THE ASYMPTOTIC NUMBER OF LABELLED WEAKLY-CONNECTED DIGRAPHS WITH A GIVEN NUMBER OF VERTICES AND EDGES

Edward A. Bender[1]

Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA

E. Rodney Canfield[1]

Department of Computer Science, University of Georgia, Athens, GA 30602, USA

Brendan D. McKay[2]

Computer Science Department, Australian National University, GPO Box 4, ACT 2601, Australia

Abstract.

In a recent paper, we determined the asymptotic number of labelled connected graphs with a given number of vertices and edges. In this paper, we apply that result to investigate labelled weakly-connected digraphs. In particular, we determine the asymptotic number of them with accuracy uniform over the full range of possibilities for the number of edges.

1. Introduction.

By a *weakly-connected digraph* we mean a directed graph without loops or multiple edges, such that the underlying undirected graph is connected. This definition does not exclude pairs of directed edges of the form (u, v) and (v, u); we will call these pairs *digons*.

As in [1], let $c(n, q)$ denote the number of labelled connected graphs with n vertices and q edges. Similarly, let $w(n, q)$ be the number of labelled weakly-connected digraphs and let $w(n, q, d)$ be the number of labelled weakly-connected digraphs with n vertices, q edges, and d digons.

The principal result of [1] was an asymptotic estimate of $c(n, q)$. In [2], we used that estimate to investigate some of the properties of random connected graphs. In the present paper, we will use it to find asymptotic estimates of $w(n, q)$ and $w(n, q, d)$, and some of the properties of the associated graphs.

This problem seems to have received little attention in the past. For some early exact enumerations, see [3].

[1] Research supported by the National Security Agency

[2] Research unintelligible to the Australian Security Intelligence Organization

We begin by recounting some of the notation and theorems from [1]. For any n and q, define $N = \binom{n}{2}$, $k = q - n$, and $x = q/n$. Define the function $y = y(x)$ by $y(1) = 0$ and implicitly for $x > 1$ by

$$x = \frac{1}{2y}\log\left(\frac{1+y}{1-y}\right) = 1 + \tfrac{1}{3}y^2 + \tfrac{1}{5}y^4 \cdots .$$

Define the function $\phi(x)$ by $\phi(1) = -1 + \log(2)$ and, for $x > 1$,

$$\phi(x) = \log\left(\frac{2e^{-x}y^{1-x}}{\sqrt{1-y^2}}\right).$$

Also define the function $a(x)$ by $a(1) = 2 + \frac{1}{2}\log(\frac{3}{2})$ and, for $x > 1$,

$$a(x) = x(x+1)(1-y) + \log(1 - x + xy) - \tfrac{1}{2}\log(1 - x + xy^2).$$

Finally, define the numbers $w_0, w_1, w_2, \ldots$ by $w_0 = \pi/\sqrt{6}$ and, for $k > 0$,

$$w_k = \frac{\pi\Gamma(k)d_k\sqrt{8/3}}{\Gamma(3k/2)}\left(\frac{27k}{8e}\right)^{k/2},$$

where

$$d_1 = \tfrac{5}{36}, \text{ and } d_{k+1} = d_k + \sum_{h=1}^{k-1}\frac{d_h d_{k-h}}{(k+1)\binom{k}{h}} \quad \text{for } k > 0.$$

The paper [1] contains a large number of facts about these functions, and we will refer to it as these are needed. The principal result we need from [1] is the following asymptotic estimate of $c(n, q)$.

Theorem 1.1. *For $n \le q \le N$ we have uniformly*

$$c(n,q) = w_k\binom{N}{q}\exp\left(n\phi(x) + a(x) + O\left(\frac{(k+1)^{1/16}}{n^{9/50}}\right)\right). \quad \blacksquare$$

2. Asymptotics for weakly-connected digraphs.

If a weakly-connected digraph has n vertices, q edges and d digons, the underlying undirected graph has n vertices and $q - d$ edges. Considering the number of places that the d digons might occur, and the possible orientations of the other $q - 2d$ edges, we easily have

$$w(n,q,d) = \binom{q-d}{d}2^{q-2d}c(n, q-d). \tag{1}$$

Theorem 2.1. *For $n \le q \le 2N$ we have uniformly*

$$w(n,q) = w_k\binom{2N}{q}\exp\left(n\phi(x) + a(x) - \tfrac{1}{2}x^2(1-y) + O\left(\frac{(k+1)^{1/16}}{n^{9/50}}\right)\right).$$

Proof. Since the total number of labelled digraphs with n vertices and q edges is $\binom{2N}{q}$, the remaining parts of the right side of the theorem can be interpretted as the probability that a randomly chosen digraph in this class is weakly-connected. This observation easily leads to a proof of the theorem for $q > n^{6/5}$, as follows.

If a random choice of q edges from the $2N$ available does not produce a weakly-connected digraph, then at least one of the weak components produced has $n/2$ or fewer vertices. Thus, the probability that the digraph is not weakly-connected is at most equal to the expected number of such components, which in turn is bounded by

$$\begin{aligned}
\sum_{m=1}^{\lfloor n/2 \rfloor} \binom{n}{m} \frac{\binom{2N-2m(n-m)}{q}}{\binom{2N}{q}} &= \sum_{m=1}^{\lfloor n/2 \rfloor} \binom{n}{m} \frac{(2N-2m(n-m))_q}{(2N)_q} \\
&\leq \sum_{m=1}^{\lfloor n/2 \rfloor} \binom{n}{m} \Big(1 - \frac{m}{n}\Big)^q \\
&\leq \sum_{m=1}^{\lfloor n/2 \rfloor} \frac{n^m}{m!} \exp\Big(-\frac{mq}{n}\Big) \\
&\leq \exp\big(n \exp(-n^{1/5})\big) - 1 \\
&= O\big(n \exp(-n^{1/5})\big).
\end{aligned}$$

Furthermore, as in the proof of Lemma 3.3 of [1],

$$\exp\big(n\phi(x) + a(x)\big) = \exp\big(-ne^{-2x} + O(nxe^{-4x}) + O(x^2e^{-2x})\big)$$

as $x \to \infty$ which, along with the facts that $x^2(1-y) = O(x^2e^{-2x})$ and $w_k = 1 + O(1/k)$ [1, (3.7), (3.20)], completes the proof of the theorem for this case.

From now on, we will assume that $n \leq q \leq n^{6/5}$. Suppose initially that d satisfies $0 \leq d \leq k$ and $d^2 = o(k+1)$. Then we find the following uniform estimates.

$$\binom{q-d}{d} = \frac{q^d}{d!} \exp\Big(O\Big(\frac{d^2}{q}\Big)\Big) \tag{2}$$

$$\binom{N}{q-d} = \binom{N}{q} \Big(\frac{2x}{n}\Big)^d \exp\Big(O\Big(\frac{dx}{n} + \frac{d^2}{q}\Big)\Big) \tag{3}$$

$$w_{k-d} = w_k \exp\Big(O\Big(\frac{d}{(k+1)^2}\Big)\Big) \tag{4}$$

$$a(x - d/n) = a(x) + O(d/n^{1/2}) \tag{5}$$

$$n\phi(x - d/n) = n\phi(x) - d\phi'(x) + O\Big(\frac{d^2}{k+1}\Big) \tag{6}$$

For (4), we need the expansion $w_k = \exp\big(-1/(4k+1) + O(1/(k+1)^2)\big)$ implied by equation (3.20) of [1]. Similarly, the proofs of (5) and (6) follow from the estimates of $a'(x)$ and $\phi''(x)$ given in Lemmas 3.1 and 3.2 of [1].

From Theorem 1.1, equations (1)–(6), and the identity $y = e^{-\phi'(x)}$, we have

$$w(n,q,d) = \frac{1}{d!}\binom{N}{q} 2^q w_k (x^2y/2)^d \exp\Big(n\phi(x)+a(x)+O\Big(\frac{d^2}{k+1}+\frac{dx}{n}+\frac{d}{n^{1/2}}+\frac{(k+1)^{1/16}}{n^{9/50}}\Big)\Big), \tag{7}$$

again under the conditions $0 \le d \le k$ and $d^2 = o(k+1)$.

Now assume that $0 \le d \le k-1$. Since w_k, $a(x)$ and the error term of Theorem 1.1 are uniformly bounded, we have uniformly

$$\frac{w(n,q,d+1)}{w(n,q,d)} = O(1)\frac{\binom{q-d-1}{d+1}}{\binom{q-d}{d}}\frac{\binom{N}{q-d-1}}{\binom{N}{q-d}} \exp\big(n\phi(x-d/n-1/n) - n\phi(x-d/n)\big).$$

Since $\phi''(x) < 0$ for $x > 1$, the value of the exponential is less than y. Hence we have uniformly

$$\begin{aligned}\frac{w(n,q,d+1)}{w(n,q,d)} &= O(1)\frac{(q-2d)(q-2d-1)y}{(d+1)(N-q+d+1)}\\ &= O\Big(\frac{x^2y}{d+1}\Big).\end{aligned} \tag{8}$$

Since $c(n,n-1) = O(n^{-3/2})c(n,n)$, we find that (8) holds also for $d = k$. From this we conclude that, for $n \le q \le n^{6/5}$ and $1 \le d_0 \le k+1$,

$$\sum_{d=d_0}^{k+1} w(n,q,d) = O\big(w(n,q,d_0)\big), \tag{9}$$

provided $d_0 > cx^2y$ for some sufficiently large constant c.

We are now ready to estimate $w(n,q)$ by summing $w(n,q,d)$ over d. For $0 \le k \le n^{1/2}$, we have $x^2y = O(k^{1/2}/n^{1/2}) = O(n^{-1/4})$. This means that the sum is dominated by the term for $d = 0$, with the term for $d = 1$ giving the order of magnitude of the error (by (9)). From (7), we immediately have

$$w(n,q) = \binom{N}{q} 2^q w_k \exp\Big(n\phi(x) + a(x) + \tfrac{1}{2}x^2y + O\Big(\frac{(k+1)^{1/16}}{n^{9/50}}\Big)\Big). \tag{10}$$

Suppose instead that $n^{1/2} < k \le n^{6/5} - n$, and define $d_0 = \lceil k^{2/5} \rceil$. Using $y = O(k^{1/2}/n^{1/2})$ for $x \le 2$ and $y < 1$ for $x > 2$, we easily find that $x^2y = o(d_0)$, and so

$$w(n,q) = \sum_{d=0}^{d_0-1} w(n,q,d) + O\big(w(n,q,d_0)\big), \tag{11}$$

where all the terms of the sum lie in the region covered by (7). Using the bounds $d! > (d/e)^d$ and $y = O\big(\min(k^{1/2}/n^{1/2}, 1)\big)$, equation (7) easily implies that

$$w(n,q,d_0) = O\big(\exp(-n^{1/5})\big)w(n,q,0). \tag{12}$$

Using $y = O(\min(k^{1/2}/n^{1/2}, 1))$ and the identities $\sum_{i=0}^{\infty} iz^i/i! = ze^z$ and $\sum_{i=0}^{\infty} i^2 z^i/i! = z(1+z)e^z$, equations (7), (11) and (12) together imply that

$$\begin{aligned}\sum_{d=0}^{d_0-1} \frac{w(n,q,d)}{w(n,q,0)} \exp\Bigl(O\Bigl(\frac{(k+1)^{1/16}}{n^{9/50}}\Bigr)\Bigr) \\ &= \sum_{d=0}^{\infty} \frac{(x^2y/2)^d}{d!}\Bigl(1 + O\Bigl(\frac{d^2}{k+1} + \frac{dx}{n}\Bigr)\Bigr) + O(1)\sum_{d=d_0}^{\infty} \frac{(x^2y/2)^d}{d!} + O\bigl(\exp(-n^{1/5})\bigr) \\ &= \exp\bigl(\tfrac{1}{2}x^2y\bigr)\bigl(1 + O(n^{-2/5})\bigr) + O\bigl(\exp(-n^{1/5})\bigr),\end{aligned}$$

which shows that equation (10) holds for $n^{1/2} < k \le n^{6/5} - n$ also.

To reconcile equation (10) with the theorem statement, it only remains to note that

$$\binom{N}{q}2^q = \binom{2N}{q}\exp\bigl(-\tfrac{1}{2}x^2 + O(x^3/n)\bigr)$$

whenever $x = o(n^{1/3})$. ∎

Theorem 2.1 immediately provides the following two corollaries, for which the calculations are exactly the same as for Corollaries 2 and 3 in [1]. Both expansions can be generalised by using the techniques of [2].

Corollary 2.2. *Uniformly for* $0 \le k \le O(n^{1/2})$, *we have*

$$\begin{aligned}w(n, n+k) = \tfrac{1}{2}w_k(3/\pi)^{1/2}\bigl(e/(12k)\bigr)^{k/2} n^{n+(3k-1)/2} 2^{n+k} \\ \times \bigl(1 + O\bigl(\min(k^{3/2}/n^{1/2}, k^2/n + (k+1)^{1/16}/n^{9/50})\bigr)\bigr).\end{aligned}$$ ∎

Corollary 2.3. *If* $\epsilon > 0$ *is fixed, then*

$$w(n,q) \sim \binom{2N}{q}\exp\bigl(-ne^{-2x}\bigr)$$

uniformly for $q \ge (\frac{1}{4} + \epsilon)n\log n$. ∎

In proving Thereom 2.1, we have incidentally established the distribution of the number of digons in a random weakly-connected digraph with n vertices and q edges. For $n \le q \le n^{6/5}$, this number has a Poisson distribution with mean $\frac{1}{2}x^2y$, to the accuracy given by equation (7). For larger q, since all but a minute fraction of digraphs are weakly-connected, the binomial distribution with probability q^2/n^4 and N degrees of freedom is a more accurate approximation. We leave the details to the reader. Many other asymptotic properties of random weakly-connected digraphs can be established as well, including digraphical equivalents of all those established in [2] for undirected connected graphs.

References.

[1] E. A. Bender, E. R. Canfield and B. D. McKay, The asymptotic number of labeled connected graphs with a given number of vertices and edges, *Random Structures and Algorithms*, **1** (1990) 127–169.

[2] E. A. Bender, E. R. Canfield and B. D. McKay, Asymptotic properties of labeled connected graphs, *Random Structures and Algorithms*, to appear.

[3] F. Harary, The number of linear, directed, rooted and connected graphs, *Trans. Amer. Math. Soc.*, **78** (1955) 445–463.

(Received 21/8/91)