Rational Analysis of Exploratory Choice

Richard M Young

Department of Psychology, University of Hertfordshire, Hatfield, Herts AL10 9AB, UK

R.M.Young@herts.ac.uk

Telephone: +44 (1707) 28 5051

Fax: +44 (1707) 28 5073

July 1997

Chapter to appear in

M. Oaksford & N. Chater (Eds), *Rational Models of Cognition*.
Oxford University Press, in press.

Rational Analysis of Exploratory Choice

Richard M Young

Department of Psychology, University of Hertfordshire, Hatfield, Herts AL10 9AB, UK
R.M.Young@herts.ac.uk

Abstract

This chapter extends Anderson's (1990) rational analysis of problem solving to a class of exploratory search situations that involve selecting one of a number of possible options, where little is known about the options before exploration begins. We formulate the situation as one of single-move, multi-stage search, where the probabilities of success of the different options are initially unknown, but where a range of assessment methods is available to provide information about each option. The assessment methods differ in their costs and in the quality of the information they deliver. The analysis defines an optimal strategy, which is applied to an experimentally studied task where a subject has to use an unfamiliar computer package. Simulation of the optimal strategy shows that it exhibits a number of features characteristic of the empirical data, such as repeated scanning of the menus, progressive focussing on a subset of the options, and iterative deepening of attention.

1. Introduction

One of the virtues of rational analysis is that it offers a way to make predictions of the behaviour of a complicated system — the human cognitive processing system — from the outside, i.e., without having to know anything about the internal mechanisms. Even if one is primarily interested in investigating cognitive processing, we can identify a number of potential benefits that rational analysis might provide. First, rational analysis is a powerful tool for analysing a task, to help us understand it better, and essentially to specify what it is that the cognitive processing has to do. Second, rational analysis allows us to derive an optimal strategy for performing the task. Although people will generally not actually follow the optimal strategy, whether because of cognitive limitations or for other reasons, they will often come close to it. Third, knowing the optimal strategy often helps us to interpret empirical phenomena, for example to understand why people find certain aspects of a task difficult and other aspects easy.

This chapter addresses the topic of exploratory search (which we illustrate below in section 1.1 and define in section 2). Its inspiration and starting point is Anderson's (1990: Chapter 5) rational analysis of (non-exploratory) problem solving, part of which we extend to the exploratory case. To concretise the discussion, we introduce first a specific exploratory task making use of a computer software package called Cricket Graph, on which people's behaviour has been empirically studied.

1.1 The Cricket Graph task

Cricket Graph is a commercially available software package running on the Apple Macintosh computer. It is intended for drawing charts and graphs from files of data, and has been the subject of intensive recent study and analysis (e.g., Franzke, 1994, 1995; Rieman, 1994; Rieman, Lewis, Young & Polson, 1994; Kitajima & Polson, 1995; Rieman, Young & Howes, 1996). The experimental data come from subjects who are experienced and fluent users of the Macintosh and are therefore familiar with its conventions for the use of the mouse and menus, but who have never used Cricket Graph or any other graphing package before. They are given a two-part task: first to get Cricket Graph to plot a graph from a prepared data file, and then to edit the resulting graph so that it conforms in visual style to a printed example. Subjects are not given a manual or other instruction, the intention being that they should discover for themselves how to perform the task by exploring the menus and other options on the screen, and trying things out.

In order to perform the task, users must first open the data file within the Cricket Graph program, a part of the task it is assumed they do by analogy with data files in other programs (Rieman *et al*, 1994). Figure 1 shows the appearance of the Cricket Graph screen once the data file has been opened. The row of labels across the top of the screen, from the Apple symbol at the left to Windows at the right, is the *header bar*. Each of the nine items is a *menu header*. If the mouse button is held down when the cursor is over one of the items, a vertical *pulldown menu* appears, from which a selection can be made. The pulldown menus in Cricket Graph contain from 3 to 12 menu entries. Releasing the mouse over a menu entry either causes some action to be taken or else leads to a dialogue box, allowing further options and parameters to be set.

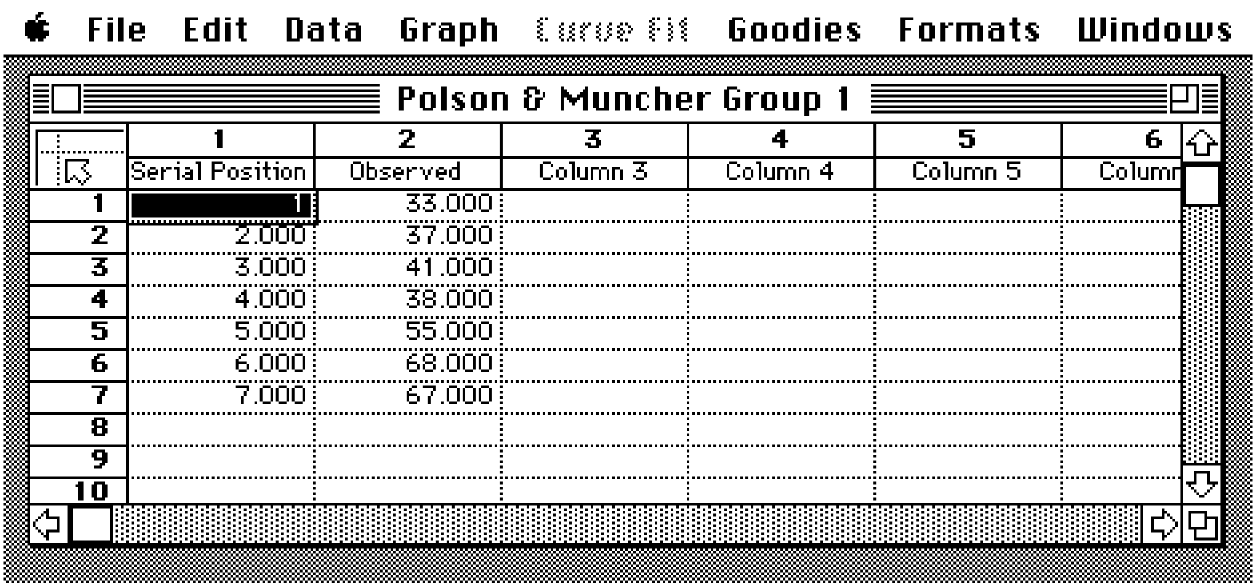

Figure 1. The appearance of Cricket Graph after the data file has been opened.

The correct sequence of actions to draw a graph of the data is to press the mouse button with the cursor over the Graph menu header, and then to select the item ‘Line Graph’ from the pulldown menu that appears. This chapter does not concern itself with later parts of the task, which require the identification of dependent and independent variables in a dialogue box and the editing of the resulting graph.

1.2 Empirical phenomena

Because subjects are not familiar with Cricket Graph, they are unable simply to proceed to execute the correct sequence of moves. Instead, they spend time pondering the menu headers examining the pulldown menus, and often invoking dialogue boxes which they then cancel, before converging on the solution. This period of exploration has been studied, and a number of empirical regularities have been found. These regularities serve to constrain any theoretical account or model of the exploratory search.

Franzke (1994, 1995) analysed the success rates and especially the solutions times in this exploratory phase, and reports two main factors that determine how long subjects take to complete this part of the task. One factor is the appropriateness of the label on the header or menu item. Franzke distinguishes four levels of closeness between the label and the task. Level 1 consists of labels that are identical to key words in the task, such as “graph”. Level 2 has labels that are semantically close to task words, such as “chart”. Level 3 labels are semantically related but require some inference to be made in order to connect them to the task, such as “drawing tools”. And level 4 labels have no clear connection to the task, such as “file”. Franzke’s data show that the time to discover correct actions increases monotonically with the level of the label. The other factor affecting exploration time is the number of potentially relevant items on the display: the more items, the longer the subjects take. While this is not surprising, more interesting is the interaction between the number of items and label quality. Although in general search times increase with the number of items, a good label can be identified relatively quickly even in a complex display.

Rieman (1994) examined subjects’ search in greater detail, again revealing empirical regularities, this time at the level of move-by-move exploratory behaviour. The first result concerns “label following”, and confirms Franzke’s and previous reports (e.g. Lewis, 1988; Polson, Lewis, Rieman & Wharton, 1992) that subjects more easily choose items with labels closely related to the task, i.e. at levels 1 and 2 of Franzke’s classification. However, although such items are readily found, subjects do not typically move directly to them. In the Cricket Graph task, most subjects will look at other headers and pulldown menus before moving to Graph; and they may examine the Graph pulldown menu, then move away and look at other items, before later returning to Graph.

Rieman’s (1994) other results provide more information about the nature of this extended menu search. Most striking was a phenomenon of “iteratively deepening attention”, in which subjects make repeated passes through a set of items, with later passes focussing on a smaller and smaller subset of the items, and typically spending longer and longer considering each item. For example, subjects might move the mouse item by item through a pulldown menu, while on an earlier pass they had merely examined that same menu visually. Rieman also observed that much of the searching conformed to the spatial layout of the items, moving left-to-right (or right-to-left) sequentially along the header bar, or vertically through a pulldown menu.

2. Rational Analysis of Exploratory Search

In this section we present a rational analysis of the kind of exploratory search exhibited by subjects on the Cricket Graph task. The analysis is carried out in much the same spirit as Anderson’s (1990: Chapter 5) analysis of problem solving.

We analyse a simplification (and generalisation) of the actual Cricket Graph search environment, by addressing the general case where someone has to choose one out of the N possible options in a choice set. Later on (in section 3) we will return to the question of applying the results of the analysis to the original Cricket Graph problem. We idealise the problem as a *single-move, multi-stage* exploratory search. That means, firstly, that the immediate goal can be reached in a single move, such as by invoking one of Cricket Graph’s menu items in order to display the desired dialogue box. The analysis therefore focusses on the person's decision of what (first) move to make. Secondly, however, the exploratory gathering of information happens in stages. For each possible option, several exploratory actions may be needed in order to gather sufficient information about the option and evaluate how appropriate it is for the current goal. Those exploratory steps correspond, for example, to the kinds of tests that emerge from Franzke's (1994, 1995) analysis and subsequent attempts at modelling (e.g. Rieman *et al*, 1996): evaluating the match between an option label and a task, assessing the semantic overlap between them, using analogy to see if a plausible action can be found, performing internal lookahead to foresee the consequences of a particular move, and so on.

We begin by considering in the first subsection (2.1) what we mean by exploratory search. Section 2.2 discusses the structure of the task environment. We then consider (section 2.3) how the problem is to be framed in formal terms. Section 2.4 discusses how options are assessed for their relevance to the goal, and section 2.5 considers how information of high quality can be used to improve the likelihood of a successful search. We are then in a position (2.6) to propose and analyse the optimal strategy.

2.1 Exploratory search

Our starting point for the analysis is Anderson's (1990) rational analysis of problem solving, together with some ideas from Rehder, Lewis, Terwilliger, Polson & Rieman's (1995) preliminary analysis of the Cricket Graph problem. Anderson casts optimal problem solving as the need to choose a route from a starting situation to a goal that maximises the expected value of the payoff

$$Y_i = P_i G - C_i,$$

where P_i is the probability that the ith possible move leads directly to the goal, G is the value to the problem solver of achieving the goal, and C_i is the cost of making the move. In the problem solving situation, the various P_i (as well as the C_i) are assumed known. It should be noted that G is measured, like the C_i, in units of cost.

In the exploratory situation, by contrast, the P_i are initially unknown. Instead, the problem solver has available a set of *assessment procedures* which can be applied to each possible option to estimate its relevance to the goal, from which in turn an (estimated) value of P_i can be derived. The assessment procedures vary in their costs and also in their *quality*, i.e. in the accuracy of the relevance estimates they deliver. More than one assessment may have to be made of an option before the P_i are known with sufficient certainty to serve as the basis for a decision. The problem solver must repeatedly choose a candidate option and an assessment procedure to apply to it in order to gather information about promising options, until enough is known about the various P_i for one of the options to be taken. The primary task for rational analysis is to identify an optimal strategy for choosing assessments (i.e. options to be assessed and assessment procedures to apply to them) and then deciding when to act, in order to maximise the expected value of a payoff function to give the greatest chance of success for the least cost.

We further assume that there is exactly one correct solution, i.e. one move that leads to the goal. The job for the person is therefore to identify and then take this one move. In what follows, we shall see that the combination of unknown probabilities with the restriction to a single correct move makes the analysis differ significantly from Anderson's (1990) for non-exploratory problem solving.

2.2 Structure of the task environment

The paradigmatic task we are analysing is that of picking one out of N possible moves, where various means are available for assessing the relevance of each move. Although the analysis does not depend on it, for clarity we will draw our illustrations from the domain of deliberately designed environments, such as computer menus, rather than naturally occurring choice sets.

2.2.1 Patterns of probabilities

The values of the probabilities of the possible moves often fall into one of a small number of characteristic patterns. Almost by definition, a well designed choice set leads to a set of probability estimates in which one — the correct one — is high, say around 0.9 or higher, while the rest are low. (As specified by equation [2] below, the assumption of one correct move means that the probabilities sum to 1.) Consider for example the choice set offered by an automatic bank teller machine (ATM) which offers services identified as *inter A/C tranfer, statement by post*, *change PIN, withdraw cash, cash with receipt*, and *balance enquiry*. When the task is to change the personal identification number (PIN), there is just one obviously relevant option, “change PIN”, while the rest are just as obviously irrelevant. We will refer to such a pattern as a *single spike*, because of the shape of the pattern if the probabilities are plotted against the options.

Not all decisions are as clear cut, however. Consider the choice set in Figure 2, which relies upon the common knowledge that Britain — actually the United Kingdom — consists (roughly) of England, Scotland, Wales, and Northern Ireland. If we want to find tourist information about Edinburgh, the capital of Scotland, which option do we choose? Choices 2, 4, and 5 are obvious non-contenders, since Edinburgh is not in England, Wales, or Northern Ireland. But both choices 1 and 3 look plausible, and it takes careful study to notice that option 1 leads to an alphabetical list of places, which is probably what we want, whereas option 3 is organised geographically and is likely to be harder for us to use. Such a pattern of probabilities, where — at least at one stage of the exploration — there are two more or less equally plausible contenders, will be called a *double spike*.

Choice sets can differ widely in their implicit structure (Young & Hull, 1981, 1983). For present purposes, however, the single and double spike can be taken as representative of a broader class of patterns.

```
  LOCAL INFORMATION

1 PLACES IN BRITAIN  Alphabetic list of
  cities, towns, & places of interest

2 REGIONS OF ENGLAND  Information about
  regions, counties & major towns &
  cities indexed according to region

3 REGIONS OF SCOTLAND  Information about
  Scotland indexed by region

4 REGIONS OF WALES  A limited amount of
  info on Wales is now available

5 NORTHERN IRELAND & ISLE OF MAN
  Tourist advice,hotels,travel,holidays
```

Figure 2. A "double spike" choice set. See text for discussion. (Adapted from Young & Hull, 1983.)

2.2.2 Interdependence of probability estimates

In Anderson's (1990) analysis of problem solving, the probabilities of different moves are assigned independently. In our exploratory situation, on the other hand, there are inter-dependencies between the probabilities, sometimes strong ones. In part, these interdependencies are due to the normalisation constraint that the probabilities sum to 1, a constraint that does not apply in the problem solving context. The constraint obviously means that if one choice has a high probability, then the others must have low ones.

It is important to realise, however, that the dependencies are not a mathematical artifact caused by the normalisation. They reflect real cross-relationships between the judgements about choices made by a person, and cannot be avoided by mathematical or terminological devices such as by defining a new measure — "appropriateness" or whatever — not subject to normalisation. Rather, the reality is that people are often forced to make rapid and radical revisions of their estimates of the correctness of particular options as they work their way through and further ponder the choice set.

To return to the case of the bank ATM, for example, suppose we are trying to use an unfamiliar machine to withdraw cash from our account. When we encounter an option labelled "Withdraw Cash" we recognise it as being a good match to our goal, good enough indeed that we are tempted to choose it without reading further. Say we assign it a probability of 0.95. But if we do read on, we find that the next option is labelled "Cash with Receipt", and now we realise that this present option is the one we want, not the previous one, because we do want a record of our transaction and it simply did not occur to us when assessing the previous option that we would be offered cash without a receipt. So we revise the previous probability down to near zero, and make this new one very high, say at 0.98.

Another form of interdependence is that the structure of typical choice sets legitimises a form of what is usually described as the "gambler's fallacy", which is that a good item is more likely to occur after a string of bad ones. When using a reasonably designed choice set, we expect at least one of the options to provide at least a fairly good match to our goal. So if we work through a menu and keep finding poor candidates for selection, as we approach the end of the menu we increasingly expect one of the remaining candidates to be good. It is possible for that expectation to be disappointed, in which case we may continue to find only poor candidates, and we then have to return to the earlier candidates, reconsider them, and perhaps revise their probabilities upward. Nonetheless, the expectation is a legitimate one (the "menu-user's legitimate belief"?) justified by the structure of typical menu designs and our experience with them.

It should be noticed that these interdependencies between the probabilities lead to situations very different to those that arise in Anderson's (1990) analysis of (non-exploratory) problem solving. In problem solving, when we find two or more good moves of roughly equal promise, there is nothing special about that situation. As always, we choose the one which maximises the payoff PG–C, or if the contenders are about equivalent on that score, simply pick either of them. In our exploratory situation, by contrast, a double spike pattern represents a situation in which we are maximally uncertain about which of two choices is the right one. It becomes important to attempt to resolve the situation, by gathering further information about the contenders and hoping that one of the probabilities will be revised upwards and the other downwards, to bring us closer to an unambiguous choice.

2.3 Framing the problem

2.3.1 The payoff function

Before we re-state the problem in more formal terms, it is worth adopting some modifications to the form of the payoff function. As noted, Anderson (1990) takes an optimal solution as one which maximises the expected value of $P_iG–C_i$. In our situation, because there is only one correct solution — which therefore has a fixed and unavoidable cost — there is no need to include the cost in the payoff function to be maximised. After all, in our exploratory setting the person's task is to find the (one) correct move, regardless of whether it is low-cost or high-cost. We therefore redefine the payoff to reflect only the cost of making an *incorrect* move, not the correct one. Furthermore, since the move will eventually be taken in the belief that it is the correct one, we need consider only a single expected cost C for making an incorrect move, independent of the option chosen:

$$Y_i = P_iG - (1–P_i)C = P_i(G+C) - C.$$

This form of the payoff emphasises that C is regarded as a penalty only for moves other than the correct one. It also shows that high-cost moves are more sensitive to the value of P, a point which is obscured by the original form. Finally, the constant trailing "– C" can be dropped from the expression, to give us simply

$$Y_i = P_i(G+C). \qquad [1]$$

This is the form of payoff we will use in the analysis.

2.3.2 Statement of problem

We now introduce some terminology to let us state the problem more precisely than has been possible so far.

The exploratory problem solver is to make one of N possible moves, M_1, M_2, … M_N. Exactly one of the moves leads to the goal, of value G. Choosing the wrong move involves a penalty of cost C. For each move M_i there is a set of assessment procedures A_{ij}, each with a cost B_{ij}. For many purposes we will find it convenient to treat all the possible assessments as a single set not differentiated by which move they apply to. In that case we will write the kth assessment as A_k with associated cost B_k.

When assessment A_k is applied to move M_i, it yields an estimate of the move's relevance to the goal, written $R_i(k)$ or more commonly just R_i. The set of $R_i(k)$ determines — by means of a transformation to be discussed later — a set of values $P_i(k)$, where $P_i(k)$ is the estimate (after the kth assessment) of the probability P_i that move M_i is the correct move, i.e. the one that leads to the goal. Because only one move is correct, and because this constraint is known to (or assumed by) the person, we assume that the P_i are normalised so that they sum to 1,

\i\su(i=1,N,P_i) = 1. [2]

Associated with each move M_i is a payoff function, Y_i, as given by equation [1]. After a series of n assessments, $A_1, A_2, \ldots A_n$, at a total cost of $\sum B_k$, the problem solver chooses the move M_m with maximum payoff Y_m. The worth of that move is defined as

W = $Y_m - \sum B_k$ = $\max_i[P_i(n).(G+C)]$ – \i\su(k=1,n,B_k). [3]

Provided that W > 0, the move is finally made.

The primary task for rational analysis is to identify a strategy for choosing the optimal sequence of assessments A_k $(1 \leq k \leq n)$, and also the stopping point n, to maximise the expected value of W.

2.4 Assessment

The approach we are adopting in this analysis is that the person performs a series of assessments, each of a single option. After each assessment A_k, the probabilities of all the moves are updated to the most recent estimates $P_i(k)$.

2.4.1 Relevance and probability

Because of the various interdependencies between the probabilities for the different moves — including but not confined to the normalisation constraint that the probabilities sum to 1 — it is extremely awkward to conceptualise the assessment of a move M_i as yielding directly an estimate of its probability P_i. Instead, we regard the assessment as delivering an estimate of the move's *relevance*, R_i, a quantity which reflects the inherent merit or plausibility of the move's appropriateness for attaining the goal. The set of R_i are then mapped into a set of probabilities P_i, in a way intended to capture most of the interdependencies between them. The relevances R_i are therefore, up to a point, independent of each other. R_i encodes an assessment obtained by considering just the one move M_i.

Rehder *et al* (1995) treat relevance as a discrete variable on a 7-point scale. Instead of that, we treat relevance as a probability-like quantity, and in the analysis R_i will enter into probability-like calculations. As with all cases of subjective probability, there seems little point in trying to provide a precise definition of what relevance means. It can perhaps best be thought of as an idealised version of P_i, the probability of move M_i, if it were not for the effects of the other moves. So R_i can be regarded as providing an estimate of P_i, but "disregarding" or somehow "averaged across" the other possible moves.

Treating R_i as a probability allows us to assign all unexamined options an initial default or neutral relevance of value 1/N.

2.4.2 Mapping relevance to probability

One can imagine many more or less elaborate schemes for mapping relevances into probabilities. For our purpose we adopt the expedient of normalising the relevances by dividing the odds of each relevance by the sum of the odds,

$$P_i \;=\; \backslash f(\mathrm{odds}(R_i),\textstyle\sum \mathrm{odds}(R_j)) \;, \qquad [4]$$

where the odds are given by the standard definition

$$\mathrm{odds}(R) \;=\; \backslash f(R,1–R) \;. \qquad [5]$$

This mapping has the merits of reasonable simplicity and mathematical tractability. It reflects the idea that the R_i are intended as approximations to the P_i, albeit one which ignores their inter-dependencies. In addition, it exhibits most if not all of the interactions we expect to see between the probabilities. For example:

- A single spike pattern of relevances, which we can define as having only one relevance with a non-low value, will be mapped into a pattern with one high probability (near 1) and all the rest low.
- A double spike pattern of relevances, meaning that exactly two of the relevances are other than low and that they are roughly equal, will be mapped into a pattern with two probabilities around the 0.5 mark and all the rest low.
- The mapping exhibits a form of the “menu-user’s legitimate belief”. If most of the possible moves have been assessed and have had their relevances downgraded from their initial neutral value of 1/N to something near zero, the probability of the remaining moves will be increased.

The relationships between assessment, relevance, and probability can get a little confusing at times. Figure 3 is intended to help keep the relations straight. It shows how an assessment of move M_i updates the estimate of its relevance, R_i, and how that relevance is combined with all the others to yield an estimate of the probability P_i.

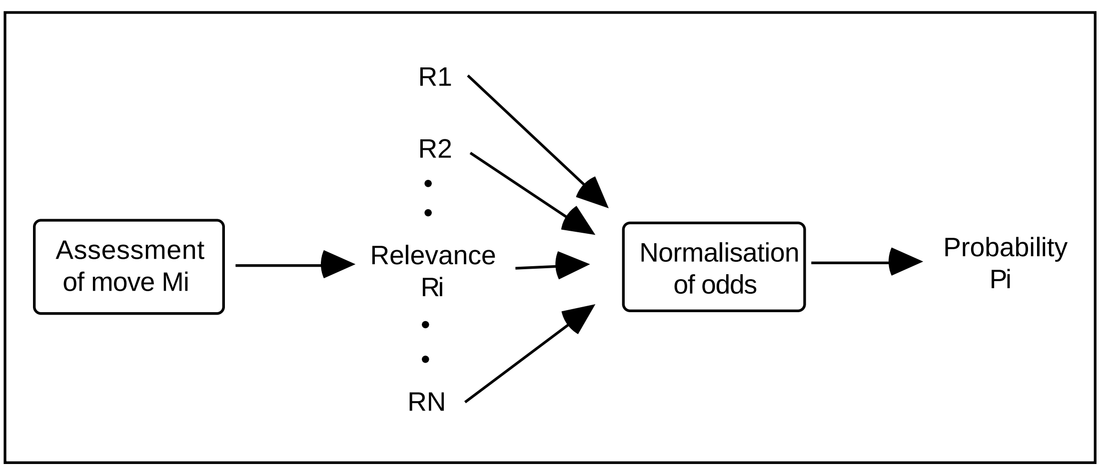

Figure 3. The relationships among assessment, relevance, normalisation, and probability.

It is worth spending a moment longer examining the nature of this mapping from relevances to probabilities. We can re-write [4] by separating the denominator into the odds of R_i itself and the sum of the rest of the odds:

$$P_i \;=\; \backslash f(\mathrm{odds}(R_i),\mathrm{odds}(R_i)+\textstyle\sum \mathrm{odds}(R_{j\neq i})) \;, \qquad [6]$$

a form which will be useful to us later on. In the case where the 'rest' of the odds sum to approximately 1, $\sum \text{odds}(R_{j\neq i}) \approx 1$, we have

$$P_i \approx \frac{\text{odds}(R_i)}{\text{odds}(R_i)+1} \approx R_i \quad [7]$$

so that in these circumstances the probability has approximately the same value as the relevance. It is worth noting that in the initial situation, with all R_i having value 1/N, this condition holds: the sum of the rest of the odds is 1. Consequently, the result of the very first assessment is that the new P_i has the same value as the new R_i.

2.4.3 Quality of assessment

Central to our approach is the idea that assessments vary in the quality of the information they deliver, with higher quality assessments, in general, costing more. But how are we to conceptualise and to represent the quality of an estimate of relevance, R_i?

In contexts dealing with probabilities, quality of assessment usually translates into a matter of *precision.* If some quantity P is thought of as having a true value such as 0.7, then a high quality assessment of P would be one whose distribution of values centres on 0.7 with low variance. However, such an approach is problematic for our situation. For one thing, it would involve us in consideration of probability distributions of probability values, which is difficult to handle both conceptually and analytically. More crucially, it becomes extremely unobvious how a high quality assessment contributes to increasing the value of the payoff. We are trying to maximise a function involving the probability values (equation [3] again). If we know that some probability P is approximately 0.7, we can plug that value into the function to compute some maximal payoff. If we are then told that P has value 0.7 *very precisely*, that may mean that we know the value of the function more precisely, but it does not change its actual value.

We solve this puzzle by adopting a different view. Each move M_i either is or is not the right one, i.e. the one to achieve the goal. So we regard each R_i as an approximation to an underlying true value of the relevance which is either 0 or 1. High quality assessments are therefore extreme assessments, i.e. near 0 or 1. A high quality assessment procedure delivers estimates of R_i which are bimodally distributed, tending towards 0 and 1. A low quality assessment procedure delivers estimates which are closer to the neutral value 1/N, and may not even be bimodal.

Given such an approach, we can re-interpret the relevance R_i as an indication of how *confident* we are, on the grounds of considering just M_i itself, that M_i is the right move. High quality assessments tend to return highly confident estimates, near 1 or 0. Low quality assessments tend to return low confidence estimates, closer to the neutral value of 1/N.

2.4.4 Assessment as Bayesian updating and the unbias constraint

We can of course say nothing detailed about the nature of the assessment procedures, which in general can involve arbitrary cognitive processing of any aspects of the option being considered. We will, however, assume that the outcome of an assessment is equivalent to a Bayesian updating procedure which computes a revised relevance in the light of its previous value and of the conditional probabilities of the evidence considered by the assessment. Specifically, we assume that the assessment delivers a revised value which can be expressed as

$$R_i(k) = \frac{r_i.P(E_+)}{r_i.P(E_+) + (1-r_i).P(E_-)} \quad [8]$$

where

$R_i(k)$	is the relevance of move M_i after the kth assessment;
r_i	is $R_i(k-1)$, the relevance of move M_i just before the kth assessment;
$P(E_+)$	is the conditional probability of the evidence given that M_i is the right move;
$P(E_-)$	is the conditional probability of the evidence given that M_i is the wrong move.

Treating assessment as a Bayesian updating offers several advantages to the analysis. One of them is that the assessment obeys what we will call the *unbias constraint.* A Bayesian computation such as [8] has the property that the expected value of the expression, $E(R_i(k))$, coincides with the prior value, r_i. In other words: although once we perform the assessment and have a particular value for the evidence, the updated value $R_i(k)$ will usually differ from the prior value, nevertheless *before* the assessment is carried out, and averaged across all the possible values for the evidence in the light of their conditional probabilities, the expected mean value for $R_i(k)$ is the same as its prior value, r_i. That means that the value of R_i *before* the assessment is an "unbiased" estimate (in the statistical sense) of what its value will be *after* the assessment.

The unbias constraint is crucial for a coherent analysis. Actually doing an assessment will usually change the estimate of a move's relevance in some way, of course, but the estimate should not change simply because of the *possibility* of performing a further assessment and considering further evidence. The expected value of the assessment must coincide with the current estimate. Otherwise we will be in the position of having two contradictory beliefs, namely (a) that our current estimate is the best we can make, based upon the evidence so far considered, and (b) that it will be, say, increased by considering some further evidence. To avoid this situation, we require our assessments to be unbiased.

A second advantage of the Bayesian computation for our analysis is that the result of a series of updatings using formula [8] is independent of the order in which the updatings are performed. The end result simply expresses the posterior estimate of R_i in the light of all the evidence considered by the assessments.

A third, and relatively minor, advantage is the computational convenience following from the well-known result that the odds of the posterior value can be found by multiplying the prior odds by the ratio of the conditional probabilities. Equation [8] can be re-written as

$$\text{odds}(R_i(k)) = \text{odds}(r_i) \;.\; \backslash f(P(E_+),P(E_-)) \;. \qquad [9]$$

2.5 The value of information

In this section we begin to explore the mechanism by which improved assessments of the relevance of moves are converted to increases in the expected value of the payoff function. The story will be completed in the next section.

There are some puzzles to be explained. Intuitively, it is clear that getting better quality estimates of the R_i (and therefore the P_i) improves our chances of choosing the right move. If we learn that a particular move has a very high P value and is therefore almost certainly the right one, that is obviously better than being less certain. On the other hand, if we learn that some move is almost certainly not the right one, that is also helpful because it means we can rule it out of consideration. But the worth of the best move is given (see equation [3]) by a function that involves maximising over the P_i, so how can finding a 'poor' move increase the value of the function? Furthermore, the unbias constraint tells us that any assessment of move M_i leaves the expected value of R_i, and therefore of P_i, unchanged. How, then, can applying an assessment procedure increase the expected worth of the best move?

The answer follows from the combination of two things, (a) the logic of the maximisation function used in equation [3], and (b) the arithmetic of the normalisation that carries R values into P values (equation [4]). We know that a high quality assessment procedure tends to produce extreme values for R (section 2.4.3). So if we submit some move M_i to a high quality assessment, although the possible outcome values for $R_i(k)$ weighted by their probability of occurrence coincide with the prior $R_i(k–1)$ — that's the unbias constraint — nevertheless the resulting estimate of $R_i(k)$ is likely to be more extreme than the prior one, i.e. closer to 1 or 0. And the trick is that getting extreme values for R_i, either high or low, helps to improve our position.

Consider the situation of a double spike pattern. Suppose that two moves, M_u and M_v, have relevances R_u and R_v that are nearly equal (though say $R_u > R_v$), and are much higher than any other relevances. Then the corresponding probabilities P_u and P_v will both be close to 0.5. If we perform a high quality assessment of M_v, say, then although the expected value of R_v is the same as its present value, most likely R_v will either increase significantly in value or decrease significantly — and the better the assessment, the more likely that is to happen. In either case, the value of the maximal P_i will be increased. If R_v decreases significantly, then the current best probability P_u increases to greater than 0.5, because of the reduced denominator in the normalisation process of [4]. If R_v increases significantly, on the other hand, then it becomes greater than R_u, and P_v becomes the new best probability, again at greater than 0.5. So, either way, the worth of the best move, as given by [3], is improved. A similar argument applies if we assess M_u rather than M_v.

Of course, this purely qualitative argument does not tell the whole story. Nonetheless, even this preliminary and informal analysis suggests some guidance as to where assessment effort is best spent. In the case of a double spike pattern, assessing either of the two contenders is worthwhile, because *any* significant change to either of their values will yield an improvement in the worth of the best move. In the case of a single spike, there is relatively little to be gained by assessing the leading option, since equation [6] shows that where R_i dominates the other relevances, the probability P_i is fairly insensitive to changes in the value of R_i. That same equation shows, though, that reassessing a poor option may be worthwhile, because any change in a low R_i translates into an almost proportional change to its corresponding P_i.

This thread of the story gets taken up again in section 2.6.2.

2.6 Optimal search

We are finally in a position to start pulling the strands together in order to compute the optimal strategy and stopping point. We begin by reviewing the structure of the search.

2.6.1 Structure of the search: the choice phase

For reasons that will become clearer later, the search divides into two main stages which we call the *choice* phase and the *confidence* phase. We focus first on the choice phase.

Following Anderson (1990, 1993) as well as Rehder et al (1995), we cast the search as a repeated choice between (a) making the most promising move now, or (b) performing another assessment. If we choose course (a) and make the most promising move now, after n assessments, then the worth is, in accord with equation [3],

$$W(n) = \max_i(Y_i(n)) - \sum B_k, \qquad [10]$$

where Y_i is given by equation [1] (or whatever payoff function we care to use). If, on the other hand, we decide to perform another assessment A_a, since we have not yet done the assessment we cannot know the worth of the resulting situation, but we can estimate its value as

$$Est(W(n+1)) = opt_a(Est(max_i(Y_i(n+1)) - \sum B_k - B_a), \quad [11]$$

where the opt_a means that A_a is chosen to optimise the expression following it. Exactly what this optimal choice is, is a topic we deal with in Section 2.6.3 below.

Which of the two courses is better depends on the relative values of [10] and [11]. The improvement to be gained by continuing one further round of assessment is computed as the difference Est(W(n+1)) – W(n), which we can re-write as

$$E(W(a)) - W(n) = [opt_a(Est(max_i(Y_i(a))) - max_i(Y_i(n))] - B_a \quad [12]$$

where we have replaced 'n+1' by 'a' to emphasise that we are considering the situation following a further assessment A_a, and we consider all occurrences of 'a' to be governed by the optimisation. The first part of the right hand side of [12], namely the expression in square brackets, is simply the estimated increase in payoff to be gained by doing a further assessment, so we can re-write the equation once more as

$$Est(W(a)) - W(n) = Est(\Delta_a max_i(Y_i)) - B_a, \quad [13]$$

which says that the estimated benefit of continuing with a further assessment is simply the estimated value of the increase in payoff reduced by the cost of performing the assessment.

We are nearly home. In order to compute [13], we need to take one final preparatory step, to see how to work out the expected increase in payoff resulting from an assessment.

2.6.2 The value of an assessment

After n assessments, the maximum payoff is $max(Y_i(n))$. If we perform a further assessment, the estimated value of the maximum payoff rises by $Est(\Delta_k max(Y_i))$. How big is that estimated increase, and how does it come about?

To answer those questions, we have to begin by considering the possible effects of the assessment, not directly on the maximal payoff, but on the maximal relevance. Let M_u be the move with maximal payoff, M_v be the move with the second best payoff, and M_w any move other than M_u. M_u, M_v, and M_w have relevances R_u, R_v, and R_w. We need to consider separately the cases where the assessment is of the maximal move M_u or of some other move M_w.

Case: M_u

If assessment A_k assesses the maximal move M_u, the possible outcomes can be divided into two regions of interest:

1. R_u increases or decreases, but remains above the second-best value R_v.
2. R_u decreases to below the value of R_v. In this case R_v becomes the new highest relevance.

Before we compute the expected increase in R_u as the result of assessment A_k, it is worth understanding qualitatively what is happening, as we are now beginning to complete the story started in Section 2.5 about how further assessment can lead to an increase in expected payoff. The unbias constraint tells us that the expected value of R_u after assessment is exactly the same as it was before. However, outcome 2 shows us that the possible drop in value is effectively limited by the second best value R_v, which acts as a floor below which the value of the maximal R cannot fall. So the effect of re-assessing the best R_u is an increase in the expected maximal value, because the bottom part of the distribution of outcomes has been removed from consideration. The actual computation closely follows this same logic.

Let $S_{ijr}(x)$ be the probability density function governing the distribution of values for R_i resulting from the assessment $A_k = A_{ij}$, where R_i has prior value r. Put $T = R_v$, the second highest R, a threshold or floor below which the maximal R cannot fall. Then, for an assessment of the maximal R_u, the expected value of the maximal R is given by

E(max(R_i(k))) = \i(0,T,T.S(x).dx) + \i(T,1,x.S(x).dx)
= \i(0,T,T.S(x).dx) – \i(0,T,x.S(x).dx) + \i(0,T,x.S(x).dx) +
\i(T,1,x.S(x).dx)
= \i(0,T,(T–x).S(x).dx) + \i(0,1,x.S(x).dx) [14]

The second integral in [14] is just the new expected value of R_u, which by the unbias constraint is the same as the present value. The first integral therefore represents the expected *increase* in the value of the maximal R. Writing R_v for T, we have

E(Δmax(R)) = \i(0,R_v,(R_v–x).S(x).dx) [15]

That integral is simply the expected amount by which R_u would fall below R_v, given that it does, weighted by the probability that it does.

Case: M_W

If assessment A_k assesses a move M_w other than the maximal move M_u, the outcomes can again be divided into two regions:

1. R_w increases or decreases, but does not overtake the maximal R_u.
2. R_w increases to become greater than R_u, and therefore the new greatest R.

This aspect of the analysis forms a further part of the explanation of how high quality assessment lead to an increase in the expected maximal payoff. No matter what happens, the maximal R_i cannot decrease, because it will have at least its present value R_u. If we apply a high quality assessment to a non-maximal R, then its value is likely to become more extreme. If it increases enough to exceed the current best R, then we have a new best choice, better than the previous one. Thus, the maximal relevance can only increase, and therefore its expected value increases. As before, the actual computation follows this same logic.

This time we set the threshold T = R_u, the highest relevance, which provides a floor below which the maximal relevance cannot fall. For an assessment of a currently non-maximal R_w, the expected value of the maximal R is given by

E(max(R_i(k))) = \i(0,T,T.S(x).dx) + \i(T,1,x.S(x).dx)
= \i(0,T,T.S(x).dx) + \i(T,1,T.S(x).dx) – \i(T,1,T.S(x).dx) +
\i(T,1,x.S(x).dx)
= T. \i(0,1,S(x).dx) + \i(T,1,(x–T).S(x).dx)
= T + \i(T,1,(x–T).S(x).dx) [16]

The first term, T, is equal to R_u, the current best relevance. The integral therefore represents the expected *increase* in the value of the maximal R. Writing R_u for T, we have

E(Δmax(R)) = \i(R_u,1,(x–R_u).S(x).dx) [17]

That integral is simply the expected amount by which R_w would rise above R_u, given that it does, weighted by the probability that it does.

Implications of the computations

Equations [15] and [17] show the expected increase in the value of the maximal relevance following an assessment, respectively, of the current maximal R_u or of some other relevance R_w. The details of how these expected increases in maximal R, E(Δmax(R)), are converted into the estimates of increases in maximal payoff, Est(Δmax(Y)), are complex and involve approximations, and are presented in Appendix 1. However, before we leave this topic of the expected increase in relevance, it is worth understanding what equations [15] and [17] tell us about which assessments are likely to be most worthwhile. The following points are implied by the computations:

- With a double or multiple spike pattern, i.e. with two or more plausible contenders of roughly equal relevance, assessing any of the contenders further is likely to be of value. Furthermore, the closer the candidates, the greater is the expected improvement. For assessing the current best relevance R_u, equation [15] shows that the higher the value of the second best relevance R_v, the greater is the expected increase in the maximal R. Similarly, for assessing any other relevance R_w, equation [17] shows that the lower the value of the best relevance R_u, the greater is the expected increase in the maximal R. This account of the value of obtaining further information about plausible candidate moves is in line with our informal discussion of the situation (section 2.5).
- With a single spike, i.e. with a single clear best candidate, there is only limited value in further assessment. If its nearest rival has only a low relevance, then equation [15] shows that the expected increase in relevance is small, and of course the increase in payoff is further restricted by the already high relevance (equation [6]).
- Candidates other than the best, but with still clearly non-zero relevances, can be worth assessing further. If their relevance is reduced then they improve the best payoff through normalisation (equation [4]), but they also have a respectable chance of being upgraded to become a new plausible choice.
- Further assessments of moves of very low relevance are unlikely to improve the situation.

2.6.3 The optimal strategy

Equation [13], repeated here,

$$Est(W(a)) - W(n) = Est(\Delta_a max_i(Y_i)) - B_a, \qquad [13, repeated]$$

states that the estimated increase in worth of the situation resulting from the optimal assessment A_a is given by the excess of the estimated increase in best payoff over the cost of the assessment, where the estimated increase in payoff, $Est(\Delta_a max_i(Y_i))$, is specified by equations [A2] and [A8] of Appendix 1.

Which assessment is the "optimal" one? The fact that all the probabilities and hence payoffs are updated after each assessment makes the choice hard to analyse rigorously, but in analogous situations the optimal strategy is to choose actions in order of the ratio of their benefit to their cost, in this case by the ratio $Est(\Delta_a max_i(Y_i))/B_a$.

One such situation arises in Anderson's (1990: Chapter 5) analysis of problem solving, where moves are tried in order of the ratios P_i/C_i. Such an ordering causes a graph of probability-of-success against cumulative effort to grow as rapidly as possible, leading to a monotonically increasing, negatively accelerated curve. The proof of that optimal ordering (Anderson, personal communication) depends on the observations that for a given set of moves, the probability of success does not depend on their ordering, while the expected total cost is minimised by trying the most efficient moves first. Ahlswede & Wegener (1987: Chapter 11) provide a more general investigation and proof of the optimal ordering, drawing upon earlier results by Kadane & Simon (1977), and again give an ordering by the "efficiency" measure P/C.

The key assumption, that the worth W(n) of a situation after n assessments is independent of the order in which the assessments are made, holds for our exploratory task. Even though the probabilities change with each assessment, the pattern of probabilities after n assessments simply reflects the best estimates that can be made of the relevances in the light of the evidence considered by those assessments, and is therefore independent of their order. The general argument about choosing assessments in order of their efficiency would therefore seem to apply, with the result that the expected worth grows as quickly as possible with effort expended. We will therefore take it that:

> *The optimal strategy is to choose at each stage the assessment A_a with the greatest efficiency, as measured by the ratio $Est(\Delta_a max_i(Y_i))/B_a$.* [18]

The consequences of that strategy for our exploratory task follow straightforwardly from the properties of the ratio. The strategy will tend to favour assessments which are expected to give a good boost to the best payoff (as discussed in the bulleted points at the end of the previous section), or which are cheap to make (such as label identity tests), or of course both.

2.6.4 Stopping rule

Equation [13] also shows clearly when the assessments should stop and the best move be taken. Further assessment is worthwhile so long as the expected gain in payoff exceeds the cost of the assessment. When the efficiency ratio $\mathrm{Est}(\Delta_a \mathrm{max}_i(Y_i))/B_a$ falls below 1, assessment should stop.

We note that, as usual with such a stopping rule, assessment will tend to continue if there are still considerable gains in payoff to be had, or if there are still cheap assessments to be made.

2.6.5 Constraints on assessments

The analysis has so far asssumed that all assessments are available for application at any time. In practice, this is not quite the case.

Firstly, of course, once an assessment has been made, there is no point in applying it again, since doing so would yield no further information. (The view of assessment as a Bayesian updating procedure makes that point clear.) But there are other kinds of dependencies and constraints between assessment procedures. Some assessments can include others as part of the procedure. Assessing the semantic overlap between a label and the task, for example, necessarily includes testing for label identity, which itself counts as an assessment method. So it is impossible to test the semantic overlap without also testing label identity. In other cases, assessments yield information which is needed by other assessment procedures in order to apply. At the other end of the the spectrum, for example, one can apply the assessment procedure “envision the outcome” (also called “internal lookahead”) only if one has already applied the procedure “find an action for an object”, i.e. decide whether to press, click, or double-click the mouse.

Ahlswede & Wegener’s (1987) analysis tells us how to handle such constraints among the assessment procedures. Where there is an assessment A_y that requires some other assessment A_x to be done first to provide it with information, or in general to satisfy a precondition, then if A_x has not yet been done we consider for the analysis the compound assessment $\{A_x;A_y\}$, i.e. A_x followed by A_y. That compound assessment, with its total cost, is used in the calculations in place of A_y.

2.6.6 Confidence phase

The optimal strategy [18] and the stopping rule together complete the first major part of the exploratory search, the choice phase. When the optimal search stops, the move M_m with the highest probability indicates the choice of move to make. The optimal strategy implicitly recognises that further search might lead to a different choice with a greater payoff; however, the anticipated improvement is outweighed by the cost of performing the additional assessments.

In environments conducive to exploration, the choice phase will usually end with a clear preference for one chosen move. In other words, the final probabilities will usually exhibit what we have termed a single spike pattern, with one clear winning option and all the rest low. However, this winning probability may itself not be especially high: it can easily be not more than around 0.5. We have seen that the optimal strategy tends not to re-assess options that are already in a winning position. This fact puts the problem solver in a difficult position. It may know that M_m is clearly the best move. But if its probability P_m is only around say 0.6, it may be unwilling actually to make the move, because of the 0.4 probability of incurring the cost C of making a wrong move. With P_m as low as 0.6, and with a considerable penalty C, the risk of making the wrong move may well exceed the anticipated gain from making the right move.

To overcome this problem, the exploratory search enters a second phase, the confidence phase, the main purpose (and usual effect) of which is to increase the estimated probability P_m of the chosen move M_m. It does so by tentatively assuming that M_m is indeed the right move, and then (perhaps repeatedly) performing further assessments. The appropriate assessment is selected and applied using the same mathematical apparatus as in the main choice phase. However, the estimates of the gain to be had from performing an assessment, Est(Δmax(Y)), come out differently because the assumed distribution of evidence is different. During the choice phase, the expected distribution of values for the evidence concerning some move M_i is determined by the conditional distributions, depending upon whether M_i is the right move or the wrong move, and on the prior value of the relevance R_i, and is computed by weighting the distributions by R_i and $(1–R_i)$ respectively before combining them. However, in the confidence phase we are asssuming (tentatively) that M_m is indeed the right move, so only the 'right' conditional distribution is used.

Usually, during the confidence phase, the value of P_m increases rapidly to around 0.9 or more. Because the confidence phase uses the same mathematical apparatus as the choice phase, the same stopping rule applies and the confidence phase ends when the estimated further improvement to Y_m becomes less than the cost of the assessment needed to achieve it.

Although the intended (and usual) effect of the confidence phase is to build up the probability P_m, it can of course happen that one of the reassessments yields evidence against M_m, indicating that it is after all the wrong choice. But that presents no deep difficulty. The problem solver simply drops back into the choice phase, with values for the R_i and P_i reflecting the new evidence.

3. Applying the Analysis to Cricket Graph

The rational analysis has abstracted away from the particulars of our Cricket Graph task in several ways. In order to apply the analysis, we need to bridge back from the assumptions of the analysis to the concrete circumstances of the task.

3.1 Accessing the item to be assessed

Subject to the constraints mentioned in section 2.6.5, the rational analysis assumes that any assessment procedure of any item can be chosen at any stage. An alternative assumption (made, for example, by Rieman *et al*, 1996) is that the person has access only to the item being attended: the person either further assesses that item, or scans left or right to one of its immediate neighbours.

The effect of such a physical proximity constraint is to add a further cost to the assessment of items other than the current one. On this view, all the items are accessible, but at the cost of the scanning moves required to reach them. So the view can be made compatible with the rational analysis by adding to the cost B_{ij} of assessment A_{ij} of item M_i an additional access cost D_i, proportional to the number of shifts of attention needed to reach M_i from the current focus of attention. There seem to be two reasonable ways to incorporate these access costs into the analysis:

(a) One way is to argue that in Cricket Graph, all the relevant items are accessible at very low cost, so we can simply ignore the access costs D_i.

(b) The other way is temporarily to add D_i to the B_{ij}, as outlined above. An interesting property of this approach is that the D_i push the rationally optimal strategy towards local, sequential scanning. The simulation to be presented shows that the main effect of small D_i is to act as a *tie-breaker*, creating a preference between otherwise equally attractive alternatives. In this role, the D_i again encourage the emergence of local scanning.

3.2 Headers and menu items

Our analysis deals with a single set of choices, arranged as a single-move exploratory search. In Cricket Graph itself, however, the choice of an item is structured as a two-step exploration: choosing the header, then choosing the menu item.

To apply the rational analysis to Cricket Graph, we could run through the optimal strategy twice, once to pick a header and once to pick a menu item. The assessment methods and costs differ slightly between headers and menu items. To increase the realism, for the headers we include a moderately expensive assessment method which involves displaying and scanning the associated pulldown menu. This assessment returns a value based on a general estimate of the match between the menu items and the task. We do not apply the rational strategy recursively within the assessment of the header.

3.3 Numbers, assessments, and distributions

In order to simulate the application of the rational strategy to the concrete case of Cricket Graph, a host of numerical parameters and probability distributions have to be specified. The details are given in Appendix 2. Here we summarise just the main points.

We set the value of G to 300 and of C to 600. If we interpret these numbers as seconds, they say that the goal of achieving the task is worth about five minutes work to the person, while making a wrong move can require as much as ten minutes extra work to recover from. So those numbers specify a situation in which it is important to make the right move.

We model four assessment methods for the menu headers. In order of increasing cost they are:

- label match: whether the header label matches a task term (cost = 1);
- semantic overlap: to what extent the meaning of the header label overlaps the task (cost = 3);
- pulldown: as described above, a broad estimate of the match between the items on the pulldown menu and task (cost = 10);
- anticipation: internal envisionment of the result of the action and comparison with the task; this procedure provides a high-quality assessment, returning extreme values (i.e. near 1 or 0) of high accuracy (cost = 20).

We include an access cost D_i as discussed earlier, in order to break the tie between otherwise equivalent items. The cost of shifting to an adjacent item is taken to be 1 unit.

3.4 Simulation results

Figure 4 shows the sequence of assessments generated by the optimal strategy. Starting at the extreme left, the simulation starts scanning rightwards across the headers, performing tests for lexical match. The headers File, Data, and Graph all match a term in the task, and in each case the simulation follows the successful lexical match with a test for semantic overlap. Following the semantic test on Graph, the pulldown menu under Data is scanned, but the inappopriateness of those items greatly reduces the estimate of the header's relevance (from .42 to .08). The simulation next examines the remaining headers, none of which provides a lexical match. After a few further assessments, the choice phase finishes with Graph as the chosen header at a P value of .68, as against the second best choice of Data with a P of just .06. The simulation then enters the confidence phase, where it finds it well worth examining the pulldown menu of Graph to confirm its choice. Doing so boosts the P value to .95. The simulation marginally considers a further assessment of Graph to be not worth while (the estimated efficiency of performing an anticipation assessment on Graph is around 0.94, compared with a threshold of 1.0), so the process terminates with the — correct — selection of the Graph header. The total cost of the exploration is 100, meaning that it is regarded as having taken about 100 seconds.

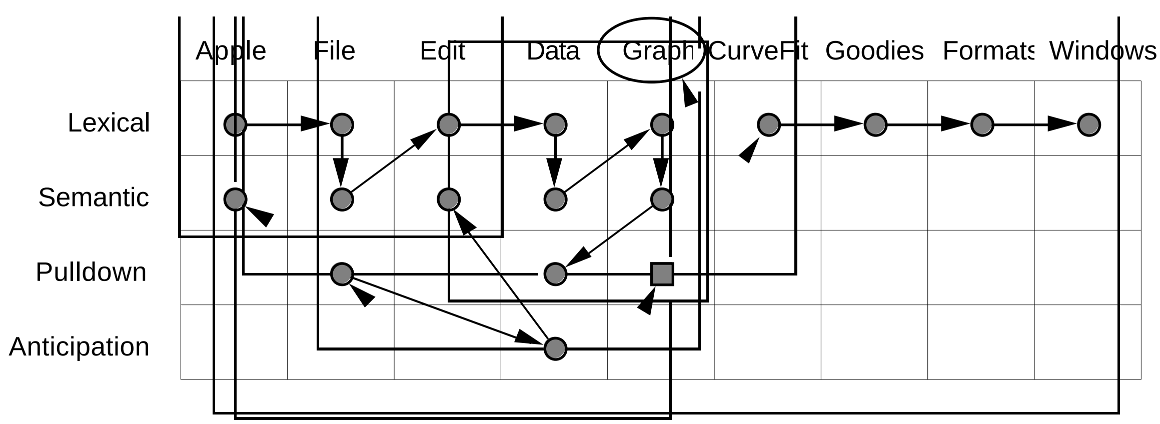

Figure 4. The sequence of assessments generated by the simulation. The square marker in the Graph/Pulldown cell indicates an assessment performed as part of the confidence phase.

A number of features of the simulation are worth further comment:

- *Repeated scanning*. Figure 4 shows clearly that the optimal search makes multiple scans though the options. For example, Data is visited on three separate occasions. Interestingly, these scans do not take the form of rigid, linear passes at successively increasing depth, but exhibit far more flexibility. Initially, there is a mainly left-to-right scan, but with some deeper side-exploration on the way and one backtracking. This is followed by a back-and-forth scan of the four leftmost items, and then by a final visit to Graph.
- *Selective focussing*. The scanning increasingly focusses on just a subset of the available options. The three scans just described visit respectively 9, 4, and 1 of the headers. Looked at another way, the three 'attractive' items (which in this case provide a lexical match) receive either 3 or 4 assessments each, while the other items are given either 1 or 2.
- *Progressive deepening of attention*. Successive visits to an item involve assessments of increasing cost. Although in other cases this pattern might be forced by the ordering constraints between methods of assessment (see section 2.6.5), on this run it arises purely from the way the optimal strategy chooses assessments in order of their estimated efficiency.
- *Partial scanning*. Although it does not occur on this run, it is possible for the optimal strategy to make a choice without having examined all the options.
- *Sensitivity to numerical assumptions*. An important finding, not shown in Figure 4 but immediately apparent from examining a detailed trace of the simulation and from running the simulation several times with differing data, is that the behaviour of the optimal strategy (i.e. the sequence of assessments it performs) is extremely sensitive to the actual numbers assumed in the simulation. For example, if the cost of the 'anticipation' assessment were 18 instead of 20, then the confidence phase would include such an assessment of the Graph option before stopping. If the cost of the 'pulldown' assessment, instead of 10, takes some other value in the range say 7-13, quite different trajectories result. This happens because, on many cycles, the optimal strategy is faced with a choice between several plausible assessment with closely similar efficiencies. The general characteristics of the exploration, as summarised in these bulleted points, are robust and continue to hold, but the detailed sequence of assessments can vary dramatically.

4. Discussion

4.1 Rational analysis of exploratory search

This chapter has extended Anderson's (1990) style of rational analysis to a class of exploratory search situations that involve choosing one out of N possible options about which little is initially known. We have assumed that information about the relevance of options to the task at hand is accumulated by performing a sequence of assessments, varying in their cost and in the quality of the information they deliver. The rational analysis identifies an optimal strategy for selecting a sequence of assessments, which — in a sense defined by the analysis — yields the greatest probability of choosing the right option at the least total cost.

Simulation of the optimal strategy shows that it exhibits a number of properties in common with people's behaviour on the Cricket Graph task, and more generally in exploratory search (this list should be compared with Table 1 of Rieman *et al*, 1996):

- it takes account of the labels of options;
- exact matches between the label and the task are acted on soonest;
- poor labels and large number of options interact to produce longer search times;
- obviously incorrect labels are avoided;
- initial scanning is predominantly guided by spatial layout of options (e.g. left-to-right);
- repeated scans are made with iterative deepening of attention;
- the scans progressively focus on a subset of the options;
- the optimal strategy may make a choice without having examined all the items.

As with most rational analyses, we take this correspondence to human behaviour as meaning not that people are “doing” a rational analysis “in the head”, but rather as evidence that people’s behaviour is quite close to the theoretically optimum — certainly close enough for the analysis to make useful predictions.

The rational analysis brings out clearly the role of certain aspects of an optimal strategy which are not necessarily obvious beforehand. One of these is the importance of the double spike pattern, i.e. the situation where there are two (or it could be more) roughly equally promising options, with the rest not nearly as good. We turned to that situation repeatedly during the analysis. In the case of a double spike, the optimal strategy — in accord with commonsense — places high value on assessing the contenders further, in the hope of resolving the contest and finding a clear winner. A second such aspect is the sensitivity of the optimal strategy to the details of the numerical assumptions made for the simulation. We return to that topic in section 4.3 below.

4.2 Extension to multi-move exploration

The analysis in this chapter applies to single-move (though multi-stage) exploration, where the immediate goal is reached in just one move and interest therefore focuses on the selection of a single, correct option. With ingenuity, though, we were able to apply the analysis to a situation including short exploratory forays (examining the pulldown menus in Cricket Graph), by regarding the forays simply as another method of assessment and summarising their effect in the result of those assessments.

Real multi-move exploration, by contrast, does not divide neatly into assessment on the one hand and making moves on the other. An obvious aim for further analysis of exploratory search is to extend the work presented here to the case of genuine multi-move exploratory search, where making moves itself serves as a means for gathering further information about the options being explored, information that will be used when a foray is wholly or partially abandoned and search returns to an earlier point. We are encouraged in this aim by Anderson’s (1990: Chapter 5) success in extending an initial analysis of single-move problem solving to the general multi-move case. We note also, though, that that extension requires generous helpings of approximations and additional assumptions, and we would expect to have to do the same.

4.3 Under-determination of exploratory behaviour

Finally, we turn to the observation (from section 3.4) that the detailed behaviour generated by the optimal strategy is highly sensitive to the numerical values assumed for the costs and probability distributions. This sensitivity arises because, on many cycles, the strategy is faced with two or more possible assessments of closely similar efficiency, so that small differences in the assumed numbers alter which one is actually best. Moreover, once an assessment is selected and applied, the new situation is different to what it would have been if a different assessment had been chosen, making it more likely still that the next assessment is different to what it would otherwise have been. In this way, a small change in the assumed numbers can lead to a gross change in the sequence of assessments selected by the optimal strategy. It should be noted, however, that although the detailed trajectory of assessment is sensitive to the numerical assumptions, the main findings of the analysis — such as the properties discussed in section 3.4 and 4.1 above — are robust against perturbations of the numbers.

It would seem that this sensitivity to assumptions is telling us something important about the structure of the exploratory task, namely, that *the detailed course of behaviour in exploratory search is under-determined by the task environment*. Of course, given definite numbers, the optimal strategy will produce a definite sequence of assessments. However, as we have just seen, a slightly different set of numbers can lead to a very different sequence. And it would seem to be a mistake to seek the exact "true" value of a cost or a probability distribution: those numbers are theoretical entities defined for the sake of deriving a concrete trajectory from the rational analysis. An assumed value of .35 for some cell of a conditional probability table cannot be meaningfully more "correct" than a value of say .30 or .40, and similarly an assumed cost of 20 could just as well be 18 or 22. Thus, the existence of a well-defined optimal strategy does not necessarily imply the existence of a unique, well-defined "optimal trajectory" of assessments. In empirical terms, this means that we should expect certain (predictable!) aspects of exploratory search behaviour to vary from one person to another, and for a person from one occasion to another. On the other hand, we expect the general properties of the exploratory search — such as the repeated scanning and the iterative deepening of attention — to hold constant.

Acknowledgements

I am grateful to John R Anderson for a proof of the optimality of the "efficiency first" strategy in problem solving, and to Ian Nimmo-Smith for pointing me towards the book by Ahlswede & Wedener (1987). Nick Chater, Andrew Howes, Stephen Payne, Peter Polson, John Rieman, and Frank Ritter commented on an earlier version of the manuscript. Mike Page helped to clarify the analysis by arguing for a quite different approach that I mean to return to some day.

References

Ahlswede, R. & Wegener, I. (1987) *Search Problems.* Wiley.

Anderson, J. R. (1990) *The Adaptive Character of Thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Franzke, M. (1994) Exploration and experienced performance with display-based systems. Ph.D. thesis, University of Colorado at Boulder, Institute of Cognitive Science: Technical Report 94-05.

Franzke, M. (1995) Turning research into practice: Characteristics of display-based interaction. In I. R. Katz, R. Mack & L. Marks (Eds), *Proceedings of CHI'95: Human Factors in Computing Systems,* 421-428. ACM Press.

Kadane, J. B. & Simon, H. A. (1977) Optimal strategies for a class of constrained sequential problems. *Annals of Statistics, 5*, 237-255.

Kitajima, M. & Polson, P. G. (1995) A comprehension-based model of correct performance and errors in skilled, display-based human-computer interaction. *International Journal of Human-Computer Studies, 43*, 65-99.

Lewis, C. H. (1988) Why and how to learn why: Analysis-based generalization of procedures. *Cognitive Science, 12*, 211-256.

Polson, P. G., Lewis, C., Rieman, J. & Wharton, C. (1992) Cognitive walkthroughs: A methodology for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies, 36*, 741-773.

Rehder, B., Lewis, C., Terwilliger, B., Polson, P. & Rieman, R. (1995) A model of optimal exploration and decision making in novel interfaces. In I. Katz, R. Mack & L. Marks (Eds), *Human Factors in Computing Systems: CHI'95 Conference Companion*, 230-231. New York: ACM.

Rieman, J. (1994) Learning strategies and exploratory behavior of interactive computer users. PhD dissertation, Technical Report CU-CS-723-94, Department of Computer Science, University of Colorado, Boulder, Colorado.

Rieman, J. (1996) A field study of exploratory learning strategies. *ACM Transactions on Computer-Human Interaction, 3*, 189-218.

Rieman, J., Lewis, C., Young, R. M. & Polson, P. G. (1994) "Why is a raven like a writing desk?": Lessons in interface consistency and analogical reasoning from two cognitive architectures. In B. Adelson, S. Dumais & J. R. Olson (Eds), *CHI'94 Conference Proceedings: Human Factors in Computing Systems*, 438-444. New York: ACM Press.

Rieman, J., Young, R. M. & Howes, A. (1996) A dual-space model of iteratively deepening exploratory learning. *International Journal of Human-Computer Studies, 44*, 743-775.

Russell, S. & Wefald, E. (1991) *Do the Right Thing: Studies in Limited Rationality*. Cambridge, MA: MIT Press.

Young, R. M. & Hull, A. J. (1982) Cognitive aspects of the selection of viewdata options by casual users. In *"Pathways to the Information Society", Proceedings of the Sixth International Conference on Computer Communication*, 571-576. London.

Young, R. M. & Hull, A. J. (1983) Categorisation structures in hierarchical menus. In L. Kirves (Ed.), *Proceedings of Tenth International Symposium on Human Factors in Telecommunications*, 111-118. Helsinki, Finland.

Appendix 1: Calculation of the Estimated Value of an Assessment

If we perform an assessment A_k, the estimated value of the maximum payoff rises by $Est(\Delta_k max(Y_i))$. How do we calculate that estimated increase?

Let M_u be the move with the highest R, and therefore the highest P and the highest payoff Y. Let M_v be the move with the second best payoff, and M_w any move other than M_u. Let $S_{ijr}(x)$ be the probability density function governing the distribution of values for R_i resulting from the assessment $A_k = A_{ij}$, where R_i has prior value r.

Equations [15] and [17] show the expected increase in maximal R following an assessment, respectively, of the currently highest R_u or of some other R_w. Our immediate task is to map that expected increase in maximal R into an estimated increase in maximal P. Note that it will *not* do simply to compute the mathematical expectation of the increase in best P. Because of the non-linear mapping from R values to P values (see [4]), the expected value $E(P_i)$ of a P_i after an assessment is not the same as its value before. In other words, reassessments of P_i do *not* meet the unbias constraint. For that reason, using $E(P_i)$ in the calculations would lead to misleading decisions about the choice of assessment — for example, assessments could appear to reduce the anticipated maximal P instead of increasing it. We need instead to calculate an estimate of maximal P consonant with the aims of the analysis, and this will involve making approximations. As before, we consider separately the cases of assessing the current best M_u or some other M_w.

Case: M_u

For assessing the currently best move, the estimate is straightforward. From [14] or [15] we know that the expected new maximal R is given by

$$R' = E(\max(R_i(k))) = R_u + \int_0^{R_v} (R_v - x).S(x).dx \ . \qquad [A1]$$

We approximate by treating R' as the anticipated new value for R_u. We can then calculate the estimated value of P_u by substituting this new value of R_u for the old one in the normalistion:

$$Est(\max(P)) = \frac{odds(R')}{\sum odds(R) + odds(R') - odds(R_u)} \ .$$

By subtracting the previous value and using the formula for the payoff [1], we have

$$Est(\Delta_k \max(Y_i)) = \left(\frac{odds(R')}{\sum odds(R) + odds(R') - odds(R_u)} - P_u\right) . (G+C) \qquad [A2]$$

Case: M_w

For assessing a move other than the current best, the situation is more complicated. Recall that a reassessment of M_w can increase the maximal P in either of two different ways. If R_w decreases, then because of the normalisation this will lead to an increased value for the currently best P_u. If R_w increases, it may rise above the current best R_u to give a new, higher maximal R, leading to a higher maximal P. We take the calculation in two parts, corresponding to those two possibilities.

It is convenient to define T(x) to be the cumulative distribution corresponding to S(x), so that

$$T(x) = \int_0^x S(y).dy \qquad [A3]$$

is the probability that a new value of R will fall below some threshold x. As with equations [16] and [17], we treat the currently best R_u as a threshold for the new value of R_w. The new value of R_w will fall below R_u with probability $T(R_u)$, and will rise above R_u with probability $1-T(R_u)$.

If R_w remains below R_u, we have for the expected new value of R_w

$$R_s = \frac{1}{T(R_u)} \int_{0}^{R_u} x.S(x).dx \quad \text{[A4]}$$

The best move remains M_u, so we have the new maximal P given by

$$P_s = \frac{odds(R_u)}{\sum odds(R) - odds(R_w) + odds(R_s)} . \quad \text{[A5]}$$

On the other hand, if R_w rises above R_u then the new value of R_w becomes the new best R. The expected new value of R_w is

$$R_d = \frac{1}{1-T(R_u)} \int_{R_u}^{1} x.S(x).dx . \quad \text{[A6]}$$

The tricky question in this case is what to use for the denominator of the normalisation. If we simply replace the old value of R_w by the new, it can sometimes happen that even though the new maximal R is greater than the old, the new greatest P is less than the old! But being told that the best move is different to the one we previously thought is an important outcome from an assessment, even if the associated payoff is slightly reduced. (Russell & Wefald [1991: pp.85-86] deal with this same possibility by saying that we are "better off" if we change the choice of best move, even if the payoff is reduced.) To avoid under-estimating the importance of this outcome, for the purpose of normalisation we treat the new best R as if it were an increased value of R_u rather than of R_w. The new best P value is then given by

$$P_d = \frac{odds(R_d)}{\sum odds(R) - odds(R_u) + odds(R_d)} . \quad \text{[A7]}$$

We can combine the two possibilities, weighted by their respective probabilities, to yield for the estimate of new best P

$$Est(max(P)) = T(R_u).P_s + (1-T(R_u)).P_d$$

where P_s and P_d are given by [A4] to [A7]. Since the previous best P was P_u, we get

$$Est(\Delta_k max(Y_i)) = [T(R_u).P_s + (1-T(R_u)).P_d - P_u] . (G+C) \quad \text{[A8]}$$

Appendix 2: Details of the Simulation

The simulation of the optimal strategy is embodied in a computer program written in the language Tcl and available from the author on request* . Whereas the mathematical treatment in the body of the paper deals with continuous probability distributions, in the simulation each assessment results in one of a small number of discrete values and the program operates in terms of the corresponding discrete distributions. Table 1 shows the possible values for each assessment and their conditional probabilities. The actual assessments used in the simulation are given in Table 2.

* The program is also accessible via the World Wide Web from the author's home page,

Lexical			Semantic			Pulldown			Anticipation		
	P+	P–		P+	P–		P+	P–		P+	P–
mismatch	.50	.80	unrelated	.05	.25	unrelated	.01	.20	no	.05	.80
match	.40	.20	distant	.10	.33	distant	.05	.40	unsure	.15	.15
super	.10	.00	moderate	.20	.25	moderate	.20	.27	yes	.75	.05
			close	.35	.12	close	.39	.10	super	.05	.00
			identity	.20	.05	identity	.30	.03			
			super	.10	.00	super	.05	.00			

Table 1. Possible values for each assessment and their conditional probabilities. The columns marked P+ show the probabilities if the option being assessed is the right one, and P- if the option is wrong. For explanation of the value 'super', see text below.

	Apple	File	Edit	Data	Graph	CurveFit	Goodies	Formats	Windows
Lexical	mismatch	match	mismatch	match	match	mismatch	mismatch	mismatch	mismatch
Semantic	unrelated	moderate	distant	close	identity	close	distant	distant	unrelated
Pulldown	unrelated	distant	unrelated	distant	identity	moderate	distant	unrelated	unrelated
Anticipation	no	no	no	unsure	yes	unsure	no	no	no

Table 2. Values returned by each of the assessments.

The use of discrete distributions leads to an unexpected artifact in the simulation. Permeating the rational analysis is the assumption that assessing an option can result in a new option becoming the best one. So when we assess some option M_w, there is some chance that its relevance might rise above that of the currently best option M_u. Even if the likelihood of that outcome is small, it plays a crucial role in the calculation of the estimated increase in best relevance, and indeed is responsible for the possibility that it can increase at all. With a small number of discrete values, however, there is usually no value for which the 'wrong' conditional probability is zero, and this imposes a strict upper bound on the greatest value of R that can result from an assessment. In practice, as soon as some M_u has a relevance R_u much in excess of the others, it becomes impossible for any other option to overtake it and the search stops after just one or two cycles. Our *ad hoc* remedy to the problem is to give all assessments an extra value of "super", which is never actually assigned but for which the 'wrong' conditional probability is zero and the 'right' one non-zero. The result is that the calculations over the discrete distributions then behave in a way corresponding to the mathematical analysis.