentalism, and the intentional stance. *Cognitive Science, 9*, 473-497.

5. Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction* . Hillsdale, NJ: Lawrence Erlbaum Associates.

6. Card, S. K., Robertson, G. G., & Mackinlay, J. D. (1991). The Information Visualizer. In Proceedings of the *CHI '92 Conference* (pp. 181-188), New Orleans.

7. Charnov, E. L. (1976). Optimal foraging: The marginal value theorem. *Theoretical Population Biology, 9*, 129-136.

8. Cutting, D. R., Karger, D. R., & Pedersen, J. O. (1993). Constant interaction-time Scatter/Gather browsing of very large document collections. In Proceedings of the *SIGIR '93 Conference.*

9. Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992). Scatter/gather: A cluster-based approach to browsing large document collections. In Proceedings of the *SIGIR '92 Conference* (pp. 318-329).

10. Hames, R. (1992). Time allocation. In E. A. Smith & B. Winterhalder (Ed.), *Evolutionary ecology and human behavior* (pp. 203-235). New York: de Gruyter.

11. Mangel, M. & Clark, C. W. (1988). *Dynamic modeling in behavioral ecology* . Princeton, NJ: Princeton University Press.

12. Marr, D. (1982). *Vision* . San Francisco: W.H. Freedman.

13. Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

14. Olson, J. R. & Olson, G. M. (1990). The growth of cognitive modeling in human-computer interaction since GOMS. *Human-Computer Interaction, 5*, 221-265.

15. Reitman, W. R. (1965). *Cognition and thought: An information-processing approach* . New York: Wiley.

16. Repo, A. J. (1986). The value of information: Approaches in economics, accounting, and management science. *Journal of the American Society for Information Science, 40*, 68-85.

17. Resnikoff, H. L. (1989). *The illusion of reality* . New York: Springer-Verlag.

18. Schamber, L., Eisenberg, M. B., & Nilan, M. S. (1990). A re-examination of relevance: Towards a dynamic situational definition. *Information Processing and Management, 26*, 755-776.

19. Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99-118.

20. Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4*, 181-204.

21. Smith, E. A. & Winterhalder, B. (1992). *Evolutionary ecology and human behavior* . New York: de Gruyter.

22. Stephens, D. W. & Krebs, J. R. (1986). *Foraging theory* . Princeton, NJ: Princeton University Press.

23. Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy, 69*, 213-225.