Appendix: Indexing the binary voting trees under amendment agendas

29th November 2013

This is an appendix to Lehtinen (forthcoming): "Strategic voting and the degree of path-dependence", *Group Decision and Negotiation*

Ordering numbers

An *ordering number* is used to identify the nodes and the corresponding probabilities in the voting trees.[1] Consider all the pairs of alternatives for four alternatives as shown in Table 1:

ordering number	first alternative	second alternative
1	1	2
2	1	3
3	1	4
4	2	3
5	2	4
6	3	4

Table 1: ordering numbers with four alternatives

The first pair of alternatives is always $\{1, 2\}$. Adding a third alternative yields the pairs $\{1, 3\}$ and $\{2, 3\}$ because the n:th alternative is put against

[1]We do not use the so called *indexed traversal order* here. This is the ordering imposed on the nodes of a game tree by a lexicographic ordering of the nodes when each node is identified by the sequence of branch numbers necessary to reach it. If the number of alternatives is large, there are a large number of branches in which the alternatives n and n-1, or n and n-2 etc. are put against each other. There are thus usually several different paths to a given pair of alternatives. The indexed traversal order is not used because expected utilities would have to be calculated several times for identical sub-branches in a voting tree. The benefit in CPU-time and memory from using an ordering number instead of the indexed traversal order is considerable in computer simulations. Using an ordering number implies the assumption that irrespective of the path with which a branch is obtained, any branch with identical alternatives is to be evaluated in the same way. It implies the assumption that there is no learning because if the agents could learn in a significant manner from previous voting rounds, it would not be justifiable to assume that the agents always evaluate the expected utility of a branch only on the basis of the alternatives in that branch.

each alternative in the last round. By the same principle, adding a fourth alternative yields$\{1,4\}$,$\{3,4\}$, $\{2,4\}$,$\{3,4\}$, and so on. These sequences of pairs of alternatives provide us with a corresponding sequence of ordering numbers.

The ordering number determines the probability p_{jk} used in a given branch, and the corresponding label for the expected utility expression EU_{jk}.[2] A node always corresponds to a comparison of a pair of alternatives. With n alternatives there are $\binom{n}{2}$ different pairwise comparisons. Given any two alternatives j and k such that $k > j$, and a number of alternatives n, an ordering number $o(n,j,k)$ is given by the following formula:

$$o(n,j,k) = (n - \frac{j}{2})(j-1) + k - j \tag{1}$$

The formula can be derived as follows. The pairwise comparisons are ordered such that all pairs in which alternative 1 is involved obtain the first ordering numbers in an ascending order, then those in which alternative 2 is involved but not alternative 1, then those in which 3 is involved but not alternatives 1 and 2, and so on. Alternative 1 is involved in $n-1$ pairs in which it is the first alternative, alternative 2 in $n-2$ such pairs and so on. There are 0 pairs in which alternative 1 is involved before alternative 1, $(n-1)$ pairs before alternative 2 is the first alternative, $(n-1)+(n-2) = 2n-3$ pairs before alternative 3 is the first alternative, $(n-1)+(n-2)+(n-3) = 3n-6$ comparisons before alternative 4 is the first and so on. The number of pairs before alternative j can thus be written as a sum of $(j-1)n$, and the sum of an arithmetic series $0, -1, -2, ..., -(j-1)$ $= -\sum_{i=0}^{j-1} i$. The sum of this series is $\frac{(j)[0-(j-1)]}{2} = \frac{-j(j-1)}{2}$ so that the number of pairs before j is $(j-1)n - \frac{j(j-1)}{2} = (n - \frac{j}{2})(j-1)$. $k-j$ expresses the number of pairs between alternative j and alternatives from $j+1$ to k. For example, with n=4, the comparison $\{1,4\}$ corresponds to the ordering number $o(4,1,4) = (1-1)*4 - \sum_{i=0}^{1-1} i + 4 - 1 = 3$, and $\{3,4\}$ corresponds to $o(4,3,4) = 6$, and $\{2,4\}$ corresponds to $o(4,2,4) = 5$. The sequence $\{1,4\}$, $\{3,4\}$, $\{2,4\}$, $\{3,4\}$, which gives the pairwise comparisons in the last round of voting, thus corresponds to (3,6,5,6). Such sequences and pairs of alternatives allow writing the expected utility expressions for any number of alternatives.

As in Lehtinen (2007b), voters are assumed to obtain perturbed signals $S_i(j,k)$ concerning the number of voters who prefer one alternative to another in a pairwise comparison. The signals and the probabilities p_{jk} in a vote between any two branches, as well as the corresponding expected utilities $EU_i(j,k)$, are labeled with the ordering numbers. p_1 thus corresponds to p_{12}, p_2 corresponds to p_{13},..., and $p_{[\binom{n}{2}-1]\binom{n}{2}}$ corresponds to $p_{\binom{n}{2}}$. Signals and expected utilities are similarly labeled. Each player i thus obtains a set of perturbed signals $\{s_1, s_2, ..., s_{\binom{n}{2}}\}$.[3]

[2]Each sub-branch in a voting tree does not have a unique ordering number, because there are several identical subtrees in any binary agenda with at least four alternatives. The number of repetitions increases with the number of alternatives so that with six alternatives, for example, the ordering numbers for the last round are 5, 15, 14, 15, 12, 15, 14, 15, 9, 15, 14, 15, 12, 15, 14, 15, and 4, 13, 11, 13, 8, 13, 11, 13, for the penultimate round etc.

[3]Another possibility would be to assume that each player obtains one observation for each

Virtual voter types

The indexing system is designed for a single agenda but studying path-dependence requires being able to formulate voters' behaviour under any possible agenda. Nevertheless, the indexing method based on ordering numbers can be used to study all agendas. The reason for this is based on the following observation. Suppose that we are interested in studying the behaviour of a given voter type t under some agenda A_c. One may think of this as a 'current' agenda. Let A_f denote the fixed agenda employed by the indexing method. Depending on the number of alternatives, it is (123), (1234), (12345), etc.[4] For each voter type t, there is a *virtual type* t' whose behaviour under agenda A_f is *equivalent* to the behaviour of type t under agenda A_c. In order to study voting under agenda (2341), for example, voters' utilities are converted such that they correspond to those they would have if alternative 1 were alternative 2, 2 were 3, 3 were 4, and 4 were 1. Consider, for example, how a voter of type 14 ($4 \succ^i 1 \succ^i 3 \succ^i 2$) would vote under this agenda. Figure 2 shows agendas (2314) and (1234). The numbers in parenthesis in agenda (2314) indicate the position of the labeled alternative in the rankings of type 14 voters, and the corresponding numbers in agenda (1234) indicate the position of the labeled alternative in the rankings of the voters of its virtual type (type 22: $3 \succ^i 4 \succ^i 2 \succ^i 1$). For example, 2(4.) in agenda (2341) means that type 14 voters consider 2 the least preferred alternative, and 1(4.) in agenda (1234) means that type 22 voters consider 1 the least preferred. Given that the numbers in the parenthesis are the same, the behaviour of type 14 voters under (2314) is equivalent to the behaviour of type 22 voters under (1234) (if their utilities are also the same).

In order to study voters' behaviour under some agenda A_c, voters' utilities and preference orderings are re-ordered according to their virtual types, and their behaviour is then analysed under A_f. Thus, rather than changing the indexing method when the voting order changes, voters' types are changed and their voting is analysed as if they were voting under A_f.

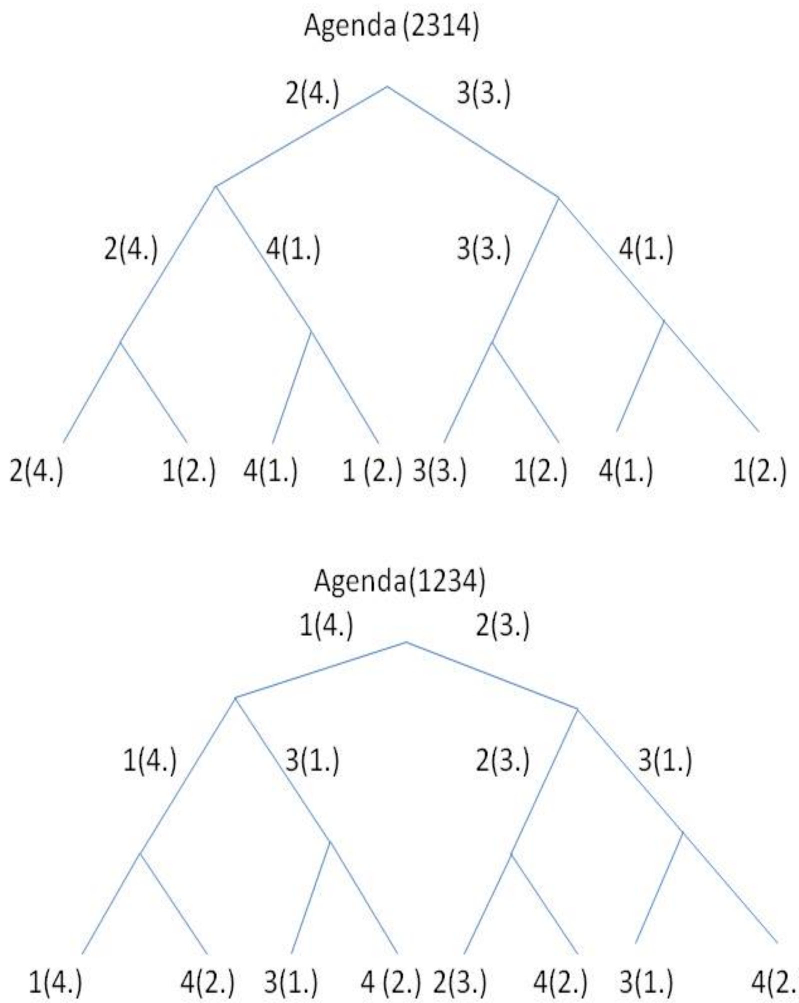
How do we know what is the relevant virtual type for any given voter type and agenda? Consider Table 2.

The virtual ordering is constructed as follows. List the fixed agenda A_f and the current agenda A_c in two adjacent columns as in Table 2. Then list the ordering of type 14 voters in a third column. The ordering of the virtual type (22) is found as follows. Find the most preferred alternative (4) on the top row of type 14 voter's ranking. Search for this alternative in the second column.

individual and each possible pair of alternatives. With n alternatives, there are $\binom{n}{2}$ different pairwise comparisons. Each player would then obtain $\binom{n}{2} * (N-1)$ observations in each simulated game g. If a setup has M elections, there would then be $M * N * \binom{n}{2} * (N-1)$ observations in one setup. With $M = 1000$, $N = 201$, and $n = 7$, this is 844200000 observations in each setup. We have not used this latter approach because computing the beliefs on the basis of them requires a considerable amount of memory and CPU-time. Most of the time is spent in calculating the probabilities from the normal distribution. This is why it is done only once.

[4] The ordering numbers are not unique because they depend on the number of alternatives.

Figure 1: Agendas (2341) and (1234)



A_f	A_c	type (14)	virtual type (22)
(1234)	(2341)		
1	2	4	3
2	3	1	4
3	4	3	2
4	1	2	1

Table 2: An example of virtual types for (2341)

The number displayed in the first column on this row (row 3) provides the corresponding alternative (3) for the virtual type. Then find the second-most preferred alternative (1), and repeat the procedure to find that the corresponding alternative is 4 for the virtual type. Go through all the alternatives in this way. A similar conversion is conducted for all voters. Then, once the outcomes of voting are calculated with these modified types *under agenda* (1234), the winning alternative is converted back to what it was with the original alternatives and voter types. This is done by finding the winner in the fourth column, and seeing which alternative corresponds to it in the third. For example, if the virtual alternative 3 (1) wins under agenda A_f when voter types are converted into virtual ones according to the agenda (2341), this means that alternative 4 (2) would win under agenda (2341).

Virtual ordering numbers

In order to use the ordering numbers for studying any agenda, it is necessary to find the virtual ordering numbers. The condition for voting for the lower branch under agenda (1234) was

$$\begin{aligned} & p_{13}\left[p_{14}U_1 + (1-p_{14})U_4\right] + (1-p_{13})\left[p_{34}U_3 + (1-p_{34})U_4\right] && (2) \\ \geq\ & p_{23}\left[p_{24}U_2 + (1-p_{24})U_4\right] + (1-p_{23})\left[p_{34}U_3 + (1-p_{34})U_4\right]. && (3) \end{aligned}$$

Expressing this in terms of ordering numbers yields

$$\begin{aligned} & p_2\left[p_3U_1 + (1-p_3)U_4\right] + (1-p_2)\left[p_6U_3 + (1-p_6)U_4\right] && (4) \\ \geq\ & p_4\left[p_5U_2 + (1-p_5)U_4\right] + (1-p_4)\left[p_6U_3 + (1-p_6)U_4\right]. && (5) \end{aligned}$$

Consider now agenda 2341 and eq. (12). Replacing alternative 2 with 1, 3 with 2, 4 with 3 and 1 with 4, we get

$$\begin{aligned} & p_{24}\left[p_{21}U_2 + (1-p_{21})U_1\right] + (1-p_{24})\left[p_{41}U_4 + (1-p_{41})U_1\right] && (6) \\ \geq\ & p_{34}\left[p_{31}U_3 + (1-p_{31})U_1\right] + (1-p_{34})\left[p_{41}U_4 + (1-p_{41})U_1\right]. && (7) \end{aligned}$$

if utilities are expressed in terms of the virtual ordering numbers. The computer is able to transform the labels for probabilities into virtual ordering numbers by using the following technique. Consider Table 5. The six lowest rows express ordering numbers for alternative pairs when the ordering of the items in the pair is interchanged. The virtual ordering numbers are found by listing the virtual alternatives in the pairwise comparisons as in columns '1. virtual' and '2. virtual', and by finding which ordering number (from the first column) correspoinds to them. Thus, for example, if 1. virtual is 4, and 2. virtual is 1, the corresponding virtual ordering number is 9. Replacing the labels of all pairs of alternatives in (14) with virtual ordering numbers then yields

$$\begin{aligned} & p_5\left[p_7U_2 + (1-p_7)U_1\right] + (1-p_5)\left[p_9U_4 + (1-p_9)U_1\right] && (8) \\ \geq\ & p_6\left[p_8U_3 + (1-p_8)U_1\right] + (1-p_6)\left[p_9U_4 + (1-p_9)U_1\right]. && (9) \end{aligned}$$

We can also write this in terms of expected utilities:

$$\begin{aligned} & p_5EU_7 + (1-p_5)EU_9 && (10) \\ \geq\ & p_6EU_8 + (1-p_6)EU_9. && (11) \end{aligned}$$

If there are five alternatives, the expression for voting for the lower branch can be obtained by replacing the utilities U_j with expected utilities from contests between j and 5. Thus, it is helpful to calculate the expected utilities for all last-round contests. These expected utilities are then put in place of utilities in equation (12), using appropriate ordering numbers.

ordering number	first	second	1. virtual	2. virtual	virtual number
1	1	2	2	3	4
2	1	3	2	4	5
3	1	4	2	1	7
4	2	3	3	4	6
5	2	4	3	1	8
6	3	4	4	1	9
7	2	1			
8	3	1			
9	4	1			
10	3	2			
11	4	2			
12	4	3			

Table 3: virtual ordering numbers for agenda 2341 with four alternatives

We will also need to know how the ordering number for the *next* round of voting is identified for calculating how the voters will vote after a result from a previous round has been revealed. This can be done using table 4 above. In order to construct this table, take Table 3, pick the ordering numbers from the rows in which the alternative that is introduced in the last round of voting (i.e. $n = 4$) is found on the 'second' column, and arrange the ordering numbers in a

1. round	**2. round**	**3. round**
		6
	4	5
1	2	3

Table 4: ordering numbers

single column to obtain a column $\begin{bmatrix} 3 \\ 5 \\ 6 \end{bmatrix}$. Then pick the rows with $n-1=3$ from this column to get $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$. Continuing like this until alternative 2, putting the results in consecutive columns, and flipping them up and down yields Table 4.

The columns were flipped up and down because the ordering numbers are now positioned in the same way as the branches in the voting tree. The leftmost column in this table displays the ordering number in the first round of voting, the second column the ordering numbers in the second, and the $(n-1)^{th}$ column from the left the $(n-1)^{th}$ (i.e. the last) round of voting. If the lower branch obtains a majority of votes, the ordering number for the next round is found on the next column directly to the right of the present one, and if the upper branch obtains more votes, the ordering number is found on the next column to the right on the south-west - north-east diagonal.

For example, with four alternatives, if 1 beats 2 in the first round, we know that we will need to calculate the expected utilities for the two branches emanating from the branch corresponding to ordering number 2, because 2 lies directly to the right of 1, the ordering number for the vote between alternatives 1 and 2. The ordering number for one branch that emanates from branch 2 must be directly to the right from 2 (3), and the other must be on the diagonal (6).

Using these indexing methods, we may thus express the expected utilities for all branches in a voting tree, and the corresponding probabilities for any number of alternatives in any voting round under an amendment agenda.