XX IMEKO World Congress
Metrology for Green Growth
September 9–14, 2012, Busan, Republic of Korea

CHAOTIC RESULTS OF MULTIDIMENSIONAL ORDINAL MEASUREMENTS

S. Muravyov/Presenter

Department of Computer-aided Measurement Systems and Metrology,
National Research Tomsk Polytechnic University, Tomsk, Russia, Pr. Lenina, 30, 634050

Abstract: Multidimensional ordinal measurement in a form of problem of a single consensus ranking determination for *m* rankings of *n* alternatives is considered in the paper. The Kemeny rule is one of deeply justified ways to solve the problem allowing to find such a linear order (Kemeny ranking) of alternatives that a distance (defined in terms of a number of pair-wise disagreements between rankings) from it to the initial rankings is minimal. But computational experiments outcomes show that the approach can give considerably more than one optimal solutions what argues instability of the measurement procedure. Hence, special efforts to avoid this phenomenon are needed.

Keywords: ordinal scale measurement, consensus relation, Kemeny ranking problem, multiple optimal solutions.

1. INTRODUCTION

In this paper we use a representational approach to measurement (see, for example, [1]). Let an empirical relational system $\langle A, \Lambda\rangle$ be given where $A = \{a_1, a_2, ...\}$ is a set of the manifestations of some property and $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_N\}$ is a family of relations over A. Let also a numerical relational system $\langle B, \mathrm{B}\rangle$ be given where $B = \{b_1, b_2, ...\}$ is a set of natural or real numbers (usually, symbols) and $\mathrm{B} = \{\beta_1, \beta_2, ..., \beta_N\}$ is a family of relations over B. Then measurement is an objective empirical operation f: $A \to B$, which maps the property manifestation onto numbers in such a way that the relations between numbers correspond to the relations between empirical elements, i.e. $\lambda_j = \beta_j$ for all $j = 1, ..., N$. It means that the mapping f is *homomorphic* [2].

It is well known that the representational treatment does not involve notation of uncertainty. However, introducing this concept into this kind of analysis can be quite easily done in case of *unidimensional quantitative* measurements, that is those in absolute, ratio and interval scales. Essentially larger complexity turns up in case of qualitative measurement (in ordinal or nominal scales) which are inherently multidimensional [2].

After short discussion how uncertainty of a qualitative measurement can be included into representational framework, in the paper we focus on quantitative (mainly ordinal) measurement uncertainty. We demonstrate that the uncertainty can be resulted from chaotic results of consensus relation determination and, consequently, special efforts to avoid this phenomenon are needed.

2. QUANTITATIVE MEASUREMENT UNCERTAINTY

Under homomorphism f the set A is broken up into non-intersected sets of preimages:

$$A = \bigcup_{b\in B} f^{-1}(b)\,, \tag{1}$$

where $f^{-1}(b)$ consists of all elements from A, having the same image in $f(A)$ [3]. This fact can serve as a mean to describe a *measurement uncertainty*. Fig. 1 illustrates this for a simple case of length measurement using a ruler, where the empirical elements form the series of values of quantity

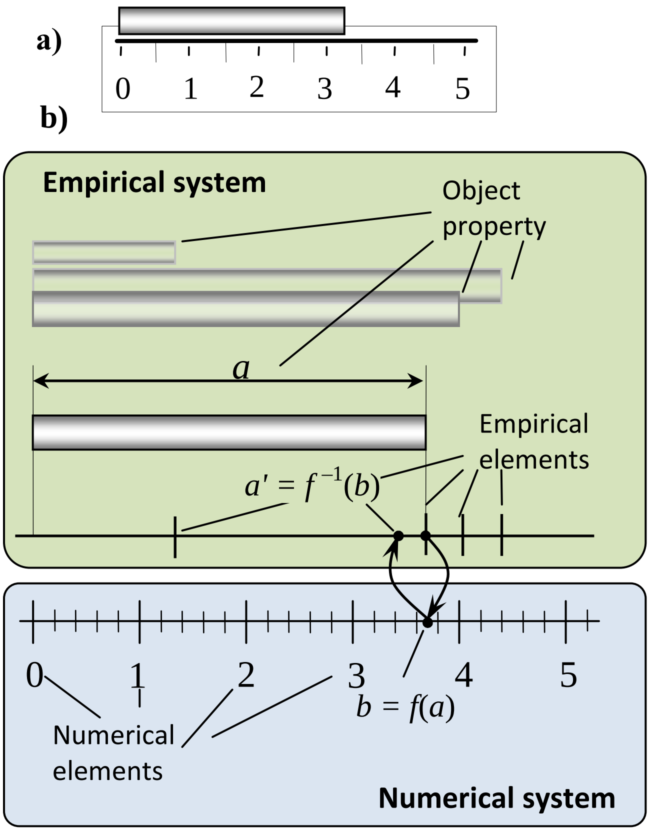

Fig. 1. Rod length measurement by means of a ruler (a), the length measurement as homomorphic mapping (b)

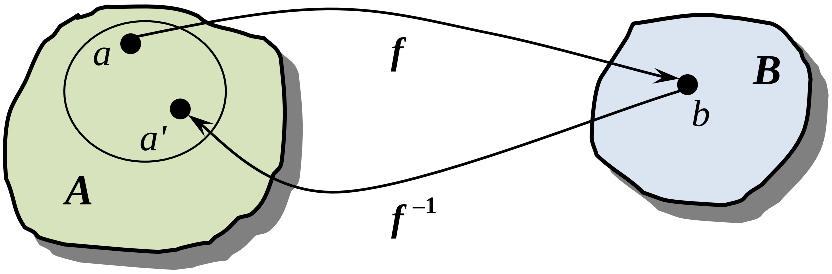

Fig. 2. Inverse mapping $a' = f^{-1}(b)$ never coincides with preimage a

and the numerical objects are the scale numberings. Here the function f is a correspondence of each ruler division to a definite numerical score.

It follows from consideration of Fig. 1 that inverse mapping $a' = f^{-1}(b)$ never coincides with preimage a since there are no empirical conditions able to guarantee validity of the hypothesis $a' = a$, see Fig. 2. Clear that from this point of view *measurement is a typical irreversible process*. This concept arises most frequently in thermodynamics [4,5].

All considerations stated above are about a *single* measured property which is described by a discrete variable A and its values a_i. When considering measurement of multiple heterogeneous properties [2], the uncertainty treatment should be implemented in different way.

3. MULTIDIMENSIONAL MEAUREMENT UNCERTAINTY

Let us deem the multiple heterogeneous properties measurement as a single consensus ranking determination for m rankings (voters), possibly including ties, of n alternatives (candidates). This is a classical problem that has been intensively investigated firstly as a Voting Problem in the framework of Social Choice Theory.

Condorcet rule: Condorcet in 1785, see [6], proposed a very natural rule for the consensus ranking determination: if some alternative obtains a majority of votes in pair-wise contests against every other alternative, the alternative is chosen as the winner in the consensus ranking. The Condorcet approach is widely recognised as the best rule for the consensus ranking determination, however, the binary relation defined by the *Condorcet rule* is not necessarily transitive, i.e. it can be for some consensus ranking β that $a_i \succ a_j$ and $a_j \succ a_k$ while $a_k \succ a_i$, $a_i, a_j, a_k \in \beta$ (see Fig. 3). This *Condorcet paradox* may occur rather frequently, for example its chances are higher than 50 % at $3 \le m \le \infty$ and $2 \le n \le 10$, if m is odd; ties reduce the probability, see, e.g., [7].

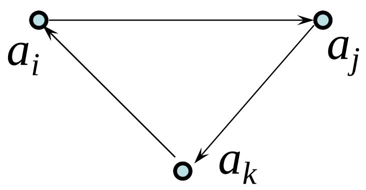

Fig. 3. Circular ambiguity in Condorcet rule

If the paradox occurs, the consensus relation does not exist and corresponding outcome is deemed to be *chaotic* [6,8-11].

Kemeny rule: The *Kemeny rule* [12] is a reasonable way to get over the paradox. Let us have m rankings on set $A = \{a_1, a_2, ..., a_n\}$ of n alternatives and the relation set $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_m\}$, where each of m rankings (also called preference relations or weak orders) $\lambda = \{a_1 \succ a_2 \succ ... \sim a_s \sim a_t \succ ... \sim a_n\}$ may include $\succ$, a strict preference relation π, and ~, an equivalence relation (or tie) ν, so that $\lambda = \pi \cup \nu$. The relation set Λ is called a *preference profile* for the given m rankings. Let Π be a set of all n! linear (strict) order relations $\succ$ on A. Each linear order corresponds to one of permutations of first n natural numbers $\mathbf{N_n}$.

Kemeny rule allows to find a consensus relation as such a linear order (Kemeny ranking) $\beta \in \Pi$ of alternatives that a distance $D(\beta, \Lambda)$ (defined in terms of a number of pair-wise disagreements between rankings) from β to the profile Λ is minimal, that is

$$\beta = \arg\min_{\lambda \in \Pi} D(\lambda, \Lambda)\ . \qquad (2)$$

But the approach has two drawbacks:

- the Kemeny Ranking Problem (KRP) had been proven to be *NP*-hard and
- it may have *considerably more than one* optimal solutions.

The former is not so disturbing since, for reasonable problem sizes (up to n < 30…50), there are exact algorithms for them to be effectively applied.

Strangely enough, the latter blemish has been given short shrift by researchers despite its importance for the problem applicability. In fact, multiple optimal solutions may rank the alternatives in significantly different ways what produces .

The Kemeny rule ambiguity can be shown by a simple example for $n = 3$, $m = 3$. Let the following three rankings be given:

$$\begin{aligned} &\lambda_1\!: a_1 \sim a_3 \succ a_2 \\ &\lambda_2\!: a_3 \succ a_1 \sim a_2 \\ &\lambda_3\!: a_2 \sim a_3 \succ a_1 \end{aligned} \qquad (2)$$

They belongs to the space of 13 weak orders as shown in Fig. 4. Integers at edges of the graph in Fig. 4 are Kemeny

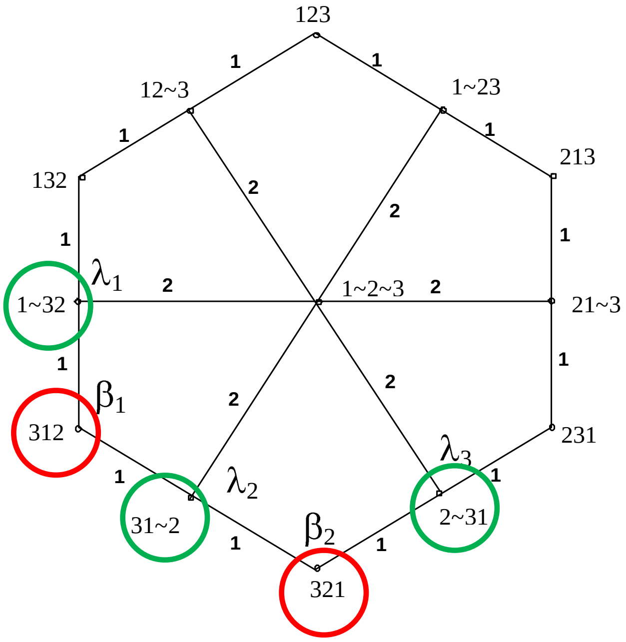

Fig. 4. The space of all weak orders (rankings) for $n = 3$; green circles denote initial rankings; red circles denote Kemeny rankings

distance $d(\lambda_i, \lambda_j)$ values. One can see that, in this case, the KRP has two solutions: $\beta_1 = a_3 \succ a_1 \succ a_2$ and $\beta_2 = a_3 \succ a_2 \succ a_1$. To avoid the ambiguity, evident way can be used, namely, to admit that $a_1 \sim a_2$ and, therefore, the final consensus relation is $\beta = a_3 \succ a_1 \sim a_2$, which coincides with one of initial rankings λ_2. These considerations are illustrated by Fig. 5, from comparison of which and Fig. 2 it is clear that multidimensional ordinal measurement uncertainty must be treated in a special way.

Instability of Kemeny rule: In order to investigate instability degree of the Kemeny rule a computational experiment has been conducted using modified branch-and-bound algorithm described in [13]. The modification allows to find all Kemeny rankings for a given profile [14]. The initial profile matrices were calculated by rankings obtained by uniting pseudo-random strict orders and ties generated separately on the basis of the C++ library function `randomize()`. Strict orders represented by $\mathbf{N_n}$ permutations were generated on the basis of the *uniform distribution* of integers in a specified range 1, ..., n. Thereby, in our experimentation, we stick the so called *impartial culture* condition implying just the uniform distribution of choices what is recognised to be a worst case for modelling preference profiles [7]. Ties were produced in similar way, and some additional measures were taken to reduce their density in each ranking.

This way, 50 profiles were generated at different values of m = 4...20 and n = 8...25, each served as input for the B&B algorithm. In most of the runnings, the algorithm outputs included relatively small number of optimal solutions: from 2 to 10. Relatively seldom (in nearly 10 % of cases) N_{nds} (that is a total number of the search tree nodes generated) was equal to several tens. However, in approx. 5 % of cases, there were outstanding solution numbers (that is number of Kemeny rankings) N_{kem} approaching to one million. One of the cases is shown in Tab. 1.

Tab.1 Profile 1 at m = 4, n = 20 and 6 ties

1	11	4	14	10	8	5~12		16	9	2~15		6	18	13	20	17	19	7	3
20	9	3	1	17	6	8	14	13	7	2	15	4	18	16	12	5	11	10	19
18	14	20	16	~9	3	7	~4	11	15	8	10	17	19~13		2	5	1	6	12
8	2	13	1	17	14	4	15	20	12	9	7	18	19	5	~3	10	6	11	16

$D(\beta, \Lambda) = 472$ $N_{\text{kem}} = 447\,614$ $N_{\text{nds}} = 6\,356\,082$

The preference profile, consisting of four rankings of 20 alternatives, having totally six ties, has 447 614 Kemeny rankings. The first seven optimal solutions for the profile are shown in Tab. 2. Notice that the B&B algorithm provided considerable reducing the solutions search space since it checked only 6 356 082 nodes while the cardinality of space of linear order relations is 20! = 2 432 902 008 176 640 000.

After removing all the six ties, the same profile brought to doubling of N_{kem} that became equal to 811 918, $D(\beta, \Lambda)$ = 476, N_{nds} = 10 440 879.

Tab.2 A fragment of the optimal solutions set for Profile 1

1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	11	10	6	16	19
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	11	10	16	6	19
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	11	10	16	19	6
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	11	16	10	6	19
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	11	16	10	19	6
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	16	11	10	6	19
1	8	14	4	20	9	2	13	15	17	7	12	18	3	5	16	11	10	19	6

Two example profiles above were found to be transitive

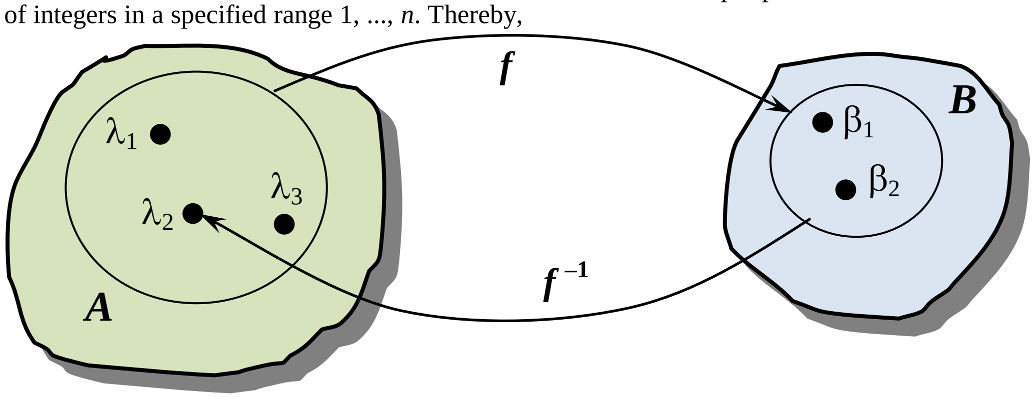

Fig. 5. Uncertainty treatment of multidimensional ordinal measurement for the example in Fig. 4

and, nevertheless, resulted in the paramount number of solutions, evidently due to essential inconsistency between initial rankings that, evidently, is the reason of so unstable and chaotic behaviour.

It is interesting to study a case of intransitive profile. Corresponding example is shown in Tab. 3 and Tab. 4. One can see that the set of solutions comprises a cycle (shown by grey background in Tab. 4) indicative of Condorcet paradox. However, the second optimal ranking saves the situation, and the final optimal solution could be {1, 4, 9, 3, 8, 2, 10, 6, 5, 7}.

Tab.3 Profile 2 at m = 5, n = 10, no ties

1	9	6	4	8	3	2	5	10	7
2	8	4	1	6	3	9	5	7	10
4	9	5	2	3	10	1	8	7	6
3	10	8	1	4	9	6	7	2	5
9	3	1	10	8	4	7	2	6	5
$D(\beta, \Lambda)$ = 128	N_{kem} = 4	N_{nds} = 1 376							

Tab.4 The optimal solutions set for Profile 2

1	4	9	3	2	10	8	6	5	7
1	4	9	3	8	2	6	5	10	7
1	4	9	3	8	2	10	6	5	7
1	4	9	3	10	8	2	6	5	7

The last example has been obtained not from pseudo-random generation and from an application the KRP to analysis of uncertainty intervals provided by m laboratories for some reference value of a measurand [15]. In corresponding profile, each ranking includes one pair of alternatives with the strict order relation and n – 2 pairs with the tie. Thus, Profile 3 presented by Tab. 5 has a high density of ties and rather minor differences in rankings, i.e., the condition of impartial culture is not valid here. This profile brought to 1440 optimal solutions, though they were all similar in positions of first through eighth places, namely: {10, 11, 9, 8, 12, 2, 3, 1, …}.

Tab.5 Profile 3 at m = 10, n = 13, 120 ties

10	~11	1	~2	~3	~4	~5	~6	~7	~8	~9	~12	~13
10	1	~2	~3	~4	~5	~6	~7	~8	~9	~11	~12	~13
10	~11	1	~2	~3	~4	~5	~6	~7	~8	~9	~12	~13
9	~10	~11	1	~2	~3	~4	~5	~6	~7	~8	~12	~13
8	~9	~10	~11	12	~1	~2	~3	~4	~5	~6	~7	~13
8	~9	~10	1	~2	~3	~4	~5	~6	~7	~11	~12	~13
11	~12	1	~2	~3	~4	~5	~6	~7	~8	~9	~10	~13
9	~10	~11	1	~2	~3	~4	~5	~6	~7	~8	~12	~13
2	~3	1	~4	~5	~6	~7	~8	~9	~10	~11	~12	~13
8	~9	1	~2	~3	~4	~5	~6	~7	~10	~11	~12	~13
$D(\beta, \Lambda)$ = 568	N_{kem} = 1 440	N_{nds} = 83 237										

4. CONCLUSION

It was shown in the paper that treatment of multidimensional ordinal measurement uncertainty must take into account unstable and chaotic behavior of consensus relation determination rule. It was demonstrated using the exact recursive B&B algorithm to determine all Kemeny rankings for different preference profiles. Even if a preference profile is transitive (that is, all rankings in it are consistent) a number of multiple solutions of the KRP may be large in spite of small amounts of m and n. The multiple solutions require to develop special measures to build some appropriate final convoluting solution. The issue can be a subject of further investigations.

5. REFERENCES

[1] L. Finkelstein, Measurement, information, knowledge-fundamental concepts, philosophical implications, applications, Proc. of the XIII IMEKO World Congress, (Turin, 5-9, September 1994), Vol. 1, Turin, Italy, pp. 11-18.

[2] S.V. Muravyov, V. Savolainen, Special interpretation of formal measurement scales for the case of multiple heterogeneous properties, Measurement, vol. 29, pp. 209-223, 2001.

[3] S.V. Muravyov, V. Savolainen, Discrete-mathematical approach to formal description of measurement procedure, Measurement, vol. 18, no. 2, pp. 71-80, 1996.

[4] P.W. Atkins, The Second Law, W.H. Freeman & Co., New York, 1997.

[5] P.W. Atkins, The Laws of Thermodynamics. A Very Short Introduction, Oxford University Press, New York, 2010.

[6] D.G. Saari, "Unsettling aspects of voting theory," Economic Theory, vol. 22, pp. 529–555, 2003.

[7] A.M.A. Van Deemen, "The probability of the paradox of voting for weak preference orderings," Social Choice and Welfare, 1999, vol. 16, pp. 171-182.

[8] D.A. Meyer, "Towards the global: complexity, topology and chaos in modelling, simulation and computation," arXiv:chao-dyn/9710005v2.

[9] D.A. Meyer, T.A. Brown, "Statistical mechanics of voting," Physical Review Letters, vol. 81, no. 8, pp. 1718-1721, 24 August 1998.

[10] T. Schwartz, "From Arrow to cycles, instability, and chaos by untying alternatives," Social Choice and Welfare, vol. 18, pp. 1-22, 2001.

[11] D. Richards, "Mutual knowledge structures and social coordination: a knowledge-induced equilibrium," University of Minnesota, The Center for Political Economy, Bulletin 98-2, 1998.

[12] J.G. Kemeny, J.L. Snell, Mathematical Models in the Social Sciences, New York, Ginn, 1962.

[13] S.V. Muravyov, "Rankings as ordinal scale measurement results," Metrology and Measurement Systems, vol. 13, no. 1, pp. 9-24, 2007.

[14] S.V. Muravyov, E.V. Tarakanov, "Multiple solutions of an exact algorithm for determination of all Kemeny rankings: preliminary experimental results," in Proceedings of ICIMCS, Hong Kong, vol. 1, ASME Press, New York, pp. 17-20, 2011.

[15] S.V. Muravyov, I.A. Marinushkina, Largest consistent subsets in interlaboratory comparisons: preference aggregation approach, in: Proceeding of IMEKO TC1+TC7+TC13 Symposium, Jena, Germany, 2011, urn:nbn:de:gbv:ilm1-2011imeko-036:0.