Proceedings of the World Congress on Engineering and Computer Science 2012 Vol I
WCECS 2012, October 24-26, 2012, San Francisco, USA

Taxonomy and Nomenclature of Preferential Voting Methods

Sung-Hyuk Cha and Yoo Jung An

***Abstract*— Voting plays an important role in many decision making problems. Various conventional preferential voting methods where the voters rank candidates in order of preference are reviewed for syntactic patterns and categorized. Several other new voting methods are devised from the conventional procedural patterns and metrics as well. Explicit formulae for over fifty different voting methods are presented and the hierarchical clustering technique is adopted to reveal semantic similarities among them. A nomenclature for voting methods is suggested to reveal their syntactic patterns. All preferential voting methods perform significantly different from the simplest plurality method.**

***Index Terms*—decision, hierarchical clustering, preference, voting, nomenclature**

I. Introduction

Consensus of a group plays an important role in decision making such as elections [1-3] and combining multiple classifiers [4]. It is essential in most democratic societies and has received great attention in artificial intelligence and computer science communities as well [5].

Consider an ordered set of four candidates, C = {'A', 'B', 'C', 'D'} and they received the corresponding votes, V = {11, 6, 7, 6}. The notations in Table I shall be used throughout the rest of this article. The most widely used and simplest voting method is called the '*plurality*', i.e. the winner is one who has the most votes as defined in (1). The majority voting method in (2) is the same as the plurality but rejects if the winner does not receive more than half votes.

$$plurality(V) = \arg\max_{x \in C}(V_x) \quad (1)$$

$$majority(V) = \begin{cases} \arg\max_{x \in C}(V_x) & \text{if } \max(V) > \frac{m}{2} \\ void & \text{otherwise} \end{cases} \quad (2)$$

TABLE I
Basic Notations

notation	meaning and/or example
C	an ordered set, e.g., {'A', 'B', 'C', 'D'} and C_2 = 'B'.
V	the corresponding votes, e.g., {11, 6, 7, 6} and V_3 = 7.
c	the number of candidates, $c = \|C\| = 4$ in the example.
m	the total number of voters, e.g., 30 in the example.
n	the number of unique preference order ballots. 5 in Table II.
$p(i, j)$	the candidate in the ith ballot and jth choice, $p(2, 1)$ = 'A'
$r(i, x)$	the choice rank for the candidate x in the ith ballot, $r(4,\text{'A'}) = 3$

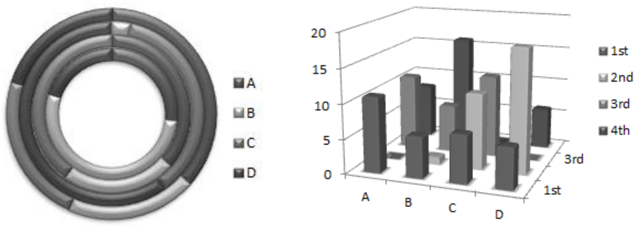
Manuscript received July 23, 2012; revised August 10, 2012.

S.-H. Cha is with Computer Science department, Pace University, New York, NY 10038 USA (e-mail: scha@pace.edu).

Y.J. An is with the Division of Engineering Technologies and Computer Sciences, Essex County College, Newark, NJ 07102 USA, (e-mail: yan@essex.edu). Starting September 2012.

TABLE II
Sample Preference Ballot Table

Choice \ votes	7	11	1	6	5
1	C	A	D	B	D
2	D	D	B	C	C
3	B	C	C	A	A
4	A	B	A	D	B



(a) stacked pies (b) multiple bars
Fig. 1. Distribution of preference votes.

Due to flaws in the simple top choice voting system, the *preferential voting system* in which the voter ranks candidates in order of preference has been proposed [1-3,5-7].

Voters are asked to rank the candidates where omissions and ties are not allowed and quantities are not important but only the strict order matters as exemplified in Table II. Fig 1 shows the distribution of preference voting in stacked pies and multiple bars. The outer most shell is the first choice and the inner shells are the next choices and so on.

The winner of Table I case differs depending on voting methods used and there is a zoo of diverse methods. Various voting methods are used in diverse social groups and different regions. There are so many variation and alternatives and thus a comprehensive study is necessary because even names for certain voting methods are fluid and promulgated differently. Also, a nomenclature for voting methods is necessary as even today another new voting method is invented.

The rest of the paper is organized as follows. In section 2, various conventional preference voting methods are given to reveal their syntactic similarities. In section 3, conventional methods are generalized and other new methods are devised from the existing methods' patterns. In order to provide a better perspective on similarity among different methods, section 3 presents the hierarchical cluster tree of over fifty different preference voting methods. Finally, section 4 concludes this work.

II. Conventional Voting Methods

In this section, different syntactic patterns of conventional voting methods are examined and expressed in as generic forms as possible. Perhaps, the most common paradigm for many methods is finding the *argmax* of certain measurable score values for each candidate as given in (3).

$$sb(p) = \begin{cases} \arg\max_{x \in C} \left(f_s(p,x) \right) & \text{if unique} \\ \text{void} & \text{otherwise} \end{cases} \quad (3)$$

The plurality method can be referred to as (3-4), i.e., it follows the general form in (3) with (4) as its score function: $f_s(p, x) = pl(p, x)$.

$$pl(p,x) = \sum_{p(i,1)=x} v_i = \sum_{i=1}^{n} \begin{cases} v_i & \text{if } p(i,1) = x \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Note that the summation notation with the subscript only shall be used frequently in this article and is defined as the summation of only $f(i,x)$ from $i = 1$ to n such that the subscript condition is met as exemplified in (4).

The *rank* method is very similar to the paradigm in (3) but uses the score function with certain weights.

$$f_w(p,x,w) = \sum_{i=1}^{n} w_{r(i,x)} v(i) \quad (5)$$

If the weights are (1,0,0,0), it is the plurality method. If the weight is (c–1,…1,0) as in (6), e.g., (3,2,1,0) in our example, then it is called *borda* method [1-4] attributing Jean-Charles de Borda [6]. (5) with (6) is called the *borda* score.

$$w_{r(i,x)} = c - r(i,x) \quad (6)$$

Similar to the paradigm in (3), finding the *argmin* of certain measurable *penalty* values for each candidate as given in (7).

$$nb(p) = \begin{cases} \arg\min_{x \in C} \left(f_n(p,x) \right) & \text{if unique} \\ \text{void} & \text{otherwise} \end{cases} \quad (7)$$

As opposed to the simplest score function in the plurality method which considers only the pluralities of the top choice, a simplest penalty function would be the plurality of the last choice as in (8).

$$pf(p,x) = \sum_{p(i,c)=x} v_i \quad (8)$$

The candidate who is disliked by most voters would have the highest penalty value. Hence, we rank the candidates in order of ascending penalty values.

Suppose we would like to compare two candidates, 'B' and 'D'. Either score or penalty function can be used in (9) to find the winner between them, the more popular and specific pairwise comparison function is given in (10).

$$pairwin(p,x,y) = \begin{cases} x & \text{if } f_p(x) > f_p(y) \\ y & \text{if } f_p(x) < f_p(y) \\ \text{void} & \text{if } f_p(x) = f_p(y) \end{cases} \quad (9)$$

$$f_p(p,x,y) = \sum_{r(i,x)<r(i,y)} v(i) \quad (10)$$

For example of Table II, pc('B') = 6 which is the fourth ballot and pc('D') = 7 + 11 + 1 + 5 = 24 where 'D' precedes 'B'. Hence $pairwin(p, \text{'B'}, \text{'D'})$ is 'D'. Let's denote this pairwise $c \times c$ victory score matrix based on (10) M_v as given in Table III.

If there exists a candidate which wins all other candidates in pairwise comparison as defined in (11), this winner is called the *Condorcet* winner [1-3] attributed to Marquis de Condorcet [7]. This concept dates back at least to Ramon Llull in the thirteenth century though [3].

TABLE III
pairwise victory score matrix, M_v.

$x \backslash y$	A	B	C	D	*Wt*
A	0	16	11	17	44
B	14	0	7	6	27
C	19	23	0	13	55
D	13	24	17	0	54
Lt	46	63	35	36	

$$condorcet(p) = \begin{cases} x & \text{if } \exists x \forall y (pairwin(p,x,y) = x) \\ & \quad \text{where } x, y \in C \ \& \ x \neq y \\ \text{void} & \text{otherwise} \end{cases} \quad (11)$$

In the example of Table I and most of cases, there is no Condorcet winner as shown in Table III. The Condorcet concept is often used as a property of other methods, i.e., whether or not a certain method x always selects the Condorcet winner if there exists one.

The borda method does not have the Condorcet property. Duncan Black suggested to select the Condorcet winner if one exists and use the borda method if not [8] and this method is referred to as the *black* method in [1]. Hence, the Condorcet concept can be used generically as an ensemble with other methods which do not have the Condorcet property as in (12).

$$cx(p) = \begin{cases} w = condorcet(p) & \text{if } w \neq void \\ methodx(p) & \text{otherwise} \end{cases} \quad (12)$$

The minimum number of pairwise swaps before they become a Condorcet winner is called the *dodgson* method which is an NP-hard problem [9]. In [1], however, a simplified version is attributed to the dodgson method which uses (8) in (12). We shall refer it as the *dodgson-s* method.

Another popular and widely used voting concept uses the two round system. First, it selects the top two candidates by a certain way and then uses the pairwise comparison between those as given in (13).

$$sp(p) = pairwin(p, x_1, x_2) \quad \text{where } (x_1, x_2) = top2(p) \quad (13)$$

Albeit any ranking function such as the borda score can be used to select the top two, the simple plurality (4) is used in (13) and this particular method is known as the *run-off* method.

Similar to the usage of the Condorcet method in (12), the majority method in (2) can be used as an ensemble with other methods as well.

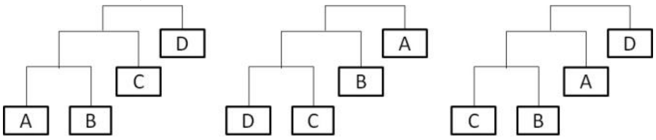
$$my(p) = \begin{cases} w = majority(p) & \text{if } w > \frac{m}{2} \\ methody(p) & \text{otherwise} \end{cases} \quad (14)$$

When (13) is used in (14), this ensemble method is called the *contingent* method or runoff method interchangeably in [1,2]. Note that the runoff in (13) is the same as (14) with (13) when the plurality (4) and pairwin (9-10) are used to find and compare the top two candidates, respectively. We make distinction here because the results may differ when other score functions and/or pairwise comparison methods are used in the later section 3.

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Suppose pairwise matches are scheduled with a fixed sequential agenda as depicted in Fig 2.

$$q(p,a)=\begin{cases} pairwin(p,a_1,a_2) & \text{if } leng(a)=2 \\ q(p,tail(a)) & \text{otherwise} \end{cases} \quad (15)$$



(a) <A,B,C,D> (b) <D,C,B,A> (b) <C,B,A,D>

Fig. 2. Sequential pairwise match agendas.

This method is called the sequential pairwise match method [3]. The winner depends on the order in the agenda, a. In Fig 2 (a), the winner of 'A' vs. 'B' match will be against 'C'. And then the winner of that round will be in the final round with 'D'. This method is defined recursively in (15).

Suppose that the candidate 'B' withdraw from the election. Then the preference ballot table in Table II is updated to the one in Table IV due to the elimination process of $p = p - \{\text{'B'}\}$. Numerous methods utilize this process.

TABLE IV
$p - \{\text{'B'}\}$ preference ballot table

I	1	2	3	4	5
Choice\votes	7	11	1	6	5
1	C	A	D	C	D
2	D	D	C	A	C
3	A	C	A	D	A

In 1861, Thomas Hare proposed the *instant runoff* voting method or simply *IRV* which eliminates the candidate with the lowest plurality recursively until there exists a majority winner [3]. The *IRV* is also called *hare* [1], *Cincinnati rule* [2], or *single transferable* [2,3] method.

The eqn (16) is a recursive generic form of the IRV where many possible elimination functions like (17) can be used. The elimination function outputs a set of candidates to be removed. Even if a specific method like the alternative or IRV method requires eliminating a single candidate per step, multiple candidates could be removed as a bulk if there are ties.

$$rmb(p)=\begin{cases} majority(p) & \text{if exists} \\ rmb(p-E) & \text{if } c>1 \,\&\, E\neq\{\} \\ void & \text{if } c=0 \text{ or } E=\{\} \end{cases} \quad (16)$$

$$E = \underset{x\in C}{\arg\min}\big(pl(p,x)\big) \quad (17)$$

Instead of eliminating the candidate(s) with the fewest first place votes, Clyde Coombs proposed to eliminate those with the most last place votes (18) [10].

$$E = \underset{x\in C}{\arg\max}\big(pf(p,x)\big) = \underset{x\in C}{\arg\max}\left(\sum_{p(i,c)=x} v_i\right) \quad (18)$$

In [1-3], the generic form (19) with (18) as the elimination function is referred to as the *Coombs* method. While the Hare method may terminate the elimination when there exists a candidate with majority, the Coombs method keeps eliminating candidates until only one remains.

$$rb(p)=\begin{cases} C_1 & \text{if } c=1 \\ rb(p-E) & \text{if } c>1 \,\&\, E\neq\{\} \\ void & \text{if } c=0 \text{ or } E=\{\} \end{cases} \quad (19)$$

In 1882, Edward J. Nanson proposed a hybrid method of the generic form (19) with the borda score where all candidates whose borda scores are below the average are eliminated per recursive step as in (20) [11].

$$E = \left\{ x \,\middle|\, x\in C \,\&\, bs(x) < \sum_{i=1}^{c} bs(c_i) \Big/ c \right\} \quad (20)$$

Joseph M. Baldwin referred (19-20) as the *Nanson* method and proposed to eliminate only candidate(s) with the lowest borda score as in (21) [12].

$$E = \{x \,|\, x\in C \,\&\, x = \arg\min(bs(*))\} \quad (21)$$

The following conventional methods require a $c \times c$ matrix produced by a pairwise comparison function, e.g., M_v in Table III produced by (10). The *minimax* method, also known as Simpson-Kramer or successive reversal method uses the generic form in (7) with a certain penalty function involving the M_v matrix [2]. The most popular penalty functions include the pairwise opposition (22), winning votes (23), and margins (24).

$$f_{nx\_op}(p,x) = \max(M_v(*,y)) \quad (22)$$

$$f_{nx\_wv}(p,x) = \max(Wv)$$

$$Wv(y)=\begin{cases} M_v(y,z) & \text{if } M_v(z,y) < M_v(y,z) \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

$$f_{nx\_mg}(p,x) = \max(Mg)$$

$$\text{where } Mg(y)=\begin{cases} M_v(y,z) - M_v(z,y) & \text{if } z\neq y \\ -\infty & \text{if } z=y \end{cases} \quad (24)$$

A H. Copeland suggested a pairwise aggregation method [13] which is simply called the Copeland method in [3]. It involves the pairwise winner matrix produced by (25) in Table V.

$$M_b(x,y)=\begin{cases} 1 & \text{if } x = pairwin(p,x,y) \\ 0 & \text{if } y = pairwin(p,x,y) \\ \tfrac{1}{2} & \text{if void} = pairwin(p,x,y) \end{cases} \quad (25)$$

TABLE V
pairwise winner matrix.

x\y	A	B	C	D	tot_W
A	0	1	0	1	2
B	0	0	0	0	0
C	1	1	0	0	2
D	0	1	1	0	2
tot_L	1	3	1	1	

Copeland method follows the standard form (3) with the score function given in (28), i.e., the number of pairwise victories (26) minus the number of pairwise defeats (27).

$$S_W(x) = \sum_{y=C_1}^{C_c} M(x,y) \quad (26)$$

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$S_L(y) = \sum_{x=C_1}^{C_c} M(x, y) \quad (27)$$

$$f_{sx}(x) = S_W(x) - S_L(x) \quad (28)$$

III. Composition of Voting Methods

This section attempts to generalize the patterns of conventional voting methods and provides a scheme to generate new composed voting methods. The conventional and composed voting methods are named as strings using the symbols in Table VI and enumerated in Table VII.

First, the conventional methods can be categorized into whether it requires any pairwise comparison and two or more rounds as shown in Fig 3. If it does not, the single symbol '*b*' is used and if so, the pair symbol '*p*' or '*q*' shall be used. The generic form in (3) can be named *sb*(f_s) which uses a certain score function to select the single winner. Instead of the linear weight in the borda score (6), the quadratic triangular number weights in (29), e.g., (6, 3, 1, 0), can be used in which the closer to the top choice, the higher weights it gets.

$$w_{r(i,x)} = T_{r(i,x)} = \frac{(c - r(i,x))(c - r(i,x)) + 1)}{2} \quad (29)$$

The concepts of Condorcet and Majority are often used as mixture of other methods such as black and dodgson-s. The symbol '*c*' and '*m*' shall be used as a prefix for (14) and (16), respectively. The black method can be stated as *csb*(6). The prefix, '*c*' can be redundant for those methods which have the Condorcet property, e.g., sequential pairwise method, Copeland, etc. The prefix, '*m*' can be also redundant, e.g., plurality = *sb*(4) = *msb*(4).

The generic form (7) which requires a certain penalty function, f_n is represented as *nb*(f_n). The dodgson-s method is *cnb*(8). If the reverse borda penalty, (4, 3, 2, 1) gives the penalty function in (30) and applied to (7), i.e., *nb*(30).

$$pf_b(p, x) = \sum_{i=1}^{n} r(i,x) v(i) \quad (30)$$

The symbol '*q*' is for the sequential pairwise voting methods where the order agenda and a pairwise comparison function are required. For example, *q*(-,10) has the agenda <'A', 'B', 'C', 'D'> and uses (10) as the pairwise comparison. The ascending or descending order from any score or penalty function can serve as the agenda. The first argument indicates the score or penalty function, e.g., *sq*(4,10) and *nq*(8,10).

The second argument is the pairwin function like (10)

TABLE VI
Symbols in Nomenclature in preferential voting methods.

	meaning	arguments / usages.
b	single	argmax/ argmin/ leave one, etc.
s	score function	f_s (4), (5-6), (5-27)
n	penalty function	f_n (8), (30)
c	Condorcet	prefix
m	majority	postfix for recursion or prefix for others
p	pairwise comparison	f_p (9-10), (31), (32), (33)
q	sequential pairwise	f_p
r	recursive elimination	f_e (17), (18), (20), (21)
x	matrix	f_p, f_{sx} (26), (28), (35)/ f_{nx} (22), (23), (24), (27)

TABLE VII
Conventional and composed preferential voting methods.

generic	conventional	composed
sb(f_s) (3)	plurality = *sb*(4), borda = *sb*(6)	rank_triang = *sb*(29)
msb(f_s) (3,14)		*msb*(6), *msb*(29)
csb(f_s) (3,12)	black = *csb*(6)	*csb*(3), *csb*(29)
nb(f_n) (7)		*nb*(8), *nb*(30)
mnb(f_n) (7,14)		*mnb*(8), *mnb*(30)
cnb(f_n) (7,12)	dodgson-s = *cnb*(8)	*cnb*(30)
q(-,f_p) (15)	seq_pair = *q*(-,10)	*q*(-,31), *q*(-,32), *q*(-,33)
sq(f_s,f_p) (15)		*sq*(4,10), *sq*(4,31), *sq*(29,32)
nq(f_n,f_p) (15)		*nq*(8,10),*nq*(8,31),*nq*(30,32)
sp(f_s,f_p) (13)	runoff = *sp*(4,10)	*sp*(4,31), *sp*(4,32), *sp*(6,32)
msp(f_s,f_p) (13)	conting- = *msp*(4,10)	*msp*(4,31), *msp*(6,33)
np(f_n,f_p) (34)		*np*(8,10), *np*(30,32)
mnp(f_n,f_p) (34)		*mnp*(8,31), *mnp*(30,33)
rb(f_e) (19)	coombs = *rb*(18), nanson = *rb*(20), baldwin = *rb*(21)	*rb*(17), *rb*(35)
rmb(f_e) (16)	hare = *rmb*(17)	*rmb*(18), *rmb*(21), *rmb*(35)
rp(f_e,f_p) (36)		*rp*(17,10), *rp*(21,33)
rmp(f_e,f_p) (37)		*rmp*(17,10), *rmp*(18,31)
sbx(f_p,f_{sx}) (37)	copeland = *sbx*(28,10b)	*sbx*(26,10b), *sbx*(28,31b), *sbx*(28,31), *sbx*(26,33)
nbx(x,y)	minimax=*nbx*(22,10), *nbx*(23,10),*nbx*(24,10	*nbx*(27,10), *nbx*(27,10b), *nbx*(23,31), *nbx*(27,33b)
rbx(f_e)		*rbx*(39,10), *rbx*(40,10)
rmbx(f_e)		*rmbx*(39,10), *rmbx*(40,10)

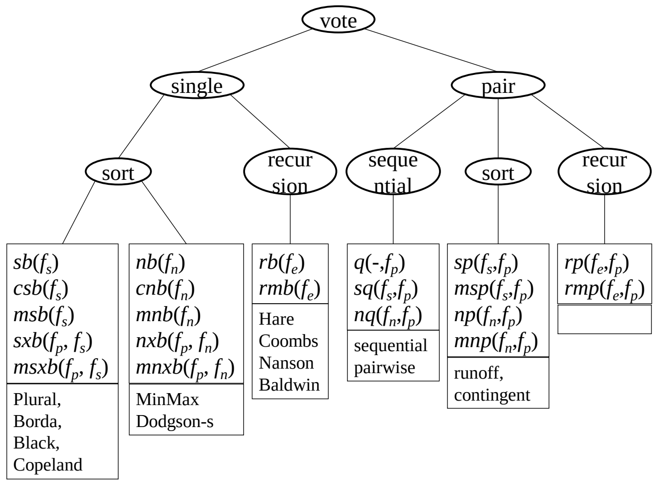


Fig. 3. Categorization tree of preference voting methods

which is winning pluralities as scores. One possible pairwise comparison function f_p takes the winning position into account as scores as in (31) just like the borda concept.

$$f_p(p, x, y) = \sum_{r(i,x)<r(i,y)} (c - r(i,x) + 1) v(i) \quad (31)$$

Another possible pairwise comparison function takes the difference between winning and losing positions into account as scores as in (32).

$$f_p(p, x, y) = \sum_{r(i,x)<r(i,y)} \left| r(i,y) - r(i,x) \right| v(i) \quad (32)$$

Another possible pairwise comparison function takes the difference between winning and losing triangular number positions in (29) into account as scores as in (33).

$$f_p(p, x, y) = \sum_{r(i,x)<r(i,y)} \left| T_{r(i,y)} - T_{r(i,x)} \right| v(i) \quad (33)$$

These new pairwin functions in (31-33) can be applied to not only sequential pairwise voting, but also contingent and IRV instead of the conventional function (10).

The conventional two round system in (13) can be

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

expressed *sp*(4,10) or *msp*(4,10) for the contingent method. The two arguments are the score function and the pairwise comparison function and thus *sp*(4,31), *sp*(6,10), *msp*(4,32), *msp*(6,33) can be composed as new methods.

Similarly penalty functions such as (8) or (30) can be used to find the lowest two candidates and then any pairwise comparison functions can be applied as in (34).

$$np(p) = pairwin(p, x_1, x_2) \quad \text{where } (x_1, x_2) = bottom2(p) \tag{34}$$

Next, the symbol '*r*' stands for the recursive elimination where the IRV family methods can be represented. $rb(f_e)$ means keeping eliminating recursively until one single winner remains as in (19), e.g., Coombs = *rb*(18), Nanson = *rb*(20), and Baldwin = *rb*(21). Instead of the argmin and average of borda scores in Baldwin and Nanson methods, one can use the argmin (17) and average of the plurality (35) as the elimination function.

$$E = \left\{x \middle| x \in C \;\&\; pl(1,x) < \sum pl(1,*)/c\right\} \tag{35}$$

Any score or penalty functions can be used to find argmin or argmax and average to compose another new elimination function.

The '*m*' symbol is used after '*r*' for the hare method (16). Note that $rmb(f_e) \neq mrb(f_e)$. In $mrb(f_e)$, the majority is checked only once at the beginning as in (14) and then (19) is executed.

The concept of the runoff (13) can be used with the recursive elimination. Candidates can be eliminated until two candidates remain as in (36) or (37) instead of a single candidate (19). The symbol '*p*' is used instead of '*b*'.

$$rp(p) = \begin{cases} C_1 & \text{if } c = 1 \\ pairwin(C_1, C_2) & \text{if } c = 2 \\ rp(p - E) & \text{if } c > 1 \;\&\; E \neq \{\} \\ void & \text{if } c = 0 \text{ or } E = \{\} \end{cases} \tag{36}$$

$$rmp(p) = \begin{cases} w = majority(p) & \text{if exists} \\ pairwin(C_1, C_2) & \text{if } c = 2 \\ rmp(p - E) & \text{if } c > 2 \;\&\; E \neq \{\} \\ void & \text{if } c = 0 \text{ or } E = \{\} \end{cases} \tag{37}$$

Finally, the symbol '*x*' stands for the $c \times c$ matrix such as the pairwise victory score matrix and its binary matrix in Tables III and V. The *minimax* and *copeland* methods can be stated as *nbx*(22, 10) and *sbx*(28, 10b), respectively. The first argument is either score or penalty functions and the second argument is the pairwise comparison function which makes the matrix. The matrix (10b) is the binary version (25) of the pairwise comparison function (10).

The score function with a matrix, f_{sx} includes (26) and (28) and the penalty function with a matrix, f_{nx} includes (22), (23), (24) and (27). As opposed to the minimax, one can compose the *maxmin* method with the following score function (38).

$$f_{sx}(p, x) = \min(M(x, \{C - x\})) \tag{38}$$

Several webpages such as [14] describe some voting methods without any reference and where definitions vary from other sources. While some are the same as the original source, others are different but promulgated the same or produce the same results but more complex algorithm involving matrices.

In [14]. The Borta method is defined as the score function in (28) with the matrix (10) and let's denote it as bs_2. For the Baldwin method, bs_2 is used instead of *bs* in *rb*(21) to make *rbx*(39, 10). The nanson method is defined differently where (40) is used instead of (20).

$$E = \left\{x \middle| x \in C \;\&\; x = \arg\min(bs_2(*))\right\} \tag{39}$$

$$E = \left\{x \middle| x \in C \;\&\; bs_2(x) < 0\right\} \tag{40}$$

IV. Hierarchical Clustering of Voting Methods

Hitherward, the focus is moved from the syntactic patterns to the semantic similarity between voting methods. So as to assess how similar voting methods are, the following experiments were conducted using the cluster analysis. For m = 100 voters and c = 4 candidates, (nt = 100) number of preference ballot test cases are randomly generated. The distance between two voting methods is the number of the mismatches in (41).

$$d(vm_x, vm_y) = \sum_{i=1}^{nt} \begin{cases} 0 & \text{if } vm_x(i) = vm_y(i) \\ 1 & \text{if } vm_x(i) \neq vm_y(i) \end{cases} \tag{41}$$

If two methods are identical, the distance is zero and if they do not agree on many cases, the distance is high. The hierarchical cluster tree in Fig 4 reveals the similarities among various voting methods.

Prior to rigorous mathematical proof, the dendrogram provides the intuitive identity, similarity, correctness, etc. The simple plurality method and the dodgson-s method are quite different from most preference voting methods.

In this experiment, the *borda sb*(6), *nbx*(27,10), *q*(-,32), *sq*(4,32), *sbx*(28,32b), and *sbx*(28,10) turned out to be identical. The *Baldwin rb(*21) and *rbx*(39,10) are also identical to each other. Another identical group has *sb*(29), *q*(-,33), *sq*(4,33), *nq*(8,33), *sbx*(28,33b). Also, *rmb*(21) is the same as *rmbx*(39,10).

V. Conclusion

While most voting method survey works in literature focus on mathematical properties and flaws of voting methods, this article attempts to reveal their syntactic and semantic relationships. Albeit it needs community based consensus and further refinement, an initial naming convention is suggested.

This article reviewed several popular conventional methods but there are numerous other voting methods are in use. Further comprehensive survey is necessary as a future work.

References

[1] M. Samuel III, *Making Multicandidate Elections More Democratic.* Princeton University Press, Princeton, NJ, 1988.

[2] J. Levin and B. Nalebuff, "An Introduction to Vote-Counting Schemes," Journal of Economic Perspectives, vol. 9, no. I, Winter. 1995, pp. 3–26

[3] A. D. Taylor and A. M. Pacelli, *Mathematics and Politics: Strategy, Voting, Power and Proof*, Springer-Verlag, 1995.

[4] T.K. Ho, J.J. Hull, and S.N. Srihari, "Decision Combination in Multiple Classifier Systems," IEEE Trans. J. PAMI, vol. 16, no. I, Jan. 1984, pp. 66–75

[5] P. Faliszewski, E. Hemaspaandra, L. A. Hemaspaandra, and J. Rothe, "Llull and Copeland Voting Computationally Resist Bribery and Constructive Control," Journal of Artificial Intelligence Research. vol. 35, 2009, pp. 275–341

[6] J.-C. de Borda, *Mémoire sur les Élections au Scrutin.* Paris: histoire del l'Académie Royale des Sciences, 1781.

[7] M. de Condorcet, *Essay on the Application of Mathematics to the Theory of Decision-Making.* Paris, 1785.

[8] D. Black, *The Theory of Committees and Elections.* Cambridge University Press, London, 1953.

[9] J. Bartholdi III, C. A. Tovey, and M. A. Trick, "Voting schemes for which it can be difficult to tell who won the election," Social Choice and Welfare, Vol. 6, No. 2, 1989, pp. 157–165.

[10] C. Coombs, *A Theory of Data.* Wiley, New York, 1964.

[11] E. J. Nanson, "Methods of election," Transactions and Proceedings of the Royal Society of Victoria 19, 1882, pp. 197–240.

[12] J. M. Baldwin, "The technique of the Nanson preferential majority system of election," in *Proc. 4th the Royal Society of Victoria*, n.s. 39: 1926, pp. 42–52.

[13] A. H. Copeland, "A 'reasonable' social welfare function," presented at the Seminar on Mathematics in Social Sciences, University of Michigan, 1951.

[14] R. LeGrand, "Descriptions of ranked-ballot voting methods," http://www.cs.wustl.edu/~legrand/rbvote/desc.html as of June 2012.

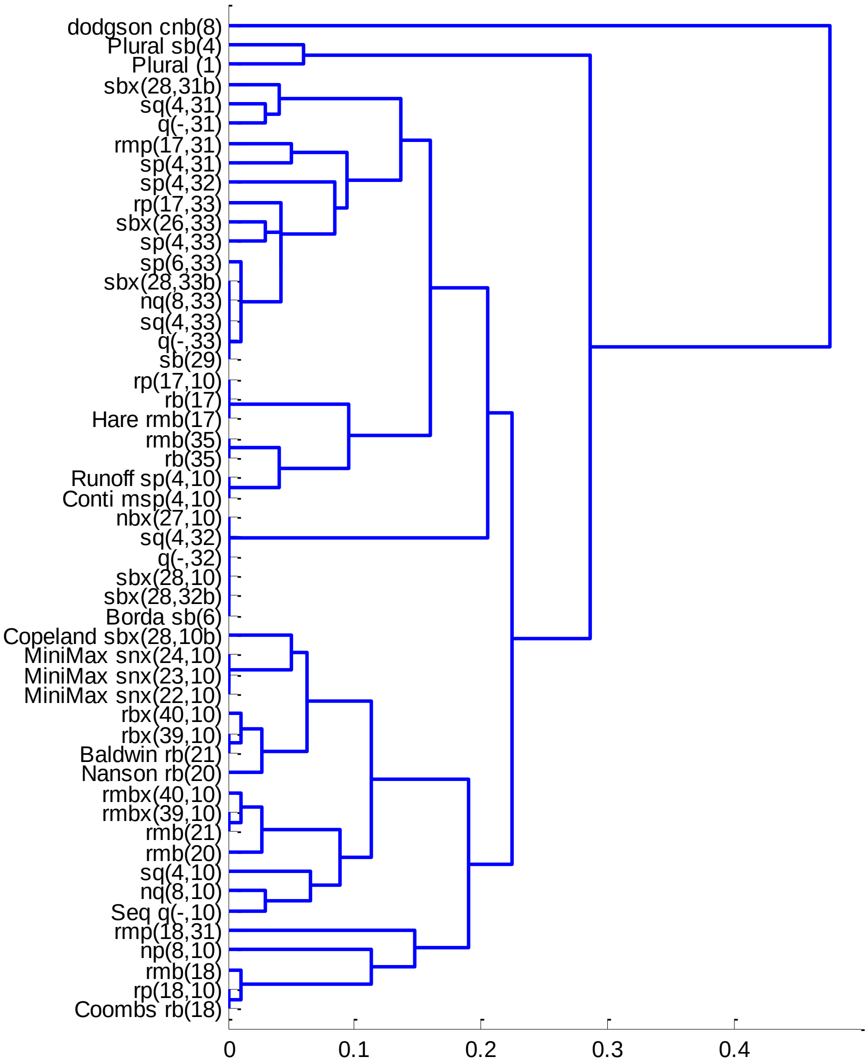


Fig. 4. Hierarchical Clustering of 51 conventional and devised voting methods.

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)