AN INFORMATIONAL BASIS FOR VOTING RULES

Alexander Karpov

National Research University Higher School of Economics

Motivation

Fishburn's data requirements classes

- C1 — voting rules that only depend on pairwise majority comparisons,
- C2 — voting rules that only depend on weighted pairwise majority comparisons,
- C3 — voting rules that require more information than weighted pairwise majority comparisons.

Fishburn's classification do not represent the relation "need more information". Intuitively, the plurality rule requires less information than the Borda rule, but according to Fishburn's classification, the plurality rule belongs to C3, and the Borda rule belongs to C2.

Related Literature

- Communication Complexity of voting rules (Conitzer, Sandholm 2005)
- Message functions and informational requirements (Sato 2011, 2016)

Framework

- Let a finite set $X = \{1, \dots, m\}, m \geq 2$ be the set of alternatives and a finite set $\mathcal{N} = \{1, \dots, n\}, n \geq 2$ be the set of agents (voters). Each agent $i \in \mathcal{N}$ has a strict preference P_i over X (linear order). Let $\mathcal{L}(X)$ be the set of all possible linear orders over X. An *n*-tuple of preference orders generates preference profile $\mathcal{P} = (P_1, \dots, P_n) \in \mathcal{L}(X)^n$.
- Within this model, the names of voters (anonymity) and names of alternatives (neutrality) do not matter. An anonymous and neutral equivalence class (ANEC) is a set of preference profiles that could be obtained from each other by permuting preference orders and renaming alternatives.

ANECs, m=3, n=3.

$$\mathcal{P}_1 = \begin{pmatrix} a & a & a \\ b & b & b \\ c & c & c \end{pmatrix}; \mathcal{P}_2 = \begin{pmatrix} a & a & a \\ b & b & c \\ c & c & b \end{pmatrix}; \mathcal{P}_3 = \begin{pmatrix} a & a & b \\ b & b & a \\ c & c & c \end{pmatrix}; \mathcal{P}_4 = \begin{pmatrix} a & a & c \\ b & b & a \\ c & c & b \end{pmatrix};$$

$$\mathcal{P}_5 = \begin{pmatrix} a & a & b \\ b & b & c \\ c & c & a \end{pmatrix}; \mathcal{P}_6 = \begin{pmatrix} a & a & c \\ b & b & b \\ c & c & a \end{pmatrix}; \ \mathcal{P}_7 = \begin{pmatrix} a & a & b \\ b & c & a \\ c & b & c \end{pmatrix}; \mathcal{P}_8 = \begin{pmatrix} a & a & b \\ b & c & c \\ c & b & a \end{pmatrix};$$

$$\mathcal{P}_9 = \begin{pmatrix} a & b & c \\ b & a & a \\ c & c & b \end{pmatrix}; \mathcal{P}_{10} = \begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix}.$$

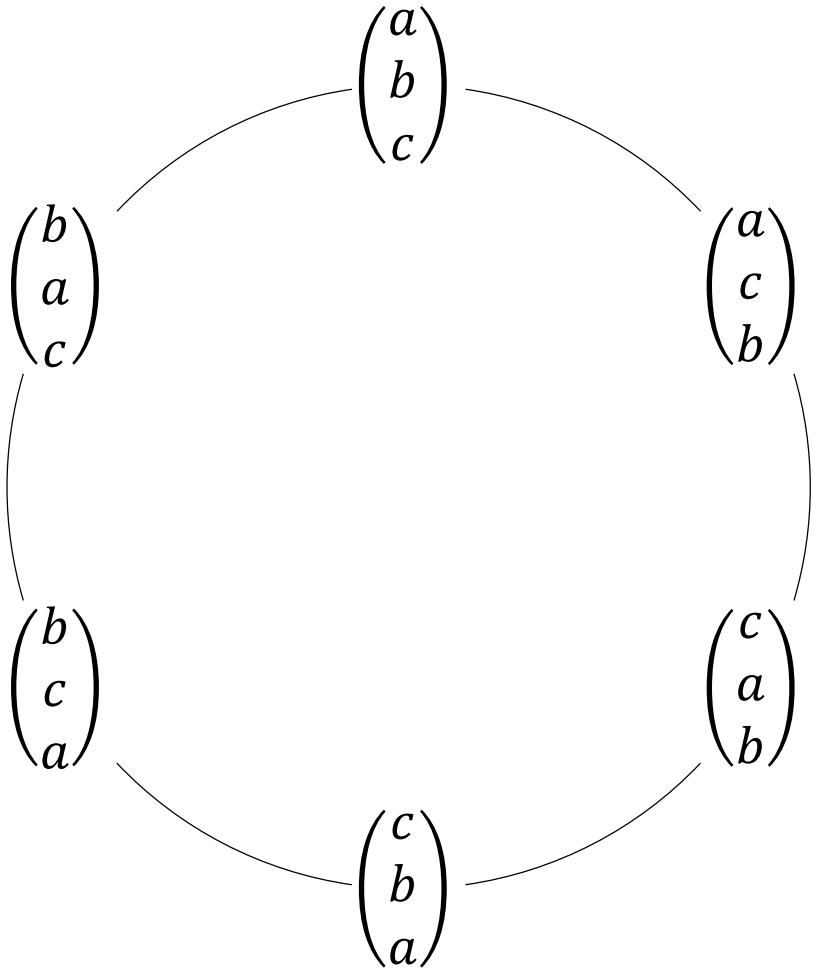
Reversal symmetry

- A preference profile $\mathcal{P}'$ is the **reversal of preference profile** $\mathcal{P}$ if ($\forall x \in X, \forall i \in \mathcal{N}, pos(P_i, x) = m + 1 - pos(P'_i, x)$).
- An ANEC is **self-symmetric** if for every $\mathcal{P}$ from the ANEC, the reverse profile $\mathcal{P}'$ belongs to the same ANEC.
- A pair of ANECs is **reverse symmetric** if for every $\mathcal{P}$ from one ANEC, the reverse profile $\mathcal{P}'$ belongs to the other ANEC.
- Considering symmetric ANECs as equivalent, we obtain a set of **reverse invariant ANECs (RIANECs)** and binary relation $\sim_{RIANEC}$, which contains $\sim_{ANEC}$.

Circle representation of preference orders.

$$P_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, P_1 = \begin{pmatrix} a \\ c \\ b \end{pmatrix}, P_2 = \begin{pmatrix} c \\ a \\ b \end{pmatrix},$$

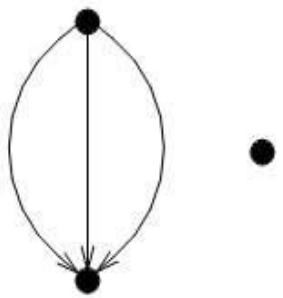
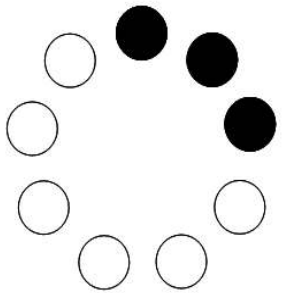
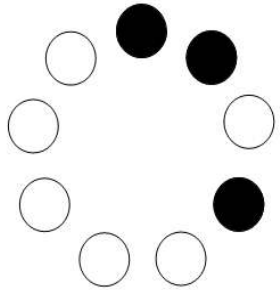
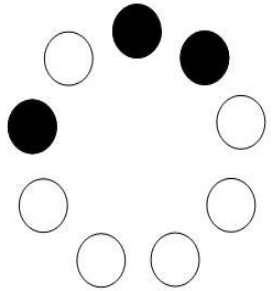
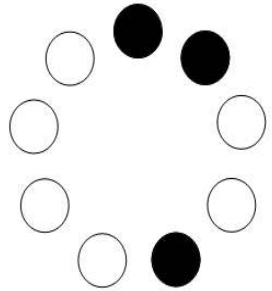
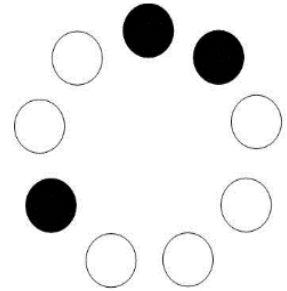
$$P_3 = \begin{pmatrix} c \\ b \\ a \end{pmatrix}, P_4 = \begin{pmatrix} b \\ c \\ a \end{pmatrix}, P_5 = \begin{pmatrix} b \\ a \\ c \end{pmatrix}.$$



Multigraph representation of preference profile. *Having 3 alternatives as vertices of a graph, for each preference order in the profile, we define an arc from the best alternative to the worst alternative.*

Bracelet representation of preference profile. *According to the circle representation of preference orders (formula (6) and figure 1), we numerate preference orders from 0 to 5 in a clockwise manner. Every preference order in a profile is represented by a black bead. The space between preference orders in the circle representation of preference orders is represented by a white bead. A preference profile is a bracelet with $n + 6$ beads, where exactly 6 beads are white. We take n_0 black beads, where n_0 is the number of type 0 preference orders, then one white bead, than n_1 black beads, where n_1 is the number of type 1 preference orders, etc. Adding a white bead between n_5 black beads, where n_5 is the number of type 5 preference orders, and n_0 black beads, where n_0 is the number of type 0 preference orders, we complete the bracelet.*

Multigraph and bracelet representations

Number	$\mathcal{P}_1$	$\mathcal{P}_2$	$\mathcal{P}_3$	$\mathcal{P}_4$	$\mathcal{P}_5$
ANEC	$\begin{matrix} a & a & a \\ b & b & b \\ c & c & c \end{matrix}$	$\begin{matrix} a & a & a \\ b & b & c \\ c & c & b \end{matrix}$	$\begin{matrix} a & a & b \\ b & b & a \\ c & c & c \end{matrix}$	$\begin{matrix} a & a & c \\ b & b & a \\ c & c & b \end{matrix}$	$\begin{matrix} a & a & b \\ b & b & c \\ c & c & a \end{matrix}$
Multigraph representation					
Bracelet representation					

Enumeration. m=3

- $\#ANEC(3,n) = \begin{cases} \frac{1}{6}\binom{n+5}{5} + \frac{1}{16}(n+4)(n+2) + \frac{1}{9}(n+3), \text{if } n \equiv 0\ (mod\ 6); \\ \frac{1}{6}\binom{n+5}{5}, \text{if } n \equiv (1\ or\ 5)(mod\ 6); \\ \frac{1}{6}\binom{n+5}{5} + \frac{1}{16}(n+4)(n+2), \text{if } n \equiv (2\ or\ 4)(mod\ 6); \\ \frac{1}{6}\binom{n+5}{5} + \frac{1}{9}(n+3), \text{if } n \equiv 3\ (mod\ 6). \end{cases}$
- $\#RIANEC(3,n) = \begin{cases} \frac{1}{12}\binom{n+5}{5} + \frac{1}{96}(n+7)(n+4)(n+2) + \frac{1}{18}(n+6), \text{if } n \equiv 0\ (mod\ 6); \\ \frac{1}{12}\binom{n+5}{5} + \frac{1}{96}(n+5)(n+3)(n+1), \text{if } n \equiv (1\ or5)\ (mod\ 6); \\ \frac{1}{12}\binom{n+5}{5} + \frac{1}{96}(n+7)(n+4)(n+2), \text{if } n \equiv (2\ or\ 4)(mod\ 6); \\ \frac{1}{12}\binom{n+5}{5} + \frac{1}{96}(n+5)(n+3)(n+1) + \frac{1}{18}(n+6) - \frac{1}{6}, \text{if } n \equiv 3(mod\ 6). \end{cases}$
- $\#SSANEC(3,n) = \begin{cases} \left\lceil \frac{1}{48}(n+4)^2(n+2) \right\rceil, \text{if } n \text{ is even}; \\ \frac{1}{48}(n+5)(n+3)(n+1), \text{if } n \text{ is odd}. \end{cases}$

Voting situations

- Voting situation, we mean a class of ANECs that are indiscernible, having information obtained from a procedure.
- For the case of 4 agents and 3 alternatives, we have only four types of voting situations for plurality rule described by top alternatives: (a, a, a, a), (a, a, a, b), (a, a, b, b), and (a, a, b, c).
- The run-off rule utilizes more information than the plurality rule. In this case, we have five voting situations described by the first-best and the second-best alternatives
- (a, a, a, a), (a, a, a, b), (a, a, b, b), $\begin{pmatrix} a, a, b, c \\ ?, ?, ?, b \end{pmatrix}$, $\begin{pmatrix} a, a, b, c \\ ?, ?, ?, a \end{pmatrix}$.

Information utilization index

$\#Rule(m,n)$ is the number of anonymous and neutral equivalent voting situations induced by a rule.

The information utilization index of a rule is defined as follows:

$$I(Rule,m,n)=\frac{\#Rule(m,n)-1}{\#ANEC(m,n)-1}$$

$I(Rule,m,n)=1$ corresponds to the full discernibility power of a voting rule. A higher values of the index means that a rule utilizes more information and it is in some sense more complicated.

Plurality rule

The number of voting situations induced by the plurality rule is equal to

$$\#Plurality(m,n) = \sum_{i=1}^{m} p_i(n)\,,$$

Where $p_i(n)$ is the number of partition n on i parts.

Proposition 1. *For $m = 3$, the number of voting situations generated by the plurality rule is equal to*

$$\#Plurality(3,n) = round\left[\frac{1}{12}(n+3)^2\right]$$

Note that ties in rounding in this formula never arise.

Run-off rule

For the 3 alternatives and 3 agents case, we have the following voting situations (ANECs and corresponding partitions of the top choices and partition of votes in the second round):

$\mathcal{P}_1, \mathcal{P}_2$: $(3,0,0)$;

$\mathcal{P}_3, \mathcal{P}_4, \mathcal{P}_5, \mathcal{P}_6, \mathcal{P}_7, \mathcal{P}_8$: $(2,1,0)$;

$\mathcal{P}_9, \mathcal{P}_{10}$: $(1,1,1), (2,1)$.

Proposition 2. *For $m = 3$, the number of voting situations induced by the run-off rule is equal to*

$$\begin{cases} \left\lceil \frac{1}{1728}(7n^3 + 192n^2 + 1008n + 1088) \right\rceil, if\ n \equiv (0\ or\ 4)\ (mod\ 12); \\ \left\lceil \frac{1}{1728}(7n^3 + 165n^2 + 729n + 827) \right\rceil, if\ n \equiv (1\ or\ 9)\ (mod\ 12); \\ \frac{1}{1728}(7n^3 + 192n^2 + 804n + 1024), if\ n \equiv 2\ (mod\ 12); \\ \left\lceil \frac{1}{1728}(7n^3 + 165n^2 + 837n + 935) \right\rceil, if\ n \equiv (3\ or\ 7)\ (mod\ 12); \\ \frac{1}{1728}(7n^3 + 165n^2 + 633n + 475), if\ n \equiv 5\ (mod\ 12); \\ \left\lceil \frac{1}{1728}(7n^3 + 192n^2 + 900n + 1088) \right\rceil, if\ n \equiv (6\ or\ 10)\ (mod\ 12); \\ \frac{1}{1728}(7n^3 + 192n^2 + 912n + 1024), if\ n \equiv 8\ (mod\ 12); \\ \frac{1}{1728}(7n^3 + 165n^2 + 741n + 583), if\ n \equiv 11\ (mod\ 12). \end{cases}$$

Borda rule

The Borda rule is a scoring rule in which the worst alternative has a score of 0, and the best alternative has a score of $m-1$. The Borda rule utilizes information only about the sum of scores for each alternative.

For the 3 alternatives and 3 agents case, we have the following voting situations (ANECs and corresponding sums of scores vector in decreasing order):

$$\mathcal{P}_1\colon \begin{pmatrix}6\\3\\0\end{pmatrix}; \mathcal{P}_2\colon \begin{pmatrix}6\\2\\1\end{pmatrix}; \mathcal{P}_3\colon \begin{pmatrix}5\\4\\0\end{pmatrix}; \mathcal{P}_4\colon \begin{pmatrix}5\\2\\2\end{pmatrix}; \mathcal{P}_5\colon \begin{pmatrix}4\\4\\1\end{pmatrix}; \mathcal{P}_6,\mathcal{P}_8,\mathcal{P}_9\colon \begin{pmatrix}4\\3\\2\end{pmatrix}; \mathcal{P}_7\colon \begin{pmatrix}5\\3\\1\end{pmatrix}; \mathcal{P}_{10}\colon \begin{pmatrix}3\\3\\3\end{pmatrix}.$$

Voting situations correspond to different partitions of the sum of scores.

Proposition 3. *For $m=3$ and $n\geq 2$, the number of voting situations induced by the Borda rule is equal to*

$$\#Borda(3,n)=\left\lceil \frac{1}{2}(n+1)^2 \right\rceil.$$

Kemeny rule (1)

For the 3 alternatives and 3 agents case, we have the following voting situations (ANECs and the corresponding sums of distances in the circle form, which is invariant up to rotating and turnover):

P_0: $-0-3-6-9-6-3-$;

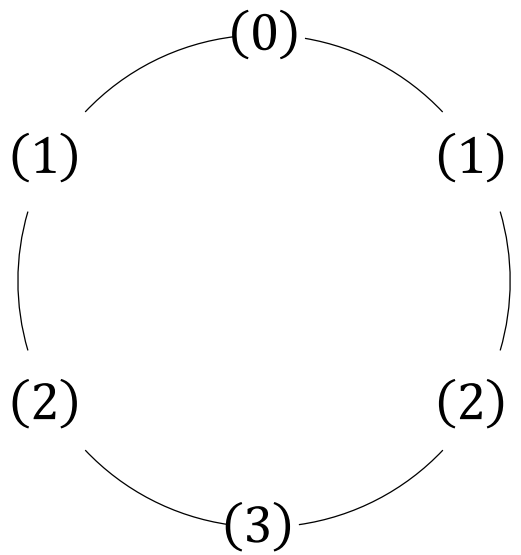
P_2, P_3: $-1-2-5-8-7-4-$;

P_4, P_5: $-2-3-4-7-6-5-$;

P_6, P_8, P_9: $-3-4-5-6-5-4-$;

P_7: $-2-3-6-8-6-3-$;

P_{10}: $-4-5-4-5-4-5-$.



Circle of distances from order P_0.

Kemeny rule (2)

Proposition 4. *For $m = 3$, the number of voting situations induced by the Kemeny rule is equal to*

$$\#Kemeny(3,n) = \begin{cases} \left\lceil \frac{1}{72}(4n^3 + 21n^2 + 54n + 56) \right\rceil, \text{if } n \text{ is even;} \\ \left\lceil \frac{1}{72}(4n^3 + 21n^2 + 36n + 11) \right\rceil, \text{if } n \text{ is odd.} \end{cases}$$

Discernibility potential of scoring rules (1)

Scoring rules utilize information only about the sum of scores for each alternative. In an extreme case of irrational scores (e.g., for $m = 3$ scores $0, 1, \sqrt{2}$), each combination of the scores (e.g., for $m = 3, n = 5$ scores $0, 0, 1, \sqrt{2}, \sqrt{2}$) leads to the unique sum of scores ($1 + 2\sqrt{2}$ in our example). In this case, we can unambiguously derive the combination of scores from the sum of scores vector. In other words, we can derive the scoring matrix. The element at the *i*-th row and *j*-th column of this matrix equals the number of times that the *i*-th candidate is ranked at the *j*-th position

$$a_{ij} = |\{k | pos(P_k, i) = j\}|.$$

Discernibility potential of scoring rules (2)

10 scoring matrices for the 3 alternatives and 3 agents case, which corresponds to the voting situations (ANECs and corresponding scoring matrices):

$$\mathcal{P}_1: \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \mathcal{P}_2: \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}; \mathcal{P}_3: \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \mathcal{P}_4: \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}; \mathcal{P}_5: \begin{pmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix};$$

$$\mathcal{P}_6: \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}; \mathcal{P}_7: \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}; \mathcal{P}_8: \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}; \mathcal{P}_9: \begin{pmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}; \mathcal{P}_{10}: \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Discernibility potential of scoring rules (3)

Proposition 5. (MacMahon 1918) *For $m = 3$, the number of voting situations induced by the scoring rules in the extreme case is equal to*

$$\#Scoring(3,n) = \begin{cases} \left\lceil \frac{1}{48}(n^4 + 6n^3 + 18n^2 + 36n + 32) \right\rceil, \text{if } n \text{ is even;} \\ \left\lceil \frac{1}{48}(n^4 + 6n^3 + 18n^2 + 18n + 5) \right\rceil, \text{if } n \text{ is odd.} \end{cases}$$

For $m = 3$, the simplest example of indiscernibility is the following:

$$\hat{\mathcal{P}} = \begin{pmatrix} a & a & b & c \\ b & b & c & a \\ c & c & a & b \end{pmatrix}; \tilde{\mathcal{P}} = \begin{pmatrix} a & a & b & c \\ b & c & a & b \\ c & b & c & a \end{pmatrix}$$

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Polino mial degree
#ANEC	5	10	24	42	83	132	222	335	511	728	1047	1428	1956	2586	5
#RIANEC	4	7	16	26	50	76	126	185	280	392	561	756	1032	1353	5
#SSANEC	3	4	8	10	17	20	30	35	49	56	75	84	108	120	3
share in #ANEC	0.60	0.40	0.33	0.24	0.20	0.15	0.14	0.10	0.10	0.08	0.07	0.06	0.06	0.05	
Plurality	2	3	4	5	7	8	10	12	14	16	19	21	24	27	2
I(Rule,3,n)	0.25	0.22	0.13	0.10	0.07	0.05	0.04	0.03	0.03	0.02	0.02	0.01	0.01	0.01	
Run-off	2	3	5	5	9	10	14	15	21	22	31	31	40	43	3
I(Rule,3,n)	0.25	0.22	0.17	0.10	0.10	0.07	0.06	0.04	0.04	0.03	0.03	0.02	0.02	0.02	
Kemeny	4	6	12	17	28	37	54	69	93	115	148	178	221	261	3
I(Rule,3,n)	0.75	0.56	0.48	0.39	0.33	0.27	0.24	0.20	0.18	0.16	0.14	0.12	0.11	0.10	
I^{IR}(Rule,3,n)	1	0.83	0.73	0.64	0.55	0.48	0.42	0.37	0.33	0.29	0.26	0.23	0.21	0.19	
Borda	5	8	13	18	25	32	41	50	61	72	85	98	113	128	2
I(Rule,3,n)	1	0.78	0.52	0.41	0.29	0.24	0.18	0.15	0.12	0.10	0.08	0.07	0.06	0.05	
Scoring	5	10	23	40	73	114	180	262	379	521	712	938	1228	1567	4
I(Rule,3,n)	1	1	0.96	0.95	0.88	0.86	0.81	0.78	0.74	0.72	0.68	0.66	0.63	0.61	

Conclusion

- Only the run-off rule has a non-monotonic function of the number of voting situations.
- The plurality rule has the lowest number of voting situations.
- The scoring rules in the extreme case have the highest number of voting situations.
- For the other rules, the highest information utilization index occurs with the Kemeny rule, which is more complicated in comparison with the other rules.

Thank you for attention!