

INFORMATION FORAGING THEORY

1

Information Foraging Theory

Framework and Method

Knowledge is power.
 —Sir Francis Bacon,
Meditationes Sacrae,
De Hæresibus (1597)

Modern mankind forages in a world awash in information, of our own creation, that can be transformed into knowledge that shapes and powers our engagement with nature. This information environment has coevolved with the epistemic drives and strategies that are the essence of our adaptive toolkit. The result of this coevolution is a staggering volume of content that can be transmitted at the speed of light. This wealth of information provides resources for adapting to the problems posed by our increasingly complex world. However, this information environment poses its own complex problems that require adaptive strategies for information foraging. This book is about *Information Foraging Theory*, which aims to explain and predict how people will best shape themselves for their information environments and how information environments can best be shaped for people.

Information Foraging Theory is driven by three maxims attributable in spirit, if not direct quotation, to Allen Newell's (1990) program of Unified Theories of Cognition:¹

1. *Good science responds to real phenomena or real problems.* Human psychology has evolved as an adaptation to the real world. Information foraging theory is concerned with understanding representative problems posed by the real-world information environment and adaptive cognitive solutions to those problems.
2. *Good science makes a difference.* Information Foraging Theory is intended to provide the basis for application to the design and evaluation of new technologies for human interaction with information, such as better ways to forage for information on the World Wide Web.
3. *Good science is in the details.* The aim is to produce working formal models for the analysis and prediction of observable behavior.

Like much of Newell's work, the superficial elegance and simplicity of these maxims unfurls into complex sets of entailments. In this book I argue that the best approach to studying real information

4 INFORMATION FORAGING THEORY

foraging problems is to adopt *methodological adaptationism*, which directs our scientific attention to the ultimate forces driving adaptation and to the proximate psychological mechanisms that are marshaled to produce adaptive solutions. Thus, the methodology of Information Foraging Theory is more akin to the methodology of biology than that of physics, in contrast with the historical bulk of experimental psychology. To some extent, this choice of methodology is a consequence of the success with which Information Foraging Theory has been able to draw upon metaphors, models, and techniques from optimal foraging theory in biology (Stephens & Krebs, 1986). The concern with application (Newell & Card, 1985) drives the theory to be relevant to technological design and evaluation, which requires that models be truly predictive a priori (even if approximately so) rather than a “good fit” explanation of the data a posteriori, as is the case with many current psychological models. Being concerned with the details drives the theory to marshal a variety of concepts, tools, and techniques that allow us to build quantitative, predictive models that span many levels of interrelated phenomena and interrelated levels of explanation. This includes the techniques of task analysis through state-space and problem-space representations, rational analysis and optimization analysis of adaptive solutions, and production system models of the cognitive systems that implement those adaptive solutions.

Audience

The intent of this book is to provide a comprehensive presentation of Information Foraging Theory, the details of empirical investigations of its predictions, and applications of the theory to the engineering and design of user interfaces. This book aims primarily at an interdisciplinary audience with backgrounds and interests in the basic and applied science aspects of cognitive science, computer science, and the information and library sciences. The theory and methodology have been developed by drawing upon work on the rational analysis of cognition, computational cognitive modeling, behavioral ecology, and microeconomics. The crucible of empirical research that has shaped Information Foraging Theory has been application problems in human–information interaction, which is emerging as a new branch in the

field traditionally known as human–computer interaction. Although the emphasis of this book is on theory and research, the insights and results are intended to be relevant to the practitioner interested in a deeper understanding of information-seeking behavior and guidance on new designs. Chapter 9 is devoted entirely to practical applications of the theory.

By its nature, Information Foraging Theory involves the use of technical material such as mathematical models and computational models that may not be familiar to a broad audience. Generally, the technical aspects of the theory and models are presented along with succinct discussion of the key concepts, insights, and principles that emerge from the technical parts, along with illustrative examples, metaphors, and graphical methods for understanding the key points. The aim of this presentation is to provide intuitive understanding along with technical precision and insight.

Frameworks, Theories, and Models

Like other programs of research in the behavioral and cognitive sciences, Information Foraging Theory can be discussed in terms of the underlying framework, the theory itself, and the models that specify predictions in specific situations. *Frameworks* are the general pools of concepts, assumptions, claims, heuristics, and so forth, that are drawn from to develop theories, as well the methods for using them to understand and predict the world. Often, frameworks will overlap. For instance, information processing psychology is a broad framework that assumes that theories about human behavior can be constructed out of information processing concepts, such as processes that transduce physical sensations into sensory information, elements storing various kinds of information, and computational processes operating over those elements. A related framework, connectionism, shares these assumptions but makes additional ones about the nature of information processing being neuronlike. Although bold claims may be made by frameworks, these are typically not testable in and of themselves. For instance, whether the mind is mostly a general purpose learning machine or mostly a collection of exquisitely evolved computational modules are not testable claims in and of themselves.

Theories can be constructed within frameworks by providing additional assumptions that allow one to

make predictions that can be falsified. Typically, this is achieved by specifying a *model* for a specific situation or class of situations that makes precise predictions that can be fit to observation and measurement. For instance, a model of information seeking on the Web (SNIF-ACT) is presented in chapter 5 that predicts the observed choice of Web links in given tasks. It includes theoretical specifications of the information processing model of the user, as well as assumptions about the conditions under which it applies (e.g., English-speaking adults seeking information about unfamiliar topics). The bulk of this book is about Information Foraging Theory and specific models. The aim of this introductory chapter is to provide an outline of the underlying framework and methodology in which Information Foraging Theory is embedded. However, before presenting such abstractions, a simple example is offered in order to illustrate the basic elements and approach of Information Foraging Theory.

Illustration

The basic approach of Information Foraging Theory can be illustrated with a simple example that I hope is familiar to many, involving the task of finding a good, reasonably priced hotel using the World Wide Web (Pemberton, 2003). A typical hotel Web site (see figure 1.1) will allow a user to search for available hotels in some specified location (e.g., “Paris”) and then allows the user to sort the results by the hotel star rating (an indicator of quality) or by price (but not both). The user must then click-select each result to read it, because often the price, location, and features summaries are inaccurate. Lamenting the often poor quality of such hotel Web sites, Pemberton (2003) suggested that improved “usability is about *optimizing the time* you take to achieve your purpose, how well you achieve it, and the satisfaction in doing it. . . . How fast can you find the perfect hotel?”^a This notion of usability is at the core of Information Foraging Theory.

For illustration, consider the somewhat simplified and idealized task of finding a low-priced, two-star hotel in Paris.² This example shows (in much simplified form) the key steps to developing a model of information foraging: (a) a rational analysis of the task and information environment that draws on optimal foraging theory from biology and (b) a production system model of the cognitive structure of task.

Rational Analysis of the Task and Information Environment

Figure 1.2 presents an analysis of results of search for two-star Paris hotels that I conducted on a popular hotel Web site. The Paris hotel descriptions and prices were returned as a vertical list presented over several Web pages. I sorted the list by star rating and went to the page that began to list two-star hotels. In figure 1.2, the x-axis indicates the order of two-star hotel listings in the search result list when sorted by star rating, beginning at the first two-star hotel through the last two-star hotel, and the y-axis indicates price. Prices fluctuate as one proceeds down the list of Paris hotels. As noted above, this particular hotel Web site, like many others, does not allow the user to sort by both quality (star rating) and price—one must choose one or the other sorting. Assume a rational (and perhaps somewhat boring) hotel shopper who was concerned only with being frugal and sleeping in a two-star hotel. If that shopper methodically scanned the two-star hotel listings, keeping track of only the lowest priced hotel found so far, the lowest price encountered would decrease as plotted in figure 1.3. That is, the shopper would at first find a relatively rapid decrease in lowest price, followed by fewer improvements as the scan progressed. Figure 1.4 shows the savings attained (compared with the very first hotel price found on the list) by continuing to scan down the list. Figure 1.4 is a typical *diminishing returns* curve in which additional benefits (returns) diminish as one invests more resources (in this case, scan time).

A diminishing returns curve such as figure 1.4 implies that the expected value of continuing to scan diminishes with each additional listing scanned. If the list of search results were very long—as is often the case with the results produced by Web search engines—there is usually a point at which the information forager faces the decision of whether it is worth the effort of continuing to search for a better result than anything encountered so far. In the particular example plotted in figure 1.2, there were no additional savings for the last 18 items scanned. Figure 1.3 includes a plot of the expected minimum price encountered as a function of scanning a search result list, and figure 1.4 includes a plot of the expected savings as a function of scanning. These expectations were computed assuming that observed hotel prices in figure 1.2 come from a standard

^a[AQ: pls cite pg no for quote, if this work has > 1 pg overall (i.e., no need to add pg no if the whole piece is only 1 pg) - pls see query in ref list]

6 INFORMATION FORAGING THEORY

The screenshot shows a web browser window displaying search results for hotels in Paris, France. The page includes a search bar, filters for arrival and departure dates, and a list of six hotels. Each hotel listing includes a star rating, hotel name, location, description, average nightly rate, and prices for specific dates (Oct 1 and Oct 2). A 'SELECT' button is provided for each hotel.

| Hotel Name | Star Rating | Average Nightly Rate | Oct 1 | Oct 2 |
|----------------------------------|-------------|----------------------|--------|--------|
| 1. Villa Beaumarchais | ★★★★ | \$ 599.95 | 599.95 | 599.95 |
| 2. Amarante Elysees Star | ★★★★ | \$ 269.95 | 249.95 | 269.95 |
| 3. Best Western Astra Opera | ★★★★ | \$ 189.95 | 189.95 | 189.95 |
| 4. Best Western Beau Manoir | ★★★★ | \$ 229.95 | 229.95 | 229.95 |
| 5. Best Western Bradford Elysees | ★★★★ | \$ 209.95 | 209.95 | 209.95 |
| 6. Best Western Derby Alma | ★★★★ | \$ 179.95 | 179.95 | 179.95 |

FIGURE 1.1 A typical Web page from a hotel search site.

distribution of commodity prices (see the appendix for details). Assuming that our hypothetical rational hotel shopper valued time (time is money), the question would be whether the savings expected to be gained by additional scanning of hotel results was worth the time expected to be expended.

In contrast to this simple illustration, typical information problems solved on the Web are more

complicated (Morrison, Pirolli, & Card, 2001), and the assessments of the utility of encountered information foraging depend on more subtle cues than just prices. However, the basic problem of judging whether continued foraging will be useful or a waste of valuable time is surely familiar to Web users. It turns out that this problem is very similar to one class of problems dealt with in optimal foraging theory.

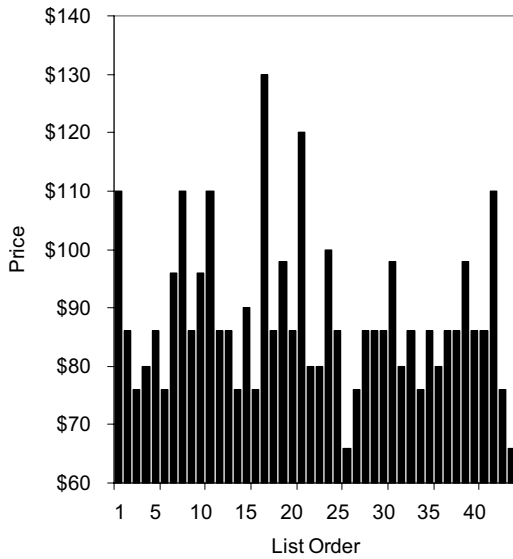


FIGURE 1.2 Prices of two-star Paris hotels in the order encountered in the results of a search of a hotel Web site.

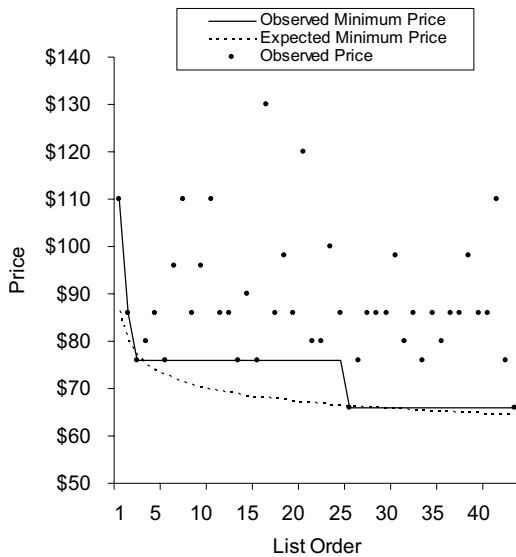


FIGURE 1.3 The minimum two-star Paris hotel price as a function of order of encounter. The observed prices are the same as those in figure 1.2. The observed minimum is the least expensive hotel price found so far in a process that proceeds through the prices in the order listed. The expected minimum is a prediction based on the assumption that prices are being sequentially and randomly sampled from a fixed distribution of prices (see the appendix for details).

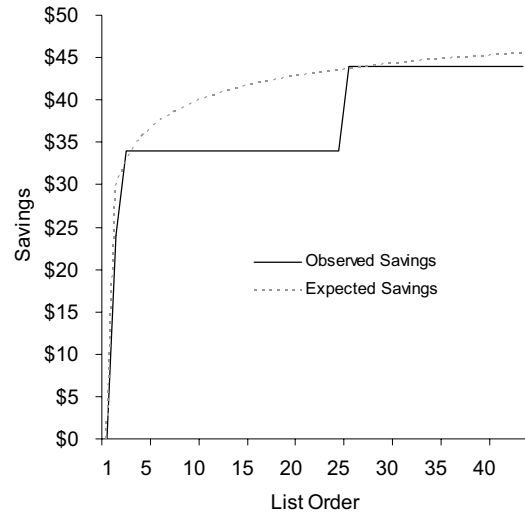


FIGURE 1.4 Diminishing returns of savings as a function of list order. The observed savings is the difference between the observed minimum price found so far and the first price encountered (\$110), presented in figure 1.3. The expected savings is the difference between the expected minimum price and first price encountered.

An Optimal Foraging Analogy

Many animals forage in patchy environments, with food arranged into clumps. For instance, a bird that feeds on berries in bushes will spend part of its time searching for the next bush and part of its time berry picking after having found a bush. Often, as an animal forages in a patch, it becomes harder to find food items. In other words, foraging within a food patch often exhibits a diminishing returns curve similar to the one in figure 1.5. Such diminishing returns may occur, for instance, because prey actively avoid the forager as they become aware of the threat of predation. Diminishing returns may also occur because the forager has a strategy of picking off the more highly profitable items first (e.g., bigger berries for the hypothetical bird) from a patch with finite resources. Like the hypothetical Web shopper discussed above, the problem for a food forager facing diminishing returns in a patch is whether to continue investing efforts in getting more out of the patch, or to go look for another patch.

Figure 1.5 is a graphical version of a simple *conventional patch model* (Stephens & Krebs, 1986) based on *Charnov's Marginal Value Theorem* (Charnov,

8 INFORMATION FORAGING THEORY

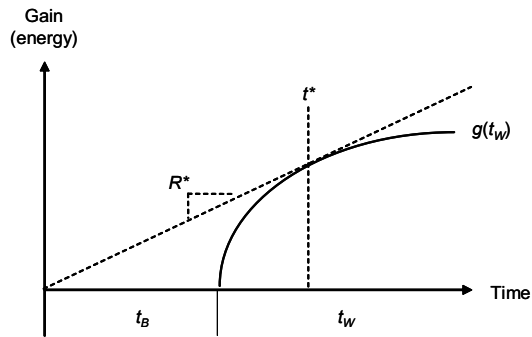


FIGURE 1.5 Charnov's Marginal Value Theorem states that the rate-maximizing time to spend in patch, t^* occurs when the slope of the within-patch gain function g is equal to the average rate of gain, which is the slope of the tangent line R^* .^a

^a[AQ: see query in text re t_s vs. t_B - pls correct here or there]

1976). The model depicted in figure 1.5 assumes that an animal foraging for food encounters only one kind of food patch at random that is never reencountered. When searching for the next food patch, it takes an average of t_B amount of time to find the next patch (between-patch time).^b Once a patch is encountered, foraging within the patch returns some amount of energy (e.g., as measured by calories) that increases as a function, g , of the time, t_w ,^c spent foraging within the patch. Figure 1.5 shows a diminishing returns function, g , for within-patch foraging. The problem for the forager is how much time, t_w , to spend within each patch before leaving to find the next patch.

^b[AQ: pls also define t_s here from figure 1.5 - is this time searching, = t_B ?]

^c[AQ: correct to use uppercase W, as on figure and in later chapters? Pls confirm throughout (changed also in equations)]

The conventional patch model assumes that the animal forager optimizes the overall rate of gain, R , that characterizes the amount of energy gained per unit time of foraging:

$$R = \frac{g(t_w)}{t_B + t_w}, \quad (1.1)$$

or the amount of energy (calories) gained from an average patch divided by the time spent traveling from one patch to the next (t_B) plus the time spent foraging within a patch (t_w).^d The optimal amount of time, t^* , to spend in a patch is the one that yields the maximum rate of gain, R^* ,

^d[AQ: lowercase w correct, as in most cases in book?]

$$R^* = \frac{g(t^*)}{t_B + t^*}. \quad (1.2)$$

Charnov's Marginal Value Theorem (Charnov, 1976) is a mathematical solution to this problem of determining t^* . It basically says that a forager should leave a patch when the rate of gain within the patch [as measured by the slope of $g(t_w)$ or more specifically the derivative $g'(t_w)$] drops below the rate of gain that could be achieved by traveling to, and foraging in, a new patch. That is, the optimal forager obeys the rule,

if $g'(t_w) \geq R^*$, then continue foraging in the patch; otherwise, when $g'(t_w) < R^*$, then start looking for a new patch.

Charnov's Marginal Value Theorem can be illustrated graphically in figure 1.5 for this simple problem (one kind of patch, randomly distributed in the world). First, note that the gain function g begins to climb only after t_B ,^e which captures the fact that it takes t_B time to go from the last patch to a new patch. If we draw a line beginning at the origin to any point on the gain function, g , then the slope of that line will be the overall rate of gain R , as specified in equation 1.1. Figure 1.5 shows such a line drawn from the origin to a point just tangent to the function g . The slope of this line is the optimal rate of gain R^* as computed in equation 1.2. This can be verified graphically by imagining other lines drawn from the origin to points on the function g . None of those lines will have a steeper slope than the line plotted in figure 1.5. The point at which the line is tangent to g will be the point at which the rate of gain, $g'(t_w)$ within the patch is equal to R^* . This point also determines t^* , the optimum time to spend within the average patch.

^e[AQ: t_B does not appear on figure - do you mean t_s ? Pls correct either here or on figure]

Production System Models

The rational analyses in Information Foraging Theory, which often draw from optimal foraging theory, are used to inform the development of production system models. These rational analyses make minimal assumptions about the capabilities of foragers. Herbert Simon (1955) argued that organisms are not optimal, rational agents having perfect information and unlimited computational resources. Rather, organisms exhibit *bounded rationality*. That is, agents are rational and adaptive, within the constraints of the environment and the psychological machinery

^f[AQ: correct first name?]

available to them biologically. Production system models provide a way of specifying the mechanistic structures and processes that implement bounded rationality. On the one hand, production systems have been used in psychology as a particular kind of computer simulation formalism for specifying the information processing that theorists believe people are performing. On the other hand, production systems have evolved into something more than just a class of computer simulation languages: They have become theories about the basic information processing architecture of cognition that is implemented in human brains (Anderson, 1983; Anderson & Lebiere, 1998; Newell, 1990; Young, in press^g).

In general, as used in psychology,³ production systems are composed of a set of *production rules* that specify the dynamics of information processing performed by cognition (*how we think*). Production rules operate over memories (or databases) that contain symbolic structures that represent aspects of the external environment and internal thought (*what we think about*). The system operates in a cyclical fashion in which production rules are selected based on the contents of the data memories and then executed. The execution of a production rule typically results in some change to the memories.

The production system models presented in this book are extensions of^h ACT theory (Anderson et al., 2004; Anderson & Lebiere, 1998). ACT theory assumes that there are two kinds of knowledge, *declarative* and *procedural* (Ryle, 1949). Declarative knowledge is the kind of knowledge that a person can attend to, reflect upon, and usually articulate in some way (e.g., by declaring it verbally or by gesture). Declarative knowledge includes the kinds of factual knowledge that users can verbalize, such as “The ‘open’ item on the ‘file’ menu will open a file.” Procedural knowledge is the know-how we display in our behavior, without conscious awareness. For instance, knowledge of how to ride a bike and knowledge of how to point a mouse to a menu item are examples of procedural knowledge. Procedural knowledge specifies how declarative knowledge is transformed into active behavior.

ACT-R (the most recent of the ACT theories) has a memory for each kind of knowledge (i.e., a *declarative memory* and a *procedural memory*) plus a special *goal memory*. At any point in time, there may be a number of goals in goal memory, but the system behavior is focused to achieve just one goal at a time.

Complex arrangements of goals and subgoals (e.g., for developing and executing plans to find and use information) can be implemented by manipulating goals in goal memory.

Production rules (or *productions*) are used to represent procedural knowledge in ACT-R. That is, they specify how to apply cognitive skill (know-how) and how to retrieve and use declarative knowledge. Table 1.1 presents an example of a production system for the task of finding a low-cost hotel using a Web site. The example in table 1.1 is not intended to be a psychologically plausible model, but rather it illustrates key aspects of production system models and how they are used in this book. The productions in table 1.1 are English glosses of productions written in ACT-R 5.0, which is discussed in greater detail below.⁴ Each production rule is of the form

IF ⟨condition⟩, THEN ⟨actions⟩.

The condition of a rule specifies a pattern. When the contents of declarative working memory match the pattern, the rule may be selected for application. The actions of the rule specify additions and deletions of content in declarative working memory, as well as motor commands. These actions are executed if the rule is selected to apply. In ACT-R, each production rule has conditions that specify which goal information must be matched and which declarative memory must be retrieved. Each production rule has actions that specify behavioral actions and possibly the setting of subgoals. Typically, ACT-R goal memory is operated on as what is known in computer science as a push-down stack: a kind of memory in which the last item stored will be the first item retrieved. Hence, storing a new goal is referred to as “pushing a goal on the stack,” and retrieval is referred to as “popping a goal from the stack.”

The production rules in table 1.1 assume that declarative memory contains knowledge encoded from the external world about the location and content of links on a Web page. The productions also assume that an initial goal is set to find a hotel price, and the productions accomplish the task by “scanning” through the links keeping track of the lowest price found so far. This involves setting a subgoal to judge the minimum of the current best price and the price just attended when each link is scanned. Figure 1.6 presents a trace of the productions in table 1.1

^g[AQ: can you update?]

^h[AQ: pls define the abbr “ACT” here]

10 INFORMATION FORAGING THEORY

TABLE 1.1 A production system for the task of finding a low hotel price.

P1: Start
 IF the goal is to find a hotel
 & there is a page of Web results
 & no link location has been processed
 THEN modify the goal to specify that the first location is to be processed

P2: First-link
 IF the goal is to find a hotel
 & a link location is specified
 & no best price has been noted yet
 & the link at the location indicates a price
 & the link is followed by a link at a new location
 THEN note that the best price is the price from the link at that location
 & modify the goal to specify the new location of the next link

P3: Next-link
 IF the goal is to find a hotel
 & a link location is specified
 & there is a current best price
 & the link at the location indicates a new price
 & the link is followed by a link at a new location
 THEN create a subgoal to find the minimum of the current price and the new price
 & push the subgoal on the goal stack
 & modify the current goal to specify the new location of the next link
 & note the resulting new minimum price as the best price

P4: Minimum-price-stays-the-same
 IF the goal is to find the minimum of the current price and the new price
 & there is a current best price
 & there is a new price
 & the current best price is less than or equal to the new price
 THEN note that the current best price is the minimum
 & pop the subgoal

P5: New-minimum-price
 IF the goal is to find the minimum of the current price and the new price
 & there is a current best price
 & there is a new price
 & the current best price is greater than the new price
 THEN note that the new price is the minimum
 & pop the subgoal

P6: Go-do-something-else (Done)
 IF the goal is to find a hotel
 & there is a current best price
 THEN stop

operating to scan the list of hotel prices depicted in figure 1.1 and graphed in figure 1.2.

Production “P1: Start” in table 1.1 applies at cycle 0 in figure 1.6 when the goal is to find a hotel price. Production “P2: First-link” applies at cycle 1 to scan the first link location and set the initial minimum hotel price. Then, production “P3: Next-link” applies repeatedly to scan subsequent links (cycles 2–53). For each link scanned, P3 sets a subgoal—by creating a new

goal and making it the focus in goal memory—to compare the currently scanned price to the current minimum price. This subgoal evokes either production “P4: Minimum-price-stays-the-same” or “P5: New-minimum-price.” When either P4 or P5 applies, it pops the subgoal to determine the minimum, and control passes back to the top-level goal of finding a hotel price.

Note in figure 1.6 that the trace ends at cycle 52 with the execution of production “P6: Done” after

| <i>Level</i> | <i>Question</i> | <i>Stance</i> | <i>Analysis Elements</i> | <i>Examples</i> |
|--------------|---|------------------------|---|--|
| Rational | What environmental problem is solved? Why is this solution a good one? | Design | <ul style="list-style-type: none"> • States, resources, state dynamics • Constraints, affordances • Feasible strategies • Optimization criteria | <ul style="list-style-type: none"> • Optimal foraging theory • Information Foraging Theory |
| Knowledge | What does the system know? | Intentional | <ul style="list-style-type: none"> • Environment • Goals, preferences • Knowledge • Perception, action | <ul style="list-style-type: none"> • Knowledge-level analysis |
| Cognitive | How does the system do it? | Information processing | <ul style="list-style-type: none"> • Cognitive states • Cognitive processes | <ul style="list-style-type: none"> • ACT-R • Soar |
| Biological | How does the system physically do it? | Biophysical | <ul style="list-style-type: none"> • Neural processes | <ul style="list-style-type: none"> • Neural models |

FIGURE 1.6 Levels of explanation.

scanning the link at location 26 in the list of results. The list actually contains 44 links in the result list (figure 1.2). The production system stops at link location 26 because of the way it implements elements of the rational analysis described above. Productions “P3: Next-link” and “P6: Done” match very similar patterns in declarative memory. In fact, on every cycle that P3 or P6 fires in the trace, the other production also matches. In production system terminology, P3 and P6 form a *conflict set* when on a particular cycle they both match the current pattern in the goal stack and declarative memory. In such cases, the *utility* of each production in the conflict set is evaluated and used to perform *conflict resolution* to determine which production to execute.

Production “P6: Done” is associated with a utility that corresponds to R discussed above: the overall rate of gain. I simply assumed that this corresponds to how the production system values its time. For the trace in figure 1.6, I assumed that the production system valued its time at $R = \$10/\text{hour}$.

Production “P3: Next-link” is associated with a utility that corresponds to $g'(t)$ discussed above: the rate of savings that would be achieved by looking at the next link: expected savings from scanning next link/time to scan link (in hours). The appendix discusses how expected savings is computed assuming the distribution of hotel prices evident in figure 1.2. From self-observation, I noted that it took 30 sec (30/3600 hour) to scan a link on the Web site depicted in figure 1.1. The competition between productions P3 and P6 implements the key idea of Charnov’s Mar-

ginal Value Theorem: As long as the rate of savings expected for production “P3: Next-link” is greater than the overall rate of gain, R , associated with “P6: Done,” then the system continues to scan links; otherwise, it quits.

Summary

I have presented this simple concrete example to sketch out the overall framework and approach of Information Foraging Theory before beginning more abstract discussion of framework and method. At this preliminary stage, it was necessary to gloss over unrealistic assumptions about Web use and the technical details of the analysis and model. However, it is important to point out two realistic aspects of the example. First, as will become clear in chapter 3, the Web does have a patchy structure (e.g., Web sites and search results), and diminishing returns within those information patches is common. For instance, figure 1.7 is based on data from a study of medical information seeking (Bhavnani, 2005).⁵ Bhavnani, Jacob, Nardine, and Peck (2003)ⁱ asked melanoma experts to identify melanoma risk facts that they identified as important for a melanoma patient to understand. Figure 1.7a shows the distribution of melanoma risk facts across Web pages. Very few pages contain all 14 expert-identified melanoma risk concepts, but many contain one of the melanoma risk facts. Figure 1.7b is an estimate of the number of melanoma risk facts that a user would encounter as a function of visits to melanoma-related pages (Bhavnani et al.,

ⁱ[AQ: correct cite?]

12 INFORMATION FORAGING THEORY

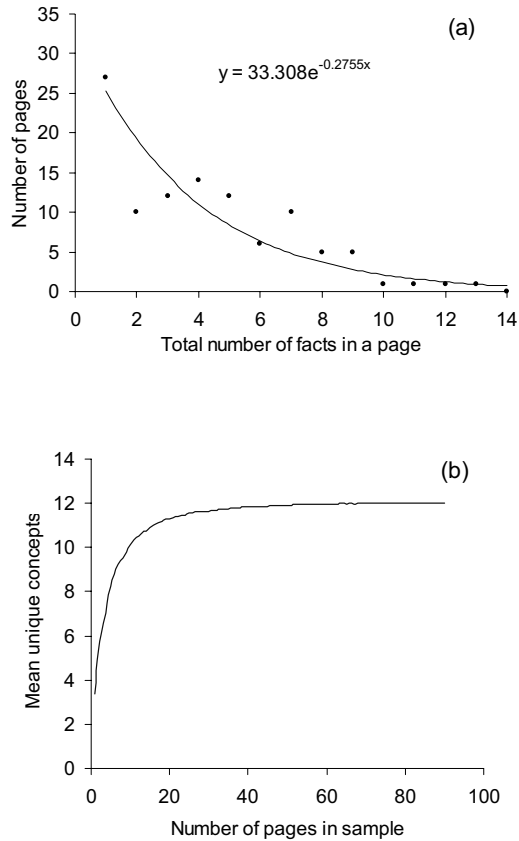


FIGURE 1.7 (a) The distribution of number of key concepts about melanoma risk across Web pages, and (b) the cumulative number of key concepts encountered as a function of size of sample of pages (Bhavnani, 2005; Bhavnani et al., 2003).

2003). Note that it is a diminishing returns curve and that the user is expected to require 25 page visits to find all expert-identified melanoma risk facts.

In the remaining sections of this chapter, I provide an overview of broader framework and method. The remainder of this book is about the empirical and theoretical details.

Man the Informavore

All men by nature desire knowledge.—*Aristotle, Metaphysics*

The human propensity to gather and use information to adapt to everyday problems in the world is a core piece of human psychology that has been largely ig-

nored in cognitive studies. George A. Miller (1983), however, recognized the centrality of this human propensity to our cognitive natures and argued that mankind might fruitfully be viewed as a kind of *informavore*: a species that hungers for information in order to gather it and store it as a means for adapting to the world. Picking up on this idea, Dennett (1991) traced out a plausible evolutionary history in which he suggested that our ancestors might have developed vigilance behaviors that required surveying and assessing the current state of the environment, much like the prairie dogs who pop up on two feet to perform their situation appraisals or the harbor seals that break the surface in the middle of a beach break to check out whether the surfers are friends, foe, or prey. Adaptive pressures to gain more useful, actionable knowledge from the environment could lead to the marshaling of available cognitive and behavioral machinery, resulting in organisms, such as primates, that have active curiosity about the world and themselves. Humans, of course, are extreme in their reliance on information, with language and culture, and now modern technology, providing media for transmission within and across generations. Humans are the *Informavores rex* of the current era.

George Miller's notion of humans as *informavores* suggests that our genes have bestowed upon us an evolving behavioral repertoire that now includes the technological aspects of our culture associated with finding, saving, and communicating information. It is common in evolutionary discussions to distinguish between *genotype* and *phenotype* (Johanssen, 1911). The genotype is the blueprint for an individual. What get passed from one generation to the next (if it survives and reproduces) are the genotypic blueprints. Phenotypes are the outward manifestation of the genotype. Typically, people think of this as the bodily structure and behavior of the individual organism. However, Dawkins (1989) introduced the notion of *extended phenotype* to clarify the observation that the genotype has extended effects on the world at large that go beyond the actual body and behavior of the individual. Not only do beavers have tails, but they use them to make dams. Not only do spiders have legs, but they use them to make webs. Humans have not only brains but also external technology for storing information, and information foraging strategies that can be invoked to call forth the right knowledge, at the right time, to take useful action. It remains an open question as to why humans have evolved such

information collection strategies—a question that I raise again at the end of this book.

The Adaptive Pressure of the Wealth of Information

Thanks to science and technology, access to factual knowledge of all kinds is rising exponentially while dropping in unit cost. . . . We are drowning in information, while starving for wisdom.—*E. O. Wilson, (1998)*

Information Foraging Theory emerges from a serious consideration of Miller's notion of *informavores*. A serious consideration of the concept leads to questions regarding the adaptive forces that drive human interaction with information. Simon (1979) articulated the basic design problem facing us: "What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it" (1).

According to statistics compiled by the University of California–Berkeley School of Information Science (Lyman & Varian, 2003), almost 800 megabytes of recorded information are produced per person per year, averaged over the estimated 6.3 billion people in the world. This is the equivalent of about 30 linear feet of books. In an information-rich world, the real design problem to be solved is not so much how to collect and distribute more information but rather how to increase the rate at which persons can find and attend to information that is truly of value to them.

The Principle of the Extremization of Information Utility as a Function of Interaction Cost

An investment in knowledge always pays the best interest.—*Benjamin Franklin*

In modern society, people interact with information through technology that more or less helps them find and use the right knowledge at the right time. In evolutionary terms, one can argue that increasing the rate of gain of valuable information increases fitness. As Sir Francis Bacon observed, "knowledge is power." Power (control over the world to achieve one's goals) can be improved by better knowledge, or lower costs of access and application of knowledge. In evolu-

tionary terms, an agent's fitness is improved to the extent that it can predict and control the environment in order to solve the problems it faces in everyday life. In psychological terms, increasing the rate at which people can find, make sense of, and use valuable information improves the human capacity to behave intelligently. We should expect adaptive systems to evolve toward states that maximize gains of valuable information per unit cost (Resnikoff, 1989, p. 97). A useful way of thinking about such adaptation is to say that

Human–information interaction systems will tend to maximize the value of external knowledge gained relative to the cost of interaction.

Schematically, we may characterize this maximization tendency⁶ as

$$\max \left[\frac{\text{Expected value of knowledge gained}}{\text{Cost of interaction}} \right]. \quad (1.3)$$

Cognitive systems engaged in information foraging will exhibit such adaptive tendencies, and they will prefer technologies that tend to maximize the value (or utility) of knowledge gained per unit cost of interaction. For instance, sensory systems appear to evolve in ways that deliver more bits of information for the amount of calories expended. Similarly, offices, with their seeming chaotic mess of piles of papers, books, and files, appear to become organized in ways that optimize access costs of frequently needed information (Case, 1991; Malone, 1983; Soper, 1976). Resnikoff (1989, pp. 112–117) presented a mathematical analysis suggesting that physical library catalog card systems would become arranged in ways that minimized manual search time. Information Foraging Theory assumes that people prefer information-seeking strategies that yield more useful information per unit cost. People tend to arrange their environments (physical or virtual) to optimize this rate of gain. People prefer, and consequently select, technology designs that improve returns on information foraging.

The Exaptation of Food Foraging Mechanisms

Natural selection favored organisms—including our human ancestors—that had better mechanisms for extracting energy from the environment and translating that energy into reproductive success. Organisms

¹[AQ: pls cite pg no for quote]

14 INFORMATION FORAGING THEORY

with better food-foraging strategies (for their particular environment) were favored by natural selection. Our ancestors evolved perceptual and cognitive mechanisms and strategies that were very well adapted to the task of exploring the environment and finding and gathering food. Information Foraging Theory assumes that modern-day information foragers use perceptual and cognitive mechanisms that carry over from the evolution of food-foraging adaptations.

If information foraging is like food foraging, then models of optimal foraging developed in the study of animal behavior (Stephens & Krebs, 1986) and anthropology (Winterhalder & Smith, 1992a) should be relevant. Figure 1.5 presents the conventional patch model and Charnov's Marginal Value Theorem as a possible analog for information foraging at a Web site. A typical optimal foraging model characterizes an agent's interaction with the environment as an optimal solution to the tradeoff of costs of finding, choosing, and handling food against the energetic benefit gained from that food. These models would look very familiar to an engineer because they are basically an attempt to understand the design of an agent's behavior by assuming that it is well engineered (adapted) for the problems posed by the environment. Information foraging models include optimality analyses of different information-seeking strategies and technologies as a way of understanding the design rationale for user strategies and interaction technologies.

Optimal foraging theorists assume that energy, originating predominantly from the sun, seeps through the food chain to be deposited in various plants and animals that are distributed variably through the environment. Food foragers may have different mechanisms and strategies available to them for navigating through the environment. Their potential sources of food may have different prevalences in different habitats and may have different profitabilities in terms of how many calories can be extracted when foraged. The optimal forager is one who has the strategies, mechanisms, diets, and so forth, that maximize the calories gained per unit of effort expended.⁷ Similarly, Information Foraging Theory assumes that information comes to be stored in various prevalences in different kinds of repositories, in various forms and media. The information forager has different means available for navigating and searching the information environment, and different information sources have different profitabilities in terms of the interaction cost

required to gain useful information. As suggested by equation 1.3, the optimal information forager is one who maximizes the value of knowledge gained per unit cost of interaction.

Application to Human–Information Interaction

The legacy of the Enlightenment is the belief that entirely on our own we can know, and in knowing, understand, and in understanding, choose wisely.—E. O. Wilson, (1998)

Human–information interaction (HII) is a nascent field that is concerned with how people interact with, and process, outwardly accessible information in service of their goals.⁸ It adopts an information-centric approach rather than the computer-centric approach of the field of human–computer interaction (HCI) (Lucas, 2000). This shift to an information-centric focus is a natural evolution for the field of HCI because of the increasing *pervasiveness* of information services, the increasing *transparency* of user interfaces, the *convergence* of information delivery technologies, and the trend toward *ubiquitous computing*.

Access to the Internet is pervasive in the developed world through land lines, satellite, cable, and mobile devices. The field of HCI, over the past two decades and more, has led to the development of computers and computer applications that are transparent to users performing their tasks. In parallel, the business world around consumer media technologies shows excitement over the convergence of television, cell phones, personal computers, PDAs (personal digital assistants), cars, set-tops, and other consumer electronic devices, as well as the convergence among the means for transporting information, such as the Internet, radio, satellite, and cable. Research on ubiquitous computing looks forward to a world in which computational devices are basically everywhere in our homes, mobile devices, cars, and so on, and these devices can be marshaled to perform arbitrary tasks for users. The net effect of these trends is to make computers invisible, just as electricity and electric motors are invisible in homes today (Lucas, 2000). As computers become invisible, and information becomes ample and pervasive, we expect to see a shift in studies from HCI to HII. Rather than focus on the structure of devices and application programs, the focus of HII research must center on content and interactive media.

Information Foraging Theory arose during the 1990s, coinciding with an explosion in the amount of information that became available to the average computer user and with the development of new technologies for accessing and interacting with information. The late 1980s witnessed several strands of HCI research that were devoted to ameliorating problems of exploring and finding electronically stored information. It had become apparent that users could no longer remember the names of all their electronic files, and it was even more difficult for them to guess the names of files stored by others (Furnas, Landauer, Gomez, & Dumais, 1987). One can see proposals in the mid- to late 1980s HCI literature for methods to enhance users' ability to search and explore external memory. Jones (1986) proposed the Memory Extender (ME), which used a model of human associative memory (Anderson, 1983) to automatically retrieve files represented by sets of keywords that were similar to the sets of keywords representing the users' working context. Latent Semantic Analysis (LSA; Dumais, Furnas, Landauer, Deerwester, & Harshman, 1988) was developed to mimic human ability to detect deeper semantic associations among words, such as "dog" and "cat," to similarly enhance information retrieval. Interestingly, the work on ME and LSA was contrasted with work in the "traditional" field of information retrieval in computer science, which had a relatively long history of developing automated systems for storing and retrieving text documents. The CHI '88 conference where LSA was introduced also hosted a panel bemoaning the fact that automated information retrieval systems had not progressed to the stage where anyone but dedicated experts could operate them (Borgman, Belkin, Croft, Lesk, & Landauer, 1988). Such systems, however, were the direct ancestors of modern search engines found on the World Wide Web.

Hypermedia also became a hot topic during the late 1980s, with Apple's introduction of HyperCard in 1987, the first ACM Conference on Hypertext in 1987, and a paper session at the CHI '88 conference. The very idea of hypertext can be traced back to Vannevar Bush's *Atlantic Monthly* article, "As We May Think," published in 1945. Worried about scholars becoming overwhelmed by the amount of information being published, Bush proposed a mechanized private file system, called the Memex, that would augment the memory of the individual user. It was explicitly intended to mimic human associative

memory. Bush's article influenced the development of Douglas Engelbart's NLS (oNLine System), which was introduced to the world in a tour-de-force demonstration at the 1968 Fall Joint Computer Conference. The demonstration of NLS—a system explicitly designed to "augment human intellect" (Engelbart, 1962)—also introduced the world to the power of networking, the mouse, and point-and-click interaction. Hypertext and hypermedia research arose during the late 1980s because personal computing power, networking, and user interfaces had evolved to the point where the visions of Bush and Engelbart could finally be realized for the average computer user.

The confluence of increased computing power, storage, and networking and information access and hypermedia research in the late 1980s set the stage for the widespread deployment of hypermedia in the form of the World Wide Web. In 1989, Tim Berners-Lee (1989) proposed a solution to the problems that were being faced by the CERN community in dealing with distributed collections of documents, which were stored on many types of platforms, in many types of formats. This proposal led directly to the development of HTML, HTTP, and, in 1990, the release of the World Wide Web. Berners-Lee's vision was not only to provide users with more effective access to information but also to initiate an evolving web of information that reflected and enhanced the community and its activities.

The emergence of the Web in the 1990s provided new challenges and opportunities for HCI. The increased wealth of accessible content, and the use of the Web as a place to do business, exacerbated the need to improve the user experience on the Web.

The usability literature that has evolved surrounding the Web user experience is incredibly rich with design principles and maxims (Nielsen, 2000; Spool, Scanlon, Schroeder, Snyder, & DeAngelo, 1999), the most important of which is to test designs with users. Much of this literature is based on a mix of empirical findings and expert ("guru") opinion. A good deal of it is conflicting. The development of theory in this area can greatly accelerate progress and meet the demands of changes in the way we interact with the Web. Greater theoretical understanding and the ability to predict the effects of alternative designs could bring greater coherence to the usability literature and provide more rapid evolution of better designs. In practical terms, a designer armed with such theory could explore and explain the effects of

different design decisions on Web designs before the heavy investment of resources for implementation and testing. This exploration of design space is also more efficient because the choices among different design alternatives are better informed: Rather than randomly generating and testing design alternatives, the designer is in a position to know which avenues are better to explore and which are better to ignore. Unfortunately, cognitive engineering models that have been developed to deal with the analysis of expert performance on well-defined tasks involving application programs (Pirolli, 1999) have little applicability to understanding foraging through content-rich hypermedia, and consequently new theories are needed.

Methodological Adaptationism

Adaptationist reasoning is not optional; it is the heart and soul of evolutionary biology.—D. C. Dennett (1992)^k

^k[AQ: pls add title of work this is from (1992 is not in ref list)]

The concept of *informavores*, and concern with the application domain of HII, leads us to reconsider the dominance of strictly mechanistic analyses of HCI. Miller, in his 1983 article about “*informavores*,” commented on the incompleteness of the mechanistic approach by using the following analogy:

Insofar as a limb is a lever, the theory of levers describes its behavior—but a theory of levers does not answer every question that might be asked about the structure and function of the limbs of animals. Insofar as the mind is used to process information, the theory of information processing describes its behavior—but a theory of information processing does not answer every question that might be asked about the structure and function of the minds of human beings. (p. 112)

Information processing (mechanistic) analyses of HCI—by themselves—give only partial explanations. They provide mechanistic explanations of the “levers” of the mind. In reaction to this inadequacy, Information Foraging Theory has been guided by the heuristics and explanatory framework of *methodological adaptationism*, and the specific version of it developed by Anderson (1990) called *rational analysis* (see also Oaksford & Chater, 1998). The illustration above concerning hotel prices on the Web involved a very simple rational analysis. Methodological adaptationism presumes that it is a good

heuristic for scientists to assume that evolving, behaving systems are rational, or well designed, for fulfilling certain functions in certain environments. There is an assumption of *ecological rationality* regarding the behavior of the system being observed (Bechtel, 1985; Dennett, 1983, 1988, 1995; Gigerenzer, 2000). The adaptationist approach involves a kind of reverse engineering in which the analyst asks (a) *what* environmental problem is solved, (b) *why* is a given system a good solution to the problem, and (c) *how* is that solution realized (approximated) by mechanism.

Versions of methodological adaptationism have shaped research programs in behavioral ecology (e.g., Mayr, 1983; Stephens & Krebs, 1986; Tinbergen, 1963), anthropology (e.g., Winterhalder & Smith, 1992a), and neuroscience (e.g., Glimcher, 2003). The approach gained currency in cognitive science during the 1980s as a reaction to ad hoc models of how people performed complex cognitive or perceptual tasks. At that time, models of cognition and perception were generally mechanistic, detailing perceptual and cognitive structures and the processes that transformed them. The Model Human Processor (MHP) and GOMS^l (Card, Moran, & Newell, 1983) are cognitive engineering examples in the field of HCI that derive from this approach. The MHP specifies a basic set of information storage and processing machinery, much like a specification of the basic computer architecture for a personal computer. GOMS specifies basic task performance processes, much like a mechanical program that “runs” on the MHP.

Around the same time that GOMS and MHP were introduced into HCI, there emerged a concern among cognitive scientists that mechanistic information processing models, by themselves, were not enough to understand the human mind (Anderson, 1990; Marr, 1982). A major worry was that mechanistic models of cognition had been developed in an ad hoc way and provided an incomplete explanation of human behavior. It had become common practice to cobble together a program that simulated human performance on some task and then claim that the program was in fact a theory of the task (Marr, 1982, p. 28). Anderson (1990) lamented that cognitive modelers “pull out of an infinite grab bag of mechanisms bizarre creations whose only justification is that they predict the phenomena in a class of experiments. . . . We almost never ask the question of

^l[AQ: pls define GOMS here]

why these mechanisms compute the way they do” (p. 7, emphasis added).

Figuring out a mechanistic account of human behavior—for instance, with MHP analysis—is no small feat. However, as the Miller quote above suggests, such accounts do not explain everything. The mind is not just any old arbitrary, cobbled-together machine; rather, it is a fantastically complex machine that has been designed by evolution to be well tailored to the demands of surviving and reproducing in the environment. The adaptationist approach recognizes that one can better understand a machine by understanding its *function*. By this I mean both that (a) adaptationist accounts make more sense and (b) the search for better understanding proceeds at a faster pace.

Levels of Explanation

The analysis of people interacting with information involves interrelated layers of explanation. This is because scientific models in this area assume that human activity is (a) purposeful and adaptive, which requires a kind of rational analysis, (b) based on knowledge, (c) computed by information processing mechanisms, which are (d) realized by physical, biological, processes. Figure 1.6 presents a summary of the relevant framework that has emerged in the behavioral sciences (see, e.g., Anderson, 1990; Cosmides, Tooby, & Barow, 1992; Gigerenzer, 2000; Winterhalder & Smith, 1992a).

Rational analysis, in the case of Information Foraging Theory, focuses on the task environment that is the aim of performance, the information environment that structures access to valuable knowledge, and the adaptive fit of the HII system to the demands of these environments. Rational analysis assumes that the structure of behavior can be understood in terms of its adaptive fit to the structure and constraints of the environment. The analysis of searching for hotel prices on Web involved a rational analysis of the expected savings to be gained from information search and an analysis of the rational choice to make when faced with decisions of whether to continue search or to give up. When performing a rational analysis the theorist may be said to take a *design stance* (Dennett, 1995) that focuses on an analysis of the functionality of the system with respect to its ostensive purpose. At this level, the analyst acts

most purely as an engineer concerned with why users’ behavior is rational given the task context in which it occurs, and it is assumed that users are optimizing their performance in achieving their goals.

Knowledge-level analysis concerns the knowledge content involved in achieving goals. Knowledge-level analysis involves descriptions of a system in intentional terms with the assumption that behavior is the product of purposes, preferences, and knowledge. The knowledge level has been important in artificial intelligence since its introduction by Newell (1982). A knowledge-level analysis of the task of searching for hotel prices on the Web was a prerequisite to the specification of the production rules and chunks involved in the cognitive simulation. Dennett (1988) defined an observer who describes a system using an intentional vocabulary (e.g., “know,” “believe,” “think”) as one taking an *intentional stance*. Typically, a task analysis focuses mainly on an analysis of users’ knowledge, preferences, perceptions, and actions, with respect to the goal and environment. At this level of analysis, it is assumed that users deploy their knowledge to achieve their goals, and the focus is on identifying what knowledge is involved.

Modern cognitive psychology assumes that the knowledge level can be given a scientific account (i.e., be made predictable) by explaining it in terms of mechanistic information processing (Newell, 1990). This is the *cognitive level* of explanation. This level of analysis focuses on the properties of the information processing machinery that evolution has dealt to humans to perceive, think, remember, learn, and act in what we would call purposeful and knowledgeable ways. This is the level of most traditional theorizing in cognitive psychology and HCI—the level at which computational models may, in principle, be developed to simulate human cognition. GOMS (Card et al., 1983), described above, is an example of an analysis method aimed at cognitive-level analysis. Cognitive architectures such as ACT-R (Anderson et al., 2004) or Soar (Newell, 1990) and the simulations developed in those architectures are developed at the cognitive level. The production system specified in table 1.1 was a simple example of a cognitive-level analysis.

Accounts at the cognitive level are assumed to be instantiated at the *biological level* by the physical machinery of the brain and body. The biological level of explanation specifies the proximal physical

mechanisms underlying behavior. For instance, Anderson et al. (2004) have recently presented results suggesting the mapping of the ACT-R architecture onto neural structure and functioning.

Phenomena at Different Time Scales of Behavioral Analysis

Many of our goals can drive our behavior for days, months, and even years. These longer term goals are typically realized by task structures composed of many shorter term goals. Card et al. (1983) suggested that there is a base level of tasks, called the *unit task level*, that controls immediate behavior. Unit tasks empirically take about 10 seconds. To an approximation, unit tasks are where “the rational rubber meets the mechanistic road.” To an approximation, the structure of behavior above the unit task level largely reflects a rational structuring of the task within the constraints of the environment, whereas the structure within and below the unit task level reflects cognitive and biological mechanisms. Phenomena occur at multiple grain sizes of time, and effects propagate in both upward and downward directions: Rational/ecological structuring goes downward from longer time scales of phenomena, and environment and proximal mechanism constraints go upward. A significant claim of the framework adopted by Information Foraging Theory from Newell (1990) and Anderson (2002) is that the phenomena of human cognition can be decomposed and modeled at many different time scales.

Newell (Newell, 1990; Newell & Card, 1985) argued that human behavior arises from a hierarchically organized system in which the basic time scale of operation of each system level increases by a factor of 10 as one moves up the hierarchy (table 1.2). The phenomena at each band in table 1.2 are largely dominated by different kinds of factors. Behavioral analysis at the *biological band* (approximately milliseconds to tens of milliseconds) is dominated by biochemical, biophysical, and especially neural processes, such as the time it takes for a neuron to fire. The *psychological band* of activity (approximately hundreds of milliseconds to tens of seconds) has been the main preoccupation of cognitive psychology (Anderson, 1983, 1993; Newell, 1990). At this time scale, it is assumed that elementary cognitive mechanisms play a major part in shaping behavior. The typical unit of analysis is a single response function,

TABLE 1.2 Time scale on which human action occurs.

| Scale (seconds) | Time Unit | Band |
|--------------------|------------------|------------|
| 10^7 | Months | Social |
| 10^6 | Weeks | |
| 10^5 | Days | |
| 10^4 | Hours | Rational |
| 10^3 | 10 minutes | |
| 10^2 | Minutes | |
| 10^1 | 10 seconds | Cognitive |
| 10^0 | 1 second | |
| 10^{-1} | 100 milliseconds | |
| 10^{-2} | 1 millisecond | Biological |

Different bands are quite different phenomenological worlds.

Adapted from Newell (1990, p. 122).

involving a perceptual input stage, a cognitive stage, and a stage of action output—for instance, finding a word in the menu of a text editor and moving a mouse to select the menu item. The mechanisms involved at this level of analysis include elementary information processing functions such as memory storage and retrieval, recognition, categorization, comparison of one information element to another, and choosing among alternative actions.

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, 1990, p. 150). This is the *rational band* of phenomena (minutes to days). The typical unit of analysis at this level is the *task*, which is defined, in part, by a *goal*. It is assumed that an intelligent agent will have *preferences* for *actions* that it *perceives* to be applicable in its *environment* and that it *knows* will move the current situation toward the goal.^m So, on the one hand, goals, knowledge, perceptions, actions, and preferences shape behavior. On the other hand, the structure, constraints, and resources of the environment in which the task takes place—called the *task environment* (Newell & Simon, 1972)—will also greatly shape behavior. Explanations at the rational band assume that behavior is governed by rational principles and that it is largely shaped by the structure and constraints of the task environment, although it is also realized that people are not infinitely and perfectly rational (Simon, 1955). The rationale for behavior at this level is its adaptive fit to its task environment.

^m[AQ: pls confirm that each of these italicized terms are being introduced here with specific definitions that differ from those given above; if not, pls remove italics]

Task Environments and Information Environments

To understand information foraging requires analysis of the environment in addition to analysis of the forager. The importance of the analysis of the environment to psychology was a more general point made by Brunswik (1952) and Simon (1981). It is useful to think of two interrelated environments in which an information forager operates: the *task environment* and the *information environment*. The classical definition of the task environment is that it “refers to an environment coupled with a goal, problem or task—the one for which the motivation of the subject is assumed. It is the task that defines a point of view about the environment, and that, in fact allows an environment to be delimited” (Newell & Simon, 1972, p. 55). The task environment is the scientist’s analysis of those aspects of the physical, social, virtual, and cognitive environments that drive human behavior.

The information environment is a tributary of knowledge that permits people to more adaptively engage their task environments. Most of the tasks that we identify as significant problems in our everyday life require that we get more knowledge—become better informed—before taking action. What we know, or do not know, affects how well we function in the important task environments that we face in life. External content provides the means for expanding and improving our abilities. The information environment, in turn, structures our interactions with this content. Our particular analytic viewpoint on the information environment will be determined by the information needs that arise from the embedding task environment. From the standpoint of a psychological analysis, the information environment is delimited and defined in relation to the task environment.

Problem Spaces

A large class of tasks may be understood as variations on problem solving. Indeed, Newell (1990) essentially argued that all of cognition could be understood by taking this stance. Newell and Simon (1972) characterized problem solving formally as a process of *search* through a *problem space*. A problem space consists of an initial situation called the *start state* and some desired situation called the *goal state*. Other situations that may occur while solving the problem are *intermediate states*. Problem-solving *operators* (e.g., actions

performed by the problem solver) transform problem states. For instance, the problem faced by a toddler seeking to eat cookies from a cupboard may have an initial state that consists of the child standing on the floor and a chair some distance away, and the child may apply problem-solving operators such as moving the chair, climbing on the chair, and opening the cupboard to transform the initial state toward the goal state. The various states that can be achieved are referred to as a problem space (or sometimes a *state space*). Often, any given problem state is a situation that affords many possible actions (operators). In such cases, each state branches to many possible subsequent states, with each branch in each path corresponding to the application of an operator. The problem is to find some path through the maze of possible states. Finding this path is a process of search through a problem space.

Ill-Structured Problems and Knowledge Search

Well-structured problems, such as puzzles and games, have well-defined initial states, goal states, operators, and other problem constraints, which contrasts with the ill-structured problems. Ill-structured problems, such as choosing a medical treatment or buying a house, typically require additional knowledge from external sources in order to better understand the starting state, to better define a goal, or to specify the actions that are afforded at any given state (Simon, 1973). People typically need to perform *knowledge search* in order to solve their ill-structured problems (e.g., to define aspects of a problem space that permit effective or efficient problem space search). The information environment is a potential source of valuable knowledge that can improve our ability to achieve our goals, especially when they involve ill-structured tasks. More generally, knowledge shapes human functionality, and consequently external access to large volumes of widely variegated knowledge may improve our range of adaptation because we can solve more problems, or solve problems using better approaches.

Knowledge-Level Systems

Knowledge, if it does not determine action, is dead to us.—Plotinus

Externally available content provides us with knowledge valuable to the achievement of our goals. Given

the central role of external knowledge to Information Foraging Theory, it is useful to review Newell's (1982) influential framework for the study of knowledge systems. This provides a way of characterizing adaptation in terms of knowledge content. This framework, which arises from the cognitive sciences, assumes that knowledge shapes the functionality of our cognitive abilities and that intelligent behavior depends on finding and using the right knowledge at the right time. This framework was largely articulated by Allen Newell (1982, 1990, 1993) and Daniel Dennett (1988, 1991). Traditionally (e.g., Dennett, 1988; Newell, 1990), the information processing system under consideration for analysis is an unaided person or computer program working in some task environment. However, we can extend the approach to understand a system that consists of a person tightly coupled with technological support and access to a teeming world of information.

Over the course of 20 years, Newell (Moore & Newell, 1973; Newell, 1982, 1990; Newell et al., 1992) developed a set of ideas about understanding how physical systems could be scientifically characterized as knowledge systems. A parallel set of ideas was developed by Dennett (1988) in his discussion of intentional systems.⁹ The notions developed by Newell and Dennett derive from the philosophical contributions of Brentano (1874/1973).¹⁰ⁿ The knowledge level was developed by Newell (1982) as a way to address questions about the nature of knowledge and the nature of scientifically ascribing knowledge to an agent.

In the frame of reference developed by Newell and Dennett, scientific observers ascribe knowledge to behaving systems. A key assumption is that knowledge-level systems can be specified completely by reference to their interaction with the external world, without reference to the mechanical means by which the interactions take place. A knowledge-level system consists of an *agent* behaving in an *environment*. The agent consists of a set of *actions*, a set of *perceptual devices*, a *goal* (of the agent), and a body of *knowledge*. The operation of such systems is governed by the *principle of rationality*: If the agent knows that one of its actions will lead to a situation preferred according to its goal, then it will *intend* the action, which will then be taken if it is possible. As Newell (1982) stated, *knowledge* is “whatever can be ascribed to an agent, such that its behavior can be computed according to

the principle of rationality” (p. 105). In essence, the basic observations at the knowledge level are statements of the form:

In situation *S*, agent *A* behaves as if it has knowledge *K*.

Value and Structure of Knowledge

New knowledge is the most valuable commodity on earth. The more truth we have to work with, the richer we become.—Kurt Vonnegut, *Breakfast of Champions*

Our ability to solve ill-structured problems such as buying a house, finding a job, or throwing a Super Bowl party is, in large part, a reflection of the particular external knowledge used to structure and solve the problem. Consequently, the value of external content may often ultimately be measured in the improvements to the outcomes of an embedding task. The value of knowledge gained may be measured in terms of what additional value it attains for the agent. Of course, a lot of external content provides no new knowledge (e.g., perhaps it is “old news” to us), or information that does not contribute to our goals.

In simple well-structured problems, the value of knowledge gained from information foraging can be generally expressed as a difference between two strategies: one that rationally uses knowledge acquired by foraging from external information sources to choose among outcomes, and another that does not use such information.¹¹ For instance, suppose a man who has a budget wants to purchase a product on the Web and knows of a price comparison Web site (e.g., as in the hotel illustration above). If blindly purchasing a product costs a certain expected amount X , but after visiting the price comparison Web site the man will be able to find a less expensive product Y , then the net value of that knowledge will be $X - Y - C$, where C is some measure of the cost of gaining the knowledge. If the analysis in the hotel price illustration above were correct, then the expected price of a hotel (without knowledge) would have been about \$86 (see the appendix), but after looking at a Web site, the price would have been \$66, and the time cost would be approximately $13 \text{ min}/60 \text{ min} \times \$10/\text{hr} = \$2$, so the value of the Web site knowledge would be $\$86 - \$66 - \$2 = \18 . In simple cases such as these, one may imagine that a person could completely construct a decision model in which all possible decision

ⁿ[AQ: pls confirm that note 10 is appropriately placed here - it does not address Brentano]

outcomes are specified, as well as the relationships among information sources, potential results from those sources, and the relation of information results gathered to decisions and the utility of those decisions. Indeed, artificial intelligence systems (e.g., Grass & Zilberstein, 2000) have been developed to use this approach to tackle problems such as purchasing a digital camera, purchasing a removable media device, or choosing a restaurant. Real-world problems, however, typically require a more complicated analysis of the value of knowledge.

Knowledge and Intelligence

Knowledge is of two kinds: we know a subject ourselves, or we know where we can find information upon it.
—Samuel Johnson

Physically instantiated cognitive systems are limited in their ability to behave as rational knowledge-level systems. Newell (1990) proposed that “intelligence is the ability to bring to bear all the knowledge that one has in service of one’s goals” (p. 90).¹² This corresponds to our everyday notion that we can behave more intelligently by being better informed. In the idealized view of the knowledge level, everything in a body of knowledge (including all possible entailments) is instantly accessible. However, people, or any physical system, can only approximate such perfect intelligent use of knowledge because the ability to bring forth the right knowledge at the right time is physically limited. The laws of physics limit the amount of information that can be stored or processed in a circumscribed portion of space and time. Within those limits, however, intelligence increases with the ability to bring to bear the right knowledge at the right time.

Dennett (1991, pp. 222–223) notes that this conception of knowledge and intelligent reasoning goes back to Plato (*Theaetetus*, 197–198a, Cornford translation^o). Plato saw knowledge as something that one could possess like a man who keeps captured wild birds in an aviary. There is a sense in which the man has the birds, but a sense in which he has none of them until he can control each bird by calling forth the bird at will. Plato saw intelligent reasoning as not only having the birds but also having the control to bring forth the right bird at the right time.

Newell’s discussions focused on unaided intelligent systems (people or computer programs) and the

knowledge that they had available in their local memories. But there is a sense in which the world around us provides a vast external memory teeming with knowledge that can be brought forth to remedy a lack on the part of the individual. We can extend Newell’s notion of intelligence and argue that intelligence is improved by enhancement of our ability to bring forth the right knowledge at the right time from the external world. Of course, the world (both physical and virtual) shapes the manner in which we can access and transform knowledge-bearing content and thus shapes the degree to which we reason and behave intelligently. The task of acquiring knowledge from external sources is itself a task that can be performed more or less intelligently.

Consider the illustration above in which a hypothetical user searches for hotel prices on the Web. From a knowledge-level perspective, the user has knowledge of how to navigate the Web, operate the Web site search engine, and perform price comparisons. The illustration assumed that the user applies this knowledge flawlessly, but the structure of the Web environment determines the rate at which new knowledge (of hotel prices) is gained. A different design could improve the rate at which the user accomplishes the task. For instance, if the Web site sorted hotels by both quality (star rating) and price, the user could accomplish the task much faster. Although the user’s navigation and calculation knowledge has not changed, it is being applied more efficiently because of a change in the information environment. In other words, a change in the information environment has made the user more intelligent.

Rational Analysis

Anderson’s rational analysis approach is a specific version of methodological adaptationism applied to the development of cognitive theory. It was inspired by Marr’s (1982) influential approach to computer vision, in which Marr argued that visual processing algorithms (and other intelligent information processes) are “likely understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is solved” (p. 92).¹³ The term “rational analysis” was inspired by rational choice theory in economics, in which people are assumed to be

^o[AQ: pls add complete entry to ref list and cite name/date here]

22 INFORMATION FORAGING THEORY

rational decision makers who optimize their behavioral choices in order to maximize their goals (utility). In rational analysis, however, it is not the person who is the agent of rational choice, but rather it is the selective forces of the environment that choose better biological and behavioral designs.

Anderson has used rational analysis to study the human cognitive architecture by assuming that natural information processing mechanisms involved in such functions as memory (Anderson & Milson, 1989; Anderson & Schooler, 1991) and categorization (Anderson, 1991) were well designed by evolutionary forces to meet the problems posed by the environment. The key assumption behind rational analysis could be stated as

Principle of rationality: The cognitive system optimizes the adaptation of the behavior of the organism.

As developed by Anderson (1990), rational analysis requires a focus on understanding the structure and dynamics of the environment. This understanding provides a rationale for the design of information processing mechanisms. Anderson proposed the following recipe for rational analysis:

1. Precisely specify the goals of the agent.
2. Develop a formal model of the environment to which the agent is adapted.
3. Make minimal assumptions about the computational costs.
4. Derive the optimal behavior of the agent considering items 1–3.
5. Test the optimality predictions against data.
6. Iterate.

Note, generally, the emphasized focus on optimal behavior under given goals and environmental constraints and the minimal assumptions about the computational structure that might produce such behavior.

Probabilistically Textured Environments

Interaction with the information environment differs in a fundamental way from well-defined task environments that have been the dominant paradigms in HCI, such as expert text editing (Card et al., 1983) or telephone assistance (Gray et al., 1993). In contrast to

such tasks—in all but the most trivial cases—the information forager must deal with a *probabilistically textured* information environment (Brunswik, 1952). In contrast to application programs such as text editors and spreadsheets, in which actions have fairly determinate outcomes,¹⁴ foraging through a large volume of information involves uncertainties—for a variety of reasons—about the location, quality, relevance, veracity, and so on, of the information sought and the effects of foraging actions. The ecological rationality of information foraging behavior must be analyzed through the theoretical lens and tools appropriate to *decision making under uncertainty*. The determinate formalisms and determinate cognitive mechanisms that are characteristic of the HCI paradigm are inadequate for the job of theorizing about information foraging in probabilistically textured environments. Models developed in Information Foraging Theory draw upon probabilistic models, and especially Bayesian approaches, and they bear similarity to economic models of decision making (rational choice) under uncertainty and engineering models.

Role of Optimization Analysis

Optimization models¹⁵ are a powerful tool for studying the design features of organisms and artifacts. Consequently, optimization models are often found in the toolbox of the methodological adaptationist (e.g., as found in Anderson's rational analyses). Optimization models are mathematical models borrowed from engineering and economics. They are used to model a rational decision process faced with a problem and constraints. In engineering, they are used as a tool for quantifying the quality of design alternatives with respect to some problem specification. In economics, they are used typically to characterize a rational decision maker choosing among courses of action in order to maximize utility (*a rational choice model*), often operating in situations of limited or uncertain knowledge about possible outcomes. Optimization models in general include the following three major components:

- *Decision assumptions* that specify the decision problem to be analyzed, such as the amount of time to spend on an activity, or whether or not to pursue a particular type of information content.

- *Currency assumptions*, which identify how choices are to be evaluated, such as time or money or other resources.
- *Constraint assumptions*, which limit and define the relationships among decision and currency variables. Examples of constraints include the rate at which a person can navigate through an information access interface, or the value of results returned by bibliographic search technology.

All cognitive agents must reason about the world with limited time, knowledge, and computational power. Consequently, the use of optimization models cannot be taken as a hypothesis that human behavior is omnisciently rational, with perfect information and infinite computational resources. Indeed, unbounded optimization models are likely to fail in predicting any complex behavior. Anderson's (1990) rational analysis approach is based on optimization under constraints. The basic idea is that the constraints of the environment place important shaping limits on the optimization that is possible.

Optimization models, such as rational choice models from economics, allow us to define the behavioral problems that are posed by the environment, and they allow us to determine how well humans (or animals or other cognitive agents) perform on those problems. This does not mean that one assumes that the cognitive agent is performing the same calculations as the optimization models. It is possible that simple mechanisms and heuristics may achieve optimal or near optimal performance once the limits of the environment are taken into account (Todd & Gigerenzer, 1999). This is the essence of *bounded rationality* and the notion that real cognitive agents make choices based on *satisficing* (Simon, 1955). In the analysis of adaptation, optimization models are^P

Generally, "One does not treat the optimization principle as a formula to be applied blindly to any arbitrarily selected attribute of an organism. It is normally brought in as a way of expanding our understanding from an often considerable base of knowledge" (Williams, 1992, p. 62). As eloquently stated by the evolutionary theorist G. C. Williams (1992),

Organisms are never optimally designed. Designs of organs, developmental programs, etc. are legacies from past and natural selection can affect them in only two ways. It can adjust the numbers of mutually exclusive designs until they reach

frequency-dependent equilibria, often with only one design that excludes alternatives. It can also optimize a design's parameters so as to maximize the fitness attainable with that design under current conditions. This is what is usually meant by optimization in biology. An analogy might be the common wooden-handled, steel-bladed tool design. With different parameter values it could be a knife, a screw driver, or many other kinds of tool—*many*, but not *all*. The fixed-blade constraint would rule out turning it into a drill with meshing gears. The wood-and-steel constraint would rule it out as a hand lens. (p. 56, emphasis original)

Activities can be analyzed according to the value of the resource currency returned and costs incurred. Generally, one considers two types of costs: (1) *resource costs* and (2) *opportunity costs* (Hames, 1992). Resource costs are the expenditures of calories, money, and so forth, that are incurred by the chosen activity. Opportunity costs are the benefits that could be gained by engaging in other activities but are forfeited by engaging in the chosen activity. For instance, junk mail incurs a resource cost in terms of the amount of money (not to mention trees) involved in delivering the junk, but it also incurs an opportunity cost for the recipients who read the junk because they have forgone gains that could have been made by engaging in other activities.

Production System Theories of Cognition

Production systems have had a successful history in psychology (Anderson et al., 2004; Neches, Langley, & Klahr, 1987) since their introduction into the field by Newell (1973a). The ACT family of production system theories has the longest history of these kinds of cognitive architectures. The seminal version of the ACT theory was presented in Anderson (1976), shortly after Newell's (1973b) challenge to the field of cognitive psychology to build unified theories of cognition, and it has undergone several major revisions since then (Anderson, 1976, 1983, 1990, 1993; Anderson et al., 2004; Anderson & Lebiere, 1998). Until recently, it has been primarily a theory of higher cognition and learning, without the kind of emphasis on perceptual-motor processing found in EPIC (Kieras & Meyer, 1997) or MHP (Card et al., 1983). The success of ACT as a cognitive theory has been historically in the study of memory (Anderson

^P[AQ: pls complete this paragraph]

& Milson, 1989; Anderson & Pirolli, 1984), language (Anderson, 1976), problem solving (Anderson, 1993), and categorization (Anderson, 1991). As a learning theory, ACT has been successful (Anderson, 1993) in modeling the acquisition of complex cognitive skills for tasks such as computer programming, geometry, and algebra and in understanding transfer of learning across tasks (Singley & Anderson, 1989^a). ACT has been strongly tested (Anderson, Boyle, Corbett, & Lewis, 1990^b) by application in the development of computer tutors, and less so in the area of HCI. The production system models presented in this book are extensions of the ACT theory.

Figure 1.8 presents the basic cognitive architecture used in this book. It couples the basic ACT-R architecture to a module that computes *information scent*,^s which for convenience I will call the ACT-Scent¹⁶ architecture. This book presents specific models of Web foraging (SNIF-ACT 1.0 and SNIF-ACT 2.0) and Scatter/Gather (Cutting, Karger, Pedersen, & Tukey, 1992) browsing (ACT-IF) that were developed within the ACT-Scent architecture. The architecture includes a declarative memory containing chunks, a procedural

memory containing production rules, and a goal stack containing the hierarchy of intentions driving behavior. The information scent module is a new addition to ACT that is used to compute the utility of actions based on an analysis of the relationship of content cues from the user interface to the user's goals. The theory behind this module is described in detail in chapter 4.

Summary

Humans are informavores. We adapt to the world by seeking and using information. As a result, we create a glut of information. This causes a poverty of attention and a greater need to allocate that attention effectively and efficiently. Information Foraging Theory is being developed to understand and improve human-information interaction. It borrows from optimal foraging theory, but it assumes that humans optimize the gain of information per unit time cost. The following chapters deal with various applications of the framework, method, and theory. This includes analyses of information foraging on the Web, in document browsers, and in social networks. In addition, I discuss design and engineering applications of the theory that illustrate its practical utility.

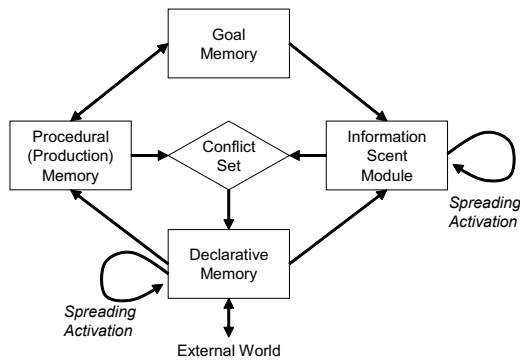


FIGURE 1.8 The ACT-Scent cognitive architecture. Information perceived from the external world is encoded into chunks in declarative memory. Goals and subgoals controlling the flow of cognitive behavior are stored in goal memory. The system matches production rules in production memory against goals and activated information in declarative memory, and those that match form a conflict set.^b The matched rule instantiations in the conflict set are evaluated by utility computations performed in the information scent module. Based on the utility evaluation, a single production rule instantiation is executed, updates are made to goal memory and declarative memory, if necessary, and the cycle begins again.^c

APPENDIX

The analysis presented in this section is provided for those readers with a background that includes exposure to basic probability theory and who are interested in the mathematics involved in calculating the expected value of searching for better hotel prices in the illustration.

The observed frequency distribution of Paris two-star hotel prices presented in figure 1.2 is presented in figure 1.A.1. Also shown in figure 1.A.1 is a best-fit lognormal distribution, which is typically found for commodity prices and would probably be characteristic of many of the things that one could buy on the Web. The estimate was performed by starting with the maximum likelihood estimates, which can be biased for small samples, and then adjusting the parameters slightly to obtain best linear fits on a Q-Q plot.

A variable X (e.g., prices) is lognormal distributed if the natural log of X , $\ln(X)$, is normal distributed.

^a[AQ: pls add to ref list or delete cite here]

^b[AQ: pls add to ref list or delete cite here]

^s[AQ: pls define "scent" here as you are using the term in this book]

^b[AQ: you appear define the term "conflict set" earlier in chapter 1 as two conflicting production rules ("In production system terminology, P3 and P6 form a *conflict set* when on a particular cycle they both match the current pattern in the goal stack and declarative memory.") Can you explain here how a rule-goal set is conflicting? (note also defined as conflicting productions in chapter 5)]

^c[AQ: pls briefly explain, here or in text, what is "spreading activation" as you use the term here]

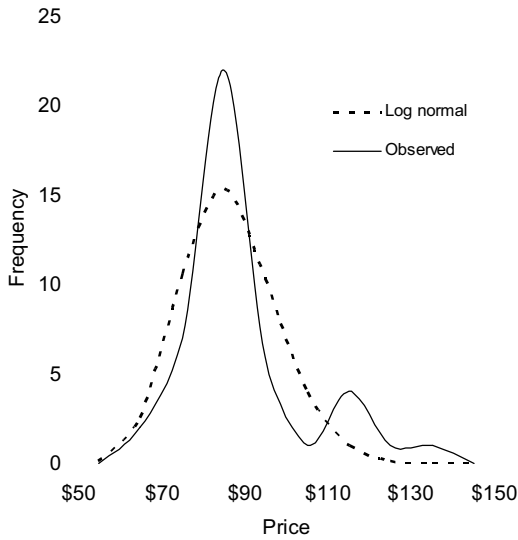


FIGURE 1.A.1 The observed distribution of Paris two-star hotel prices is approximately lognormal, which is typical of commodity prices.

The probability density function of the lognormal distribution is

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(\ln(x)-\mu)^2/2\sigma^2}, \quad (1.A.1)$$

where μ is the mean of $\ln(X)$ and σ is the standard deviation of $\ln(X)$. For the prices in figure 1.A.1, $\mu = 4.45$ and $\sigma = 0.13$. The cumulative distribution function, $F(x)$, for the lognormal is typically computed numerically using the cumulative distribution function Φ for the normal distribution,

$$F(x) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right). \quad (1.A.2)$$

The expected value of a lognormal distributed variable X is

$$E(X) = e^{\mu + \sigma^2/2}, \quad (1.A.3)$$

and the variance is

$$\text{var}(X) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \quad (1.A.4)$$

The distribution in figure 1.A.1 has an expected value of \$86.35 and a variance of \$127.09.

The expected minimum price in figure 1.3 and expected savings in figure 1.4 were computed from the probability density function of minimum values. Assume that prices are sampled n times from a random variable, such as X characterized above. The minimum value of that sample of size n can be characterized as another random variable Y_n ,

$$Y_n = \min\{X_1, X_2, \dots, X_n\}, \quad (1.A.5)$$

where the X_i are independent random draws from the random variable X . From the basic definitions of probability, the cumulative density function for the minimum of a random sample of size n , Y_n , is defined as the probability that a randomly sampled value (minimum prices in this case) will be less than some value y ,

$$G_n(y) = \Pr(Y_n \leq y), \quad (1.A.6)$$

which is equivalent to the probability that the minimum Y_n is not greater than y ,

$$G_n(y) = 1 - \Pr(Y_n > y). \quad (1.A.7)$$

The probability, $\Pr(Y_n > y)$, that the minimum value of a sample is greater than some value y would be the same as the probability that every sampled value from the random variable X was greater than y , so

$$\begin{aligned} \Pr(Y_n > y) &= \Pr(X_1 > y) \cdot \Pr(X_2 > y) \cdots \\ &= \Pr(X_n > y) \\ &= \Pr(X > y)^n. \end{aligned} \quad (1.A.8)$$

Since the meaning of the cumulative density function for X is

$$F(x) = \Pr(X \leq x), \quad (1.A.9)$$

one can define

$$\Pr(X > y) = 1 - F(y). \quad (1.A.10)$$

Now, one can substitute equation 1.A.10 into 1.A.8 into 1.A.7 to get

$$\begin{aligned} G_n(y) &= \Pr(Y_n \leq y) \\ &= 1 - \Pr(Y_n > y) \\ &= 1 - \Pr(X > y)^n \\ &= 1 - [1 - F(y)]^n \end{aligned} \quad (1.A.11)$$

The probability density function is defined as the derivative of the cumulative density function. So,

taking the derivative of equation 1.A.11, the probability density function of the random variable Y_n representing the minimum of a sample of size n drawn from variable X will be

$$g_n(y) = n[1 - F(y)]^{n-1}f(y), \quad (1.A.12)$$

where the probability density function $f(x)$ ^t and cumulative density function $F(x)$ ^u are for the sampled random variable X . The expected minimum prices and expected savings in figures 1.3 and 1.4 were computed using equation 1.A.5 assuming the probability density function and cumulative distribution function in equations 1.A.1 and 1.A.2, with the parameters $\mu=4.45$ and $\sigma=0.13$ estimated in fitting the lognormal in figure 1.A.1.

The utility of production “P3: Next-link” in table 1.1 was computed by determining the expected savings that would be attained by randomly sampling the lognormal distribution of prices in figure 1.A.1 while having a minimum price m already in hand. This expected savings can be computed by integrating over all savings achieved by prices less than m and greater than 0, weighted by the probability of getting those lower prices. So the expected savings to be achieved by a randomly sampled price x given that one has a current minimum price m in hand is

$$S(m) = \int_0^m (m-x)f(x)dx. \quad (1.A.13)$$

Given the lognormal distribution of prices in figure 1.A.1, if the lowest price found so far were \$100, then the expected savings of taking looking at the next price would be

$$S(\$100) = \$14.43.$$

Some other example expected savings would be

$$S(\$90) = \$6.62$$

$$S(\$80) = \$1.86$$

$$S(\$70) = \$0.23$$

Notes

- ^v<http://www-2.cs.cmu.edu/~hzhang/Newell.GoodScience>.

2. This example is inspired by a microeconomic analysis of the value of information in consumer purchasing by Stigler (1961).

3. For early uses of production systems in psychology, see Newell (1973^w) and Newell and Simon (1972). For overviews and history of their use in psychology, see Anderson (1993), Klahr, Langley, and Neches (1987), and Young (in press^x).

4. For those familiar with ACT-R 5.0, the productions run without the perceptual-motor modules or the subsymbolic computations. The model is available at <http://xxxxx.y>

5. Data provided courtesy of Suresh Bhavnani.

6. I purposely use the phrase “maximization tendency” because of the assumption that this is an ongoing process limited by physical and biological bounds on instantaneously achieving omniscient optimality. It is a bounded rationality process.

7. The implicit assumption is that energy translates into fitness.

8. As far as I can tell, the term “human–information interaction” first appeared in the public literature in^z the title of Gershon (1995).

9. To clarify terminology, what I am calling “knowledge” corresponds to Newell’s (e.g., 1982, 1990) use of the term. This, in turn, corresponds to Dennett’s use of “belief,” which is consistent with common philosophical usage.

10. Dennett defined an observer who describes a system using an intentional vocabulary (e.g., “know,” “believe,” “think”) as one taking an *intentional stance*.^{aa}

11. This definition is based on Pearl (1988, pp. 313–314).

12. Newell’s technical definition was that “[a] system is *intelligent* to the degree that it approximates a knowledge-level system” (Newell, 1990^{bb}). Knowledge-level systems are discussed below.

13. See Glimcher (2003) for how Marr’s work inspired a parallel rational analysis approach to understanding neuroscience.

14. Barring bugs, of course.

15. Following natural selection theorist G. C. Williams (1992), I prefer the term “optimization model” over “optimality model” to acknowledge a focus on corrective processes rather than optimal end states.

16. Pronounced “accent.”

References

- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.

^t[AQ: (x) correct here, since you use f(y) in 1.A.12?]

^u[AQ: pls confirm (x) here]

^w[AQ: 1973 a or b?]

^x[AQ: can you update?]

^y[AQ: pls complete this address]

^z[AQ: correct as edited? Else, pls explain your use of this phrase “birds of a feather” here]

^{aa}[AQ: pls confirm that this note is appropriately placed in the text - see query there; if it is properly placed, pls explain how this note relates to the sentence to which it is attached]

^{bb}[AQ: pls add pg no for quote]

^v[AQ: pls explain why you cite this web address here - is this where the list is from? Can we find Newell’s writings here or a biography of Newell? Pls also add name/date cite here for the web address and add full entry to ref list]

- Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98, 409–429.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26(1), 85–112.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review*, 11(4), 1036–1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703–719.
- Anderson, J. R., & Pirolli, P. (1984). Spread of activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 791–798.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396–408.
- Bechtel, W. (1985). Realism, instrumentalism, and the intentional stance. *Cognitive Science*, 9, 473–497.
- Berners-Lee, T. (1989). *Information management: A proposal*. Geneva, Switzerland: CERN.
- Bhavnani, S. K. (2005). Why is it difficult to find comprehensive information? Implications of information scatter for search and design. *Journal of the American Society for Information Science and Technology*, 56(9), 989–1003.
- Bhavnani, S. K., Jacob, R. T., Nardine, J., & Peck, F. A. (2003^{cc}). *Exploring the distribution of online healthcare information*. Paper presented at the CHI 2003 Conference on Human Factors in Computing Systems, Fort Lauderdale, FL.
- Borgman, C. L., Belkin, N. J., Croft, W. B., Lesk, M. E., & Landauer, T. K. (1988^{dd}). *Retrieval systems for the information seeker: Can the role of intermediary be automated?* Paper presented at the CHI 1988 Conference on Human Factors in Computing Systems, Washington, DC.
- Brentano, F. (1973). *Psychology from an empirical standpoint*. New York: Humanities Press. (Original work published 1874)
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.
- Bush, V. (1945). As we may think. *Atlantic Monthly*, 176, 101–108.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Case, D. O. (1991). The collection and use of information by some American historians: A study of motives and methods. *Library Quarterly*, 61, 61–82.
- Charnov, E. L. (1976). Optimal foraging: The marginal value theorem. *Theoretical Population Biology*, 9, 129–136.
- Cosmides, L., Tooby, J., & Barow, J. H. (1992). Introduction: Evolutionary psychology and conceptual integration. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 3–15). New York: Oxford University Press.
- Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992^{ec}, June). *Scatter/gather: A cluster-based approach to browsing large document collections*. Paper presented at the^{ff} 15th annual International ACM Conference on Research and Development in Information Retrieval, New York.
- Dawkins, R. (1989). *The extended phenotype*. Oxford: Oxford University Press.
- Dennett, D. C. (1983). Intentional systems in cognitive ethology: The “Panglossian Paradigm” revisited. *Behavioral and Brain Sciences*, 6, 343–390.
- Dennett, D. C. (1988). *The intentional stance*. Cambridge, MA: Bradford Books, MIT Press.
- Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little, Brown & Co.
- Dennett, D. C. (1995). *Darwin’s dangerous idea*. New York: Simon & Schuster.
- Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S., & Harshman, R. (1988^{gg}). *Using latent semantic analysis to improve access to textual information*. Paper presented at the CHI 1988 Conference on Human Factors in Computing Systems, Washington, DC.
- Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework* (No. AFOSR-3223). Menlo Park, CA: Stanford Research Institute.
- Furnas, G. W., Landauer, T. K., Gomez, L. W., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, 30, 964–971.
- Gershon, N. (1995, December). Human information interaction. In^{hh} *Proceedings of the Fourth International World Wide Web Conference*. Retrievedⁱⁱ from <http://www.w3.org/Conferences/WWW4/bofs/hii-bof.html>.
- Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.
- Glimcher, P. W. (2003). *Decisions, uncertainty, and the brain: The science of neuroeconomics*. Cambridge, MA: MIT Press.
- Grass, J., & Zilberstein, S. (2000). A value-driven system for autonomous information gathering. *Journal of Intelligent Information Systems*, 14, 5–27.

^{ec}[AQ: pls cite month of meeting]

^{ff}[AQ: correct to delete “Proceedings”? Else, pls cite as abstract in a published proceedings]

^{gg}[AQ: pls cite month of meeting]

^{hh}[AQ: pls cite all eds’ names/initials]

ⁱⁱ[AQ: pls cite day/month/year you retrieved this]

^{cc}[AQ: pls cite month of meeting]

^{dd}[AQ: pls cite month of conference]

- Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: A validation of GOMS for prediction and explanation of real-world task performance. *Human-Computer Interaction*, 8, 237–309.
- Hames, R. (1992). Time allocation. In E. A. Smith & B. Winterhalder (Eds.), *Evolutionary ecology and human behavior* (pp. 203–235). New York: de Gruyter.
- Johanssen, W. (1911). The genotype conception of heredity. *American Naturalist*, 45, 129–159.
- Jones, W. P. (1986^{jj}). *The memory extender personal filing system*. Paper presented at the CHI 1986 Conference Human Factors in Computing System, Boston, MA.
- Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, ^{kk}391–438.
- Klahr, D., Langley, P., & Neches, R. (Eds.). (1987). *Production system models of learning and development*. Cambridge, MA: MIT Press.
- Lucas, P. (2000^{ll}). *Pervasive information access and the rise of human-information interaction*. Paper presented at the CHI 2000 Human Factors in Computing Systems, The Hague.
- Lyman, P., & Varian, H. R. (2003). *How much information*. Retrieved February 2005 from <http://www.sims.berkeley.edu/how-much-info-2003>.
- Malone, T. (1983). How do people organize their desks? Implications for the design of office systems. *ACM Transactions on Office Systems*, 1, 25–32.
- Marr, D. (1982). *Vision*. San Francisco: W.H. Freedman.
- Mayr, E. (1983). How to carry out the adaptationist program? *American Naturalist*, 121, 324–334.
- Miller, G. A. (1983). Informavores. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 111–113). New York: Wiley.
- Moore, J., & Newell, A. (1973). How can MERLIN understand? In L. Gregg (Ed.), *Knowledge and cognition* (pp. 201–252). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Morrison, J. B., Pirolli, P., & Card, S.K. (2001). A taxonomic analysis of what World Wide Web activities significantly impact people's decisions and actions. ^{mmm}CHI 2001, *ACM Conference on Human Factors in Computing Systems, CHI Letters*, 3(1), 163–164.
- Neches, R., Langley, P., & Klahr, D. (1987). Learning, development, and production systems. In ^{mm}(Ed.), *Production system models of learning and development* (p. 1). Cambridge, MA: MIT Press.
- Newell, A. (1973a). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.
- Newell, A. (1973b). You can't play 20 questions with nature and win: Projective comments on the paper of this symposium. In W. G. Chase (Ed.), *Visual information processing*. ^{ooo}New York: Academic Press.
- Newell, A. (1982). The knowledge level. *Artificial Intelligence*, 18, 87–127.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A. (1993). Reflections on the knowledge level. *Artificial Intelligence*, 59, 31–38.
- Newell, A., & Card, S. K. (1985). The prospects for a psychological science in human-computer interactions. *Human-Computer Interaction*, 2, 251–267.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.
- Newell, A., Yost, G., Laird, J. E., Rosenbloom, P. S., & Altmann, E. (1992). Formulating the problem-space computational model. In R. F. Rashid (Ed.), *CMU Computer Science: A 25th anniversary commemorative* (pp. 255–293). New York: ACM Press.
- Nielsen, J. (2000). *Designing Web usability*. Indianapolis, IN: New Riders.
- Oaksford, M., & Chater, N. (Eds.). (1998). *Rational models of cognition*. Oxford: Oxford University Press.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Los Altos, CA: Morgan Kaufman.
- Pemberton, S. (2003). Hotel heartbreak. *Interactions*, ^{pp64}
- Pirolli, P. (1999). Cognitive engineering models and cognitive architectures in human-computer interaction. In F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H. Chi (Eds.), *Handbook of applied cognition* (pp. 441–477). West Sussex, UK: John Wiley & Sons.
- Resnikoff, H. L. (1989). *The illusion of reality*. New York: Springer-Verlag.
- Ryle, G. (1949). *The concept of mind*. London: Hutchinson.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99–118.
- Simon, H. A. (1971). Designing organizations in an information-rich world. In M. Greenberger (Ed.), *Computers, communications and the public interest*. ^{qq}Baltimore, MA: Johns Hopkins University Press.
- Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence*, 4, 181–204.
- Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.
- Soper, M. E. (1976). Characteristics and use of personal collections. *Library Quarterly*, 46, 397–415.
- Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., & DeAngelo, T. (1999). *Web site usability*. San Francisco, CA: Morgan Kaufman.

^{jj}[AQ: pls cite month of conference]

^{kk}[AQ: pls cite vol no]

^{ll}[AQ: pls cite month of meeting]

^{mmm}[AQ: pls confirm that this is the complete title of the serial publication - pls delete any part of this title that does not appear on all issues]

^{mm}[AQ: pls cite all eds' names/initials]

^{ooo}[AQ: pls cite inclusive pg nos for the chapter you are citing here]

^{pp64}[AQ: pls cite full journal title, vol no, and inclusive pg nos]

^{qq}[AQ: pls cite inclusive pg nos]

- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69, 213–225.
- Tinbergen, N. (1963). On the aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20, 410–463.
- Todd, P. M., & Gigerenzer, G. (1999). Simple heuristics that make us smart. *Behavioral and Brain Sciences*, 22, XXX–XXX.^{††}
- Williams, G. C. (1992). *Natural selection: Domain, levels, and challenges*. New York: Oxford University Press.
- Winterhalder, B., & Smith, E. A. (1992a). Evolutionary ecology and the social sciences. In E. A. Smith & B. Winterhalder (Eds.), *Evolutionary ecology and human behavior* (pp. 3–23). New York: de Gruyter.
- Young, R. M. (in press^{§§}). Production systems in cognitive psychology. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences*.^{†††} Pergamon.

^{††}[AQ: pls
cite
inclusive
pg nos]

^{§§}[AQ: can
you
update?]

^{†††}[AQ: pls
cite pg nos
and
city/state/
country of
publisher]