УДК 007.52

Furems E., Larichev O., Lotov A., Miettinen K., Roizenson G.
Institute for Systems Analysis RAS, Moscow

Human Behavior in a Multi-Criteria Choice with Individual Tasks of Different Difficulties*

Работа посвящена изучению поведения людей при решении задач многокритериального выбора. Особенностью экспериментов является использование для каждого испытуемого специально подобранных для него индивидуально трудных задач в соответствии с его/ее предпочтениями по каждому критерию. Изучение поведения людей осуществлялось на основе специально созданной прикладной системы, реализующей декомпозицию процесса решения многокритериальной задачи на несколько этапов с проведением на каждом из них соответствующих парных сравнений. Эксперименты были проведены в двух странах: России и Финляндии.

Introduction

The problems of choosing the best object from their small set is a typical human activity. Such problems, as a rule, are ill-structured ones (Simon, 1969). Examples of such problems include the selection of a good by a consumer, choosing an apartment for tenancy or purchasing, choosing the university for a student, etc. While solving such problems people consider the different features of the objects by taking into account a number of significant aspects (criteria). To what extent are the human beings consistent and reasonable in such decision-making? This question has been studied in many papers [1-5]. The results of descriptive studies may be summarized as follows: The multi-criteria object comparison, and particularly a choice of the best object, is difficult to the human system for processing information, and the more criteria, the more complexity is. While solving such problems, human beings make errors as well as use simplifying strategies to adapt the problem to their capabilities.

Experiments evidence that comparison of objects that differ with estimates only upon 2 criteria is relatively easy to human being (e.g., such technique of information obtaining is the base of the method ZAPROS [5]). When the objects are described with ordinal criteria with 3–4 verbal estimates each, individuals make 1–2 contradictions from 50–60 pair-wise comparisons of 2-criteria combinations. Human being abilities in respect to comparison of objects that differ with verbal estimates upon 3 and more criteria are known insufficiently. It seems to be reasonable to study such abilities. Due to special features of human information processing system, it should be expected that an individual most likely would, firstly, put attention on a part of estimates of two objects, previously compare such parts, and then involve other estimates and so one.

* The research was partially supported by the Russian Foundation for Basic Research grants № 0015-96053, 01-01-00514 and 02-01-06286.

To study human behavior in respect to multi-criteria object comparison that differ with verbal estimates upon some criteria we apply the following procedure. All the comparisons have a qualitative nature. The DM makes a choice during several stages. Each stage involves all necessary comparisons of parts that have the same dimension (that is, 1, 2, 3, etc. criteria). Thus, on each next stage the dimension of parts to be compared increments by one. A number of sequential stages that allow to chose the best object create a round. Once a DM has completed the first round, he/she is communicated the result of his/her choice. Then the procedure repeats as the second round, which differs from the first one in an arbitrarily order of same criteria. We would like to study out to what extent these procedures of comparisons allow people to make a reasonable choice taking into account all (or a majority of) the criteria?

The Problem of Choice and Individually Adjusted Instances of Different Difficulties

In order to find the answer on the question above we carried out the experimental study of human behavior in multi-criteria choice. For our experiments we took the problem of choosing a part-time job by a student to be engaged concurrently with his/her studies. *Salary*, *Work time*, *Position* and *Time to workplace* were used as criteria. Such a problem is typical for both Finnish and Russian students involved in our experiments as subjects. The students were of the last year of education. While solving such a problem the students were in the position of a DM.

It should be noted that in many psychological experiments the same task is presented to all the subjects to study human behavior in multi-criteria choice. However, such a task may be easy to one subject and complex to another. We, on the contrary, believe that it is of great importance to give each subject an individually adjusted instance of a general task (hereinafter «instance»), taking into account his/her desirable levels upon some criteria, communicated by him/her preliminary. Our experiments were arranged as follows.

Initially, a student was presented a description of general task (choosing a part-time job), and was prompted to enter a desirable salary and a time period of his/her study in the University, which time period he/she preferred not to omit. Then, each student had to order a set of predefined estimates upon each criterion according to his/her preferences, excluding the *Salary* (obviously, more salary is preferable to anybody). We started to generate to a student the individual set of jobs from the criteria *Salary* and *Work time* creating his/her instance of either of two different levels of difficulty: Difficult Choice (DC) or Moderately Difficult Choice (MDC). In the case of DC, initially, to each job was arbitrarily assigned a salary that was more than, equal to or less than the desirable one. Then the following principle was used to specify a work time to each job: the bigger salary, the more overlapping of work time and study time and, accordingly, the less time to study. Thus, each student was put in a situation of a difficult choice between lecture and/or seminar attendance and additional payment. Finally, each job created as above was assigned the estimates upon the criteria *Position* and *Time to workplace*; such estimates were generated on the basis of the principle: the better position to a student, the longer the time to get to a workplace was. Such pairs were assigned to the jobs in an arbitrary order. In the case of MDC, in addition to the jobs created according to DC, we added jobs that had the best estimate upon either of the two criteria (*Salary* or *Work time*) and the second best

estimate on the other, or the best estimates upon both the criteria. We retained non-dominated objects (alternative jobs) only in each instance. It should be noted that the students did not know what kind of tasks (DC, MDC) they had to solve. Each instance consisted of 4 to 10 alternative jobs.

The Procedure

A student solved his/her instance twice, in two rounds. Each round consisted of several stages. The difference between the stages was the number of criteria to be taken into account (2, 3 or 4 in our case). Let's as note that comparisons done on some part of criteria do not restrict the freedom of choice on other parts of criteria. The order of alternative pairs (correspondent parts of objects) to be compared was set with the 2 different order of criteria. **The two rounds were used to check the following hypothesis: the subjects could make a stable choice in both the rounds.**

In order to select the best object (in our case, a part-time job), each student had to carry out a series of comparisons between the alternatives available. During the first stage in either of the rounds, each student was presented all the pairs of alternatives that differed in the estimates upon 2 criteria. Each time, while making the comparison, each student could answer in one of the following ways: «the first alternative is preferred over the second one»; «I'm is indifferent to the both alternatives»; «the second alternative is preferred over the first one» or «I do not know».

As a result, a student produced an order of all 2-criteria alternatives. There were cases where such information was sufficient to compare some 4-criteria objects or, moreover, to choose the best object. For example, if the first object estimates upon both the first 2 criteria and the last 2 criteria are better than the second object estimates upon the same criteria, respectively (according to a student's comparisons), then the first object is preferable than the second one (we assume that the criteria are preference-independent). However, if 2-criteria comparisons were not sufficient to choose the best 4-criteria object, the procedure moved to the next stage. In that stage, a student had to compare alternatives that differed in the estimates upon 3 criteria. Let us note, that each 3 criteria alternative may be split on 2-criteria and 1-criterion parts by 3 different ways (see below). All of them were presented to students. If the information obtained on this stage was still not sufficient to choose the best 4-criteria alternative, the procedure moved on to the final stage. The results of the previous stage gave an opportunity to rank order all the 3-criteria parts. They could then be used along with the estimates of the fourth criterion, by 4 different ways, to make a choice of the best 4-criteria alternative. Thus, in the framework of our procedure, it was also possible to collect information on how successful and non-contradictory the subjects were in using the so-called parts. In other words, we would like to check the **following hypothesis: human beings are able to compare consistently 3-criteria alternatives, presented as 2-criteria and 1-criterion parts, when for a pair of 3-criteria alternatives the parts of each dimension (2-criteria and 1-criterion) are previously ordered by a subject with respect to his/her preferences**. We are interesting also, whether human beings are able to compare 4-criteria alternatives, presented as 3-criteria and 1-criterion parts, previously ordered by preferences.

Experiments

The experiments with the students (13 persons) of the Department of Mathematical Information Technology at the University of Jyvaskyla in Finland and the

students (25 persons) of the Department for Computational Mathematics and Cybernetics (CMC) at the Moscow State University named after M.V. Lomonosov (MSU) in Russia were conducted. A computer program was developed to implement the procedure described earlier.

Each student was presented step-by-step all the pairs of the alternative jobs to be compared. The jobs of each pair differed in the estimates upon 2, 3 or 4 criteria until the results of the comparisons provided information enough to make a choice of a most preferable job. While presenting a pair of alternatives, the computer program reminded a student, with the help of colors, which estimate or combination of estimates of the first alternative he/she indicated as a better or worse than the corresponding estimate(s) of the second one on the previous stages. Fig. 1 illustrates two alternatives with different estimates upon criteria *Position* and *Time to workplace*. In this case, we use the color display in a way that the better estimates are presented on a white background and the worse estimates on a grey one.

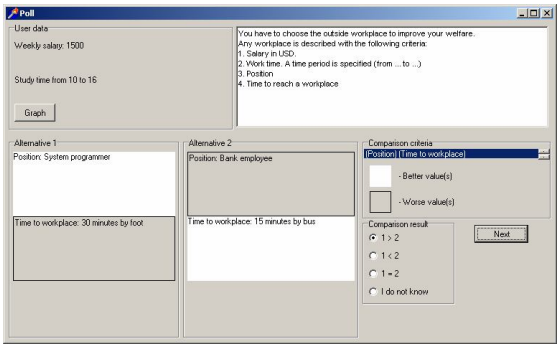

Fig. 1. Dialog example of two alternatives comparison regarding the criteria *Position* and *Time to workplace*.

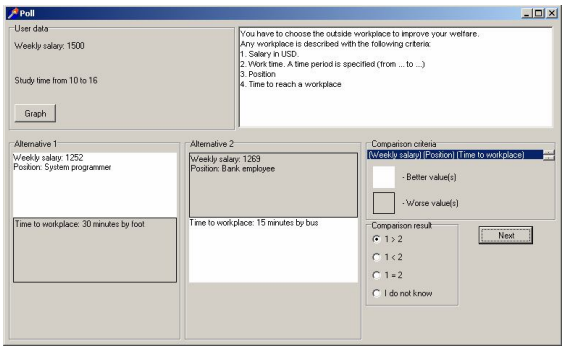

Fig. 2. Dialog example of two alternatives comparison regarding the criteria *Weekly salary*, *Position* and *Time to workplace*.

If the comparison results of all the pairs of criteria did not provide information enough in order to make a choice of the most preferred job, the student was asked to compare alternatives that differed in the estimates upon 3 criteria. Fig. 2 gives an example of the comparison of 2 alternatives differing in the estimates upon 3 criteria *Weekly salary*, *Position* and *Time to workplace*. The estimates of each alternative were divided to 2 parts for easier comparison: (*Weekly salary* & *Position*) and (*Time to workplace*). Since this particular student said in the previous stage (2-criteria alternative comparison) that the alternative 'System programmer, 1252 USD' was preferred to the alternative 'Bank employee, 1269 USD', the corresponding combinations of estimates are displayed in the dialog window on a white and on a grey background, respectively. The better estimate '*Time to workplace'* is displayed on a white background. Each 3-criteria part may be reduced to 2-criteria and 1-criterion parts in 3 different ways. In a way similar to fig. 2, a student was asked to make 3 comparisons of different 2-criteria and 1-criterion parts to check the consistency of the results. Finally, the best 4-criteria alternative selected in the round was presented to a student (the presentation of 4-criteria alternatives to be compared was analogous to 3-criteria alternatives, for the exception, that each pair of 4-criteria alternatives may be presented in four different ways).

Results

The main results of the experiments are presented in Table 1. The columns 'Round I Choice' and 'Round II Choice' contain the ordinal estimates' vectors of the objects chosen by each student in the first and the second rounds, respectively. The figures within any vector positions are the ranks of verbal estimates of the correspondent criterion scales according to the student's preferences (1 is the best and so on), while the criteria are arranged in the following order: *Salary*, *Work time*, *Position* and *Time to workplace* (notwithstanding different orders of criteria in 2 rounds). The column 'Type of preferences' contain characteristic of a student preferences' stability. «S» means stable preferences, while «US» means unstable ones. The preferences are implied to be stable if: (a) the same part-time job was selected in the both rounds; or (b) the part-time jobs selected in the first and the second rounds, respectively, have the first estimate(s) upon the same criterion/criteria.

As one can see from Table 1, 14 of the 38 students selected the same job both in the first and in the second round. Furthermore, according to the definition above 20 from 38 students were absolutely stable. So, we are not able either to confirm or to reject the hypotheses on a subject stability in 2 rounds.

Table 1

Choices of students

# of the student	Round I Choice	Round II Choice	Type of instance	Type of preferences
1.	1, 1, 2, 4	1, 4, 5, 1	MDC	S
2.	4, 2, 3, 3	4, 1, 5, 1	DC	US
3.	2, 1, 2, 4	1, 2, 5, 1	MDC	US
4.	2, 4, 2, 4	1, 3, 4, 2	DC	US
5.	1, 2, 1, 5	1, 2, 1, 5	MDC	S
6.	3, 1, 1, 5	2, 2, 1, 5	MDC	S
7.	2, 2, 3, 3	1, 4, 5, 1	MDC	US
8.	1, 1, 1, 5	1, 1, 1, 5	MDC	S
9.	2, 1, 1, 5	2, 1, 1, 5	MDC	S
10.	1, 3, 1, 5	1, 3, 2, 4	DC	S
11.	1, 1, 5, 1	1, 1, 5, 1	MDC	S
12.	2, 4, 1, 5	1, 3, 3, 3	DC	US
13.	2, 2, 1, 5	2, 2, 4, 2	MDC	US
14.	6, 1, 5, 1	1, 4, 3, 3	DC	US
15.	2, 3, 1, 5	1, 3, 2, 4	DC	US
16.	5, 2, 4, 2	1, 4, 2, 4	DC	US
17.	3, 1, 2, 4	1, 2, 3, 3	MDC	US
18.	1, 3, 3, 3	1, 3, 3, 3	DC	S
19.	3, 1, 1, 5	1, 1, 2, 5	MDC	S
20.	2, 1, 2, 4	2, 1, 2, 4	MDC	S
21.	1, 3, 4, 2	1, 3, 1, 5	DC	S
22.	1, 2, 2, 4	1, 2, 2, 4	MDC	S
23.	6, 1, 2, 4	2, 2, 4, 2	MDC	US
24.	4, 1, 5, 1	1, 3, 4, 2	DC	US

Confirmation of the table 1

# of the student	Round I Choice	Round II Choice	Type of instance	Type of preferences
25.	2, 3, 1, 5	2, 3, 1, 5	DC	S
26.	1, 4, 1, 5	1, 4, 1, 5	DC	S
27.	3, 2, 1, 5	2, 3, 2, 4	DC	US
28.	3, 2, 4, 2	1, 4, 1, 5	DC	US
29.	1, 4, 3, 3	1, 4, 3, 3	DC	S
30.	4, 1, 3, 3	1, 4, 1, 5	DC	US
31.	2, 3, 3, 3	2, 3, 3, 3	DC	S
32.	1, 4, 1, 5	1, 4, 1, 5	DC	S
33.	4, 1, 2, 4	2, 3, 4, 2	DC	US
34.	1, 4, 4, 2	1, 4, 4, 2	DC	S
35.	1, 4, 1, 5	3, 2, 1, 5	DC	S
36.	2, 3, 1, 5	1, 4, 2, 4	DC	US
37.	4, 1, 4, 2	1, 4, 2, 4	DC	US
38.	1, 4, 3, 3	1, 4, 3, 3	DC	S

Let us then consider the cases of multi-criteria choices where it was necessary to compare the 3- and 4-criteria alternatives. In most of the cases, the students compared 3-criteria alternatives in their different format of presentation in an identical way. 3-criteria alternatives were presented to N=34 students. N_1=5 students gave either only 2 identical or even 3 different answers on 3 different representations of the same 3-criteria alternative. N_2=29 students gave three identical answers on three different representations of the same 3-criteria alternative.

Let us introduce the following hypotheses: H_0: 3-criteria alternative comparison is difficult to subjects, i.e., subjects change their answers while comparing the same pair of alternatives of such dimension. H_1: subjects are able to compare 3-criteria alternatives.

Let us denote as P_1 a probability of subjects' ability to compare 3-criteria alternatives, and as P_2 a probability of subjects' inability to do so. Thus, formally, H_0: $P_1 \le P_2$, and H_1: $P_1 > P_2$.

Since we consider directional hypotheses, they should be checked on the base of one-side significance criterion. According to Table A from [7], in order to reject H_0 under $\alpha = 0{,}05$ (one-side significance criterion) and N=34 a value of subjects able to compare 3-criteria alternatives is to be equal at least to 23. Since N_1=29 > 23, we have to reject H_0 and to admit H_1.

Although there were a few comparisons of alternatives that differed in the estimates upon 4 criteria, the students identically compared such alternatives, if any, when they were presented by 4 different means. This is probably due to the qualitative nature of the choices and the utilization of color displays for distinguishing positive and negative aspects of the alternatives. Thus, the utilization of 2-criteria parts in the procedure was rather reliable for the majority of the students.

Discussion

The first finding of the experiments: while compared 3-criteria alternatives after preliminary comparison of 2-criteria alternatives, students made consistent choice in 85 % of cases. This result is statistically significant to confirmation of human being ability to compare 3-criteria alternatives through the presentation proposed here.

The second finding is the usefulness of two-round experiment arrangements. The first round gives an opportunity to find a compromise between the criteria, to study the problem, and thereby to form a strategy of choice (i.e., to find a job with a salary as high as possible, and/or with the most convenient work time and/or with the most preferable position and/or with time as little as possible to get to the workplace). The second round allows students to improve the selected strategy or to confirm the previous choice (in a majority of the cases).

The third finding is the importance of individually adjusted instances for the subjects. While solving such instances, they could make meaningful and multi-criteria choices. That is why we can recommend this approach for future research of human ability to solve multi-criteria tasks.

The fourth finding: about 37 % of students have selected the same objects, respectively, in the both rounds; from our point of view, it evidences on preliminary developed preferences. While about one half of students were stable, the others were unstable. We deem that it would be desirable to elaborate a Decision Support System (DSS) that could help a human being to solve the problem under consideration with better stability. Such DSS must meet the following requirements: a) qualitative technique for preference elicitation; b) on-line check of subject answers with respect to inconsistencies; c) intermediate results provide effective feedback; d) two rounds are reasonable.

Acknowledgment

The authors wish to express their gratitude to Mr. A. Kistanov, a PhD student of the Moscow State University, for his help in the organization of the experiments.

Literature

1. Russo J.E., Rosen L.D. An eye fixation analysis of multiattribute choice // Memory and Cognition. – 1975. – № 3. – P. 267-276.
2. Montgomery H, Svenson O. A think-aloud study of dominance structuring in decision processes // Process and Structure on Human Decision Making / Edit. by H. Montgomery, O. Svenson. – Chichester, 1989. – P. 135-150.
3. A constructive process view of decision making: multiple strategies in judgment and choice / J.W. Payne, J.R. Bettman, E. Coupey, E.J. Johnson // Current Themes in Psychological Decision Research / Edit. by O. Huber, J. Mumpower, J. van der Pligt, P. Koele. – Amsterdam, 1993 – P. 107-142.
4. Choice behaviour in a computer-aided multiattribute decision task / P. Korhonen, O. Larichev, H. Moshkovich, A. Mechitov, J. Wallenius // Journal of Multicriteria Decision Analysis. – 1997. – № 6. – P. 233-246.
5. Larichev O.I. Cognitive validity in design of decision-aiding techniques // Journal of Multicriteria Decision Analysis. – 1992. – № 1(3). – P. 127-138.
6. Larichev O.I. Ranking multicriteria alternatives: The method ZAPROS III // European Journal of Operational Research. –2001. – № 131(2001). – P. 550-558.
7. Runyon R.P. Nonparametric statistics: a contemporary approach // Reading. – MA: Addison Wesley, 1977. – 218 p.

Материал поступил в редакцию 09.04.02.