ELSEVIER
Discrete Applied Mathematics 119 (2002) 265–271
DISCRETE APPLIED MATHEMATICS

Note

Cooperative phenomena in crystals and the probability of tied Borda count elections

Thierry Marchant

Fac. Psychologie en Ped Wetensch Ghent University, H. Dunantlaan 1, B-9000, Gent, Belgium

Received 10 September 1998; received in revised form 14 April 2000; accepted 8 January 2001

Abstract

In 1960, C. Domb published a paper entitled *On the Theory of Cooperative Phenomena in Crystals* in which he presented an expression for the number of cycles of length l in a triangular lattice. This expression was erroneous. We present a correct expression and we show that it is linked, in social choice theory, to the probability that all candidates are tied in an election with the Borda rule. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Borda; Probability of ties; Random walk

1. Introduction

In 1960, C. Domb published a massive paper (212 pp.) entitled *On the Theory of Cooperative Phenomena in Crystals* [2]. In this paper, he addressed many different problems. One of them, related to the magnetic properties of crystals, was the following. Consider the lattice [1] presented in Fig. 1.

A cycle in this lattice is a path starting from some node, travelling along some edges and coming back to the same node. The length of a cycle is the number of edges contained in it. E.g. the shortest cycle has length 2; it starts from some node, travels along one edge and directly comes back along the same edge. The next shortest cycle has length 3. It travels along the 3 edges delimiting a small triangle. How many different cycles, with length l, starting from a given node, are there?

The answer given by Domb [2, pp. 344] was

$$\sum_{s,t}\frac{1}{s!t!}\sum_{q=0}^{s}2^{s-q}\frac{(t+q)!}{((t+q)/2!)^2}\frac{1}{q!(s-q)!}, \tag{1}$$

E-mail address: thierry.marchant@rug.ac.be (T. Marchant).

[1] In classical graph theory, lattice has a different meaning. We use it here in its crystallography meaning which is close to the concept of pavement in geometry.

PII: S0166-218X(01)00308-0

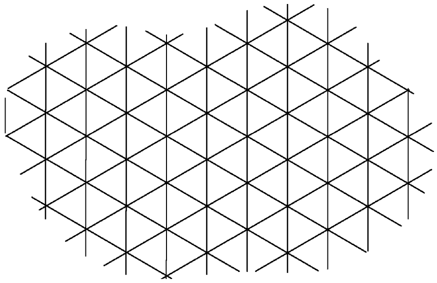

Fig. 1. Triangular lattice.

with $s,t=0,1\ldots l$, satisfying the conditions $(2s+t)=l$ and $(t+q)$ even. He called this number r_l. On p. 345, he also computed some values of r_l for $l=2$–9. Unfortunately expression (1) for r_l is incorrect although numerical values given on p. 345 are correct. In Section 2, we present a correct expression for r_l. In the last section, we present some links with social choice theory.

2. A correct expression for the number of cycles

At each node, there are six possible edges. Let us call them $x,-x,y,-y,z,-z$ as in Fig. 2.

In a cycle, an edge of any kind can be compensated by a corresponding edge of the opposite sign. E.g. a x edge can be compensated by a $-x$ edge; a $-z$ edge by a z edge, and so on. But an edge of any kind can also be compensated by two other edges of different kinds and same sign. E.g., a y edge can be compensated by a x edge and a z edge; a $-z$ edge by a $-x$ edge and a $-y$ edge and so on.

Hence, in any cycle, we can sort the edges into two parts: (a) those compensating 2 by 2 and (b) those that do not compensate 2 by 2 (thus compensating 3 by 3). In (a), we necessarily have $2s$ edges (s integer between 0 and $l/2$). In (b), we necessarily have $3t$ edges (t integer between 0 and $l/3$) and they have the same sign, otherwise some of them could compensate 2 by 2. Obviously, $2s+3t=l$. In (a), s edges are positive, while s edges are negative. Among the s positive edges, there is any repartition between

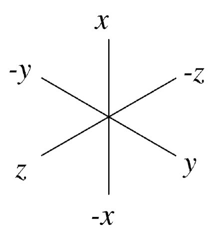

Fig. 2. Names of the edges.

x, y and z. Among the s negative edges, there is the same repartition between $-x, -y$ and $-z$. Hence, if $2s$ edges among l are chosen, the number of possible configurations of these $2s$ edges, such that they all compensate 2 by 2 (case (a)), is given by

$$\binom{2s}{s}\sum_{q=0}^{s}\sum_{r=0}^{s-q}\frac{(s!)^2}{((s-q-r)!r!q!)^2}. \tag{2}$$

In this expression, q represents the number of x edges, r the number of y edges and $(s-q-r)$ the number of z edges. Let us come back to (b). Among the $3t$ edges, we necessarily have the same number of x, y, z or (exclusive) $-x, -y, -z$. Thus, if we choose $3t$ edges, the number of possible configurations such that they compensate 3 by 3 and not 2 by 2, is given by

$$2^{[t>0]}\frac{(3t)!}{(t!)^3}, \tag{3}$$

where $[t>0]$ equals 1 if $t>0$ and 0 otherwise.

For given s and t, the number of possible cycles is not just the product of expressions (2) and (3). It would be equivalent to considering as different some cycles just because we arbitrarily separated some edges of the same kind and sign in the (a) and (b) parts. Thus we have to take into account the number of ways to choose t edges of kind x (or $-x$) among the whole number of x edges. And there are $q+t$ such edges. The number of ways to make this choice is thus $\binom{q+t}{t}$. For y and z edges, we must consider $\binom{r+t}{t}$ and $\binom{s-q-r+t}{t}$.

Hence, for given s and t, the number of possible cycles is given by

$$\binom{l}{2s}\binom{2s}{s}\sum_{q=0}^{s}\sum_{r=0}^{s-q}\frac{\frac{(s!)^2}{((s-q-r)!r!q!)^2}\,\frac{2^{[t>0]}(3t)!}{(t!)^3}}{\binom{q+t}{t}\binom{r+t}{t}\binom{s-q-r+t}{t}}. \tag{4}$$

Finally, letting vary s and t so that $2s+3t=l$, we obtain the following expression:

$$r_l=\sum_{2s+3t=l}\binom{l}{2s}\binom{2s}{s}\sum_{q=0}^{s}\sum_{r=0}^{s-q}\frac{\frac{(s!)^2}{((s-q-r)!r!q!)^2}\,\frac{2^{[t>0]}(3t)!}{(t!)^3}}{\binom{q+t}{t}\binom{r+t}{t}\binom{s-q-r+t}{t}}. \tag{5}$$

After some simplification, r_l is given by

$$l!\sum_{2s+3t=l}2^{[t>0]}\sum_{q=0}^{s}\sum_{r=0}^{s-q}\frac{1}{q!r!(s-q-r)!(q+t)!(r+t)!(s-q-r+t)!}. \tag{6}$$

Shortly after we found this expression, Domb (personal communication) found the error in his expression: a multiplicative factor $l!$ had disappeared from his formula during the typing process! Therefore, an alternate expression for (5) is

$$\sum_{s,t}\frac{l!}{s!t!}\sum_{q=0}^{s}2^{s-q}\frac{(t+q)!}{((t+q)/2!)^2}\frac{1}{q!(s-q)!}, \tag{7}$$

with $s, t = 0, 1, \ldots, l$, satisfying the conditions $(2s + t) = l$ and $(t + q)$ even. So, Domb knew the right expression in 1960. But, due to the fact that a proof of this expression has never been published and that this result finds some new applications in social choice theory (see next section), we think that it is worth publishing our proof.

3. Some links with social choice theory

A very classical problem in social choice is the following. Suppose that l voters $\{1, 2 \ldots, l\}$ must elect a president and there are k candidates $\{a, b, c, \ldots\}$. Each voter expresses his preferences about the candidates by mean of a complete ranking, from best to worst. We call profile a vector containing the rankings of each voter. E.g.,

$$\begin{pmatrix} a > b > c \\ b > c > a \\ c > b > a \end{pmatrix} \tag{8}$$

is a profile with three voters and three candidates such that voter 1's most preferred candidate is a, voter 1's last candidate is c and voter 2's most preferred candidate is b. How shall we derive from a profile which candidate should be elected? This question has been at the heart of social choice theory since the end of the 18th century. Many methods have been proposed. For example,

- choose the candidate with most first positions,
- or the candidate with least last positions,
- or compute the ranking which is at minimum distance of the l rankings in the profile (a metric needs to be defined over the set of the rankings). Then choose the candidate in first position in this new ranking.
- A very popular method is the Borda method. A candidate receives one point for each first position in the profile, 2 points for each 2nd position, 3 points for each 3rd position, … and k points for each last position. The candidate who has the fewest points is elected.

Let us illustrate the Borda method by an example. In the profile shown in (8), a has 7 points, b, 5 points and c, 6 points. Hence, b is elected. In some cases the Borda method does not help much as all candidates have the same number of points and are tied, as in profile (9) where they all have six points.

$$\begin{pmatrix} a > b > c \\ b > c > a \\ c > a > b \end{pmatrix}. \tag{9}$$

Of course, most methods that have been devised lead to different results. Which one should we choose? Many criteria have been proposed to assess the merits of a method. Hundreds of axiomatic studies have been conducted, characterizing the various methods by a set of axioms.

A possible criterion to compare different methods, is the probability that a method yields a tie (by this, we mean a complete tie of all candidates). A method with a

Table 1
Numerical values of the probability of ties for three candidates

l	2	3	4	5	6	7	8
Probability of ties	0.1667	0.0556	0.0694	0.0463	0.0437	0.0360	0.0326
Probability of Condorcet paradox		0.0556		0.0694		0.0750	
l	9	19	29	39	49		
Probability of ties	0.0288	0.0141	0.0093	0.0070	0.0056		
Probability of Condorcet paradox	0.0780	0.0832	0.0848	0.0856	0.0860		

high probability of tie might be considered as less interesting than a method with a low probability because it more often fails to designate a winner. Of course, if the difference of the probabilities is not very large, this disadvantage might be compensated by other advantages. Hence, this criterion should be taken into account only for very large differences of probabilities. We are going to show now that the probability that the Borda method yields a tie is related to r_l.

3.1. The case of three candidates

For three candidates, there are $3! = 6$ possible rankings and each voter can choose any of the 6 rankings. Let us associate each ranking to one of the 6 different kinds of edges of our triangular lattice (see Fig. 2).

$$\begin{aligned} a > b > c &: \quad x, \\ b > c > a &: \quad y, \\ c > a > b &: \quad z, \\ c > b > a &: -x, \\ a > c > b &: -y, \\ b > a > c &: -z. \end{aligned} \tag{10}$$

Then any profile corresponds to a path in the triangular lattice. For example, the profile in (9) corresponds to a path x, y, z. Remark that this path is in fact a cycle. It is not difficult to see that it is not a coincidence. A profile will yield all candidates tied (under the Borda method) if and only if the corresponding path in the triangular lattice is a cycle. Hence, the number of profiles yielding all candidates tied under the Borda method is r_l. And the probability we were looking for is just r_l divided by the number of different profiles, i.e. $(3!)^l$. Some numerical values of the probability of ties are given in Table 1.

The Condorcet method selects the candidate that beats every other candidates in pairwise comparisons. It is well known that the Condorcet method can also fail to produce a winner (Condorcet paradox) but for very different reasons: the candidates are not tied, the method just does not work. Nevertheless, from a practical point of view, if all candidates win (tied), using the Borda method, or no candidate wins, using the Condorcet one, the president of the committee where such an election happens is very

Fig. 3. Linear lattice.

Table 2
Numerical values of the probability of ties for two candidates

l	2	4	6	8	10	20	50	∞
Probability of ties	0.5000	0.3750	0.3125	0.2734	0.2461	0.1762	0.1123	0

embarrassed: he does not know what to choose. Therefore, it seems interesting to us to compare the probabilities of ties for the Borda method to those of Condorcet paradox for the Condorcet method (see Table 1), taken from [4]. The proportion of profiles such that the president of the committee is not helped is larger with the Condorcet method. Furthermore, the probability of ties decreases with l while the probability of Condorcet paradox increases with l. From this viewpoint, the Borda method seems more interesting that the Condorcet one. In fact, for very large number of voters and candidates, the probability that the Condorcet method designates no winner approaches 1 [1].

3.2. The case of two candidates

Let us consider the lattice of Fig. 3, consisting of edges aligned on a straight line.

At each node of this lattice, there are two possible edges. Let us call them x and $-x$. It is clear that we can describe any profile with two candidates by a path in our linear lattice. We just need to associate $a > b$ rankings to an edge, say x, and $b > a$ rankings to the other edge, i.e. $-x$. It is obvious as well that all profiles such that a and b are tied correspond to cycles in the linear lattice and the number of different cycles of length l is given by $\binom{l}{l/2}$ for l even and 0 for l odd. Some values of the probability of ties are given in Table 2.

For larger number of candidates, other lattices must be used but they can no longer be represented in two dimensions. Derivation of explicit formulas for r_l is much more difficult.

In our computations of the probabilities, we considered that any profile is as likely as any other one (this condition is known as the *impartial culture condition*). Therefore, it is obvious that our results must be taken with a pinch of salt for, in reality, such an assumption is clearly questionable [3]. Nevertheless, they provide some hint.

Note that I discovered the similarity between the two problems thanks to the amazing *Encyclopedia of integer sequences* [5, sequence M4101].

References

[1] C.E. Bell, A random voting graph almost surely has a hamiltonian cycle when the number of alternatives is large, Econometrica 49/6 (1981) 1597–1603.
[2] C. Domb, On the theory of cooperative phenomena in crystals, Adv. Phys. 9 (1960) 149–361.
[3] P.C. Fishburn, W.V. Gehrlein, The paradox of voting. Effects of individual indifference and intransitivity, J. Public Econom. 14 (1980) 83–94.
[4] W.V. Gehrlein, Condorcet's paradox, Theory Decis. 15 (1983) 161–197.
[5] N.J.A. Sloane, S. Plouffe, The encyclopedia of Integer Sequences, Academic Press, New York, 1995.