Identity Digraphs of Minimum Size*

Frank Harary
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003-8001

Robert W. Robinson
Department of Computer Science
The University of Georgia
Athens, GA 30602-7404

Abstract

The minimum size $\sigma(n)$ of an identity digraph of any given order n is determined in a way that provides for efficient calculation. Here the *order* is the number n of vertices and the *size* is the number of arcs. An *identity* digraph is one for which the identity is the only automorphism. It is shown that $\sigma(n) = n - \delta(n)$ where $\delta(n)$ is a positive nondecreasing function with growth rate $\Theta(n/\log n)$. The number $\nu(n)$ of nonisomorphic identity digraphs of minimum size and order n is also studied; it takes the value 1 infinitely often but is unbounded.

1 Introduction

The notation and terminology of the books [2,3] are followed. A digraph $D = (V, E)$ is specified by its sets of vertices V and arcs (oriented edges) E. The number of vertices is the *order* of D and the number of arcs is its *size*. If $e = (u, v) \in E$ then e is *incident from* u and *incident to* v. An *automorphism* of D is a pair of permutations, of the vertices and of the arcs, which preserve the incidence to and the incidence from relations.

A digraph for which the only automorphism consists of the identity permutations on its vertices and arcs is called an *identity* digraph. We determine the minimum size $\sigma(n)$ for an identity digraph of order n. It is shown that $\sigma(n) = n - \delta(n)$ where $\delta(n) \in \Theta(n/\log n)$ is a positive, non-decreasing function which can be computed exactly in $O(\log^2 n)$ time and $O(\log n)$ space. Here arithmetic operations on integers are assumed to take constant time, and integers are assumed to take constant space to store.

In [4] $\sigma(n)$ appears as $\imath o(K_n)$, the minimum number of edges which can be oriented in the complete graph K_n on n vertices in order to obtain an identity mixed graph. There it is pointed out that $\sigma(n)$ is the minimum number of arcs in an identity oriented forest. This follows from three obvious facts.

1. Any digraph is an identity digraph if and only if its weak components are different (non-isomorphic) identity digraphs.

*This is a preprint of an article accepted for publication in *Congressus Numerantium* ©(2002) Utilitas Mathematica Publishing, Inc.

2. A weakly connected identity digraph has fewer arcs than vertices if and only if it is an identity oriented tree.

3. There are identity oriented trees of all positive orders, *e.g.*, the directed paths.

It is now clear that $\sigma(n) = n - \delta(n)$ where $\delta(n)$ is the maximum number of different identity oriented trees which can be weak components in a digraph of order n . This depends straightforwardly on the numbers of identity oriented trees. Let $F(n)$ denote the number of different identity oriented trees of order n, and let

$$s(m) = \sum_{k \le m} kF(k),$$

$$d(m) = \sum_{k \le m} F(k),$$

$$p(n) = \max\{m : s(m) \le n\},$$

$$S(n) = s(p(n)),$$

$$D(n) = d(p(n)),$$

$$q(n) = \lfloor (n - S(n))/(p(n) + 1) \rfloor,$$

$$r(n) = (n - S(n)) \bmod (p(n) + 1).$$

Then $\delta(n) = D(n) + q(n)$. The idea is that $p(n)$ is the maximum order through which all of the different identity oriented trees can be included as weak components of a digraph of order n. There are $D(n)$ different identity oriented trees of order at most $p(n)$, and together they contain $S(n)$ vertices. The remaining $n - S(n)$ vertices are sufficient to accomodate $q(n)$ additional identity oriented trees of order $p(n) + 1$. This accounts for $n - r(n)$ vertices in all, where $0 \le r(n) \le p(n)$. If $r(n) > 0$ then any tree of maximum order $m(n)$ ($p(n)$ if $q(n) = 0$, or $p(n) + 1$ if $q(n) > 0$) can be replaced by any identity oriented tree of order $m(n) + r(n)$ so as to arrive at an identity oriented forest of order n having $D(n) + q(n)$ weak components. This achieves the minimum possible size for an identity digraph of order n, although in general there will be other ways in which the minimum can be achieved.

The equations above parallel almost exactly the expression obtained in the undirected case for the minimum size m_p of an identity graph of order p by Quintas [9, Thm.1]. A minor difference is that m_p is undefined if $2 \le p \le 6$. Quintas and others have gone on to prove many results on the minimum size $e(\mathcal{G}, n)$ of an undirected graph of order n which has automorphism graph isomorphic to $\mathcal{G}$ for a finite abstract group $\mathcal{G}$ drawn from various classses of groups. A good survey of these results is contained in [8]. As will be noted in Section 4, the asymptotic results obtained for $\delta(n)$ and $\sigma(n)$ have a close parallel for m_p.

In Section 2 equations are derived for calculating $F(n)$ exactly, and in Section 3 the asymptotic behavior of $F(n)$ is determined. These results allow $\delta(n)$, and hence $\sigma(n)$, to be found exactly and asymptotically. The last section contains observations on the number $\nu(n)$ of different identity digraphs of minimum size and order n. However a complete determination of $\nu(n)$ in the general case is deferred to a later paper.

2 Exact Counting

As is usual for tree counting, rooted trees are considered first. Then unrooted (free) trees are counted with the help of Otter's dissimilarity characteristic equation [10].

Let $R(n)$ denote the number of identity rooted oriented trees of order n, and let

$$R(x) = \sum_{n\geq 1} R(n)x^n$$

be the ordinary generating function. (There will be no occasion to substitute an integer value for the variable x, so this notation should cause no confusion.)

In R. Simion's paper [11], $R(n)$ appears as the number of identity rooted matched trees of order $2n$. A *matched* tree is one which contains a perfect matching (which is necessarily unique). R. Simion gave a natural $1-1$ correspondence between rooted matched trees of order $2n$ and rooted oriented trees of order n [9, Theorem 2.1]. It is easy to see that this correspondence preserves automorphisms and so applies to identity trees of the two types. For the sake of completeness the functional equation satisfied by $R(x)$ and the recurrence relations for calculating $R(n)$ are repeated here.

Theorem 1 (Theorem 1.3(a) of [9])

The ordinary generating function for identity rooted oriented trees satisfies

$$R(x) = x\exp\{2\sum_{i\geq 1}(-1)^{i+1}R(x^i)/i\}. \tag{1}$$

Corollary 1 (Corollary 1.4 of [9])

The numbers of identity rooted oriented trees satisfy

$$R(n) = \frac{2}{n-1}\sum_{1\leq k\leq n} R(k)B(n-k) \tag{2}$$

for $n\geq 2$ *with* $R(1)=1$, *where*

$$B(n) = \sum_{d|n}(-1)^{n/d+1}dR(d) \tag{3}$$

for $n\geq 1$.

Equation (1) can be obtained by viewing an arbitrary identity rooted oriented tree as a single root vertex (represented by the factor of x on the right side) along with a collection of distinct identity rooted oriented trees which are branches at the root each marked "in" or "out". The factor of 2 on the right side of (1) represents the choice of "in" or "out", which corresponds to an arc from the new root to the branch's root or the converse. A very similar functional equation for identity rooted trees can be found in [6] or [5, Section 3.3].

Applying the formal differential operator $x\frac{d}{dx}$ to both sides of (1) gives

$$xR'(x) = R(x) + 2R(x)\sum_{i\geq 1}(-1)^{i+1}x^iR'(x^i) \tag{4}$$

Now (2) and (3) follow by letting $B(n)$ denote the coefficient of x^n in the sum in (1). Then n iterations of equations (2) and (3) provide $R(1),...,R(n)$ and $B(1),...,B(n)$ using $O(n^2)$ arithmetic operations and storing $O(n)$ integers.

When free (unrooted, undirected) trees are considered there is no longer a 1-1 correspondence between oriented trees and matched trees of twice the order. However the usual dissimilarity equation

$$p^* - q^* = \begin{cases} 1 & \text{for an identity oriented tree} \\ 0 & \text{for a non-identity oriented tree} \end{cases} \tag{5}$$

holds, where p^* is the number of different ways to root an oriented tree at a vertex so that the result is asymmetric and q^* is the corresponding number of ways to root at an arc. The form of (5) is simplified by the absence of symmetry arcs, due to the orientations of the arcs.

Summing (5) over all oriented trees of order n with the factor of x^n gives

Theorem 2 *The ordinary generating function $F(x)$ for identity (free) oriented trees satisfies*

$$F(x) = R(x) - R(x)^2 \tag{6}$$

Equation (6) has exactly the same form as the relationship for all oriented trees found in [5, equation (3.3.2)]. Comparing coefficients of x^n on both sides of this equation gives

Corollary 2 *The number $F(n)$ of identity (free) oriented trees satisfies*

$$F(n) = R(n) - \sum_{1\leq k\leq n-1} R(k)R(n-k) \tag{7}$$

for $n \geq 1$.

In conjunction with (2) and (3), (7) provides an algorithm for computing $F(1),\ldots,F(n)$ using $O(n^2)$ arithmetic operations and storing $O(n)$ numbers in the process.

3 Asymptotic Analysis

The asymptotic analysis of $R(n)$ and $F(n)$ follows closely the pattern in [7, Section 4] for the corresponding numbers of undirected identity rooted and free trees. Here we indicate the differences which arise for identity oriented trees.

The power series $R(x)$ and $F(x)$ have the same radius of convergence α which satisfies $R(\alpha) = 1/2$ and

$$\alpha = \frac{1}{2} e^{-1+2\sum_{i\geq 2}(-1)^i R(\alpha^i)/i} \tag{8}$$

Applying (8) iteratively it is found that $\alpha = 0.1905112993053734417244295\cdots$.

Considering them as functions of a complex variable, both $R(x)$ and $F(x)$ have branch points of order 2 at $x = \alpha$. From the expansion at $x = \alpha$ in powers of $(\alpha - x)^{1/2}$ it is found that

$$R(n) = C_R n^{-3/2} \alpha^{-n} (1 + O(\frac{1}{n})) \tag{9}$$

where $C_R = b_1 \sqrt{\alpha/\pi}/2$ and

$$b_1^2 \alpha = \frac{1}{2} - \sum_{i\geq 2} (-1)^i \alpha^i R'(\alpha^i).$$

Computation then gives $C_R = 0.1920662886452003712378791\cdots$. Likewise

$$F(n) = C_F n^{-5/2} \alpha^{-n} (1 + O(\frac{1}{n})) \tag{10}$$

where $C_F = 2b_1^2 \alpha \cdot C_R = 0.1780710391407842464386300\ \cdots$.

Now the ordinary generating functions

$$s(x) = \sum_{m\geq 1} s(m) x^m,$$

$$d(x) = \sum_{m\geq 1} d(m) x^m$$

can be expressed as

$$s(x) = \frac{x}{1-x} F'(x),$$

$$d(x) = \frac{1}{1-x} F(x).$$

The expansion for $F(X)$ at $x = \alpha$ then gives

$$s(m) = \frac{C_F}{1-\alpha} m^{-3/2} \alpha^{-m} (1 + O(\frac{1}{m})), \tag{11}$$

$$d(m) = \frac{C_F}{1-\alpha} m^{-5/2} \alpha^{-m} (1 + O(\frac{1}{m})). \tag{12}$$

Let $\rho = \alpha^{-1} = 5.2490324912281705791649522\cdots$. For $m = p(n)$ we have $s(m) \in \Theta(n)$ so that taking the logarithm on both sides of (11) leads to

$$p(n) = \log_\rho n + O(\log\log n). \tag{13}$$

Since $D(n) = d(p(n))$ and $S(n) = s(p(n))$, equations (11) and (12) imply

$$D(n) = \frac{S(n)}{p(n)}(1 + O(\frac{1}{p(n)})). \tag{14}$$

Then $\sigma(n) = D(n) + q(n)$, and

$$q(n) = \frac{(p(n)+1)q(n)}{p(n)}(1 + O(\frac{1}{p(n)})) \tag{15}$$

is immediate. In view of

$$n = S(n) + (p(n)+1)q(n) + r(n)$$

where $0 \le r(n) \le p(n)$ we have

$$S(n) + (p(n)+1)q(n) = n(1 + O(\frac{p(n)}{n}))$$

so (14) and (15) yield

$$\delta(n) = \frac{n}{p(n)}(1 + O(\frac{p(n)}{n}))(1 + O(\frac{1}{p(n)})),$$

or

$$\delta(n) = \frac{n}{\log_\rho n}(1 + O(\frac{\log\log n}{\log n})) \tag{16}$$

when (13) is applied.

It can now be seen that the exact computation of $S(n)$ can be accomplished in $O(\log^2 n)$ time and $O(\log n)$ space, since the numbers $R(k)$, $F(k)$, $s(k)$ and $d(k)$ need only be calculated for $k \le p(n) + 1$.

As an addendum, the ratio $nF(n)/R(n)$ which gives the proportion of the identity rooted oriented trees of order n which are obtained by rooting identity oriented trees of order n in all possible ways is

$$\frac{nF(n)}{R(n)} = \frac{C_F}{C_R}(1 + O(\frac{1}{n}))$$

where $C_F/C_R = 2b_1^2\alpha = 0.92713323299400434252117830\cdots$.

For large n, then, the probability is just under 7.3% that unrooting an identity rooted oriented tree leaves an oriented tree with a nontrivial symmetry.

4 Related Results

As noted in the introduction, Quintas [9] determined the minimum size m_n of an identity graph of order n. To obtain equations equivalent to his, simply replace $F(k)$ in our equations for $\sigma(n)$ by the number $I(k)$ of identity trees of order k. The asymptotic growth rate of $I(k)$ was determined in [7, Section 4]. The dominant feature is a radus of convergence μ , for which computation gives the reciprocal τ as

$$\tau = \mu^{-1} = 2.51754035263200389079535 46\cdots .$$

Then the analysis leading to (16) converted to undirected graphs gives

$$n - m_n = \frac{n}{\log_\tau n}(1 + O(\frac{\log\log n}{\log n})). \tag{17}$$

The number $\nu(n)$ if different identity oriented trees of order n and minimum size $\sigma(n) = n - s(n)$ takes the value 1 infinitely often and yet is unbounded. To see this, note that for any m and any q with $0 \leq q < F(m+1)$ the minimum size for an identity tree of order $s(m) + (m+1)q$ can only be attained by including all identity trees of order $\leq m$ along with some q of the possible $F(m+1)$ identity trees of order $m+1$. Thus

$$\nu(s(m) + (m+1)q) = \binom{F(m+1)}{q},$$

which as $m \to \infty$ takes the value 1 when $q = 0$ and is unbounded when $q \geq 1$. The function $\nu(n)$ takes much larger values when $r(n) \geq 1$ for large n. It is planned to study this phenomenon in more detail in a future paper.

References

[1] L.V. Ahlfors, *Complex Analysis*, second edition, McGraw-Hill, New York (1966).

[2] G. Chartrand and L. Lesniak, *Graphs & Digraphs*, third edition, Chapman & Hall, New York (1996).

[3] F. Harary, *Graph Theory*, Addison-Wesley, Reading MA (1969).

[4] F. Harary and M.S. Jacobson, Destroying symmetry by orienting edges: Complete graphs and complete bigraphs, *Discussiones Math. Graph Theory* (to appear).

[5] F. Haray and E.M. Palmer, *Graphical Enumeration*, Academic Press, New York (1973).

[6] F. Harary and G. Prins, The number of homeomorphically irreducible trees and other species, *Acta Math.* **101** (1959) 141-162.

[7] F. Harary, R.W. Robinson and A.J. Schwenk, Twenty-step algorithm for determining the asymptotic number of trees of various species, *J. Austral. Math. Soc. Ser. A.* **20** (1975) 483-503; Corrigenda, **41** (1986) 325.

[8] D. J. McCarthy and L. V. Quintas, The construction of minimal-line graphs with given automorphism group, *Ann. N.Y. Acad. Sci.* **328** (1979) 144-156.

[9] L. V. Quintas, Extrema concerning asymmetric graphs, *J. Combin. Theory* **3** (1967) 57-82.

[10] R. Otter, The number of trees, *Ann. of Math.* **49** (1948) 583-599.

[11] R. Simion, Trees with 1-factors and oriented trees, *Discrete Math.* **88** (1991) 93-194.