Canad. Math. Bull. Vol. 15 (1), 1972

ALMOST ALL GRAPHS HAVE A SPANNING CYCLE

BY
J. W. MOON[1]

In memory of Leo Moser

1. **Introduction.** A *graph* is a collection of *nodes* some pairs of which are joined by a single *edge*. A *k-path*, or a path of length k, is a sequence of nodes $\{p_1, p_2, \ldots, p_{k+1}\}$ such that p_i is joined to p_{i+1} for $1 \le i \le k$; we assume the nodes are distinct except that p_1 and p_{k+1} may be the same in which case we call the path a *k-cycle* or a cycle of length k. (Notice that two nodes joined by an edge determine a 2-cycle according to this definition; it will also be convenient to regard a single node as a 1-cycle.) A *spanning* path or cycle is one that involves every node of the graph. One of the unsolved problems of graph theory is to characterize those graphs that have a spanning path or cycle.

If $0 < p < 1$, let $G(n, p)$ denote a random graph with n nodes in which each of the $\frac{1}{2}n(n-1)$ possible edges is present with probability p. Erdös and Rényi [1] have conjectured that most graphs with n nodes and $n^{1+\epsilon}$ edges contain a spanning cycle. Our object here is to prove the following weaker result.

THEOREM. *If ϵ is any positive constant and $p^2 = (1+\epsilon)(2/n)^{1/2} \log n$, then the probability that the random graph $G(n, p)$ has a spanning cycle tends to one as n tends to infinity.*

2. **Proof of theorem.** Suppose node x does not belong to a given k-cycle C in a random graph $G(n, p)$. If x is joined to two consecutive nodes of C, then x can be inserted between these nodes to form a $(k+1)$-cycle. In this case we shall say we have *extended* the k-cycle C. (Extending a 1-cycle means adjoining a new node that is joined to it.)

If $1 \le k \le n-1$, let $P(n, k)$ denote the probability that a given k-cycle C in a random graph $G(n, p)$ cannot be extended. The probability that a given node x, not in C, cannot be inserted between a given pair of consecutive nodes of C is $1-p^2$. If we only try to insert these $n-k$ nodes x between every other pair of consecutive nodes of C, then the outcomes of these attempts are independent of each other. It follows, therefore, that $P(n, 1) = (1-p)^{n-1}$, $P(n, 2) = (1-p^2)^{n-2}$, and $P(n, k) \le (1-p^2)^{\frac{1}{2}(k-1)(n-k)}$ for $k \ge 3$.

There are $\binom{n}{k} \cdot \frac{1}{2}(k-1)!$ ways to choose k nodes from a graph with n nodes and order them in a cycle if $k \ge 3$; the probability that any such ordering actually

Received by the editors November 24, 1970 and, in revised form, March 23, 1971.

[1] On leave from the University of Alberta.

https://doi.org/10.4153/CMB-1972-008-3 Published online by Cambridge University Press

determines a cycle is p^k. (The corresponding expressions for 1 and 2-cycles are obvious.) If μ denotes the expected number of cycles in $G(n, p)$ that cannot be extended and whose length is at most $n-L$ where $L=[(2n)^{1/2}]$, then

$$\mu = nP(n, 1)+\binom{n}{2}pP(n, 2)+\tfrac{1}{2}\sum_{k=3}^{n-L}\binom{n}{k}(k-1)!\,p^kP(n, k)$$
$$\leq n(1-p^2)^{n-3}+n^2(1-p^2)^{n-3}+\sum_{k=3}^{n-L} n^k(1-p^2)^{\frac{1}{2}(k-1)(n-k)}.$$

Since $p^2=(1+\epsilon)(2/n)^{1/2}\log n$, it follows that

$$1-p^2 \leq n^{-(1+\epsilon)(2/n)^{1/2}}.$$

If we split the sum into two parts, consisting of those terms for which $k \leq L$ and $k > L$, it is not difficult to see that when n is large

$$\mu \leq Ln^L(1-p^2)^{n-3}+n^n(1-p^2)^{\frac{1}{2}L(n-L-1)}$$
$$\leq (2n)^{1/2}n^{-\epsilon(2n)^{1/2}+O(n^{-1/2})}+n^{-\epsilon n+O(n^{1/2})}.$$

This tends to zero as n tends to infinity. Since $G(n, p)$ certainly has some 1-cycles, by definition, it follows that the probability that a random graph $G(n, p)$ has at least one $(n-L)$-cycle tends to one as n tends to infinity.

Now let C denote some $(n-L)$-cycle in a random graph $G(n, p)$. We split this cycle into L subpaths $P_1, P_2, \ldots, P_L$ each of length at least $[(n-L)/L] \geq (1/2n)^{1/2}-2$ in such a way that consecutive nodes of any path P_i are also consecutive nodes of C and only the first and last nodes of any path P_i belong to any other path P_j. Let $q_1, q_2, \ldots, q_L$ denote the nodes of $G(n, p)$ that are not in C. We try to find two consecutive nodes of P_i that are both joined to q_i, for $1 \leq i \leq L$. If, as before, we only try to insert q_i between every other pair of consecutive nodes of P_i we find that the probability that q_i cannot be inserted in P_i is at most $(1-p^2)^{\frac{1}{2}((1/2n)^{1/2}-2)}$. Thus the probability that at least one of the nodes q_i cannot be inserted in its corresponding path is at most

$$L(1-p^2)^{\frac{1}{2}((1/2n)^{1/2}-2)} \leq 2^{1/2}n^{-1/2\epsilon+O(n^{-1/2})}.$$

This also tends to zero as n tends to infinity. It follows, therefore, that the probability that $G(n, p)$ contains an $(n-L)$-cycle that can be successively extended to a spanning cycle tends to one as n tends to infinity. This suffices to complete the proof of the theorem. (This proof can easily be modified to establish analogous results for oriented and directed graphs; the result is undoubtedly valid for considerably smaller values of p.)

ACKNOWLEDGEMENTS. I should like to acknowledge that part of the preceding argument is due to the late Professor Leo Moser; he suggested that one should be able to extend most cycles in most graphs because, roughly speaking, when a cycle is short there are many nodes remaining to try to insert and when it is long there are many places in which to try to insert the remaining nodes. I am also indebted to Dr. Paul Erdös for some helpful remarks.

The preparation of this paper was assisted by a grant from the National Research Council of Canada.

REFERENCE

1. P. Erdös and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1970), 17–61.

UNIVERSITY OF CAPE TOWN,
CAPE PROVINCE, SOUTH AFRICA