No. 45

79-9

Informational Evaluation of Decision Criteria in Situational Decision Making Model

August 1979

Nozomu Matsubara

University of Tsukuba

abstract

"The situational decision making model" is a qualitative, non-metric approach to a decision making. Originally aimed at a more realistic application of the statistical decision theory, the model does not assume an assumptive loss function, but consists of more essential ingredients of a decision making, *i.e.*, 'decision criteria,' 'situations' and 'actions.' The paper is of a cognitive nature, dealing with how to order, *retrospectively*, the decision criteria in terms of their influences, and hence to analyse the structure of decisions.

Keywords and Phrases;

Decision making, evaluation, ordering, decision procedure, entropy, Kullback's information number

1. Introduction

Let $\{a_1,\ldots, a_I\}$, $\{b_1,\ldots, b_J\}$, $\{e_1,\ldots, e_K\}$ be the set of "decision criteria," "actions," "situations" of decision making, respectively. A decision maker is to decide, given a criterion, on the appropriate action corresponding to the situation that prevails. The outcome of a particular decision making is denoted by the combination of these three ingredients, such as (a_i, e_k, b_j). We call this triplet "outcome." It is the process that a criterion a_i is given, e_k is privailing and then b_j is decided upon.

A decision maker's behavior is based upon the preference ordering for outcomes under the situation e_k:

$$\succ_k \text{ on } (a_i, e_k, b_j) \qquad (i = 1,\ldots, I;\ j = 1,\ldots, J) \tag{1}$$

The author [1] defines, on these orderings alone, the natural class of optimal decision procedures $(d : e \to b)$ for appropriate action 'b' corresponding to the situation 'e,' given the criterion 'a.' Note that some criteria are dominated and thus deleted from our consideration. Hence without loss of generality, any criteria are not dominated.

Now given a_i, the set of situations $E = \{e_1,\ldots, e_K\}$ is partitioned into ℓ_i disjoint and exhaustive subsets, $F_s^{(i)}$ $(s = 1, 2,\ldots, \ell_i)$ of situations, which we call "strategic subsets for a_i," by the relation;

$$\text{if } e, e' \in F_s^{(i)}, \text{ then } d(e) = d(e').$$

In other words, when a_i is taken, one need not switch actions as long as a situational change is limited within $F_s^{(i)}$. We denote by p_i the partition

$$E = \bigcup_{s=1}^{\ell i} F_s^{(i)}, \quad F_s^{(i)} \cap F_{s'}^{(i)} = \phi(s \neq s') \tag{2}$$

When all criteria are applied to the decision making together, the decision maker is now acting with the superposition of all partitions $p_1, p_2, \ldots, p_I$, which is again a partition of E and denoted by

$$\bar{p} = p_1 \wedge p_2 \wedge \cdots \wedge p_I \tag{3}$$

The purpose of the paper is to order decision criteria $a_1, a_2, \ldots, a_I$, when the decision making is thought to be based upon all criteria combined together.

The principle of ordering would certainly depend upon the nature of problems to solve. Our present concerns are, however, mainly two principles, one of which will apply:

(a) the Unity Principle; it would be desirable for $\bar{p}$ to be close to the single-subset partition,

(b) the Diversity Principle; it would be desirable for $\bar{p}$ to be composed of large number of subsets, preferably with equal "likelihood."

It is assumed that each situation E_i has the probability (as likelihood) of occurance θ_k, though it can be interpreted in various ways, e.g., in Bayesian context or in the frequentists' context.

The author employs informational quantities to implement the ordering based on these two principles introduced above. They are entropy and so called "Kullback's information number" ([2]). Numerical examples are given.

For subsequent chapters, let the partition $\overline{p}$ be

$$\overline{p}: \quad E = \bigcup_{s=1}^{\ell i} \overline{F}_s, \quad \overline{F}_s \cap \overline{F}_{s'} = \phi \ (s \neq s'). \tag{4}$$

By definition, each $\overline{F}_s (s = 1, \ldots, \ell)$ is the subsets of situations within which there is no strategic need to switch one's (the decision-maker's) action according to the situation which prevails, independently of decision criteria applied. We call $\overline{F}_s$ "over-all strategic subsets."

2. The Case of Unity Principle

The Unity Principle asserts that $\overline{p}$ be composed of small number of strategic subsets with large probabilities. The principle would prevail when it would be strategically tractable for strategic subsets to become large in each size and small in the total number.

Let us generally denote the entropy of the partition p of the abstract set X;

$$X = A_1 \cup \ldots \cup A_S, \quad A_\alpha \cap A_{\alpha'} = \phi \ (\alpha \neq \alpha'), \tag{5}$$

by

$$H(p) = - \Sigma_{\alpha=1}^{S} p_\alpha \log p_\alpha \tag{6}$$

with $p_\alpha = P(A_\alpha)$. $\qquad (p_\alpha \geqq 0, \ \Sigma p_\alpha = 1)$.

Then, we can prove the general property of the entropy,

$$H(p) \equiv -\Sigma_\alpha\ p_\alpha\ \log p_\alpha \geqq -\Sigma_\beta q_\beta\ \log q_\beta\ , \tag{7}$$

$$q_\beta = \Sigma_{G_\beta}\ p_\alpha,$$

where $G = \{G_1, G_2, \dots \}$ is (any) grouping of the index set (α) or, equivalently, of $A\alpha$'s. The property is called "convexity" ([2]) and proven by the fact that

$$\text{L.H.S.} - \text{R.H.S.} = \Sigma_\beta\ q_\beta(-\Sigma_{G_\beta}\ \frac{p_\alpha}{q_\beta}\ \log \frac{p_\alpha}{q_\beta}) \geqq 0, \tag{8}$$

the last quantity being called "conditional entropy" given the grouping G, and denoted by $H(p \mid G)$.

If, in our problem, $H(\overline{p}) = 0$ (or close to 0), then the Unity Principle is already (almost) fulfilled since in that case we have $\ell \doteq 1$.
If on the contrary $H(\overline{p})$ is sufficiently large, the most influential decision criterion, a_1*(say), to cause this departure would be the one to attain the maximum decrease, Δ_1^*, from $H(\overline{p})$ to

$$H_i \equiv H(p_1 \wedge \cdots \wedge p_{i-1} \wedge p_{i+1} \wedge \cdots \wedge p_I) \tag{9}$$

with i running over $i = 1, \dots, I$;

$$\begin{aligned} \Delta_1^* &= \max_i\ [\ H(\overline{p}) - H_i\ ] \\ &= \max_i\ H(\overline{p} \mid p_1 \wedge \cdots \wedge p_{i-1} \wedge p_{i+1} \wedge \cdots \wedge p_I) \end{aligned} \tag{10}$$

The quantity in the bracket [ ] is the effect of criterion a_i, given the rest of criteria. The second most influential criterion, a_2^*, is found in similar way, that is, given $a_1{*}$ already deleted, seeking the maximum decrease, $\Delta_2{*}$, of entropy caused again by deleting any one criterion. Repeating this process, one can order the set of decision criteria a_1, a_2, ..., a_I in terms of influence;

$$a_1{*},\ a_2{*},\ \ldots\ ,\ a_I{*},$$

with each contribution (given the subsequent ones) $\Delta_1{*}$, $\Delta_2{*}$, ..., $\Delta_I{*}$ adding up to $H(\overline{p})$;

$$H(\overline{p}) = \Delta_1{*} + \Delta_2{*} + \ldots + \Delta_I{*} \tag{11}$$

Note that it does not necessarily hold that $\Delta_i{*}$ decreases monotonously

$$\Delta_1{*} \geqq \Delta_2{*} \geqq \ldots \geqq \Delta_I{*}, \tag{12}$$

though the construction process might suggest to be the case. Our evaluation of decision criteria, however, is the conditional one, i.e., the evaluation of the effect of a_i "given the rest of them", thus admitting the possible (and natural also) interaction among decision criteria. Monotonicity of $\Delta_i{*}$ could hardly be the case in such a complexity. It would be, therefore, surprising and of special interest, if the monotonicity holds.

3. The Case of Diversity Principle

The Diversity Priciple asserts that $\overline{p}$ be composed of large number of strategic subsets preferably with equal probabilities. This principle aims at more exactness to adjust actions flexibly with the situational change, in contrast with the Unity Principle which, essentially, is the expression of the strategic stability in changing situations. The assignment of equal probabilities may not be the only one for the mathematical expression of diversity, though it will serve the first-order approximation.

Let, in general,

$$\mathbb{p} = (p_1, \ldots, p_S) , \; \mathbb{p}' = (p_1', \ldots, p_S') \tag{13}$$

be two probability assignments to the purtition (5), and define "Kullback's information number" to be

$$I_{\mathbb{p}, \mathbb{p}'}(p) = \Sigma_{\alpha=1}^{S} p_\alpha \log \frac{p_\alpha}{p'_\alpha} \; (\geqq 0) \tag{14}$$

Convexity holds also for this quantity;

$$\Sigma_\alpha p_\alpha \log \frac{p_\alpha}{p'_\alpha} \geqq \Sigma_\beta q_\beta \log \frac{q_\beta}{q'_\beta},$$

$$q_\beta = \Sigma_{G_\beta} p_\alpha , \quad q'_\beta = \Sigma_{G_\beta} p'_\beta \tag{15}$$

for any grouping $G = \{G_1, G_2, \ldots\}$

Correspondingly to (8), we have for (15)

$$\text{L.H.S.} - \text{R.H.S.} = \Sigma_\beta \, q_\beta (\Sigma_{G_\beta} \; (p_\alpha/q_\beta) \; \log \frac{(p_\alpha/q_\beta)}{(p'_\alpha/q'_\beta)}) \geqq 0, \tag{16}$$

which serves the proof of (15). The quantity (16) is called "conditional Kullback's information number" and denoted by $I_{\mathbb{p}, \mathbb{p}'}(p | G)$ or briefly $I(p | G)$.

$I_{\mathbb{P},\mathbb{P}'}(p)$ signifies the degree of departure of $\mathbb{P}$ from $\mathbb{P}'$. In our problem, then, let us define the departure of $\mathbb{P} = (\pi_1, \ldots, \pi_\ell)$ from the complete diversity $\mathbb{P}' = (1/\ell, \ldots, 1/\ell)$ to be

$$I(\overline{p}) = \Sigma_{s=1}^{\ell} \pi_s \log \frac{\pi_s}{(1/\ell)} \tag{17}$$

The ordering procedure for the Diversity Principle goes in quite the same manner as in (9) - (12), with $I(\cdot)$ this time in the place of $H(\cdot)$ there.

There are two points noteworthy. First, $I(\overline{p})$ has the statistical interpretation that, if $\pi_1, \ldots, \pi_\ell$ are relative frequency counts $\hat{\pi}_1, \ldots, \hat{\pi}_\ell$ in the total n counting on the strategic subsets $\overline{F}_1, \ldots, \overline{F}_\ell$, then the likelihood ratio test statistic

$$\lambda = \Pi_{s=1}^{\ell} \hat{\pi}_s^{n\hat{\pi}_s} / (1/\ell)^n \tag{18}$$

to test the null hypothesis (of the "complete diversity" in our problem)

$$H_0 : \pi_1 = \cdots = \pi_\ell = 1/\ell \tag{19}$$

has the logarithmic expression

$$\begin{aligned} \log \lambda &= n\, \hat{I}(\overline{p}) , \\ \hat{I}(\overline{p}) &= \Sigma_{s=1}^{\ell} \hat{\pi}_s \log \frac{\hat{\pi}_s}{(1/\ell)} , \end{aligned} \tag{20}$$

whose limiting distribution is

$$2\log \lambda \sim \chi^2(\ell-1) \tag{21}$$

as $n \to \infty$([2] , p.98).

Second, the Unity Principle and the Diversity Principle is dual in the sense that

$$H(\overline{p}) + I(\overline{p}) = \log \ell \ (= \text{const.}) \tag{22}$$

4. Numerical Examples

Let us present the theory thus far developed in two typical examples. They get started with partitions as given, though it would serve more to the consistency of the paper to derive partitions themselves from the system of the situational ordering (1), which is left to the author's previous paper [1].

In Examples, we denote $e_k \equiv k$ for the sake of simplicity.

Example 1. Let $I = 10$, $K = 1000$. The assumption of such a large K would be by no means unrealistic, since the situation is usually made up of the combination of several, dependent or independent, factors. For example, situations with 4 factors $e = (f, g, h, i)$, $f, g, h, i \in \{1, 2, 3, 4, 5\}$ form the 625-element set of situations.

In our examples throughout (1 and 2), we limit, for the sake of simplicity, all partitions to "slit-type." By "slits" for the slit-type partition of $\{1, 2, \ldots, K\}$ into subsets $\{F_1, F_2, \ldots, F_\ell\}$, we mean $\ell - 1$ integers k_s $(s = 1, 2, \ldots, \ell-1)$ satisfying

$$1 \leqq k_1 < \ldots < k_{\ell-1} \leqq K \tag{23}$$

and representing F_s as

$$F_s = \{k_{s-1} + 1, \ldots, k_s\} \qquad (s = 1, \ldots, \ell) \tag{24}$$

with the convention $k_0 \equiv 0$, $k_\ell \equiv K$. Slits for 10 partitions are given in Table 1. The assignment of $\theta_k (k = 1, \ldots, K)$ is necessary only to derive $\pi_s = P(\overline{F}_s)$ $(s = 1, \ldots, \ell)$. $\overline{p}$ is also slit-type with $\ell = 92$, and represented in Table 2 by s, k_s and π_s $(s = 1, \ldots, 92$. k_{92} is not called 'slit' in our definition, however). The resulting orderings of decision criteria and their contributions are shown in Table 3.

What is the most marked with this Example is that partitions $p_1, p_2, \ldots, p_{10}$ tend to be finer increasingly and the order is almost retrieved by our procedure in Table 3 both in case of the Unity Principle and the Diversity Principle. The procedure passes the "internal check."

Contributions rather increase in the Unity case, whereas they decrease in the Diversity case. This will afterwards turn out to be a generally observed tendency. That they go in opposite direction is understood reasonable by seeing that if the R.H.S. of the identity

$$H(p \mid G) + I(p \mid G) = (H(p) + I(p)) - (H(G) + I(G))$$

remains rather stable as in (22) under mild conditions.

<u>Example 2.</u> Let $I = 10$, $K = 1000$. Slit-type partitions are considered also considered also and given in Table 4. ℓ turns out to be 94.

This Example 2 assumes, in contrast with Example 1 (Table 1), partitions of comparable degree of finess. Table 5 is the representation of $\overline{p}$ for the Example 2.

Table 6 reveals resulting orderings. Increase and decrease are likewise observed for this Example 2. However, the increase this time for the Unity case is almost typical one, whereas for the Diversity case the decrease is less marked with some irregularities. Roughly, two orderings run in the other directions. Otherwise, two Principles would have lost their independent grounds of existence, since in Example 1 two orderings go in the same direction.

Example 2 obviously presents us harder problems than Example 1, but the result of the Example 2 shows that our procedure does not fail us and furnish us a sufficient distinguishability.

5. Acknowlegement

The autor is grateful to Mr. M. Murakami of the Institute of Statistical Mathematics for his computational works and to Miss Y. Nakayama for her excellent typewriting of the manuscript.

[REFERENCES]

[1] N. Matsubara(1979), 'Invariance of a Complete Class in "Situational" Decision Making' in 'Optimizing Methods in Statistics' (edited by J. S. Rustagi), Academic Press, New York.

[2] S. Kullback(1968),'Information Theory and Statistics' (first published by John Wiley and Sons, Inc.), Dover Publications, Inc., New York.

decision criteria	slits for strategic subsets
a_1	332, 666
a_2	256, 342, 751, 992
a_3	55, 173, 222, 514, 741, 852
a_4	77, 123, 312, 415, 541, 632, 778, 892
a_5	77, 135, 289, 320, 451, 555, 641, 782, 852, 951
a_6	5, 51, 111, 222, 278, 352, 462, 512, 586, 632, 708, 986
a_7	6, 82, 142, 305, 315, 478, 492, 512, 641, 784 792, 831, 852, 890
a_8	52, 142, 196, 245, 283, 301, 370, 452, 492, 562 583, 621, 752, 790, 881, 890
a_9	7, 102, 230, 264, 305, 385, 410, 482, 512, 593 601, 652, 742, 803, 850, 912, 950, 983
a_{10}	5, 12, 76, 105, 190, 222, 285, 301, 395, 452 476, 520, 531, 666, 682, 701, 841, 892, 903, 957

Table 1.

decision criteria and strategic subsets (Example 1)

s	k_s	π_s
1	5	0.00738
2	6	0.00176
3	7	0.00016
4	12	0.00619
5	51	0.03881
6	52	0.00196
7	55	0.00241
8	76	0.02170
9	77	0.00079
10	82	0.00528
11	102	0.02163
12	105	0.00352
13	111	0.00657
14	123	0.01436
15	135	0.01377
16	142	0.00755
17	173	0.03336
18	190	0.01885
19	196	0.00379
20	222	0.02659
21	230	0.00525
22	245	0.01421
23	256	0.01549
24	264	0.00765
25	278	0.01292
26	283	0.00497
27	285	0.00123
28	289	0.00256
29	301	0.00898
30	305	0.00577
31	312	0.00651
32	315	0.00352
33	320	0.00620
34	332	0.01260
35	342	0.00897
36	352	0.00641
37	370	0.01859
38	385	0.01498
39	395	0.00999
40	410	0.01671

Table 2.

over-all strategic subsets
in slit representation and probabilities

s	k_s	π_s
41	415	0.00740
42	451	0.03991
43	452	0.00128
44	462	0.00492
45	476	0.01650
46	478	0.00220
47	482	0.00400
48	492	0.00994
49	512	0.02221
50	514	0.00051
51	520	0.00667
52	531	0.01202
53	541	0.01304
54	555	0.01204
55	562	0.00861
56	583	0.02290
57	586	0.00382
58	593	0.00378
59	601	0.01017
60	621	0.01821
61	632	0.00489
62	641	0.00794
63	652	0.01017
64	666	0.01016
65	682	0.01999
66	701	0.02016
67	708	0.00490
68	741	0.03095
69	742	0.00196
70	751	0.00772
71	752	0.00039
72	778	0.02985
73	782	0.00260
74	784	0.00077
75	790	0.00568
76	792	0.00180
77	803	0.01313
78	831	0.02615
79	841	0.01027
80	850	0.00902
81	852	0.00301
82	881	0.02724
83	890	0.01186
84	892	0.00097
85	903	0.01076
86	912	0.00757
87	950	0.03157
88	951	0.00114
89	983	0.03132
90	986	0.00217
91	992	0.00748
(92)	(1000)	(0.00632)

Table 2. (Continued)

Symbols for Ordering	Unity Princile		Diversity Principle	
	Ordered Criteria	Contributions Δ_i*	Ordered Criteria	Contributions Δ_i*
a_1*	a_9	0.189255	a_9	0.063217
a_2*	a_{10}	0.209990	a_{10}	0.060529
a_3*	a_8	0.195014	a_8	0.041787
a_4*	a_6	0.222506	a_7	0.057174
a_5*	a_7	0.222556	a_5	0.037830
a_6*	a_5	0.274405	a_6	0.032880
a_7*	a_4	0.580347	a_3	0.017415
a_8*	a_3	0.752933	a_2	0.018960
a_9*	a_2	0.436150	a_4	0.008828
a_{10}*	a_1	1.098336	a_1	0.001676
Total	——	4.181492	——	0.340296

Table 3.

orderings of decision criteria
by two principles
(Note) The logarithm is to the base e = 2.71828 ··· .

decision criteria	slits for strategic subsets
a_1	110, 156, 208, 386, 418, 463, 527, 572, 776, 926
a_2	112, 193, 228, 309, 329, 603, 782, 837, 884, 901
a_3	37, 39, 175, 241, 305, 326, 482, 611, 695, 722
a_4	41, 42, 90, 161, 187, 281, 305, 724, 753, 821
a_5	5, 18, 76, 116, 190, 256, 436, 553, 608, 922
a_6	83, 190, 324, 330, 368, 482, 731, 845, 849, 962
a_7	74, 169, 424, 477, 488, 538, 572, 946, 947, 993
a_8	104, 217, 302, 528, 633, 665, 687, 737, 955, 975
a_9	75, 344, 395, 401, 634, 670, 794, 816, 907, 927
a_{10}	45, 54, 143, 251, 368, 513, 643, 652, 724, 837

Table 4.

decision criteria and strategic subsets (Example 2)

s	k_s	π_s
1	5	0.00594
2	18	0.01283
3	37	0.01595
4	39	0.00215
5	41	0.00348
6	42	0.00122
7	45	0.00445
8	54	0.00951
9	74	0.02092
10	75	0.00158
11	76	0.00185
12	83	0.00471
13	90	0.01024
14	104	0.01368
15	110	0.00656
16	112	0.00086
17	116	0.00569
18	143	0.02813
19	156	0.00817
20	161	0.00512
21	169	0.00684
22	175	0.00821
23	187	0.01235
24	190	0.00467
25	193	0.00234
26	208	0.01147
27	217	0.00802
28	228	0.01145
29	241	0.00994
30	251	0.00725
31	256	0.00483
32	281	0.02564
33	302	0.02417
34	305	0.00249
35	309	0.00294
36	324	0.01777
37	326	0.00131
38	329	0.00366
39	330	0.00132
40	344	0.01543

Table 5.

over-all strategic subsets
in slit representation and probabilities

s	k_s	π_s
41	368	0.02324
42	386	0.01736
43	395	0.00717
44	401	0.00395
45	418	0.01756
46	424	0.00775
47	436	0.01104
48	463	0.02887
49	477	0.01347
50	482	0.00389
51	488	0.00587
52	513	0.02185
53	527	0.01157
54	528	0.00138
55	538	0.01196
56	553	0.01663
57	572	0.01978
58	603	0.03238
59	608	0.00368
60	611	0.00364
61	633	0.02210
62	634	0.00057
63	643	0.01043
64	652	0.00924
65	665	0.01551
66	670	0.00351
67	687	0.01822
68	695	0.00764
69	722	0.02730
70	724	0.00306
71	731	0.00769
72	737	0.00309
73	753	0.01972
74	776	0.02189
75	782	0.00760
76	794	0.01048
77	816	0.02190
78	821	0.00536
79	837	0.01685
80	845	0.00759
81	849	0.00445
82	884	0.03291
83	901	0.01395
84	907	0.00630
85	922	0.01441
86	926	0.00597
87	927	0.00152
88	946	0.02101
89	947	0.00041
90	955	0.00536
91	962	0.00739
92	975	0.01638
93	993	0.01569
(94)	(1000)	(0.00631)

Table 5. (Continued)

symbols for Ordering	Unity Principle		Diversity Principle	
	Ordered Criteria	Contributions Δ_i^*	Ordered Criteria	Contributions Δ_i^*
a_1^*	a_5	0.134260	a_8	0.042946
a_2^*	a_{10}	0.130650	a_7	0.037712
a_3^*	a_4	0.140497	a_2	0.035618
a_4^*	a_3	0.172319	a_5	0.045909
a_5^*	a_2	0.179558	a_6	0.024805
a_6^*	a_6	0.217718	a_9	0.028221
a_7^*	a_8	0.304661	a_3	0.022470
a_8^*	a_9	0.512323	a_4	0.013382
a_9^*	a_7	0.703153	a_{10}	0.006655
a_{10}^*	—	1.772983	—	0.017454

Table 6.

orderings of decision criteria
by two principles