The Emergence of Interactive Behavior : A model of rational menu search.

Xiuli Chen, Gilles Bailly, Duncan Brumby, Antti Oulasvirta, Andrew Howes

HAL Id: hal-02063155
https://hal.sorbonne-universite.fr/hal-02063155

Submitted on 10 Mar 2019

The Emergence of Interactive Behaviour:
A Model of Rational Menu Search

Xiuli Chen[1], Gilles Bailly[2], Duncan P. Brumby[3], Antti Oulasvirta[4], and Andrew Howes[1]

[1]University of Birmingham, Birmingham, UK. xiuli.chenuk@gmail.com
[2]CNRS LTCI, Telecom-Paristech, Paris, France.
[3]UCL Interaction Centre, University College London, UK.
[4]Aalto University, Helsinki, Finland.

ABSTRACT

One reason that human interaction with technology is difficult to understand is because the way in which people perform interactive tasks is highly adaptive. One such interactive task is menu search. In the current article we test the hypothesis that menu search is rationally adapted to (1) the ecological structure of interaction, (2) cognitive and perceptual limits, and (3) the goal to maximise the trade-off between speed and accuracy. Unlike in previous models, no assumptions are made about the strategies available to or adopted by users, rather the menu search problem is specified as a reinforcement learning problem and behaviour *emerges* by finding the *optimal policy*. The model is tested against existing empirical findings concerning the effect of menu organisation and menu length. The model predicts the effect of these variables on task completion time and eye movements. The discussion considers the pros and cons of the modelling approach relative to other well-known modelling approaches.

Author Keywords

menu search; cognitive modelling; visual search; eye movements; Markov Decision Process; reinforcement learning; interaction science

ACM Classification Keywords

H.5.2 Theory and methods: H.1.2 User/Machine Systems: Human Information Processing

INTRODUCTION

Over the past few years, a trend has been for models of human performance to encompass the *adaptive* nature of interaction [15, 22, 31, 32, 34, 41]. In these models a sequence of user actions is predicted from an analysis of what it is rational for a user to do given an interface design and given known constraints on human cognition. These analyses take into account the costs and benefits of each action to the user so as to compose actions into efficient behavioural sequences. Examples include models of multitasking in which the time spent on each of two or more tasks is determined by their relative benefits and time costs [24, 42]. They also include models of the search for information on the web in which the benefits include information gain [34] and even models of how to use a visual programming language [14]. The analysis of rationality has also informed empirical studies of tasks such as driving while using a phone or iPod [9], studies of Powerpoint use [11], and of interactive planning [30]. In each case human behaviour is shown to be an *emergent* consequence of adaptation to the task environment and psychological constraints.

The goal of the current paper is to explain menu search as an emergent consequence of adaptation to these same constraints. Understanding menu search is important for HCI because menus are the primary interaction method in most computer technologies and because there are outstanding design issues; issues concerning both organisational and layout factors.There are also many empirical studies of the use of menus and they continue to present design challenges [6, 10, 17, 20, 28]. Understanding menu search is also important, because many menu commands are used infrequently and users often have to search for commands that are new or that have forgotten locations. Sometimes the search will be localised to a particular menu and sometimes not. A predictive model of menu search could assist in research aimed at improving this ubiquitous and powerful interface technology [3].

Evidence suggests that there is substantial scope for strategic adaptation in menu search [6, 10, 17, 20, 28]. Despite deceptive simplicity, menus invite a flexible range of behaviour. Even, for example, the seemingly mundane task of searching through a vertically arranged textual menu can be achieved by starting at the top and considering each item or, alternatively, by guessing where the desired item will be in an alphabetically ordered or semantically grouped list. Sometimes users appear to skip over some items but not others [6]. They also search differently when looking for a known-word than when looking for a semantically related item [7]. There are many further possibilities and refinements and many subtle variations that determine, for example, where to look next, when to make a guess, and when to search in a different menu entirely. For many users, these are choices that are made quickly and implicitly.

In the current article we show that the range of observed behaviours emerges through psychologically constrained interaction with menus. Specifically, we report a cognitive model of rational menu search that offers the first account in which

predictions of search time and eye movements emerge from assumptions about the user's task environment and limitations.

Another goal of the paper is to advance the role of cognitive models as a method for explaining interactive behaviour. Modelling is important to HCI because it has the potential to offer more systematic, rigorous and general explanations of interaction than, for example, verbal descriptions. It is *rigorous*, because modeling uses computational methods to simulate cognitive processes and thereby reduce the potential for ambiguity and miscalculation. It is *systematic*, because otherwise disparate theoretical commitments can be brought together into a single explanatory framework. Lastly, it is more *general*, because the systematicity of theory allows a large body of facts, e.g., about performance time, sometimes on very different tasks, to be given a single explanation.

Despite their potential strengths, cognitive models have been under-exploited in HCI. While there has been a substantial effort to model menu search and related visual search tasks [8, 10, 15, 18, 25, 28], these existing models have tended require the modeler to explicitly define the set of available strategies. For example, the ACT-R model reported in [10] implemented a set of production rules to decide which item to assess next. One of these rules would direct gaze to the next item in the menu while the other would move gaze to a randomly selected (and previously unassessed) item. Because these production rules competed stochastically, the resulting behaviour was a mix of systematic top-to-bottom scanning with occasional jumps. The EPIC model reported in [18] captured item skipping in a different way. Production rules were used to implement a visual search strategy in which fixations were constrained to fall on new items that were outside the current field of view (even though they might not be directly fixated). Despite these differences in the production rule specifications, the behaviour of both of these models is restricted to the strategy alternatives that were explicitly defined (hand-coded) by the modeler.

In contrast to the previous models of menu search, the model reported in the current article has two advantages: (1) it does not require the modeller to hand-code production rules and, as a consequence, (2) there are no arbitrary restrictions on the range of possible behaviours. Instead control knowledge, and therefore predicted effects for a particular menu design, is an emergent consequence of rational adaptation to the structure of interaction, cognitive and perceptual limits, and the goal to maximise the trade-off between speed and accuracy. In other words, the model shows how user strategy might emerge from constraints. Therefore, while the proposed model builds on a number of insights concerning cognitive mechanisms that were first proposed with ACT-R and EPIC models [8, 10, 15, 18, 25, 28], it does not require any hand-coded production rules to make predictions.

In what follows we first review the background to the work. In particular, we focus on contributions to understanding Human-Computer Interaction as a rational adaptation to constraints. Subsequently, we report a model of menu search which shows how visual search behaviour is an emergent con-

sequence of adaptation. Critically, the model does not make any a priori assumptions about either eye movement strategy or when to select an item (which are usually thought of as a *stopping rule* problem). The model is tested with two studies one of which was chosen to demonstrate that the model can generate predictions for common computer application menus and the second of which was chosen so as to provide a comparison to existing data sets. The first study generates predictions for users of an Apple Mac who have experienced a number of applications and a number of the menus deployed by those applications. The second study tests the model's predictions against data from participants in a laboratory experiment reported by [2]. Detailed comparisons to human performance metrics (including performance time and eye movements) are provided.

Our primary contributions are:

1. A novel computational model, based on machine learning, of menu search showing how behaviour is an emergent consequence of environment, cognition, and utility.
2. A quantitative account of how existing empirical findings concerning menu search can be explained as rational adaptation to constraints.
3. The further development of a general computational modelling framework for explaining the emergence of rational interactive behaviour.

BACKGROUND

The idea that human interaction with technology can be understood as a rational adaptation has strong roots in the HCI and Human Factors literatures [5, 15, 26, 32, 31, 34, 33, 40, 41]. One way to summarise the framework adopted in this literature is with Figure 1 [31]. In the figure there are four components to an explanation of behaviour: Environment, Utility, Mechanism and Strategy. *Utility* concerns assumptions about what a person wants to do. *Ecology* concerns the statistical structure of the environment. *Mechanism* concerns human information processing capacities. These three components determine a space of strategies, which is the fourth component, and with a principle of rationality they then jointly determine the *strategy* (also called a *policy*) and therefore a sequence of *actions*. In the framework the strategies are an emergent consequence of rational adaptation to the other components.

The successful application of this framework requires a theory of each of the components in Figure 1. One key set of constraints are those imposed by the human visual system. In the vision research literature there has been much recent interest in explaining visual search as an adaptation to visual processing mechanisms and reward [19, 29, 37, 39]. Key
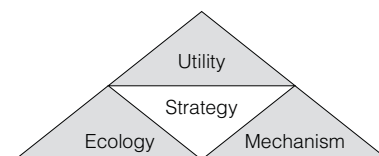

Figure 1: The Adaptive Interaction Framework [31]

constraints imposed by the mechanisms concern saccade and fixation latencies [35] and also the reduction of acuity with eccentricity from the fovea [25]. It has been shown that given these constraints, strategies can be derived through the use of reinforcement learning algorithms [12, 19, 37], though it is possible that strategies may be acquired by other learning mechanisms, for example, by cultural transmission, through instructions, or by evolution.

The approach that we take is also influenced by ideas in optimal control and Machine Learning [5, 36, 38]. A key contribution of this literature has been to provide a formal basis for learning an optimal control policy given only a definition of the reward function, the state space, and the action space. Control knowledge is simply knowledge that determines what-to-do-when. In the case of menu search it concerns where to move the eyes and when to select an item. In this framework, the expected value of an action given a state is the sum of the immediate reward plus the rewards that would accrue from subsequent actions if that action were selected. This simple assumption has provided a means of deriving human visual search strategies in well-known laboratory tasks [12]. It also provides a means by which to derive rational menu search behaviour given assumptions about utility, ecology and psychological mechanisms but only if the user's menu search problem can be defined as a reinforcement learning problem. In the following section we report a model that does just that.
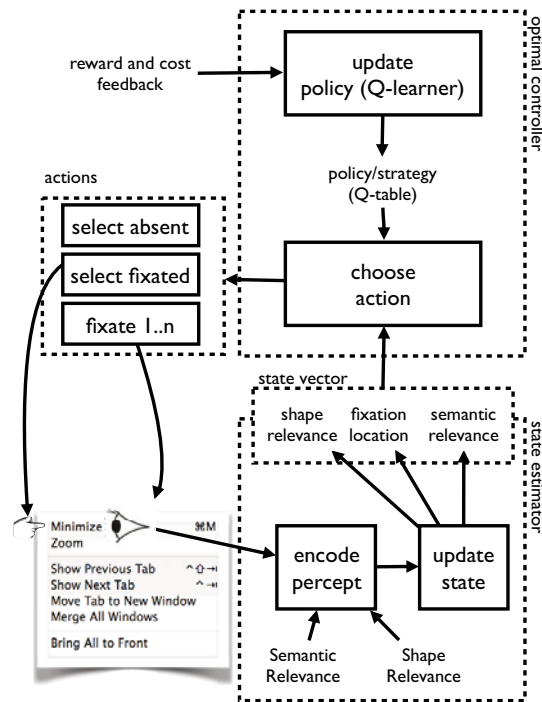

Figure 2: An overview of the adaptive menu search model.

THEORY AND MODEL

Imagine that the goal for a user who is experienced with menus, but who has never used Apple's OS X Safari browser before, is to select 'Show Next Tab' from the Safari Window menu. This task and menu are illustrated to the bottom-left of Figure 2. A user might solve this goal by first fixating the top menu item, encoding the word 'Minimize'; rejecting it as irrelevant to the target, moving the eyes to the next group of items, that begins 'Show Previous Tab', noticing that this item is not the target but is closely related and also noticing, in peripheral vision, that the next item has a similar word shape and length to the target; then moving the eyes to 'Show Next Tab', confirming that it is the target and selecting it. The aim of the modelling is that behaviours such as this should emerge from theoretical assumptions. Importantly, the aim is not to model how people learn specific menus and the location of specific items, rather the aim is to model the menu search task in general. The requirement is that the model should learn, from experience, the best way to search for new targets in new, previously unseen, menus.

To achieve this goal we use a *state estimation and optimal control* approach. In Figure 2 an external representation of the displayed menu is fixated and the state estimator encodes a percept containing information about the relevance of word shapes ('Minimise' and 'Zoom', for example have different lengths) and semantics (word meanings). This information is used to update a state vector, which has an element for the shape relevance of every item in the menu, an element for the semantic relevance of every item in the menu, and an element for the current fixation location. The vector items are null until estimates are acquired through visual perception. Updates are made after every fixation, e.g. after fixating 'Minimize' in the above example. After having encoded new information through visual perception, the optimal controller chooses an action on the basis of the available state estimate and the *strategy* (i.e., the policy that determines a state-action value function). The chosen action might be to fixate on another item or to make a selection, or to exit the menu if the target is probably absent. State-action values are updated incrementally (learned) as reward and cost feedback is received from the interaction. The menu search problem is thereby defined as a reinforcement learning problem [38].

The paragraph above offers only a very brief overview of the theory and it leaves out many of the details. In the following subsections more detail is provided about how the state estimation and optimal controller work. Subsequently a model walkthrough is provided.

State estimator

The state estimator (the bottom right of Figure 2) encodes semantic, alphabetic and shape information, constrained by visual and cognitive mechanisms.

Semantic relevance

In common with many previous models of menu search [8, 15, 28, 34, 33], our model assumes that people have an ability to determine the semantic relevance of items by matching them to the goal specification. To implement this assumption, we used average pairwise relevance ratings gathered from human participants (which are taken from [2]). These relevance ratings are described in detail below. For now, consider the following example: if the model sampled the goal Zoom and

foveated the word Minimize then it could look-up the relevance score 0.75 which was the mean relevance ranking given by participants. The level of this relevance score will only acquire meaning (whether it is considered a good or bad match) during learning. We also assume that people are able to maintain a short term representation of the semantic relevance of items that have been perceived. These are encoded in the state representation. No capacity limitations are assumed in this version of the theory.

Note that while the pairwise relevance ratings provide the model with a substantial database of information they do not specify the actions that should be taken on the basis of this information. How the best actions are found is described below in *The Optimal Controller* section.

Alphabetics and Shape relevance

The shape of each menu item was represented by its length in characters. No effort was made to model the shape of individual characters. The shape relevance had two levels, [0 for non-target length; 1 for target length]. The alphabetic relevance of two items was determined using the distance apart in the alphabet of their first letters. This was then standardised to a four-level scale between 0 and 1, i.e., [0, 0.3, 0.6, 1].

Saccade duration

The saccade duration D (in milliseconds) was determined with the following equation [4]:

$$D = 37 + 2.7A \qquad (1)$$

where A is the amplitude (in terms of visual angle in degrees) of the saccade between two successive fixations.

Fixation duration

It is known that the average fixation duration for reading is 200-250ms [35]. However, menu search involves some matching process and so some additional latency per menu item gaze is expected. In fact, in a typical menu search task the mean duration of item gazes was reported as about 400ms [7] and this is the fixation duration assumed in our model.

Peripheral vision

Visual acuity is known to reduce with eccentricity from the fovea. In our model, the acuity function was represented as the probability that each visual feature of the item was recognised. Our model made use of semantic features and shape features but could easily be enhanced with other features such as colour. Semantic acuity (the semantic relevance of the item to the target) was specified as being available within $1°$ of the current fixation [23]. The model predictions were compared with the empirical data from Bailly et al. [2]. In their experiment, the height of the items in the menu was about $0.7°$. Hence, our model assumed that the semantic of item was obtained only when it was fixated. Shape acuity was specified as a quadratic psychophysical function from [25]. This function was used to determine the availability of the shape of the item based on the eccentricity and the size of the

item (Equation 2).

$$\begin{cases} P(available) & = P(s + X > threshold) \\ threshold & = ae^2 + be + c \\ X \sim \mathcal{N}(s, v \times s) \end{cases} \qquad (2)$$

where, s is the item size; e is eccentricity; X is random noise with standard deviation $v \times s$ (v is a constant).

The parameters in Equation 2 were chosen so as to fit the materials used in Bailly et al. [2]. In their experiment, the height of an item was 0.75 cm; the distance of the users eyes from the screen was 65 cm. Participants should therefore have been able to simultaneously gather information about the shape of 3 items given a fovea of $2°$. Parameters for Equation 2 were set as follows, $v = 0.7$, $b = 0.1$, $c = 0.1$, $a = 0.3$ and $s = 0.75$. These parameter settings resulted in the following availability probabilities: 0.95 for the item fixated, 0.89 for items immediately above or below the fixated item, and 0 for items further away. On each fixation, the availability of the shape information was determined by these probabilities.

The optimal controller: Strategy/Policy Learning

The function of the optimal controller (top right in Figure 2) is to choose which action to do next (e.g. to fixate or select an item). It does so by looking up a value for each action available and picking one. These values are called *q-values*, or state-action values, and are stored in a q-table. A row in the q-table might, for example, specify that in a state with many low-relevance items ending the search by exiting the menu has a high value. One important question concerns how these q-values are learnt and why they implement the optimal menu search policy. However, to answer this question we must first briefly describe how learnt q-values are used to control search. Doing so requires us to introduce the concept of a Markov Decision Process (MDP); see Figure 3.

Recall the task illustrated in Figure 2 of selecting 'Show Next Tab'. The model starts in state $s1$ (Figure 3), which is the *initial state*. As no visual information has been encoded the state vector [semantic relevance, fixation, word shape relevance] is null (See the top row of Table 1).

The next step is to choose an action. From each state the available actions include: an action for fixating on each menu item, an action for selecting the fixated item, and an action for stopping the search because the item is believed to be absent. In the figure, the state-action values (q-values) are represented on the arcs from states (red circles) to actions (green circles). (How these q-values are learnt is explained below but for now assume the values presented in the figure.) If the menu searcher is greedy then it will select the most valuable actions as represented by larger q-values. It follows that, using the MDP in the figure, the model will first fixate on the first menu item; that is it will choose the 'fixate 1' action (with value q=8) in the uppermost row of green actions in Figure 3 and as a consequence it would fixate the 'Minimize' item in Figure 2.

Having fixated 'Minimize' the model encodes the semantic and shape relevance for the fixated item. The goal is 'Show
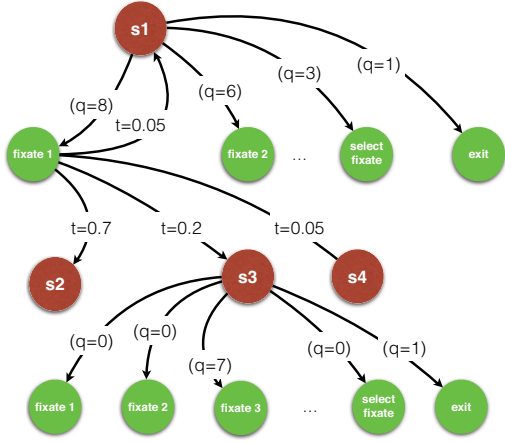
Figure 3: A part of a solution to a Markov Decision Process (MDP) for searching the Safari Window menu. Red circles labelled 's' represent states. Green circles represent actions. 'q' values represent learned state-action values. 't' values represent state-action to state transition probabilities. Action 'fixate 1' is the consequence of choosing the highest value action. The model subsequently transitions to state 's3' with probability 0.2.

Action	Semantic relevance	fixation	Shape relevance
start	N, N, N, N	N	N, N, N, N
fixate 1	$0, N, N, N$	1	$0, 0, N, N$
fixate 3	$0, N, 0.3, N$	3	$0, 0, 1.0, 1.0$
fixate 4	$0, N, 0.3, 1.0$	4	$0, 0, 1.0, 1.0$
select 4			

Table 1: Example changes in the state representation for a 4 item menu ($N = null$). The semantic relevance had 5 levels [Null, 0, 0.3, 0.6, 1].

Assumption	Description
Utility	Utility = $10000 \times correct - 10000 \times error - time$, where *correct* and *error* are Boolean variables; the unit of *time* is *ms*.
Ecology	The ecology concerns the statistical structure of menus. Menus, in the world, have a distribution of length (number of items) and a distribution of group size. Menu items have a distribution of semantic relevance and shape/length (number of chars). See Figure 4.
Mechanism	People can estimate the semantic relevance of the foveated menu item. They can also estimate the shape/length of items in the periphery, although acuity decreases with eccentricity. See Equations 1 and 2.
Strategy	A strategy (or policy) for menu search is optimised to Utility, Ecology and Mechanism assuming a state space that consists of the relevance vectors and the fixation.

Table 2: Summary of assumptions for the adaptive model of menu search.

Table 1 gives an example of how the state vector is updated as information is gathered through eye movements and visual perception.

State and action space

For a set of menus that have a maximum of n items, a state is represented as a vector V with $2 \times n + 1$ elements. This consists of n elements for the shape, n for the semantic relevance, and 1 for the fixation location. The semantic/alphabetic relevance had 5 levels [Null, 0, 0.3, 0.6, 1]. The shape relevance had 2 levels [0 for non-target length; 1 for target length]. The fixation was an integer representing one of the menu item locations [1..n]. From each state there were $n + 2$ actions, including n actions for fixating on n menu item locations, an action for selecting the fixated item, and an action to exit the menu without selection (target absent).

Learning

The q-values, as illustrated in Figure 3, represent control knowledge and are learnt with a reinforcement learning algorithm. The reinforcement learning algorithm used was a standard implementation of Q-learning. We do not present the details of the algorithm here but they can be found in any standard Machine Learning text (e.g. [38]).

Q-learning requires reward and cost feedback so as to learn the state-action values. A state-action value can be thought of as a prediction of the future reward minus costs that will accrue if the action is taken. The rewards and costs must be combined into a single utility score. We assumed a simple utility function in which there is a reward for success, a penalty for an error, and a penalty for time. The time cost is measured in milliseconds and is determined by the visual information processing and motor assumptions and by the number of fixations made. The reward for a correct menu item selection was $+10000$; the penalty for an incorrect selection

Next Tab' and so the relevance of 'Minimize' relative to this goal is likely to be low. The model also encodes shape relevance about neighbouring items in accordance with Equation 2. However, the exact values encoded are subject to noise because visual information processing is uncertain. As a consequence, the model might end up in either state $s2$, $s3$, or $s4$, or perhaps back in state $s1$ if no new information was encoded after the previous action. The transition probabilities to these states, t, are shown in the figure. These transition probabilities are a consequence of the interaction between the model of the visual system and the external environment. After any action the model is more likely to transition to a state that is (a) more likely in the environment and (b) more likely to be encoded by the perceptual model.

Subsequently, assuming that the model has transitioned to state $s3$ representing low semantic and low shape relevance, then the highest value action from this state is to fixate on item 3 (q=7). In other words, because of the low relevance of item 1, the model has skipped item 2 and focused on an item in the next semantic group.

was -10000. These numbers are large so as to emphasise accuracy over time. Indeed in the studies reported below the model achieved 99% accuracy.

Before learning, an empty Q-table was assumed in which all state-action values were zero. The model therefore started with no control knowledge and action selection was entirely random. The model was then trained until performance plateaued (requiring 20 million trials). On each trial, the model was trained on a menu constructed by sampling randomly from the ecological distributions of shape and semantic/alphabetic relevance defined below. The model explored the action space using an ϵ-greedy exploration. This means that it exploited the greedy/best action with a probability $1-\epsilon$, and it explored all the actions randomly with probability ϵ. q-values were adjusted according to the reward and cost feedback. The optimal policy acquired through this training was then used to generate the predictions described below. To do so the optimal policy was run on a further 10,000 trials of newly sampled menus, and its performance was recorded.

While we used Q-learning, any MDP solver that is guaranteed to converge on the optimal policy is sufficient to derive the rational adaptation [38]. The Q-learning process is not a theoretical commitment. Its purpose is merely to find the optimal policy. It is not to model the process of learning and is therefore used to achieve methodological optimality and determine the computationally rational strategy [16, 27].

In summary, Q-learning was used to learn (or estimate) the value of each state-action pair by simulated experience of interaction with a distribution of menu tasks, where the interaction is mediated by the theory of visual perception and knowledge. The optimal policy is then the greedy policy given the q-values. The assumptions are summarised, briefly, in Table 2 and explained fully below. (The model was implemented in Matlab and can be downloaded on request from the first author.)
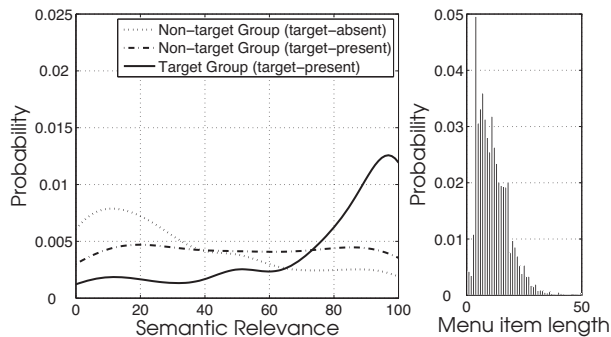
Figure 4: Menu ecology of a real-world menu task (Apple OS X menus). Left panel: The distribution of semantic relevance. Right panel: The distribution of menu length.

STUDY 1: PREDICTING REAL-WORLD MENU SEARCH

The purpose of this first study was to examine the model's predictions on commonly used computer application menus. The task was to search for specific target items in a vertically arranged menu. How items were organised in the menu was
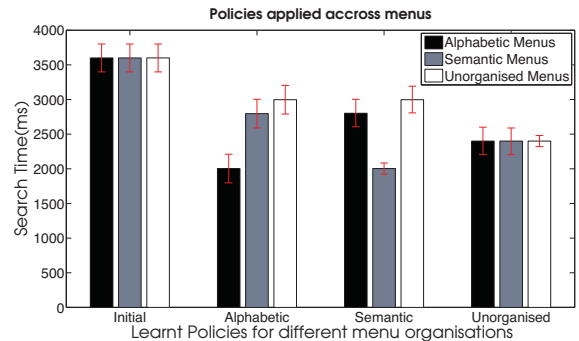
Figure 5: The search duration taken by the optimal strategy for each type of menu (95% confidence intervals (C.I.s))

varied. Items could be unorganised, alphabetically organised, or semantically organised. To determine the statistical properties of a menu search environment we used the results of a previous study [1] in which the menus of 60 applications from Apple OS X were sampled. Together these applications used a total 1049 menus, and 7802 menu items. We used these to determine the ecological distribution of menu length, item length, semantic group size and first letter frequencies (for alphabetic search). The probability of each length (number of characters) is shown in Figure 4 right panel. The mean item length was 11.5, the median was 10 and the standard deviation was 6.82. The most frequent menu item length was 4 characters. As should be expected the distribution is skewed, with a long tail of low probability longer menu items.

We then ran a study in which we asked 31 participants to rate how likely two menu items were to appear close together on a menu. Each participant rated 64 pairs that were sampled from the Apple OS X Safari browser menu items. The probability of each semantic relevance rating is shown in Figure 4 left panel. These ratings were used both to construct example menus and to implement the model's semantic relevance function.

Results

All results are for the optimal strategy (after learning), unless stated otherwise. The optimal policy achieved 99% selection accuracy. The utility of all models plateaued, suggesting that the learned strategy was a good approximation to the optimal strategy.

Search duration

Figure 5 is a plot of the duration required for the optimal policy to make a selection given four types of experience crossed with three types of test menu. The purpose of this analysis is to show how radically different patterns of behaviour emerge from the model based on how previously experienced menus were organised. Prior to training (the left most *Initial* set of three bars in the figure), the model offers the slowest performance; it is unable to take advantage of the structure in the alphabetic and semantic menus because it has no control knowledge. After training on a distribution of *Unorganised* menus (far right set in the figure), performance time is better
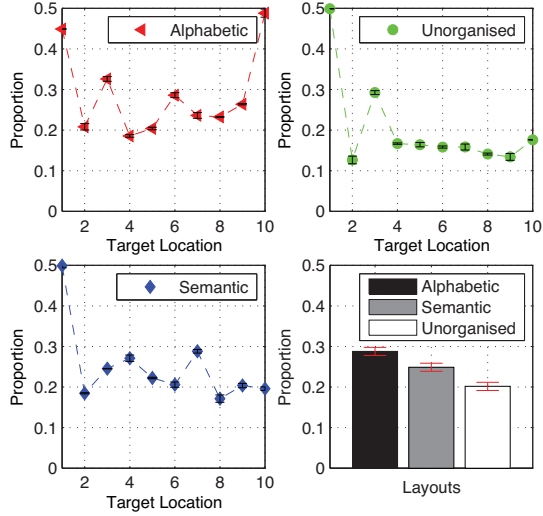
Figure 6: The proportion of gazes on the target location for each type of menu (95% C.I.s).
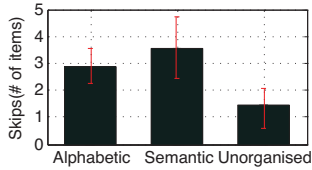

Figure 7: The model's prediction of skipping (mean gap between fixations).

Figure 8: The effect of semantic group size (95% C.I.s).

Figure 9: Menu ecology for the experimental menu search task. Left panel: The distribution of semantic relevance. Right panel: The length distribution.

than prior to training. However, there is no difference in performance time between the different menu organisations. After training on a distribution of semantically organised menus (middle right set in the figure), the model is able to take advantage of semantic structure, but this training is costly to performance on alphabetic and unorganised menus. After training on a distribution of alphabetically organised menus (middle left set in the figure), the model is able to take advantage of alphabetic ordering, but this training is again costly to the other menu types. The optimal policy must switch the policy depending on the menu type.

Gaze distribution
Figure 6 shows the effect of the menu organisation/layout on the distribution of gazes landing on target items. This is one measure of the effectiveness of each menu organisation for finding the target. A higher proportion of gazes on the target suggests a better organisation for search. The plots show the advantage of the alphabetic and semantic menus over the unorganised menus. The reason that there is higher proportion of gaze on the target item in the alphabetic and semantic menus is that in the emergent policy more low relevance items could be skipped (See Figure 7).

Effect of semantic grouping
Figure 8 shows the effect of different semantic groupings on performance time. It contrasts the performance time predic-
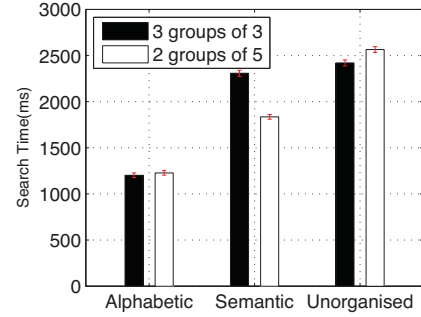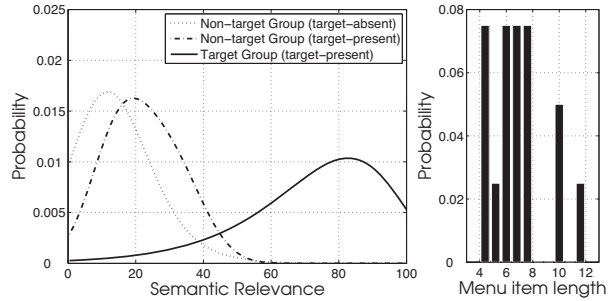
tions for menus that are organised into 3 groups of 3 or into 2 groups of 5. The contrast between these kinds of design choices has been studied extensively before (See e.g. [28]). What has been observed is an interaction between the effect of longer menus and the effect of the number of items in each semantic group (See [28] Figure 8). As can be seen in Figure 8 while the effect of longer menus ($3 \times 3 = 9$ versus $2 \times 5 = 10$) is longer performance times in the unorganised and alphabetic menus, the effect of organisation (3 groups of 3 versus 2 of 5) gives shorter performance times in the semantic condition. This prediction corresponds closely to a number of studies (See [28] Figure 8).

Discussion
The results show that deriving strategies using a reinforcement learning algorithm, given plausible assumptions about the menu search problem (Table 2), can lead to reasonable predictions about human behaviour in ecologically valid task environments. In the following section we test these predictions against data from a study of human participants using a small set of representative menus.

STUDY 2: TESTING THE MODEL AGAINST HUMAN DATA
While the previous model demonstrated the viability of the approach it did not provide an opportunity to test the predictions against human data. In this section we test the model against two previously reported data sets [2, 7]. The data set reported by Bailly et al. [2] has the advantage that it includes eye movement and task performance time data. However, it
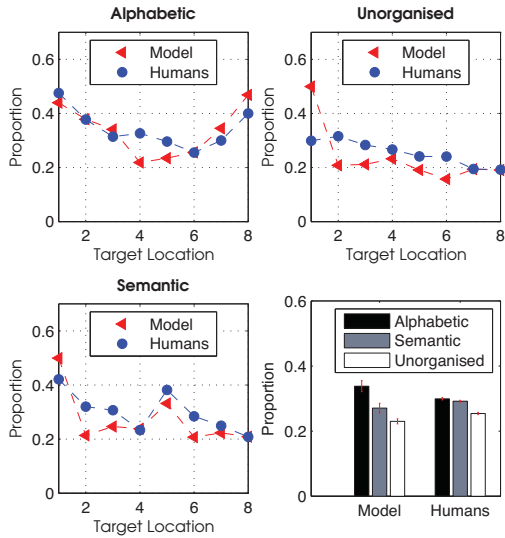
Figure 10: The proportion of gazes on the target location for each of the three types of menu (95% **C.I.s**).
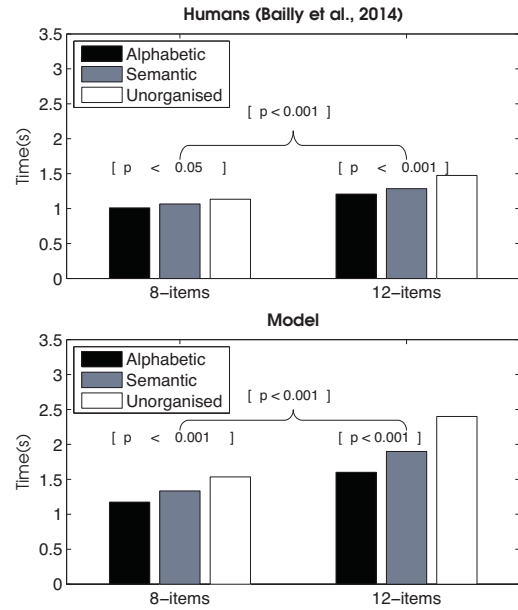

Figure 11: Search time for 8 and 12 item menus each organised in three different ways. Top panel for humans [2], bottom panel for the model. The predict effects are long but the directions are confirmed.

has the disadvantage that the distributions of menu length, menu item length and semantics do not correspond to the ecological distributions. Therefore, while this model shares all of the cognitive, visual, and utility assumptions with the above model, it does not share the same assumptions about the environment. Instead, for the experimental environment the distributions of semantic relevance and menu length was determined using relevance ratings reported in [2]. In addition, the experiment menus were 8-items (2 groups of 4) and 12-items (3 groups of 4). The distribution is plotted in Figure 9. The data set reported by Brumby et al. [7] reveals the effect on search of whether or not the exact target word is known.

As with the previous model, this model was trained on distributions of menus and menu items sampled from the ecological distributions. It was trained on 20 million samples. The optimal policy was then used to generate the predictions described in the following results section.

Results
We report a range of effects predicted by the optimal policy and compare them to the human data. As before, the optimal policy achieved 99% accuracy which corresponded with the reported participant accuracy level [2]. We report effects of menu layout on performance time and also effects on gaze distribution. The gaze distributions provide evidence that both model and people strategically adjust behaviour to the particular menu organisation.

Target location effect on gaze distribution
Figure 10 shows the effect of target position on the distribution of item gazes for each menu organisation. The model is compared to human data reported in [2]. The adjusted R^2 for each of the three organisations (alphabetic, unorganised, semantic) are $0.84, 0.65, 0.80$. In the top left panel, the model's gaze distribution is a consequence of both alphabetic anticipation and shape relevance in peripheral vision. Interestingly,

both the model and the participants selectively gazed at targets at either end of the menu more frequently than targets in the middle. This may reflect the ease with which words beginning with early and late alphabetic words can be located. In the top right panel, there is no organisational structure to the menu and the model's gaze distribution is a consequence of shape relevance only in peripheral vision. The model offers a poor prediction of the proportion of gazes on the target when it is in position 1, otherwise, as expected, the distribution is relatively flat in both the model and the participants. In the bottom left panel, the model's gaze distribution is a function of semantic relevance and shape relevance. Here there are spikes in the distribution at position 1 and 5. In the model, this is because the emergent policy uses the relevance of the first item of each semantic group as evidence of the content of that group. In other words, the grouping structure of the menu is evidence in the emergent gaze distributions. The aggregated data is shown in the bottom right panel; the model predicts the significant effect of organisation on gaze distribution, although it predicts a larger effect for alphabetic menus than was observed.

Effect of length (8 v 12) on duration
It is known that people take longer to search menus with more items. Figure 11 shows that the model predicts the effect observed by [2].

Effect of known- versus unknown-target on duration
Using the same materials as [2], Brumby et al. [7] have shown that people are slower at searching menus when given only a semantic description of the target (rather than being told the exact target word). To capture this effect, we implemented

a variant of the model that was given a semantically related word rather than the target word. The model could therefore not make use of the shape of items located in peripheral vision during search. All other assumptions remained the same. This meant that the model had to directly fixate items in order to assess them. We found that this variant of the model performed slower (M = 2183 ms, SD = 777 ms) than the standard model (M = 1458 ms, SD = 760 ms). It predicted known-versus unknown-target effect observed in [7].

DISCUSSION

We have reported a model in which search behaviours were an emergent consequence of a combination of three sources of constraint: the ecological structure of interaction, the cognitive and perceptual limits, and the goal to maximise the trade-off between speed and accuracy. The model was tested with two studies of its predictions. The first study involved applying the model to a real world distribution of menu items and in the second study the model was compared to human data from a previously reported experiment. The model was thereby tested against existing empirical findings concerning the effect of menu organisation, menu length, and whether or not the target is a known word. The predictions of the model were largely supported by the experimental evidence.

Unlike in previous models, no assumptions were made about the gaze strategies available to or adopted by users, rather the menu search problem was specified as a Markov Decision Process (MDP) and behaviour emerged from solving the reinforcement learning problem. The states of the MDP were defined by a theory of foveal and peripheral vision that encoded information about semantic and word shape relevance. Unlike production system based approaches (ACT-R and EPIC) there are no rules programmed by the modeller, rather the strategy is emergent from the constraints.

While the reported models demonstrate that it is possible to predict user behaviour with both a real world and a laboratory menu system the method also has potential to predict performance with interesting design variations. It could be used, for example, to verify performance predictions with automatically designed interfaces [3]. It could also be extended to predict visual search of icons [25] or any other means of laying out options spatially on a display. It could also be used to predict rational strategies in aimed movement required for using a mouse or a touch surface [31, 39].

Also, while we have reported a model of menu search, the theory that underpins the model and much of the way in which the problem was solved using machine learning, is potentially general to many modelling problems in HCI. Indeed elements of this argument have been made by a number of authors [12, 31, 33, 41]. A key property of the approach is that behavioural predictions are derived by maximising utility given a quantitative theory of the constraints on behaviour, rather than by maximising fit to the data. Although this *optimality* assumption is sometimes controversial, the claim is simply that users will do the best they can with the resources that are available to them. Further discussions of this issue can be found in [21, 27, 31].

Finally, there are many issues that need to be resolved. For example, the conversion of the Partially Observable Markov Decision Process (POMDP) into an MDP by assuming a psychologically plausible state estimation is known to fail to achieve an exact solution to the POMDP. Further work is required to understand the properties of a model in which, unlike ours, an optimal action is found for each possible belief over the states of the world (a belief state). Further work is also required to understand how this approach can be extended to capture the range of phenomena associated with skilled menu use [13]. Scalability is also a concern and the analyses presented in the paper required substantial computation. As noted above, the semantic relevance had 5 levels and shape had 2 levels. Therefore, the state space for an 8-item menu is $5^8 \times 2^8 \times 8$ states. But, this is an upper-bound and because the menus have structure, the state space experienced during training is much smaller.

REFERENCES

1. Bailly, G., Malacria, S., et al. Menuinspector: Outil pour l'analyse des menus et cas d'étude. In *25ème conférence francophone sur l'Interaction Homme-Machine, IHM'13* (2013).

2. Bailly, G., Oulasvirta, A., Brumby, D. P., and Howes, A. Model of visual search and selection time in linear menus. In *ACM CHI'14* (2014), 3865–3874.

3. Bailly, G., Oulasvirta, A., Kötzing, T., and Hoppe, S. Menuoptimizer: Interactive optimization of menu systems. In *ACM CHI'13* (2013), 331–342.

4. Baloh, R. W., Sills, A. W., Kumley, W. E., and Honrubia, V. Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology 25*, 11 (1975), 1065–1065.

5. Baron, S., and Kleinman, D. L. The human as an optimal controller and information processor. *Man-Machine Systems, IEEE Transactions on 10*, 1 (1969), 9–17.

6. Brumby, D., and Howes, A. Strategies for guiding interactive search: An empirical investigation into the consequences of label relevance for assessment and selection. *Human-Computer Interaction 23*, 1 (2008), 1–46.

7. Brumby, D. P., Cox, A. L., Chung, J., and Fernandes, B. How does knowing what you are looking for change visual search behavior? In *ACM CHI'14* (2014), 3895–3898.

8. Brumby, D. P., and Howes, A. Good enough but i'll just check: Web-page search as attentional refocusing. In *ICCM* (2004), 46–51.

9. Brumby, D. P., Salvucci, D. D., and Howes, A. Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while driving. In *ACM CHI'09* (2009), 1629–1638.

10. Byrne, M. D. ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies 55*, 1 (2001), 41–84.

11. Charman, S. C., and Howes, A. The adaptive user: an investigation into the cognitive and task constraints on the generation of new methods. *Journal of experimental psychology. Applied 9*, 4 (2003), 236–48.

12. Chen, X, Howes, A., Lewis, R.L., Myers, C.W., Houpt, J. Discovering computationally rational eye movements in the distractor ratio task. In *RLDM* (Princeton, 2013).

13. Cockburn, A., Gutwin, C., and Greenberg, S. A predictive model of menu performance. In *ACM CHI'07* (2007), 627–636.

14. Fu, W.-T., and Gray, W. D. Resolving the paradox of the active user : stable suboptimal performance in interactive tasks. *Cognitive Science 28* (2004), 901–935.

15. Fu, W.-T., and Pirolli, P. SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human–Computer Interaction 22*, 4 (2007), 355–412.

16. Gray, W. D., Sims, C. R., Fu, W.-T., and Schoelles, M. J. The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological review 113*, 3 (2006), 461–82.

17. Halverson, T., and Hornof, A. J. A minimal model for predicting visual search in human-computer interaction. In *ACM CHI'07*, ACM (2007), 431–434.

18. Halverson, T., and Hornof, A. J. A computational model of active vision for visual search in human–computer interaction. *Human–Computer Interaction 26*, 4 (2011), 285–314.

19. Hayhoe, M., and Ballard, D. Modeling task control of eye movements. *Current Biology 24*, 13 (2014), 622–628.

20. Hornof, A. J., and Halverson, T. Cognitive strategies and eye movements for searching hierarchical computer displays. In *ACM CHI'03*, ACM (2003), 249–256.

21. Howes, A., Lewis, R. L., and Vera, A. Rational adaptation under task and processing constraints: implications for testing theories of cognition and action. *Psychological Review 116*, 4 (2009), 717–751.

22. Howes, A., Vera, A., Lewis, R. L., and McCurdy, M. Cognitive constraint modeling: A formal approach to supporting reasoning about behavior. *Proc. Cognitive Science Society* (2004), 595–600.

23. Inhoff, A. W., and Rayner, K. Parafoveal word perception: A case against semantic preprocessing. *Perception & Psychophysics 27*, 5 (1980), 457–464.

24. Janssen, C. P., Brumby, D. P., Dowell, J., Chater, N., and Howes, A. Identifying optimum performance trade-offs using a cognitively bounded rational analysis model of discretionary task interleaving. *Topics in Cognitive Science 3*, 1 (2011), 123–139.

25. Kieras, D. E., and Hornof, A. J. Towards accurate and practical predictive models of active-vision-based visual search. In *ACM CHI'14* (2014), 3875–3884.

26. Lelis, S., and Howes, A. Informing decisions: how people use online rating information to make choices. In *ACM CHI'11* (2011), 2285–2294.

27. Lewis, R., Howes, A., and Singh, S. Computational rationality: linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science 6*, 2 (2014), 279–311.

28. Miller, C. S., and Remington, R. W. Modeling information navigation: Implications for information architecture. *Human-Computer Interaction 19*, 3 (2004), 225–271.

29. Nunez-Varela, J., and Wyatt, J. L. Models of gaze control for manipulation tasks. *ACM Transactions on Applied Perception (TAP) 10*, 4 (2013), 20.

30. O'Hara, K. P., and Payne, S. J. The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive psychology 35*, 1 (1998), 34–70.

31. Payne, S. J., and Howes, A. Adaptive Interaction: A Utility Maximization Approach to Understanding Human Interaction with Technology. *Synthesis Lectures on Human-Centered Informatics 6*, 1 (2013), 1–111.

32. Payne, S. J., Howes, A., and Reader, W. R. Adaptively distributing cognition: a decision-making perspective on human-computer interaction. *Behaviour & Information Technology 20*, 5 (2001), 339–346.

33. Pirolli, P. *Information foraging theory: Adaptive interaction with information*, vol. 2. OUP, USA, 2007.

34. Pirolli, P., and Card, S. Information foraging. *Psych Review 106* (1999), 643–675.

35. Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psych. Bulletin 124*, 3 (1998), 372.

36. Russell, S., and Subramanian, D. Provably bounded-optimal agents. *Journal of Artificial Intelligence Research 2* (1995), 575–609.

37. Sprague, N., and Ballard, D. Eye movements for reward maximization. *Advances in neural information processing systems 16* (2004), 1419–1433.

38. Sutton, R. S., and Barto, A. G. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

39. Trommershäuser, J., Maloney, L. T., and Landy, M. S. The Expected Utility of Movement. *Brain* (2009), 95–111.

40. Tseng, Y.-C., and Howes, A. The adaptation of visual search strategy to expected information gain. In *ACM CHI '08* (2008), 1075–1084.

41. Vera, A., Howes, A., McCurdy, M., and Lewis, R. L. A constraint satisfaction approach to predicting skilled interactive cognition. In *ACM CHI'04* (2004), 121–128.

42. Zhang, Y., and Hornof, A. J. Understanding multitasking through parallelized strategy exploration and individualized cognitive modeling. In *ACM CHI'14* (2014), 3885–3894.