Classification of Human Decision Behavior: Finding Modular Decision Rules with Genetic Algorithms

Franz Rothlauf, Daniel Schunk and Jella Pfeiffer

79-2005

January 2005

MEA – Mannheimer Forschungsinstitut Ökonomie und Demographischer Wandel
Gebäude L 13, 17_D-68131 Mannheim_Sekretariat +49 621 181-2773/1862_Telefax +49 621 181-1863_www.mea.uni-mannheim.de

Classification of Human Decision Behavior: Finding Modular Decision Rules with Genetic Algorithms

Franz Rothlauf
Dept. of Business Administration and Information Systems
University of Mannheim
D-68131 Mannheim/Germany
rothlauf@uni-mannheim.de

Daniel Schunk
Mannheim Research Institute for the Economics of Aging, Department of Economics
University of Mannheim
D-68131 Mannheim/Germany
dschunk@uni-mannheim.de

Jella Pfeiffer
Dept. of Business Administration and Information Systems
University of Mannheim
D-68131 Mannheim/Germany
jella@gmx.de

January 23, 2005

Abstract

The understanding of human behavior in sequential decision tasks is important for economics and socio-psychological sciences. In search tasks, for example when individuals search for the best price of a product, they are confronted in sequential steps with different situations and they have to decide whether to continue or stop searching. The decision behavior of individuals in such search tasks is described by a search strategy.

This paper presents a new approach of finding high-quality search strategies by using genetic algorithms (GAs). Only the structure of the search strategies and the basic building blocks (price thresholds and price patterns) that can be used for the search strategies are pre-specified. It is the purpose of the GA to construct search strategies that well describe human search behavior. The search strategies found by the GA are able to predict human behavior in search tasks better than traditional search strategies from the literature which are usually based on theoretical assumptions about human behavior in search tasks. Furthermore, the found search strategies are reasonable in the sense that they can be well interpreted, and generally that means they describe the search behavior of a larger group of individuals and allow some kind of categorization and classification.

The results of this study open a new perspective for future research in developing behavioral strategies. Instead of deriving search strategies from theoretical assumptions about human behavior, researchers can directly analyze human behavior in search tasks and find appropriate and high-quality search strategies. These can be used for gaining new insights into the motivation behind human search and for developing new theoretical models about human search behavior.

1 Introduction

The study of individuals' decision behavior in search situations is important for the economic and socio-psychological sciences. A common and intuitive example for search tasks is taken from consumer economics: How do people behave when they want to find the best price for an item that they want to buy? There are costs associated with visiting each store and there is an optimal number of search steps that maximizes the profit of the human searcher. The search strategy of an individual describes when the individual stops searching for a better price. Unfortunately, humans in the real world do not behave as described by analytical models since they are in most cases not able to compute the optimal solution. Therefore, search strategies that allow us to predict and explain human search behavior are important.

The different approaches for predicting and classifying human behavior in sequential decision situations can be categorized in two different groups. Traditional methods [16, 2, 3] which use a set of pre-specified decision rules and are based on theoretical explanation models for human behavior. Human behavior in decision tasks is classified according to these sets of decision rules. Finding appropriate decision rules means searching for rules (from the set of pre-specified rules) that best describe human behavior. These methods are efficient if the pre-specified rules describe human decision behavior well. Newer approaches [4, 12] choose a different way that is less restrictive concerning the nature of the search strategies. Only the basic structure of decision rules is pre-specified and decision rules that explain human behavior are derived from the observed empirical data.

This paper presents an approach on how to derive decision rules (search strategies) for human behavior in search tasks by the use of genetic algorithms. The paper assumes that only the basic structure of the search rules is pre-specified and the strategies are constructed based on the observed human behavior; therefore, the proposed approach is less restrictive concerning the character of the rules than existing approaches. For finding appropriate search strategies a traditional simple genetic algorithm [6] is used. The purpose of this paper is to present how complex search strategies can be created from a set of basic "building blocks" by the use of a genetic algorithm and to compare the prediction quality of the resulting search strategies to existing standard search strategies from the literature. Furthermore, the paper investigates the search strategies constructed from basic building blocks and examines the relationship between finding general rules that describe the search behavior of a larger group of individuals well and specific rules that are only used by a few individuals.

The paper is organized as follows: in the following section, we define human behavior in search tasks and discuss how the predictive quality of different search strategies can be measured. The basic building blocks that are used to create human search strategies are presented in section 3. Section 4 describes the laboratory experiment on human search behavior which provides the data for the investigations, and presents details of the genetic algorithm which is used for finding appropriate search strategies. The results of the experiments are shown in section 5. The paper ends with concluding remarks.

2 Human Decision Behavior in Search Tasks

Describing the behavior of humans in complex decision situations is of interest to economics, for example in marketing science for determining price behavior of consumers [23] or in labor economics for explaining human job search behavior [1].

In search tasks, humans (searchers) iteratively face different situations. In each situation the searcher gets some amount of reward and has to decide whether to stop or to continue the search. Furthermore, there are search costs implied by every search step. The goal of the searcher is to maximize its profit which is the difference between the reward resulting from the different alternatives that are observed during the search process, and the search cost, which depend on the number of search steps. A common example of a search task is comparing the price of an item in different stores. The price of the item is different in each store and search costs are associated with visiting a store.

Formally, we want to assume that a searcher sequentially observes a number of realizations x_i of a random variable X which has the cumulative distribution function $F(x)$. $F(x)$ is a discrete normal distribution with mean μ_X and standard deviation σ_X and describes for example the price x of a product in different locations. $i \in \{1, \dots, t\}$ denotes the number of the search step. The cost c of each search step is constant. We want to assume that the searcher can assess in search step t all previous situations x_i, where $i \leq t$, without additional costs. This means for the example of finding the lowest price of an item that the searcher can go back to a store visited in earlier search steps without additional costs. Therefore, the searcher has to decide in each search step t whether she wants to continue the search or to stop and choosing x_i, where $i \in \{1, \dots, t\}$. If the searcher stops after t steps, she chooses the lowest price of the item, i.e. she buys the item at the price $x_{min} = \min\{x_1, \dots, x_t\}$.

Basic search theory assumes that individuals treat the cost of each search step, once completed, as sunk costs [15, 14]. Therefore, to decide whether to continue the search process in iteration t, an individual compares the cost c of one additional search step to the expected benefit. It will only continue if the expected benefit is greater than the cost of the additional search step. Then, subjects solve the problem based on a one-step forward-induction strategy. The expected benefit G from searching one more step can be calculated as:

$$G = x_{min} - c - \underbrace{(1 - F(x_{min}))\, x_{min}}_{A} - \underbrace{\int_{-\infty}^{x_{min}} x dF(x)}_{B}, \qquad (1)$$

where $x_{min} = \min\{x_1, \ldots, x_t\}$. There are two different cases for the variable x_{t+1} observed in the next search step. x_{t+1} is either larger or lower than x_{min}. Term A describes the case that a value x_{t+1} larger than x_{min} is found (with probability $(1 - F(x_{min})))$. In this case, $x_{min} = min\{x_1, \ldots, x_t\} = min\{x_1, \ldots, x_{t+1}\}$ remains the lowest price. Term B assumes that a value x_{t+1} lower than x_{min} is found. The expected value $x_{t+1} = \min\{x_1, \ldots, x_{t+1}\}$ is calculated as $\int_{-\infty}^{x_{min}} xdF(x)$.

As G describes the expected benefit from continuing the search, a human searcher continues the search if $G > 0$ and stops otherwise. If we assume that $x_{min} = \infty$, the expected benefit G is always greater than zero ($G > 0$) and the searcher continues the search. On the other hand, if $x_{min} = -\infty$, the expected benefit $G = -c < 0$ and the searcher stops the search. As we assume that $G(x_{min})$ is continuous and monotonic, there is an unique x^*_{min}, where $G(x^*_{min}) = 0$. Therefore, the best strategy is to stop searching at search step t if $x_t < x^*_{min}$. This means the searcher should stop searching whenever a price is below a certain threshold price x^*_{min}.

In general, x^*_{min} cannot be analytically calculated and is usually determined by numerical methods. The model presented here is simple as it assumes that a searcher only plans ahead for one search step and that she fully ignores sunk costs. However, in reality, humans do not completely ignore sunk costs and also try to predict the outcome of future search steps. For an overview over more comprehensive models describing human behavior in search tasks the reader is referred to the literature [9, 11, 13, 18].

This section presented a basic model for describing human search behavior. Based on the model, an optimal stopping criterion for the search can be derived. As already mentioned in the introduction, classical decision models use a set of decision rules that are based on such theoretical models trying to model human search behavior. However, consumers in the real world do not behave according to the theoretical models. Therefore, in section 3 this paper presents basic building blocks that can be used for constructing search strategies (and stopping criteria) that are based on the observed human behavior and not on theoretical models.

2.1 Standard Strategies in Search Tasks

Although in the previous paragraphs we described that there is an optimal stopping criterion for search tasks (stopping at time t if $x_t < x^*_{min}$), individuals do not follow this rule but in reality show a different stopping behavior. Consequently, a large number of studies [8, 10, 11, 14, 7, 19, 20, 13] investigated human behavior in search tasks and tried to identify search strategies that are used by individuals in reality. All these studies used controlled laboratory experiments where individuals (subjects) search for the best price of a product and search costs are associated with every search step. Furthermore, it is assumed that the subjects do not change their search strategy over time. For the experiments, the search tasks are repeated a certain number of times for each subject.

The goal of such investigations is to find general rules that describe the search behavior of individuals. The studies revealed a few simple search strategies, which describe the observed behavior of individuals more accurately than, for example, the optimal stopping rule described in the previous section. Surprisingly, though individuals do not follow the optimal stopping rules, their search behavior is efficient in the sense that their earnings are similar to the earnings if they would follow the optimal stopping rule. This, however, does not indicate that their search strategy is close to optimal, it indicates rather that the payoff of search experiments is not sensitive to deviations in the stopping strategy.

Based on experimental research in search behavior (e.g., [10, 17]), there are three different basic search strategies that are used by individuals in search tasks. These three search strategies have subsequently been used by most of the later approaches trying to model the behavior of humans in search tasks:

- **Constant reservation price heuristic (CRPH)**: The search is stopped in iteration t if x_t is lower than or equal to the reservation price p_r ($x_t \leq p_r$). This stopping criteria is optimal ($p_r = x^*_{min}$) if the searcher ignores sunk costs and only plans one step ahead (compare the previous paragraphs).

- **Satisficer heuristic (SH)**: The search is stopped in iteration t if either the payoff is greater than a certain threshold T, or after a maximum number of search steps t_{max}. The payoff is the difference between the profit resulting from a situation x_i, where $i \in \{1, \ldots, t\}$, and the overall search cost, tc.

- **Bounce heuristic (BH)**: The search is stopped in iteration t, where $t > 1$, either only if $x_t \geq x_{t-1}$ or only if $x_t \leq x_{t-1}$.

These different models for human search behavior can be formulated by either defining reservation prices p_r for the different search steps t (CRPH and SH), or by specifying price patterns that represent falling and rising x_i (BH). For further information about models describing human search behavior we refer to the literature [8, 10, 11, 13, 18].

2.2 Measuring the Quality of Search Strategies

The quality of a search strategy is determined by how well it predicts the observed human behavior in real-world search tasks. The quality of a search strategy is high if individuals decide according to the search strategy and low otherwise. In search tasks, the individuals decision is whether to continue the search or whether to stop.

The quality of a search strategy can be measured as follows: each search strategy $c_j \in \mathcal{C}$, where $\mathcal{C}$ is the set of all possible search strategies, is a unique mapping from individual i's information set S_{it} (which usually depends on t) to his continuation decision $d_{it} \in \{0,1\} : d_{it}^{c_j}(S_{it}) \to \{0,1\}$. The continuation decision is performed in each search step t and the search is continued if $d_{it}(t) = 1$ and stopped if $d_{it}(t) = 0$. Let $d_{it}^{*}(t)$ denote the observed decision of individual i in iteration t. Then, we can define the indicator function:

$$X_{it}^{c_j}(S_{it}) = \begin{cases} 1 & \text{if } d_{it}^{*} = d_{it}^{c_j}(S_{it}) \\ 0 & \text{if } d_{it}^{*} \neq d_{it}^{c_j}(S_{it}) \end{cases} \tag{2}$$

Let t_{max} be the maximum number of decisions that we observe for individual i. Then, for each individual i,

$$\hat{T}_i = \sum_{t=1}^{t_{max}} X_{it}^{c_j}(S_{it}) \tag{3}$$

is the number of decisions that are correctly explained by search strategy c_i. Therefore, the quality of a search strategy c_j can be measured by

$$\text{fit}(c_j) = \frac{\hat{T}_i}{t_{max}}, \tag{4}$$

given the observed search behavior of an individual i.

Therefore, the fitness of a search strategy measures the number of individuals' decisions that are consistent with the search strategy c_j. The higher the fitness of a search strategy, the better it allows us to predict the individuals' behavior in search tasks.

3 Building Search Strategies from Building Blocks

Section 2.1 described standard search strategies used for explaining human behavior in search tasks. They are based on the assumption that individuals behave either conditionally on some simple preference parameters such as risk attitudes [13], or according to some pre-specified heuristics. A step towards finding reasonable search strategies that are more powerful and flexible in the sense that they better fit the data is to assume that decision rules can be constructed by using simple building blocks and combining them to form complete search strategies. From a behavioral point of view, this approach can be motivated by recent economic and psychological work [5], which claims that domain-specific heuristics are composed of cognitive building blocks.

Building blocks that can be used to construct complete search strategies can be defined based on the search strategies described in section 2.1. Therefore, either *price threshold levels* (so-called "reservation prices") or *price patterns* can be used. The price threshold level or price pattern that is used by an individual during the search can be different in different stages of the search. Consequently, these building blocks can be combined to construct search strategies that allow a more accurate modeling of human search behavior. The following three elements are necessary to construct search strategies from simple building blocks (modules):

1. **price threshold module:** For each search step t a price threshold level p_t is defined. The search is stopped at step t if p_t is activated (compare 3.) and $x_t \leq p_t$.

2. **pattern module:** The decision whether to continue the search at time-point t depends on whether a specific pattern of x_i exists in the last few search steps. A possible example for such a pattern is: rising-falling-rising. The search is stopped if the last realizations of x_i follow a pre-specified pattern. When using the pattern rising-falling-rising, the search would be stopped at t if $x_t \geq x_{t-1} \leq x_{t-2} \geq x_{t-3}$.

3. **activation module:** As there are different stopping criteria (price thresholds and patterns) at each time t available, it must be defined which modules are activated at time t and influence the individuals' decision whether to stop the search or to continue. The idea is that an individual might use a reservation price in certain search steps (for example in the first three search steps), but switch after that to a pattern-based rule.

By using these simple building blocks, a complete search strategy can be constructed that can be used for describing human behavior in search tasks. In each search step, a price threshold module or a pattern module (or both) can be activated determining whether the search stops or continues. By using these modules, all standard search strategies described in section 2.1 can be modeled. For example a constant reservation price heuristic is modeled using a price threshold level $p_i = p_r$ ($i \in \{1, \ldots, t\}$) and activating the price threshold module in all search steps.

4 Experimental Design

4.1 Measuring Human Search Behavior

The data about human behavior in search tasks was collected in extensive experimental studies. In search experiments 64 human individuals (denoted as subjects) were asked to perform 10 independent search tasks. The goal of the subjects was to purchase an item at the lowest price. The price of the item follows a normal distribution X with mean $\mu_X = 500$ and standard deviation $\sigma_X = 10$. Additionally, X is truncated at $x_{low} = 460$ and $x_{high} = 540$. The subjects knew that in each search step the price was drawn randomly from the described distribution. The subjects had 500 units of money available and each search step has cost $c = 1$. They can stop in each search step t and buy the item at a price $x_{min} = \min\{x_1, \ldots x_t\}$. Their payoff can not be negative (subjects can not loose money) and is calculated as $500 - x_{min} - tc$. There are a maximum of $t_{max} = 40$ search steps possible as the overall payoff is zero for $t \geq 40$ (the lowest possible price is 460).

To ensure that subjects were experienced with search tasks and to minimize the impact of learning, subjects were allowed to perform an unlimited number of training search tasks before performing a sequence of 10 tasks that determined their payoff. After the experiment was completed, one of these 10 rounds was selected randomly and the payoff of this round was paid (in Euro) to the subjects.

All results presented in the following sections of this paper are based on the data obtained in the experiment described above.

4.2 A Genetic Algorithm for Finding High-Quality Search Strategies

Section 3 described the building blocks that can be used to build search strategies describing human behavior in search tasks. To construct high-quality search strategies a genetic algorithm (GA) was developed. The quality of the search strategies created by the GA is measured by applying the quality measure for search strategies from section 2.2 using the experimental data from section 4.1. The following paragraphs describe the encoding, the search operators, and the fitness evaluation of the GA.

4.2.1 Encoding of Search Strategies

Each individual of a GA represents one complete search strategy[1] and each search strategy is created from the basic building blocks described in section 3. Therefore, each individual must contain for each search step $t \leq t_{max}$ a price threshold p_t, the corresponding activation a_t^{thresh} for the price threshold, a possible pattern pa, which describes whether x_i is either falling or rising in subsequent search steps, and the corresponding activation $a_t^{pattern}$ for the pattern pa. Table 1 illustrates the encoding of search strategies in the genotype. Each genotype consists of three vectors of length $l = t_{max}$ and a pattern pa of maximum length $l = 4$. Two of the vectors define the threshold components (threshold value p_T and activation a_t^{thresh}) and two define possible patterns (structure of the pattern pa and activation $a_t^{pattern}$).

		genotype					
threshold	p_t	494	494	494	494	490	$\cdots$
	a_t^{thresh}	1	0	0	0	1	$\cdots$
pattern	pa	110 (rising-rising-falling)					
	$a_t^{pattern}$	0	0	0	1	1	$\cdots$

Table 1: Encoding of search strategies

The activation variables a_t^{thresh} and $a_t^{pattern}$ indicate whether the corresponding threshold value or pattern are used as stopping criteria in the tth search step. For the threshold component, p_t defines the threshold relevant in

[1]Except for the results presented in section 5.2 where one individual encodes r different search strategies.

search step t. The activation $a_t^{thresh} = \{0, 1\}$ determines whether the threshold p_t is considered for the stopping decision of the subject. If $a_t^{thresh} = 1$, the subject stops searching if $x_t \leq p_t$; if $a_t^{thresh} = 0$, the threshold p_t is not relevant in search step t. For patterns, pa describes the structure of a pattern. Each pattern has a maximum length of four and consists of a sequence of zeros and ones. A one indicates that x_t is rising and a zero indicates a falling x_t. The activation $a_t^{pattern}$ describes whether the corresponding pattern pa is relevant for the stopping decision of the individual at time t. If both building blocks, threshold and pattern, are activated at time t ($a_t^{thresh} = 1$ and $a_t^{pattern} = 1$), the individual only stops if $x_t \leq p_t$ and the pattern pa is correct in the last few search steps (logical AND).

We want to give a brief example for the construction of a search strategy from the genotype. According to the search strategy defined in table 1 the subjects stops the search after the first search step $t = 1$ if $x_1 \leq 494$. Otherwise, it continues. In the second and third search steps ($t = 2$ and $t = 3$), the individual never stops as $a_2^{thresh} = a_3^{thresh} = 0$ and $a_2^{pattern} = a_3^{pattern} = 0$. In step $t = 4$, only the pattern 110 is considered for the stopping decision and the individual stops if $x_4 \leq x_3 \geq x_2 \geq x_1$ (rising-rising-falling). For $t = 5$ the individual stops only if $x_5 \leq 490$ and x_i is falling from $t = 4$ to $t = 5$ and rising from $t = 2$ to $t = 3$ and from $t = 3$ to $t = 4$ ($x_5 \leq x_4 \geq x_3 \geq x_2$).

4.2.2 Operators

Search operators can be defined straightforwardly for the encoding defined in the previous paragraphs. Possible crossover operators are uniform [21], or n-point crossover. For our experiments we have chosen a five-point crossover to ensure a proper mixing of the alleles [22], and to consider the fact that many subjects stop searching after a few search steps and do not search $t_{max} = 40$ search steps. Therefore, to ensure a proper mixing of the alleles, crossover operators with a high number of crossover-points are necessary. We found five-point crossover a good compromise between uniform crossover, which destroys most of the sub-structures in the genotype, and one-point crossover which results in an improper mixing of the first and most meaningful alleles. The crossover probability in the experiments was set to $p_{cross} = 0.8$.

A mutation means either flipping a bit (activation variables and pattern pa), adding a random variable drawn from a Gaussian distribution with zero mean and standard deviation of two to the threshold p_t, or adding/removing a randomly chosen bit to a pattern pa. As $t_{max} = 40$, the maximum number of alleles is 124 (40 bits for both activation variables, 40 integers for the thresholds p_t, and a maximum number of four bits for the pattern pa). The mutation operator mutates all alleles with probability $1/124 \approx 0.008$.

In all experiments a tournament selection without replacement and tournament size 3 was used.

4.2.3 Fitness evaluation

The fitness of a search strategy that is encoded as described in the previous paragraphs is calculated according to section 2.2. We want to give a brief example.

t		1	2	3	4
experiment	x_t	499	498	496	495
	decision	cont.	cont.	cont.	stop
search strategy	p_t	498	500	496	500
	a_t^{thresh}	1	1	0	0
	pa	00 (last two round falling)			
	$a_t^{pattern}$	0	0	0	1
indicator funct. X_t		1	0	1	1

Table 2: Example for fitness evaluation

Table 2 presents the experimental data observed for a subject which stops the search after $t = 4$ search steps (denoted as experiment). Furthermore, the table presents an example of a search strategy and the value of the indicator function X_t (compare equation 2). For example, the search strategy says that for $t = 1$ the subject stops if $x_1 \leq 498$. However, $x_1 > p_1$ and the search strategy correctly predicts that the user continues the search ($X_1 = 1$). According to equation 3, $\hat{T}$ is calculated for the example as $\hat{T} = 3$ as the evaluated search strategy correctly predicts the subjects behavior three times. Therefore, the fitness fit of the search strategy is 0.75 ($t_{max} = 4$).

5 Results

This section presents different types of results. In section 5.1 we use the genetic algorithm for finding high-quality search strategies and compare their fitness to the existing standard search strategies from section 2.1. In the remaining sections we extend the investigation and assume that different subjects have different preferences (i.e. human beings are heterogeneous with respect to their preferences) and therefore use different search strategies. We use a GA to identify relevant search strategies and investigate how well the found search strategies predict human search behavior.

For each of the experiments we run 10 independent GA runs and present the best found search strategy. The individuals in the initial population are chosen randomly. The population size of the GA was always set to $N = 4,000$ and the GA run was stopped either after the population was fully converged, or a maximum number of $t_{conv} = 1,000$ generations. We are aware of the fact that using such a large N and t_{conv} is computationally demanding and may not be necessary to obtain good results. However, the goal of the experiments was to identify "optimal" search strategies and the computational effort was only of minor importance.

5.1 One Search Strategy Fits All

Table 3 compares the average fitness $\mu(fit)$ (compare equation 4) of the optimal (in the sense of highest fitness) constant reservation price heuristic (CRPH), the optimal satisficer heuristic (SH), and the best search strategy that was found by the genetic algorithm for the data set described in section 4.1. The fitness $\mu(fit)$ is averaged over the fitness fit of a search strategy for all 64 subjects participating in the experiment.

	GA	CRPH	SH
$\mu(fit)$	0.902	0.883	0.886

Table 3: Average fitness of the optimal search strategies

The optimal CRPH search strategy is a reservation price $x_r = 491$, which results in an average fitness of $\mu(fit_{CRPH}) = 0.883$. Therefore, averaged over all 64 subjects, about 88% of the humans search decisions are correctly predicted by the CRPH search strategy with $x_r = 491$. The optimal SH search strategy is using a payoff of five (this means the search is stopped in step t if $500 - \min\{x_1, \ldots, x_t\} - tc > 5$) and has an average fitness of $\mu(fit_{SH}) = 0.886$. The best strategy found by the GA is to use only reservation prices and no patterns (all $a_t^{pattern}$ are zero). The reservation prices p_t are decreasing with t and are found as $p_1 = 498$, $p_2 = 494$, $p_3 = 491$, $p_4 = 488$, ... The average fitness of this search strategy is 0.902 and is significantly higher than the optimal CRPH and SH search strategy.

5.2 Different Search Strategies for Different Subjects

In this section, we assume that human subjects decide differently in the same search task due to different individual preferences. This means, there exists not only one search strategy c that is followed by all subjects S, but there are a number r of different search strategies c_r that are each used by a subset S_r of the subjects. Consequently, exactly one out of the r different search strategies is used to explain the search behavior of an individual. The average fitness of a set of search strategies is calculated as $\tilde{\mu}(fit) = 1/|S| \sum_S \max_r fit(c_r)$, where $|S|$ denotes the number of subjects participating in the search experiment. $\max_r fit(c_r)$ denotes the maximal fitness of one of the r search strategies c_r for one individual.

	$fit(c_j)$		
c_1	0.8	**0.95**	**1**
c_2	**0.85**	0.7	0.8

Table 4: Example fitness evaluation

We want to give a brief example for the calculation of $\tilde{\mu}$. We assume that there are $r = 2$ search strategies c_r and $|S| = 3$ subjects. Table 4 shows the fitness $fit(c_r)$ of the search strategies c_1 and c_2 for three different subjects. $\max_r fit(c_r)$ is shown in bold. The average fitness of the search strategies is calculated as $\tilde{\mu}(fit) = 1/3(0.85 + 0.95 + 1) = 0.93$.

	GA	CRPH	SH
$\tilde{\mu}(fit)$	0.949	0.933	0.927

Table 5: Average fitness of the optimal search strategies for r=5

Table 5 shows the average fitness of the optimal search strategies for $r = 5$, this means five different search strategies are used to explain the search behavior of the 64 subjects. When allowing five different CRPH search strategies, the average fitness of the five search strategies is $\tilde{\mu}(fit) = 0.933$ using the reservation prices $p_r^1 = 498$, $p_r^2 = 494$, $p_r^3 = 491$, $p_r^4 = 488$, and $p_r^5 = 485$. When using five different SH search strategies, the optimal strategies show a payoff of 1, 3, 5, 7, and 13. Their average fitness is $\tilde{\mu}(fit) = 0.927$. When using a GA for finding $r = 5$ search strategies, each individual of the GA consists of 5 search strategies and the fitness of an individual is $\tilde{\mu}(fit)$. The GA is able to find search strategies with $\tilde{\mu}(fit) = 0.949$. This outperforms the classification based on standard search strategies.

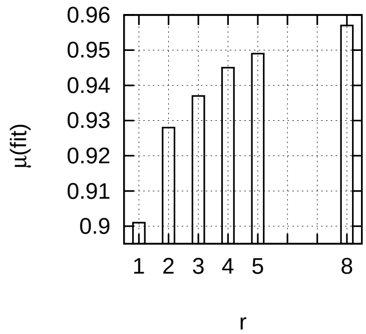

Figure 1: $\tilde{\mu}(fit)$ over the number r of search strategies

Finally, figure 1 shows the average fitness $\tilde{\mu}(fit)$ of the set of search strategies for the 64 subjects that have been found by the GA over the number r of different search strategies. The results show that with increasing r, this means using a larger number of search strategies, the behavior of the subjects can be better explained. This is per se no surprise since a larger number r of possible search strategies allows the GA to adopt each strategy to a smaller number of subjects. However, the results illustrate nicely that the GA is able to identify appropriate search strategies that are able to explain a large portion of human search behavior.

5.3 Searching for General Search Strategies

In the remainder of this section we want to examine more closely the character of the search strategies found by the GA. The question is whether the GA is able to identify characteristic search strategies that are used by a large proportion of the subjects.

When searching for r "optimal" search strategies for a group of people there is a trade-off between finding general search strategies that are used by a larger number of subjects and finding specific search strategies that more accurately describe the behavior of only a few subjects. Thus, with increasing r it is possible that either more general search strategies are found that better explain the behavior of a larger group of subjects, or we obtain very specific search strategies that are well adapted to the search behavior of only a few subjects. To find general rules that correctly predict the behavior of a large proportion of subjects is more important as such rules allow us to develop general classifications of humans' behavior.

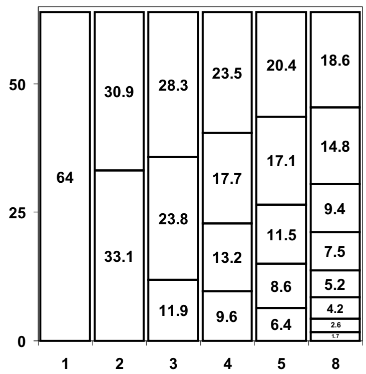

Figure 2: Average number of subjects whose search behavior is explained by the r different search strategies that are found by the GA.

Figure 2 shows the average number of subjects whose search behavior is explained by one of the r search strategies over r. For example, when using a GA for finding $r = 3$ different search strategies, on average 28.3 subjects use search strategy c_1, 23.8 of the subjects use search strategy c_2 and the behavior of only 11.9 subjects can be explained by search strategy c_3. When increasing r to $r = 8$, one of the eight search strategies still explains the behavior of, on average, 18.6 subjects. In contrast, the GA also finds very specialized search strategies that can explain the behavior of on average only 1.7, 2.6, or 4.2 subjects. Such search rules are very specific and no generalizations of these search strategies are possible.

To investigate how general the found search strategies are, figure 3 shows the number of subjects per search strategy, whose fitness is higher than 0.93 ($\max_j fit(c_j) > 0.93$). Therefore, only subjects are considered for whom a search strategy correctly predicts more than 93% of the decisions. When determining one search strategy ($r = 1$) only the search behavior of on average 22.5 subjects can be well explained (the prediction quality is on average larger than 93%). The numbers reveal that when increasing the number r of search strategies, the number of subjects whose behavior can be well explained by the two most general search strategies decreases only slightly. For $r = 2$, the be-

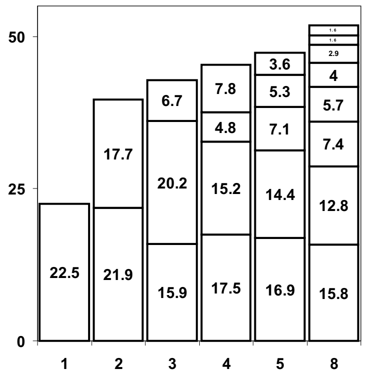

Figure 3: Average number of subjects, whose search behavior is explained by the r different search strategies better than 93% ($\max_j fit(c_j) > 0.93$)

havior of on average 39.6 subjects can be explained with $\max_j fit(c_j) > 0.93$. For $r = 8$, the behavior of on average 28.6 subjects can still be explained by the two most common search strategies. For $r = 8$, the remaining six other search strategies are able to explain the behavior of only, on average, 32.2 subjects with prediction quality $\max_j fit(c_j) > 0.93$.

The results indicate that there are only a few (about two or three) general search strategies that well explain the behavior of a large number of subjects. To assume that there are a larger number (more than three) of different and meaningful search strategies is not justified as searching for a larger number of rules only allows us to find very specific rules that only can explain the search behavior of a few subjects.

5.4 Finding General Search Strategies

In the following paragraphs, we take a closer look at the search heuristics that have been found by the GA for different r.

Section 5.1 already presented the best search strategy that is found by the GA and which on average predicts 90.2% of an individuals decision for the case $r = 1$. Table 6 presents the best search strategies found for $r = 2$. We only show the first five search steps t, as on average the subjects stop after 5.07 search steps. The search strategy c_1 is similar to the constant reservation price rule with $p_T = 494$. In addition, the pattern “falling” is relevant for $t = 2$, $t = 4$, and $t = 5$. c_2 is a combination of falling ($t = 1$, $t = 2$, and $t = 3$) and constant ($t = 4$ and $t = 5$) reservation prices. No patterns are acitvated. Very similar search strategies are found by the GA for $r = 3$ (compare table 7) as the search strategies c_1 and c_2 are similar to the case $r = 2$. There is an additional search strategy c_3 with increasing reservation prices, which starts from a low threshold $p_T = 484$.

Examining the search strategies found by the GA reveals that the search strategies are reasonable and can be used for interpreting human search behavior. Although there are no pre-specified rules available and only the basic

	search strategy c_1					search strategy c_2				
p_t	494	494	494	493	493	491	490	488	489	489
a_t^{thresh}	1	1	1	1	1	1	1	1	1	1
pa	0 (last step falling)					no pattern activated				
$a_t^{pattern}$	0	1	0	1	1					

Table 6: Found search strategies for $r = 2$

search strategy c_1					search strategy c_2					search strategy c_3				
494	494	494	494	493	492	490	491	489	488	484	485	486	488	489
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0 (last step falling)					no pattern activated					no pattern activated				
0	1	0	1	1										

Table 7: Found search strategies for $r = 3$

building blocks of the search strategies are pre-defined, some of the found rules are similar to the existing rules from the literature (like the CRPH search heuristic). In addition, the GA is able to identify unexpected search strategies (like c_3 for $r = 3$) that can help us to gain a better understanding of human search behavior. Summarizing the results, the GA is able to reproduce search strategies that are commonly used in the literature and to create new search strategies that can be used for developing better models to explain human behavior in search tasks.

6 Conclusions

This paper develops a new, modular approach for describing the behavior of humans in search tasks. In search tasks, individuals are confronted in sequential search steps with different situations and they have to decide in each step whether they want to continue or stop the search. The human behavior in search tasks (continuing or stopping) is described by a search strategy. The economic and socio-psychological sciences developed a variety of theoretical models that try to describe human behavior and from which optimal search strategies can be derived. However, in the real world, humans behave differently due to limited cognitive abilities and the search strategies derived from theoretical models do not often well predict human behavior. This paper presents a different approach where search strategies are not derived from models about human behavior but the search strategies are directly derived from the observed human behavior. Only the basic structure of decision rules are pre-specified and decision rules that explain human behavior are constructed from the observed empirical data by a genetic algorithm (GA).

To present the new approach this paper has done a variety of different things. It discussed human behavior in search tasks and exemplary illustrated how an optimal search strategy can be derived from some theoretical assumptions about human behavior. Furthermore, the paper presented the basic elements (building blocks) that can be used to construct search strategies. The building blocks used characteristic elements of standard search strategies and consisted of price thresholds and price patterns. Finally, the paper compared the decision rules, that are directly constructed from the observed human behavior by a GA, to standard search strategies from the literature. Various results are presented for the predicting quality which describes how well a search strategy predicts human behavior in search tasks.

In summary, this paper presented a GA-based approach that allows us to construct search strategies directly from the observed experimental data. A comparison to existing standard search strategies revealed that the new, modular approach resulted in search strategies with higher prediction quality. In addition, the found search strategies are general in the sense that they describe the behavior of a larger group of individuals well, and therefore, allow a categorization of human search behavior. Furthermore, the results show that the GA is able to reproduce search strategies that are similar to commonly used strategies in the literature as well as to create new search strategies which can be used as a basis for gaining new insights into human behavior in search tasks.

In the past, the most common approach in economic and socio-psychological sciences was to construct a theoretical model that explains human behavior. Based on the theoretical model and the underlying assumptions, rules describing the behavior of humans in decision situations, like search tasks, are derived. The results presented in this paper show that with the help of optimization methods like GAs, models about human behavior can be derived directly from the observed human behavior. Due to the observed high quality of the modular search strategies found by the GA, we recommend using heuristic optimization methods like GAs for the identification of human decision rules. A greater use of such optimization methods in economic and social sciences would allow us to keep the focus on human behavior, and validate the meaningfulness of theoretical models.

References

[1] Z. Eckstein and G. J. van den Berg. Empirical labor search: A survey. *Journal of Econometrics, forthcoming*, 2005.

[2] M. El-Gamal and D. M. Grether. Are people bayesian? uncovering behavioral strategies. *Journal of the American Statistical Association*, 90:1137–1145, 1995.

[3] M. A. El-Gamal and T. R. Palfrey. Vertigo: Comparing structural models of imperfect behavior in experimental games. *Games and Economic Behavior*, 8:322–348, 1995.

[4] J. Engle-Warnick. Inferring strategies from experimental data: A nonparametric, binary tree classification approach. *Journal of Economic Dynamics and Control*, 27:2151–2170, 2003.

[5] G. Gigerenzer. *The adaptive toolbox*, pages 37–50. Cambridge, MA: MIT Press, 2001.

[6] D. E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading, MA, 1989.

[7] G. W. Harrison and P. Morgan. Search intensity in experiments. *Economic Journal*, 100:478–486, 1990.

[8] J. D. Hey. Are optimal search rules reasonable? and vice versa? (and does it matter anyway?). *Journal of Economic Behaviour and Organization*, 2:47–70, 1981.

[9] J. D. Hey. *Economics in Disequilibrium*. Martin Robertson, Oxford, 1981.

[10] J. D. Hey. Search for rules for search. *Journal of Economic Behaviour and Organization*, 3:65–81, 1982.

[11] J. D. Hey. Still searching. *Journal of Economic Behaviour and Organization*, 8:137–144, 1987.

[12] D. Houser, M. Keane, and K. McCabe. Behavior in a dynamic decision problem: An analysis of experimental evidence using a Bayesian type classification algorithm. *Econometrica*, 72(3):781–822, 2004.

[13] D. Houser and J. Winter. How do behavioral assumptions affect structural inference? Evidence from a laboratory experiment. *Journal of Business and Economic Statistics*, 22(1):64–79, 2004.

[14] C. A. Kogut. Consumer search behaviour and sunk costs. *Journal of Economic Behaviour and Organization*, 14:381–392, 1990.

[15] S. A. Lippman and J. J. McCall. The economics of job search: A survey. *Economic Inquiry*, 14:155–189, 1976.

[16] R. D. McKelvey and T. R. Palfrey. An experimental stud y of the centipede game. *Econometrica*, 4:806–836, 1992.

[17] P. Moon and A. Martin. Better heuristics for economic search: Experimental and simulation evidence. *Journal of Behavioral Decision Making*, 3:175–193, 1990.

[18] D. Schunk and J. Winter. The relationship between risk attitudes and heuristics in search tasks: A laboratory experiment. *SFB-Working Paper 04-23*, 2004.

[19] J. Sonnemans. Strategies of search. *Journal of Economic Behavior and Organization*, 35:309–332, 1998.

[20] J. Sonnemans. Decisions and strategies in a sequential search experiment. *Journal of Economic Psychology*, 21:91–102, 2000.

[21] G. Syswerda. Uniform crossover in genetic algorithms. In J. D. Schaffer, editor, *Proceedings of the Third International Conference on Genetic Algorithms*, pages 2–9, San Mateo, CA, 1989. Morgan Kaufmann.

[22] D. Thierens. *Analysis and design of genetic algorithms*. Katholieke Universiteit Leuven, Leuven, Belgium, 1995.

[23] R. Zwick, A. Rapoport, A. K. C. Lo, and A. V. Muthukrishnan. Consumer sequential search: Not enough or too much? *Marketing Science*, 22(4):503–519, 2003.

Nr.	Author	Title
05-07	Hendrik Hakenes Isabel Schnabel	Bank Size and Risk-Taking under Basel II
05-06	Thomas Gschwend	Ticket-Splitting and Strategic Voting under Mixed Electoral Rules: Evidence from Germany
05-05	Axel Börsch-Supan	Risiken im Lebenszyklus: Theorie und Evidenz
05-04	Franz Rothlauf Daniel Schunk Jella Pfeiffer	Classification of Human Decision Behavior: Finding Modular Decision Rules with Genetic Algorithms
05-03	Thomas Gschwend	Institutional Incentives for Strategic Voting: The Case of Portugal
05-02	Siegfried K. Berninghaus Karl-Martin Ehrhart Marion Ott	A Network Experiment in Continuous Time: The Influence of Link Costs
05-01	Geschäftsstelle	Jahresbericht 2004
04-70	Felix Freyland	Household Composition and Savings: An Empirical Analysis based on the German SOEP data
04-69	Felix Freyland	Household Composition and Savings: An Overview
04-68	Anette Reil-Held	Crowding out or crowding in? Public and private transfers in Germany.
04-67	Lothar Essig Anette Reil-Held	Chancen und Risiken der Riester-Rente
04-66	Alexander Ludwig Alexander Zimper	Rational Expectations and Ambiguity: A Comment on Abel (2002)
04-65	Axel Börsch-Supan Alexander Ludwig Joachim Winter	Aging, Pension Reform, and Capital Flows: A Multi-Country Simulation Model
04-64	Axel Börsch-Supan	From Traditional DB to Notional DC Systems&61482; Reframing PAYG contributions to ìnotional savingsî

Discussion Paper Series

Mannheim Research Institute for the Economics of Aging Universität Mannheim

To order copies, please direct your request to the author of the title in question.

Nr.	Autoren	Titel	Jahr
66-04	Alexander Ludwig Alexander Zimper	Rational Expectations and Ambiguity: A Comment on Abel (2002)	04
67-05	Axel Börsch-Supan Anette Reil-Held	Die ökonomischen Auswirkungen der Alterung in Hessen	05
68-05	Axel Börsch-Supan Alexander Ludwig Anette Reil-Held	Projection methods and scenarios for public and private pension information	05
69-05	Axel Börsch-Supan	Risiken im Lebenszyklus	05
70-05	Hendrik Jürges	Die ökonomische Theorie der Familie und die Erklärung von Erbschaftsregeln – ein problemorientierter Überblick	05
71-05	Marcel Erlinghagen Karsten Hank	Participation of Older Europeans in Volunteer Work	05
72-05	Anette Reil-Held	Crowding out or crowding in? Public and private transfers in Germany	05
73-05	Axel Börsch-Supan Ismail Düzgün Matthias Weiss	Altern und Produktivität: Zum Stand der Forschung	05
74-05	Karsten Hank Marcel Erlinghagen Anja Lemke	Ehrenamtliches Engagement in Europa: Eine vergleichende Untersuchung am Beispiel von Senioren	05
75-05	Axel Börsch-Supan Karsten Hank Hendrik Jürges	A New Comprehensive and International View on Ageing: The Survey of Health, Ageing and Retirement in Europe	05
76-05	Karsten Hank Hendrik Jürges	Gender and the Division of Household Labor in Older Couples: A European Perspective	05
77-05	Daniel Schunk Joachim Winter	The Relationship Between Risk Attitudes and Heuristics in Search Tasks: A Laboratory Experiment	05
78-05	Daniel Schunk Cornelia Betsch	Explaining heterogeneity in utility functions by individual differences in decision modes	05
79-05	Franz Rothlauf Daniel Schunk Jella Pfeiffer	Classification of Human Decision Behavior: Finding Modular Decision Rules with Genetic Algorithms	O5