The Category of Simple Voting Games

Simon Terrington

July 11th, 2004

In this paper I describe a notation for simple voting games that has three properties:

- It facilitates the calculation of the measures of voting power
- It allows easy or intuitive proofs of certain theorems
- It offers us new insights into the so-called paradoxes of voting power

This notation arose from my efforts, over the last two years, to represent simple voting games as a category. In the beginning of the paper the notation and its uses are explained independently of category theory. The last section describes the category of simple voting games.

There are three reasons for presenting simple voting games as a category. Firstly, it seems that some of the reasoning in simple game theory is fundamentally of a category-theoretic nature (for example the construction of the constant-sum extension in Taylor and Zwicker's book). If that is the case then the machinery of category theory should facilitate understanding and discovery. Secondly, the theory of simple voting games is currently something of an island: separate from the rest of mathematics. Most of the significant classes of structures studied in mathematics (for example groups, vector spaces and topological spaces) can be expressed as categories. If the same could be done for SVGs then they could be connected to other areas of mathematics using functors (which map between categories in a way which preserves the structure). Last, I wondered if a category-theoretic presentation might aid calculation of power indices and related measures and this has indeed turned out to be the case.

These ideas were developed during preliminary work for my PhD, which is being supervised by Moshe Machover.

1 A Notation for the Theory of Simple Voting Games

For a simple voting game with n voters, the notation consists of a vector of 2^n zeros and ones. Each entry in the vector corresponds to the outcome of the SVG that results from one of the 2^n possible divisions. Each entry is labelled by a subscript. The vector might look like this $(x_0x_1...x_{2^n-1})$. In fact, I write the subscripts in the binary number system like this $(x_0x_1x_{10}...x_{111...1})$. For any of the 2^n outcomes the binary subscript makes clear the division to which the outcome corresponds. For example x_{100101} is the outcome when the first, third and sixth voter vote 'yes' and all the others vote 'no'. What is represented is a simple voting game with the voters labelled in a defined order.

The definition of the notation for an SVG with n voters, and a relation $<$ between SVGs, is by recursion (on n):

- (0) and (1) are defined as representations of SVGs with 0 voters with $(0) < (0)$, $(0) < (1)$ and $(1) < (1)$
- If (A) and (B) are representations of SVGs with n voters and $(A) < (B)$ then (AB) is a representation of an SVG with $n+1$ voters
- If $(A) < (B)$ and $(C) < (D)$ then $(AC) < (BD)$. Of course for (AC) and (BD) to be valid as games with $(n+1)$ voters we need $(A) < (B)$ and $(C) < (D)$ as games with n voters

For illustration, I will show how five SVGs translate into this notation:

- The unanimity game with three voters looks like this (00000001)
- The game with three voters in which the second voter is a dictator looks like this (00110011)
- The games with three voters where the first or third voter is a dictator look like this (01010101) and this (00001111) respectively
- The game with minimum winning coalitions {1,2} and {3} looks like this (00011111)

I will now show how various measures associated with simple voting games can be calculated using this notation.

1.1 The Bz Score

The Bz Score for the n^{th} voter is the number of coalitions that move from yielding a '0' to yielding a '1' when the n^{th} voter is added to them:

$\eta_n = \sum x_{1d_{n-1}d_{n-2}...d_0} - \sum x_{0d_{n-1}d_{n-2}...d_0}$. In other words, the sum of all the entries on the right half minus the sum of all the entries on the left half.

The Bz scores for the third voter in the examples given above ((00000001), (00110011), (01010101), (00001111) and (00011111)) are 1, 0, 0, 4 and 3 respectively

The Bz score for the first voter is:

$\eta_n = \sum x_{d_n d_{n-1}d_{n-2}...1} - \sum x_{d_n d_{n-1}d_{n-2}...0}$

For the example games, the scores for the first voter are 1, 0, 4, 0 and 1 respectively.

For the r^{th} voter the Bz score is:

$\eta_r = \sum x_{d_n d_{n-1}...d_{r+1}1d_{r-1}...d_1} - \sum x_{d_n d_{n-1}...d_{r+1}0d_{r-1}...d_1}$

For the example games, the scores for the second voter are 1, 4, 0, 0 and 1

1.2 The Penrose Measure

Moving from the Bz score to the Penrose measure is relatively simple. We need to divide by 2^{n-1}, so the formulas are:

$\psi_r = (\sum x_{d_{n-1}d_{n-2}...d_{r+1}1d_{r-1}...d_0} - \sum x_{d_{n-1}d_{n-2}...d_{r+1}0d_{r-1}...d_0})/2^{n-1}$
$\psi_1 = (\sum x_{d_{n-1}d_{n-2}d_{n-3}...1} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...0})/2^{n-1}$
$\psi_n = (\sum x_{1d_{n-2}d_{n-3}...d_0} - \sum x_{0d_{n-2}d_{n-3}...d_0})/2^{n-1}$

For our example games, the Penrose measures work out as follows:
For (00000001) $\psi_1 = \frac{1}{4}$, $\psi_2 = \frac{1}{4}$, $\psi_3 = \frac{1}{4}$
For (00110011) $\psi_1 = 0$, $\psi_2 = 1$, $\psi_3 = 0$
For (01010101) $\psi_1 = 1$, $\psi_2 = 0$, $\psi_3 = 0$
For (00001111) $\psi_1 = 0$, $\psi_2 = 0$, $\psi_3 = 1$
For (00011111) $\psi_1 = \frac{1}{4}$, $\psi_2 = \frac{1}{4}$, $\psi_3 = \frac{3}{4}$

1.3 Sensitivity

I start with some definitions:

$H[G] := \sum_{r=1}^{n} \eta_r$
$\Sigma := \sum_{i=1}^{n} \psi_i = (\sum_{i=1}^{n} \eta_i)/2^{n-1} = H/2^{n-1}$

$E(Z)$=the expected value of the number of voters who agree with the outcome of G minus the number that do not.

Adding up all the formulae for the η_i we have:

$$H[G] = nx_{111\ldots1} + (n-2)\sum_{S=n-1} x_{d_n\ldots d_1} + (n-4)\sum_{S=n-2} x_{d_n\ldots d_1}\cdots$$
$$-(n-2)\sum_{S=1} x_{d_n\ldots d_1} - nx_{000\ldots0}$$

Here I have written S as an abbreviation for $\sum d_i$

We can see from this why mean majority deficit pulls in the opposite direction to sensitivity. If the game allows lots of small coalitions to win (i.e coalitions which have a small number of 1s in their (binary) index) then this is likely to make a large negative contribution to the sensitivity and, of course a positive contribution to the mean majority deficit. Alternatively, if large coalitions with a lot (more than $n/2$) of 1s in their index do not win then this will be a missed opportunity to push up the sensitivity (and push down the mean majority deficit). It is also easy to see from this that the game with the minimum mean majority deficit (maximum sensitivity) is the one in which all the coalitions greater in size than $n/2$ win and all smaller than $n/2$ lose. If n is even then the coalitions with exactly $n/2$ voters voting 'yes' have a zero coefficient so it does not matter if they are winning or losing.

We are able to calculate the sensitivity for our example games

For (00000001), $\Sigma = \frac{3}{4}$

For (00110011), $\Sigma = 1$

For (01010101), $\Sigma = 1$

For (00001111), $\Sigma = 1$

For (00011111), $\Sigma = \frac{5}{4}$

1.4 Coleman Indices

There are two Coleman indices: γ_r and γ_r^*. They measure the ability to inhibit action and take action respectively.

To calculate these I will need:

$\omega := \sum x_i$ the number of winning coalitions

$\omega_r := \sum x_{d_1\ldots d_{r-1}1d_{r+1}\ldots d_n}$ the number of winning coalitions in which the r_{th} voter votes 'yes'

The number of losing coalitions is $2^n - \omega$. With these definitions we can calculate γ_r and γ_r^* as follows:

$$\gamma_r = (\sum x_{d_1\ldots d_{r-1}1d_{r+1}\ldots d_n} - \sum x_{d_1\ldots d_{r-1}0d_{r+1}\ldots d_n})/\omega$$

$$\gamma_r^* = (\sum x_{d_1,\ldots d_{r-1}1d_{r+1}\ldots d_n} - \sum x_{d_1\ldots d_{r-1}0d_{r+1}\ldots d_n})/(2^n - \omega)$$

For (00000001)
$\gamma_1 = \frac{1}{1}$ $\gamma_2 = \frac{1}{1}$ $\gamma_3 = \frac{1}{1}$
$\gamma_1^* = \frac{1}{7}$ $\gamma_2^* = \frac{1}{7}$ $\gamma_3^* = \frac{1}{7}$
For (00110011)
$\gamma_1 = \frac{0}{4}$ $\gamma_2 = \frac{4}{4}$ $\gamma_3 = \frac{0}{4}$
$\gamma_1^* = \frac{0}{4}$ $\gamma_2^* = \frac{4}{4}$ $\gamma_3^* = \frac{0}{4}$
For (01010101)
$\gamma_1 = \frac{4}{4}$ $\gamma_2 = \frac{0}{4}$ $\gamma_3 = \frac{0}{4}$
$\gamma_1^* = \frac{4}{4}$ $\gamma_2^* = \frac{0}{4}$ $\gamma_3^* = \frac{0}{4}$
For (00001111)
$\gamma_1 = \frac{0}{4}$ $\gamma_2 = \frac{0}{4}$ $\gamma_3 = \frac{4}{4}$
$\gamma_1^* = \frac{0}{4}$ $\gamma_2^* = \frac{0}{4}$ $\gamma_3^* = \frac{4}{4}$
For (00011111)
$\gamma_1 = \frac{1}{5}$ $\gamma_2 = \frac{1}{5}$ $\gamma_3 = \frac{3}{5}$
$\gamma_1^* = \frac{1}{3}$ $\gamma_2^* = \frac{1}{3}$ $\gamma_3^* = \frac{3}{3}$

2 Proving Some Existing Theorems

This notation allows new proofs to emerge for many existing theorems. I will give two illustrations. The first is due to Penrose and is given by Felsenthal and Machover as 3.2.16:

Theorem

P(the n^{th} voter agrees with the outcome)=(1+β'_n)/2

Proof

The probability that the n^{th} voter agrees with the outcome is equal to the number of 'yes' outcomes that result from the n^{th} voter voting 'yes' plus the number of 'no' outcomes that result from the n^{th} voter voting 'no' divided by the total number of outcomes (2^n) or

$(\sum x_{1,d_{n-1}\ldots d_1} + \sum(1 - x_{0,d_{n-1}\ldots d_1}))/2^n =$
$(2^{n-1} + (\sum x_{1,d_{n-1}\ldots d_1} - x_{0,d_{n-1}\ldots d_1})/2^n =$
=1/2+β'_n/2

Theorem

For any two distinct voters a and b of a simple voting game W
$\psi_{a\&b}[W|a\&b] = \psi_a[W] + \psi_b[W]$

Proof

Without loss of generality, let a be the first voter and b the second.
$\psi_a[W] = (\sum x_{d_{n-1}d_{n-2}d_{n-3}\ldots d_1 1} - \sum x_{d_{n-1}d_{n-2}d_{n-3}\ldots d_1 0})/2^{n-1}$
$\psi_b[W] = (\sum x_{d_{n-1}d_{n-2}d_{n-3}\ldots 1 d_0} - \sum x_{d_{n-1}d_{n-2}d_{n-3}\ldots 0 d_0})/2^{n-1}$

In $W|a+b$, a and b vote together so we only need consider outcomes that have two 1s or two zeros in the first two places of the subscript. i.e. we need

to think about the $x_{d_{n-1}...11}$ when they both vote 'yes' and $x_{d_{n-1}...00}$ when they vote 'no'.

$\psi_{a+b}[W] = (\sum x_{d_{n-1}d_{n-2}d_{n-3}...11} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...00})/2^{n-2}$

We are dividing by 2^{n-2} rather than 2^{n-1} because there are only $n-1$ voters in $W|a+b$.

$\psi_a[W] + \psi_b[W] =$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...d_1 1} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...d_1 0})/2^{n-1}+$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...1d_0} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...0d_0})/2^{n-1} =$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...11} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...10})/2^{n-1}+$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...01} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...00})/2^{n-1}+$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...11} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...01})/2^{n-1}+$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...10} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...00})/2^{n-1} =$
$(2\sum x_{d_{n-1}d_{n-2}d_{n-3}...11} - 2\sum x_{d_{n-1}d_{n-2}d_{n-3}...00})/2^{n-1} =$
$(\sum x_{d_{n-1}d_{n-2}d_{n-3}...11} - \sum x_{d_{n-1}d_{n-2}d_{n-3}...00})/2^{n-2} =$
$\psi_{a+b}[W]$

If we let the x_is take values other than 0 and 1, this proof can be extended to games with a general worth function.

3 The Category of Simple Voting Games

A category consists of objects and arrows that satisfy these axioms:

- Every arrow is associated with a domain ($Dom(A)$) and codomain ($Cod(A)$). Both are objects
- For any two arrows (A and B) such that $Cod(A) = Dom(B)$ there is a composite arrow with domain equal to $Dom(A)$ and codomain equal to $Cod(B)$. We represent this composite by BA
- Composition of arrows is associative. That is, assuming that the arrows can be composed, $A(BC) = (AB)C$
- For any object (A) we have an identity arrow i_A. i_A has A as its domain and codomain. When the identity is composed with another arrow F, with domain A, the result is F so $i_AF = F$. If the codomain of F is A then we have $Fi_A = F$

So what does the category of simple voting games look like? Actually we will define a sequence of categories. There will be a category of games with n voters for each n. This sequence will exactly mirror the recursion in the first section.

To start with I want to describe a relation on all games with n voters. (A game) A is bigger than (a game) B if and only if for any division of the voters B passes whenever A passes. The blunt version of this relation (i.e. the relation of being at least as big) will turn out to be realised by $<$ as defined in the first section.

The first category that I will describe (C_0) is the category whose objects are SVGs with no voters. How can an SVG have no voters? Well these games could be described as degenerate. Another way to think of the objects in this category is as outcomes. C_0 contains two objects: 0 and 1. These correspond to passing the motion and not passing it, respectively.

In all of the categories of SVG an arrow from A to B will say 'B is at least as big as A'. In this category we have three arrows: from 0 to 0, from 0 to 1 and from 1 to 1. These arrows will be denoted by (00), (01) and (11) respectively.

In every SVG with one voter, the voter switches between two outcomes. The monotonicity condition says that the outcome if he votes yes must be at least as big as the outcome if he votes no. So for the objects of C_1 (the category of games with one voter) we need pairs (AB) of outcomes (objects of C_0) where B is at least as big as A. These are exactly the arrows of C_0. We can define the objects of C_1 to be the arrows of C_0. This is a recognised construction of one category from another (and so we know that the axioms for a category must hold). C_1 is referred to as the arrows category of C_0. Given two objects of C_1 (arrows of C_0) there is an arrow between them in C_1 if and only if, as arrows in C_0, they form part of, what is referred to in category theory as, a commuting square.

We can now go on to define all of the C_n. If C_n has all of the games with n voters as objects then a game with n+1 voters can be seen as a decision, made be the $(n+1)^{th}$ voter, to switch between two games with n voters. The monotonicity condition implies that when the $(n+1)^{th}$ voter votes 'yes' we end up with a bigger game then when he votes 'no'. So the objects of C_{n+1} are just pairs (A, B) of objects with B at least as big as A. These correspond to the arrows of C_n.

For the sake of illustration, I have calculated the objects of the first few categories:

- In C_0 the objects are 0 and 1
- In C_1 there are three: (00), (01) and (11)
- In C_2 there are six: (0000),(0001),(0011),(0101),(0111) and (1111)
- Finally, in C_3 there are twenty:

(00000000),(00000001),(00000011),(00000101),
(00000111),(00001111),(00010001),(00010011),
(00010101),(00010111),(00011111),(00110011),
(00110111),(00111111),(01010101),(01010111),
(01011111),(01110111),(01111111),(11111111)

The categories of simple voting games have some special properties which suggest that we may be able to deploy category-theoretic techniques to learn more about SVGs:

- Each has a terminal object and an initial object (the game that always passes and the games that always blocks respectively)
- The dual category consists exactly of the duals of the games in the original category
- The simple game-theoretic product is coincident with the category-theoretic product