

THE PROBABILITY OF A GIVEN 1-CHOICE STRUCTURE*

FRANK HARARY

UNIVERSITY OF MICHIGAN

AND

RON READ

UNIVERSITY OF THE WEST INDIES

A 1-choice structure arises whenever each person in a group chooses exactly one other person according to some criterion. Our purpose is to study the situation in which these choices are made at random. As noted in Harary, Norman and Cartwright [2], such a structure can be represented by a directed graph in which the points represent people and the directed lines their choices. We first describe the shape of such a 1-choice structure, and define its symmetry number. With the help of these properties we are then able to develop and prove a formula which gives a probability of obtaining a given structure in the random choice situation. In order to supply data for these results, we include in the Appendix the diagrams of all 1-choice structures with at most 7 points and the probability of each.

Introduction

A 1-choice structure will always be represented by a directed graph (digraph) in which there is exactly one directed line from each point. These will henceforth be called *functional digraphs* as in [2]. In Fig. 1, all the functional digraphs with 4 points are shown, up to isomorphism. This means that any functional digraph with 4 points that anyone may ever draw will be identical with exactly one of these by a suitable identification of points.

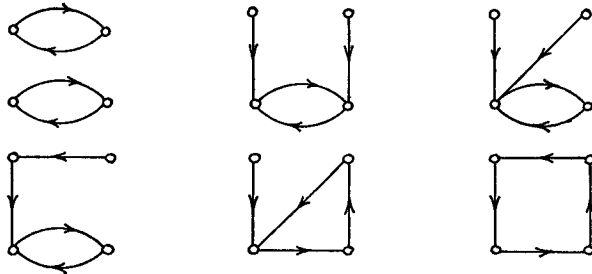


FIGURE 1

*The preparation of this article was supported by Grant MH 10834 from the National Institute of Mental Health and Grant 1026-66 from the U. S. Air Force Office of Scientific Research.

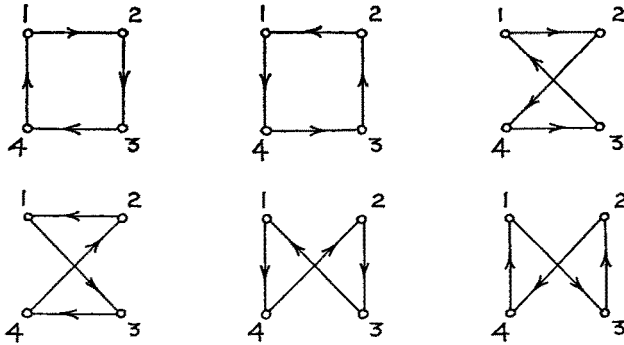


FIGURE 2

When a line is drawn at random from each point to another point, each of the functional digraphs shown in Fig. 1 can occur in more than one way. In Fig. 2, we show the six different ways in which the last functional digraph in Fig. 1 can so occur; these are called *labelled* functional digraphs, since we are now distinguishing the four points from each other.

Although the number of functional digraphs with a given number of points has been derived (see Harary [1] and Read [4]), the formulas give only the numbers, not the structures themselves. The Appendix lists the diagrams of all the functional digraphs with at most 7 points, and therefore, of course, includes those shown above in Fig. 1. In addition to functional digraphs themselves, the symmetry number (defined below) and the probability of each is given.

The Shape

How can one recognize a functional digraph when it is given? It is really very simple. Each connected part (weakly connected component) consists of a collection of rooted trees (trees with a distinguished point) in which every line is directed towards the root, together with one directed cycle containing all the roots. This result is proved in [1] and [2]. In Fig. 3, we show all the different functional digraphs, in which the cycle is of length 4, and the rooted trees are paths with 1, 2, 2, and 3 points respectively.

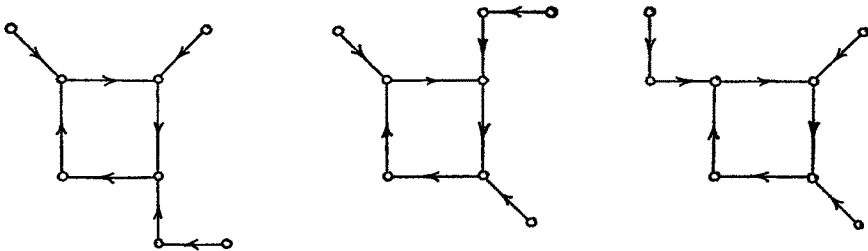


FIGURE 3

The Symmetry Number

An *automorphism* or a *symmetry* of a digraph is a mapping of the points onto themselves which preserves the directed lines. It is well known that the collection of all the symmetries of a digraph forms a (mathematical) group. We will denote the order of (number of elements in) this group by $s(D)$ and call it the *symmetry number* of D . This invariant of a digraph may be illustrated with the functional digraph D shown in Fig. 4. There are $6 = 3!$ symmetries involving only the end points of the tree in the upper right corner, 2 symmetries involving only the end points at the upper left, and similarly 2 at the lower right. Since these symmetries are independent of each other and exhaustive, $s(D) = 6 \times 2 \times 2 = 24$.

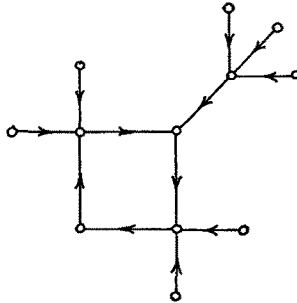


FIGURE 4

The Probability

We now have the auxiliary concepts required to express the probability of a given functional digraph. Consider the probability of obtaining a given functional digraph D having n points, by drawing directed lines at random, one from each point. The total number of possibilities is the number of labelled functional digraphs on n points. As there are $n - 1$ other points to which the line from each point may go, we see that the total number of functional digraphs on n points is $(n - 1)^n$. Since this is the magnitude of the sample space, it is the denominator in the expression for the probability $\text{Pr}(D)$ of the occurrence of the given digraph D . The numerator will be the number of functional digraphs that can be obtained from D by labelling its points; in other words it is the number of labellings of D , a number which we shall denote by $l(D)$, so that

$$\text{Pr}(D) = \frac{l(D)}{(n - 1)^n}.$$

We now obtain an expression for $l(D)$. If n points are drawn, they can be labelled with integers $1, 2, \dots, n$ in $n!$ ways before drawing the lines.

When the lines are drawn, not all of these labellings of the points will result in distinct labellings of D . For example, in Fig. 2 there are six labellings of the same functional digraph, rather than $4! = 24$. The symmetry number of D now plays a role. For each labelling of D gives rise, under the symmetries of D , to $s(D)$ equivalent labellings (one of which is the original labelling). Therefore we find that

$$l(D) = \frac{n!}{s(D)}.$$

For example, in Fig. 2, $l(D) = 4!/4 = 6$. Substituting this into the preceding equation,

$$\Pr(D) = \frac{n!}{s(D)(n-1)^n}.$$

This formula is used to obtain the probability values in the Appendix. This formula can easily be generalized to apply to a digraph D of any particular kind. If $T(n)$ denotes the total number of digraphs of the given kind having n labelled points, then

$$\Pr(D) = \frac{n!}{s(D)T(n)}.$$

Expressions for $T(n)$ for some common types of digraphs are as follows.

All digraphs	$2^{n(n-1)}$
Digraphs with q lines	$\binom{n(n-1)}{q}$
Undirected graphs	$2^{n(n-1)/2}$
Undirected graphs on q lines	$\binom{\binom{n}{2}}{q}$
Tournaments	$2^{n(n-1)/2}$

Appendix

Here is a catalogue of all functional digraphs with up to 7 points. With each digraph is given its symmetry number, $s(D)$, the number of labellings, $l(D)$, and the probability of obtaining it by a random choice of one line from each point.

These diagrams were obtained in the following systematic manner. For each choice of k , the length of the unique cycle in a connected functional digraph, a set of k rooted trees was chosen, whose total number of points did not exceed 7. A catalogue of rooted trees appears in [3]. These trees were

TWO POINTS



2, 1, 1-0000

FIVE POINTS



1, 120, 0-1172



2, 60, 0-0586



1, 120, 0-1172

THREE POINTS



1, 6, 0-7500



3, 2, 0-2500

FOUR POINTS



1, 24, 0-2963



2, 12, 0-1481



2, 12, 0-1481



1, 24, 0-2963



4, 6, 0-0740



8, 3, 0-0370



1, 120, 0-1172



2, 60, 0-0586



1, 120, 0-1172



2, 60, 0-0586



1, 120, 0-1172



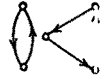
1, 120, 0-1172



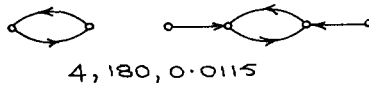
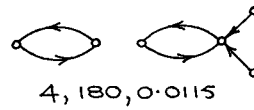
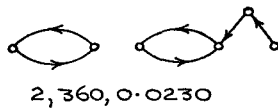
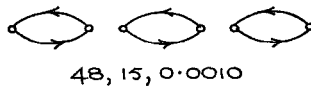
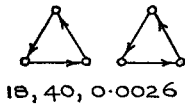
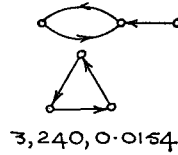
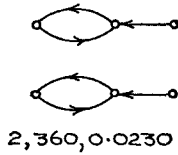
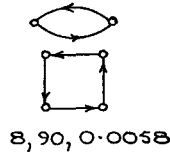
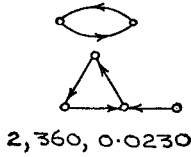
5, 24, 0-0234



2, 60, 0-0586



6, 20, 0-0195



placed in all possible ways around the cycle of length k , bearing in mind the possible symmetries of the cycle. In this way a list of all the connected functional digraphs with not more than 7 points was obtained. By multiplying the symmetry numbers of the rooted trees together with the number of symmetries of the cycle as accompanied by these rooted trees, the symmetry numbers of these digraphs were computed. By taking several of these connected functional digraphs at a time, the disconnected functional digraphs were obtained, and their symmetry numbers calculated.

To check the accuracy of the list, the functional digraphs with $n = 2, 3, \dots, 7$ points were counted and the numbers were found to agree with the coefficients in the counting series [1, 4] for these digraphs. The number $l(D)$ of labellings of each digraph was found, and a check was made that the sum of the labellings for a given value of n , i.e., that the number of labelled functional digraphs on n points, agreed with the formula $(n - 1)^n$.

From the data in this appendix, the probabilities of obtaining functional

digraphs of a particular type can be extracted. Thus the probabilities p_n of obtaining a connected functional digraph with n points are given in the following table.

n	2	3	4	5	6	7
p_n	1.0000	1.0000	0.9630	0.9219	0.8832	0.8482

The probabilities of obtaining connected functional digraphs with n points in which the cycle has a given length k are given by the following table.

n	k					
	2	3	4	5	6	7
2	1.0000					
3	0.7500	0.2500				
4	0.5926	0.2963	0.0740			
5	0.4883	0.2930	0.1172	0.0234		
6	0.5157	0.2765	0.1382	0.0461	0.0077	
7	0.3602	0.2573	0.1470	0.0630	0.1800	0.0026

REFERENCES

- [1] Harary, F. The number of functional digraphs. *Math. Annalen*, 1959, **138**, 203-210.
- [2] Harary, F., Norman, R., and Cartwright, D. *Structural models: An introduction to the theory of directed graphs*. New York: Wiley, 1965.
- [3] Harary, F. and Prins, G. The number of homeomorphically irreducible trees and other species. *Acta Math.*, 1959, **101**, 141-162.
- [4] Read, R. C. A note on the number of functional digraphs. *Math. Annalen*, 1961, **143**, 109-110.

Manuscript received 6/24/65

Revised manuscript received 2/15/66