The Fast and the Not-So-Frugal: Human Heuristics for Optimization Problem Solving

Genovefa Kefalidou (Genovefa.Kefalidou@nottingham.ac.uk)
Horizon Digital Economy Research / Human Factors Research Group, The University of Nottingham, Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

Thomas C. Ormerod (t.ormerod@surrey.ac.uk)
Department of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Abstract
In this paper, human heuristics have been identified that provide close to optimal solutions when solving Capacitated Vehicle Routing Problems (CVRPs). Results from previous experiments showed humans can produce good solutions relatively fast that compete with computer-based methods giving further support to previous research on Traveling Salesman Problems (TSPs). Multiple Regression analyses have been conducted to show the best heuristics adopted by participants and that lead to better CVRP solutions. Identified heuristics are categorized in visuospatial and arithmetic heuristics. Visuospatial heuristics (e.g. Clustering, Anchoring) performed better than the arithmetic (e.g. Balancing). Strategy switching appears to be a critical step within CVRP solutions suggesting that heuristics adopted are not yet not-so-frugal, complimenting the fast and frugal toolkit. Results are discussed under the light of problem-solving theories and in terms of how best human heuristics can inform the current state-of-art computational algorithms used in optimization problem solving.

Keywords: optimization; problem solving; weak methods; fast and frugal heuristics; capacitated vehicle routing problems.

Introduction
A key issue in developing problem-solving theories is to identify human heuristics adopted when solving complex problems. An example of such problems are the Capacitated Vehicle Routing Problems (CVRPs), which are hard optimization problems, where the best solution must be discovered from a set of candidate solutions too large to allow exhaustive search.

In CVRPs, one has to discover a number of shortest routes taken by a capacity-limited vehicle from one or more depots to deliver to customers (represented as nodes) distributed in Euclidean space (Figure 1). Each station (node) must be visited once only and the vehicle must not exceed a weight limit (in Figure 1, it is 100) for each route. Computational algorithms exist for solving CVRPs each of which maintain limitations (i.e. unable to tackle with dynamic environments or local optimum - Michalewicz & Fogel (2002)). Consequently, CVRPs have considerable practical importance (e.g., in transportation and logistics) and theoretical importance as they provide an excellent testbed to investigate problem complexity and heuristics. Relevant psychological theories of problem-solving and bounded rationality include Newell and Simon’s (1972) weak methods approach (i.e. means-ends analysis and hill-climbing) where people select solution attempts according to the progress made towards the problem goal, and which tend to produce satisfactory but sub-optimal solutions. Gigerenzer’s ‘fast and frugal’ heuristics, build upon Newell and Simon’s theories and are of low cognitive effort (Gigerenzer & Goldstein, 1996). Fast and frugal heuristics aim to explain near-perfect performance by humans on a range of seemingly computationally intractable tasks by capitalizing upon the characteristics of the task environment, without, however, determining any specific heuristics under this category. While ‘satisficing’ heuristics apply sequential searching for available alternatives, ‘fast and frugal’ heuristics necessitate little information and computational resources in order to make different decisions (Gigerenzer, Todd & ABC Research Group, 1999). The research reported in this paper aims to offer an insight as to what heuristics are involved in solving hard yet widely applied problems such as CVRPs and identify their nature and qualities.

Human heuristics employed when solving optimization problems are of interest for two reasons: for furthering the progress of problem-solving theories by understanding how people arrive at good solutions and for informing computational algorithms. McGregor & Ormerod (1996) examined drawn solutions to Traveling Salesman Problems (TSPs) – similar to CVRPs without weight constraints and multiple routes – and found that, with problems of up to 100 nodes, human solutions were comparable with heuristic computer methods. They suggested that solutions are guided by the conceptualization of a convex-hull boundary, which coincides with the detection of natural object boundaries in human vision (e.g., Marr, 1980), moving from a global to a

Figure 1: The optimal solution of a CVRP 39-6 problem.
local conceptualization of the solution. Others suggest a more local-to-global heuristics generation approach: the ‘crossing avoidance’ approach (Van Rooij, Stege & Schactman, 2003) where lines crossed indicate tour inefficiency and the nearest neighbor heuristics (Lee & Vickers, 2000) according to which people select nodes to complete a tour according to their distance from the current node. As CVRPs carry additional performance requirements (i.e. the need to calculate total weights per route to provide valid solutions), the reliance on global or local feature detection through visual perception becomes insufficient. As such, it is unclear as to whether convex-hull-type heuristics or cross avoidance may lead to near optimal solutions.

Under a weak methods account, one may exhibit an incremental testing of route alternatives leading to purely satisficing solutions while under a ‘fast and frugal’ account, rapid solutions with very limited alternatives may be considered that promote optimality in human solutions.

The Experiment

The Experiment asked people to solve CVRPs, while thinking aloud in order to identify heuristics and strategies employed during the problem solving process. Task complexity varied by increasing the number of nodes tours required in each problem. Human heuristics were identified based on participants’ verbalizations, on their video-recordings while solving the problems and on participants’ paper notes.

Method

Participants

Forty-nine students of Lancaster University (M age = 24 years, SD age = 1.86) were paid £6 each to participate.

Design

Participants were randomly assigned to either a Verbalization (24 people) or a Non-Verbalization group (25 people).

Materials and Procedure

In an individual setting, participants were asked to solve 4 randomized CVRPs after solving one practice problem.

All the sessions were video-recorded. There was no time limit and the problems were A4 paper-based. Students had to use different color pencils to draw the different routes. Each route had to start and end to the depot (a node represented with a lighter color shade than the others). Participants were instructed to use either separate papersheets or the same paper stimuli as notepads for any calculations.

Problem-specific instructions were presented at the bottom of each problem page indicating the required number of routes, the total weight to be collected and the number of nodes. In the Verbalization group, participants were asked to express all their thoughts (including calculations and strategies adopted) during the problem solving process. If they remained silent for more than 1 minute they were prompted to think-aloud.

Stimuli consisted of two problems with 33 nodes with 4 and 5 routes to draw respectively, one problem with 39 nodes and 6 routes to draw and a problem with 45 nodes and 7 routes to draw. An example of a problem given with a participant’s solution can be seen in Figure 2.

Figure 2: A participant’s solution (33-5 problem).

The CVRPs used in the Experiment and their optimal solutions (generated using a Branch and Cut algorithm (Augerat et al., 1995) are adopted from the Operational Research literature in which they have been widely used for optimization research. These CVRPs were downloaded (alongside with their published optimal solutions) in text file format from the ‘VRP Website’ (VRP Web – http://www.bernabe.dorronsoro.es/vrp/). These problems were chosen because they have published available optimal solutions and therefore, comparisons can be drawn between human solutions and algorithmic solutions. Experimental stimuli was prepared using software developed by Simon Slavin (Ormerod & Slavin, 2005) (Figure 2). The quality of the participants’ solutions was measured in Percentage Above Optimal (PAO) values (adapted from MacGregor & Ormerod, 1996) using the following formula: (x / y – 1) * 100, where x is the total problem solution distance and y is the optimal distance. The PAO values correspond to how far a participant’s solution is from the optimal solution, which is always 0. The total distance of solutions and the optimal distances were calculated using software that the first author of this paper developed.
Results

Research has shown that participants are able to reach close to optimal CVRP solutions quickly and that verbalization did not impair their performance (Kefalidou, 2011).

The verbal protocols were recorded in written form (typewritten transcripts) where all the verbalizations (both participants’ and experimenter’s), calculations and time logs of the participants’ responses were included in the verbal protocols. An inter-rater reliability was conducted and Cohen’s Kappa statistic (k-statistic) was calculated for the qualitative verbalization data obtained from the experimental data, a high Kappa statistic was obtained (k = .72) showing that the level of agreement between the analysts regarding the identified strategies was 72%, which is considered to indicate good agreement (Bakeman & Gottman, 1997).

Strategies identified

Human strategies are identified through the verbal and visual analyses of the videos taken during the experiment. Verbal protocols were retrieved from the videos of the participants from the experimental-Verbalizing group. Strategies were categorized in super-ordinate groups based on dominant features of each strategy (visuospatial route construction (V) vs. arithmetic route construction (A)) as described below. Participants’ performances (PAO) as grouped by the super-ordinate categories are presented in Table 1.

Major identified strategies are the Calculating and the Clustering. Calculators (A) are the participants that summed the weights of the nodes as nodes were added to each route ensuring the weight constraint per truck was not violated. Calculators have similarities to those using the Nearest Neighbor (V) heuristic with the exception that those using Nearest Neighbor targeted close by nodes while calculators selected nodes based on their weight. Clusterers (V) were the participants who solved the problems based on their visual perception by clustering the neighborhood nodes into separate groups each one of which was representing a different route. Clusterers are differentiated from nearest neighbor in that clustering nodes together was the first step prior to generating a solution while for Nearest Neighbor strategists including nodes based on proximity occurred during the construction of the solution.

Table 1: Super-ordinate classifications of strategies used with respective Percentage Above Optimal (PAO) across the two groups (Verbalization and non-Verbalization).

<table>
<thead>
<tr>
<th>Super-ordinate Classification</th>
<th>PAO Verbalization</th>
<th>PAO Non-Verbalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visuospatial</td>
<td>13.8</td>
<td>10.6</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>18.0</td>
<td>15.7</td>
</tr>
</tbody>
</table>

Other participants exhibited the behavior of Averaging (A) where they were designing routes based on the calculated average capacity per vehicle. Others were following an Anchoring (V) approach where they were designing routes starting from the more distant nodes while adding up weights of nodes to the route while returning back to the depot. Maximizers (A) were strategists focusing on handling the nodes with the largest demand first and then including the smaller ones in their routes. Remainder (V) was the strategy where participants were planning ahead to pick up nodes on their way back to depot. Nodes to be visited on the way back were identified visually by participants. Balancing (A) participants ensured that each route they drew retained similar total weight – another strategy relying on calculations. Finally, Remainder-Filling (V and A) was the same as Remainder except that participants were accounting for the weight of the nodes to be added. A brief extract from a verbal protocol showing application of strategies is as follows:

“So...I’m gonna try and get the routes...for each truck...hmm...closer to each other...and then...and closer to the average...and then for the last truck...I’m going to leave what’s left...few bits...ok...I’m starting from the depot.”

The participant above started with the Calculator strategy and then switched to Averaging suggesting a ‘switching’ phase, which appears to be a prevailing part of problem solving. Table 1 shows that participants that followed Visuospatial strategies performed better than participants that adopted Arithmetic strategies independently of whether they were Verbalizing or not.

Seventeen participants were categorized as adopting a predominately Visuospatial strategy while 24 adopted a predominately Arithmetic strategy. For example, if a participant was using the Clustering strategy throughout the completion of the problem (even if he had been switching to other strategies), they were considered to be a Visuospatial/Clusterer strategy that followed predominantly the Clustering Visuospatial strategy.

Optimal Heuristics by Problem

33-4 Problem (P2) One participant from the Silent group (0.51 PAO) and two from the Think Aloud group (0.26 and 0.35 PAO) performed close to optimal. The participant from the Silent group who used many strategies, including Clustering, Calculation, Nearest Neighbor, and Maximizing (they began their routes with the latter two), Anchoring and Remainder. They switched strategies frequently. The same pattern was observed with the two participants from Think Aloud group who used different strategies interchangeably. The majority of the participants’ pool performed well (69%). The worst performance in the Silent group was 26.90 PAO and in the Think Aloud group was 26.27 PAO. Both participants used Calculation as a main strategy. In contrast to the participants with a good performance, these participants used only Anchoring as an alternative strategy.

A multiple forced entry regression was performed to examine the impact of strategies use on solving all the problems in the Experiment. The forced entry method was chosen to test which strategies promote good performance when solving CVRPs because the qualitative data analysis
and the participants’ performance (PAO scores) suggested that certain strategies may improve or may impair the human performance.

As the qualitative analyses conducted though indicated the high frequency of different strategies usage from both the Visuospatial and Arithmetic groups, the first author decided to conduct further regression analyses using the low-level and more detailed classification of the strategies adopted instead of the high-level classification.

To investigate further the influence of strategies on performance a linear multiple regression analyses were conducted using the forced entry method testing all the possible combinations of low-level strategies compared to the participants’ performance.

The regression model that was fit included two predictor variables. The reason for fitting two predictor variables in each model is because the sample size of the participants in the verbalization groups did not allow the fitting of more than 2 predictors in the model. According to Field (2005), the number of predictors to fit in a regression model highly depends on the sample size. All the possible permutations of strategies’ couples were fit into separate regression models (two only strategies were fit in each model using the forced entry method for the reasons explained above). The best model was chosen based on the best AIC value. The goodness of an AIC value is determined by how low it is. A good and robust model has a low AIC value (Wagenmakers and Farrell, 2004). The dependent variable was the PAO scores of participants’ performance.

Using the enter method, a significant model emerged F(2,20) = 5.50, p < .05, Adjusted R² = .310, AIC = 143.62. Regression results for the 33-4 problem show that the use of Anchoring, which is a Visuospatial strategy predicted good performance. Table 2 shows the regression results on P2 (33-4).

Most of the participants who performed between 10 and 20 PAO had at least two different strategies that they used more than twice interchangeably. The majority of the solvers were serial planners, which means that they were designing their routes one after the completion of other (70%), while only 30% were parallel planners. The most popular strategy in the Silent group was Calculating (used throughout the problem) as well as Clustering and Anchoring with 90% frequency of use each. The most popular strategies in the Think Aloud group used throughout this problem were both Calculating and Clustering, followed by Remainder (70%), Averaging (10%), Nearest Neighbor (20%) and Balancing (20%).

Table 2: Linear Regression results for P2 Strategies.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>16.41</td>
<td>2.25**</td>
<td></td>
</tr>
<tr>
<td>P2ANCH</td>
<td>-2.94</td>
<td>1.11</td>
<td>-.49*</td>
</tr>
<tr>
<td>P2REM-FILL</td>
<td>-2.54</td>
<td>1.32</td>
<td>-.36</td>
</tr>
</tbody>
</table>

Note R² = .38; adjusted R² = .31. *p < .05, **p < .001

33-5 Problem (P3) Fewer participants performed well than with 33-4 problem (22 participants < 10 PAO vs. 34 participants in 33-4). Three participants performed in the 30-40 PAO range, one who performed in the 40-50 PAO range and one in the above 60 PAO range. One participant from the Silent group performed well with 0.62 PAO. No participant from the Think Aloud group performed at < 1 PAO. However, two participants performed in the 1-2 PAO range. Again, good-performing participants employed many strategies throughout, notably the Clustering strategy. They also used Nearest Neighbor, Maximizing, Anchoring and Remainder. Switching between strategies was less common in this problem. The worst performance in this problem came from a participant from the Think Aloud group with 66.47 PAO. The participant was using Calculating strategy throughout completion of the problem, only switching to Anchoring twice.

Using the enter method, a significant model emerged F(2,20) = 5.08, p < .05, Adjusted R² = .290, AIC = 172.273. Regression results on verbalization data for the 33-5 problem showed that an Arithmetic strategy (Balancing) is a predictor of poor performance. As the number of routes to draw increases, an Arithmetic strategy such as Balancing impairs the performance. Table 3 shows the regression results for this P3 (33-5) problem. Similar to 33-4 problem, the most popular strategies were the Calculating and Clustering (100% and 90% frequency of use respectively). However, the strategy Anchoring appears to gain in popularity (70%). The least used strategies are again Balancing (10%) and Averaging (10%).

Table 3: Linear Regression results for P3 Strategies.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>11.90</td>
<td>3.72**</td>
<td></td>
</tr>
<tr>
<td>P3BALA</td>
<td>8.09</td>
<td>2.77</td>
<td>.56**</td>
</tr>
<tr>
<td>P3REM</td>
<td>3.57</td>
<td>1.95</td>
<td>.35</td>
</tr>
</tbody>
</table>

Note R² = .36; adjusted R² = .29. *p < .05, **p < .001

39-6 Problem (P4) Similarly to the previous problems, a forced entry linear multiple regression was performed for P4 (39-6) problem (Table 4). Using the enter method, a significant model emerged F (2,20) = 3.61, p < .05, Adjusted R² = .207, AIC = 191.803. The results are displayed in Table 4.

Table 4: Linear Regression results for P4 Strategies.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>41.19</td>
<td>7.53***</td>
<td></td>
</tr>
<tr>
<td>P4CLU</td>
<td>-2.79</td>
<td>1.11</td>
<td>-.53*</td>
</tr>
<tr>
<td>P4NEA</td>
<td>-4.99</td>
<td>2.95</td>
<td>-.35</td>
</tr>
</tbody>
</table>

Note R² = .29; adjusted R² = .21. *p < .05, **p < .001, ***p < .0001.
As it can be seen, a Visuospatial strategy such as Clustering was found to be predictor of good performance. While performance was poorer than with the 33-4 and 33-5 problems, some participants performed <10% over optimal. However, the majority of participants (70%) performed 10–20% over optimal, and 18 participants performed above 20 PAO. The best performance (2.4% over optimal) had a clear design and used Clustering and Remainder strategies, occasionally drawing upon Calculating, Maximizing and Anchoring. Other participants who performed well used Clustering and Calculating, occasionally drawing upon Balancing and Anchoring. One participant performed very badly (56.45 PAO). Poor performers continued to use Calculating as a dominant strategy.

Clustering, Calculating and Remainder again are prominent (90%). Other strategies used are the Nearest Neighbor (50%), Averaging (20%), Maximizing (30%) and Balancing (10%).

Those who used Anchoring in their solutions performed worse (in some cases with 5 percentage points difference in percentage over optimal) than those who did not. However, combinations that generated good performance included Clustering with Anchoring, Clustering and Nearest Neighbor, Anchoring and Nearest Neighbor and Clustering with Anchoring and Balancing. Problem 39-6 involves more nodes to use and more routes to draw and the participants started to use more strategies.

45-7 Problem (P5) Three participants exhibited their worst individual performance in this problem. Regression results showed that participants who used Calculating, Clustering and switching between strategies performed better than those that did not. As the size of the problem increased, the regression model showed that certain interaction of strategies and strategies’ switching could predict the performance of participants. Using the enter method, a significant model (at the 10 % level though) emerged F (2,20) = 2.90, p = .081, Adjusted R² = .160, AIC = 155.385. Table 5 shows the regression results for this P5 (45-7) problem. As it can be seen, similarly to P2 and P4 problems, Visuospatial strategies – Clustering strategy specifically – can predict good performance though this finding is only marginally significant. The Clustering strategy was retained in the model because it improved the AIC value of the model compared to the model without Clustering included as a predictor.

Table 5: Linear Regression results for P5 Strategies.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>25.51</td>
<td>3.00***</td>
<td></td>
</tr>
<tr>
<td>P5CLU</td>
<td>-0.65</td>
<td>0.33</td>
<td>-.41*</td>
</tr>
<tr>
<td>P5REM</td>
<td>-1.90</td>
<td>1.63</td>
<td>-.24</td>
</tr>
<tr>
<td>FILL</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note R² = .24; adjusted R² = .16. *p < .05, **p < .001, ***p < .0001.

Participants who performed well used at least 2 strategies per route but mainly used Clustering. Participants who had no dominant strategy showed worse performance. A considerable increase in the use of a number of different strategies was observed compared with other problems. Fifty percent used the Nearest Neighbor strategy, while Calculating, Clustering and Anchoring were used by 100%. Remainder was used by 80%, Averaging and Maximizing 40%, while Balancing is used the least (20%).

Discussion

Results are congruent with both ‘fast and frugal’ heuristics (Gigerenzer & Goldstein, 1996) and weak methods (Newell & Simon, 1972) as both more ‘optimizing’ strategies (e.g. visuospatial) and ‘satisficing’ strategies (e.g. arithmetic) were identified to be interplaying within individuals’ problem solving phase. This also aligns with MacGregor and Ormerod’s (1996) suggestions that people capitalize upon visual forms (e.g. convex hull boundary), which tends to correlate closely with the path taken by good or optimal TSP solutions.

Furthermore, we have identified that despite participants exhibit strong switching behavior (thus engaging in an incremental testing of strategies) this has not occurred on a global level (e.g. incremental testing of alternative routes) rather it occurred on a local (subroute) level, especially for solutions that reached near-optimality. This suggests that even though this behavior may share elements with weak methods (i.e. checking for alternatives), this does not happen on a global level. Participants reached good solutions quickly employing successful visuospatial strategies such as Clustering (innate to human perception), thus aligning to a more ‘fast and frugal’ approach, especially for the problems that included a higher number of nodes and routes to draw. Yet, their ‘fast and frugal’ approach was employing a ‘not-so-frugal’ set of strategies (as they were rich in numbers) while switching was apparent throughout, leading to close-to-optimal solutions. An explanation to this could be that the complexity of the task environment re-enforces humans to combine weak and ‘fast and frugal’ heuristics in order to overcome thresholds of cognitive overload.

More particularly, Clustering strategy (Visuospatial strategy) and Anchoring were found to be predictors of better human performance compared to the Balancing strategy (Arithmetic strategy), which was found to be impairing human performance in CVRPs. This may provide some support to the ‘fast and frugal’ theory of Gigerenzer and Goldstein (1996) that ‘quick and dirty’ approaches to problem solving (e.g. strategies that rely in innate human abilities and require lower cognitive effort such as Clustering) can reach or even outperform other more exhaustive and satisficing means of problem solving (e.g. strategies that necessitate a higher cognitive demand such as performing calculations and Balancing). Indeed, such performance may be subject to the task environment or the type of problem involved. However, in order to test such
speculation, further research must be conducted. Furthermore, the ‘intuitive’ reliance on ‘fast and frugal’ heuristics (visually-inspired such as Visuospatial set of strategies) solely does not promote ‘optimal’ or close to ‘optimal’ performance. This finding contrasts with the general notion that ‘fast and frugal’ heuristics lead to better quality (Gigerenzer & Goldstein, 1996) and the ‘weak methods’ are usually indicators of ‘satisficing’ accounts (Newell & Simon, 1972) associated with the production of lower quality yet error-free solutions. Results suggest that humans adopt both visuospatial and arithmetic strategies while attempting to solve CVRPs. However, when employing visuospatial strategies (such as Clustering and Anchoring), participants appear to generate better solutions compared to their counterparts that adopt arithmetic strategies (such as Balancing).

Visuospatial heuristics appear to be more efficient than arithmetic. This may be because they employ behaviors that are ordinarily inhibited within humans. Furthermore, visual techniques are highly dependent on the environment and thus their successful use may be suggesting that ‘fast and frugal’ heuristics through intrinsic ecological rationality can trigger better performance in CVRP problem solving. Further research is required though to discern how people decide that a heuristic is a good one and whether a solution is a good solution. Strategy switching appears to be a recurrent theme in a CVRP solving process, thus constituting an important element of the heuristic generation. Participants are constantly changing their plan of action by judging the environment and the constraints given. Computational algorithms can be enriched by applying to them better human heuristics while incorporating critical switching thresholds between heuristics to optimize their performance and avoid the trappings of local optimum and dynamic environments.

Acknowledgments
This research was conducted as part of Genovefa Kefalidou’s PhD thesis at Lancaster University under the sponsorship of EPSRC/ESRC and under the supervision of Prof. Ormerod and Prof. Eglese. Many thanks are due to all the participants of the study and to Horizon Research Institute and the Human Factors Research Group for providing the resources for the completion of this paper.

References
Gigerenzer, G., Todd, P. M. & ABC Research Group (1999). Simple Heuristics That Make Us Smart, Oxford University Press, USA.