COUNTING UNLABELED ACYCLIC DIGRAPHS*

R.W. ROBINSON

ABSTRACT

The previously known ways to count acyclic digraphs, both labeled and unlabeled, are reviewed. Then a new method of enumerating unlabeled acyclic digraphs is developed. It involves computing the sum of the cyclic indices of the automorphism groups of the acyclic digraphs, achieving a considerable gain in efficiency through an application of the inclusion-exclusion principle. Numerical results are reported on, and a table of the numbers of unlabeled acyclic digraphs on up to 18 points is included.

1. INTRODUCTION

An *acyclic digraph* is a finite oriented graph with no loops, no multiple arcs, and no directed cycles. It is easy to see that every nonempty acyclic digraph has at least one point of in-degree 0. Such points are called *out-points*. It is the existence of out-points which allows us to bring standard enumeration techniques to bear on the problem of counting acyclic digraphs.

Two acyclic digraphs are *isomorphic* if there is a one-to-one map from the points of one onto the points of the other which preserves the adjacency relation, that is, the arcs. Our object is to find the number A_p of nonisomorphic acyclic digraphs on p points. These are called *unlabeled* acyclic digraphs because there is no structure on the points besides the adjacency relation to be preserved under isomorphisms. A *labeled* acyclic digraph has the additional structure of a linear ordering on the points. In a labeled graph one can associate the numbers $1,2,\ldots,p$ with the points so as to indicate the given linear order. These numbers are the "labels" in the labeling. Equivalently, one can take $\{1,2,\ldots,p\}$ to be the point set of any labeled digraph on p points, and this is the point of view taken in the present paper. We let a_p be the number of different labeled acyclic digraphs on p points.

The four labeled acyclic digraphs of Figure 1 illustrate the relation between labeled and unlabeled digraphs. There are only two unlabeled digraphs among them, since β, γ and δ are isomorphic to each other. Let S_p denote the symmetric group of all p! permutations of the point set $\{1,2,\ldots,p\}$. Any permutation $g \in S_p$ is considered to act on any labeled p-point digraph by relabeling it according to g. This

* The author is grateful to the Australian Research Grants Committee for providing the support necessary for the programming of all the numerical work. This was performed by Dr. Paul Butler. The running times reported were observed on a PDP-11/45 with secondary storage on an RK05 disc.

gives a representation S_p^* of S_p on the p-point acyclic digraphs. For instance, if $g = (1)(2\ 3)$ then $g^*(\alpha) = \alpha$, $g^*(\beta) = \beta$, $g^*(\gamma) = \delta$ and $g^*(\delta) = \gamma$. Since any two p-point acyclic digraphs are isomorphic just if one can be mapped to the other by a member of S_p^*, it is seen that the number of orbits of S_p^* is exactly the number A_p of unlabeled p-point acyclic digraphs.

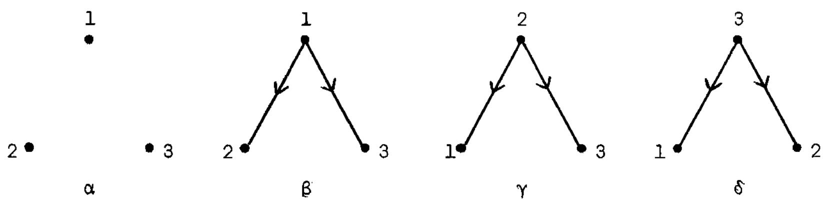

FIGURE 1. Four labeled acyclic digraphs.

The orbit of S_p^* containing a digraph η contains all of the labeled digraphs isomorphic to η. These are called the *labelings* of η. The stabilizer of η in S_p is denoted $\Gamma(\eta)$, and is called the *automorphism group* of η since it consists of all isomorphisms of η onto itself. The order of $\Gamma(\eta)$ determines the number of labelings of η, for the left cosets of $\Gamma(\eta)$ are in an obvious one-to-one correspondence with the members of the orbit of η in S_p^*. Thus the number of labelings of η is $p!/\ \Gamma(\eta)$, the index of $\Gamma(\eta)$ in S_p. For instance, in Figure 1 the digraph α with no arcs has all of S_3 as its stabilizer and so has just one labeled version. The digraph β has a stabilizer of order two, so it has three labeled versions, namely β, γ, and δ.

The key to finding the number of unlabeled graphs or digraphs of any kind is the following lemma [1, Theorem VII on p.191].

LEMMA (Burnside). *If G is a finite group represented as permutations on a finite object set, then*

$$(1) \qquad |\textit{orbits of } G| = \frac{1}{|G|} \sum_{g\in G} |\textit{fixed points of } g|.$$

Here | | denotes "the number of". In many situations, such as Pólya's fundamental *Hauptsatz* [4, p.163], the evaluation of the numbers on the right hand side of (1) can be accomplished directly. An example is the determination of the number g_p of all unlabeled graphs on p points; see the exposition in [2, Section 1 of Chapter 4].

Applying Burnside's lemma to the representation S_p^* of S_p on the labeled p-point acyclic digraphs, we have

$$(2) \qquad A_p = \frac{1}{p!} \sum_{g\in S_p} |\alpha \text{ such that } g^*(\alpha) = \alpha|.$$

The numbers appearing on the right hand side of (2) must be evaluated in a recursive fashion, as we shall see later. For the identity permutation e of S_p, all labeled acyclic digraphs are left fixed by e* and so the required number is a_p.

In the next section we describe briefly the two known methods, from [7] and [8], for computing a_p. These illustrate a simple case of the recursive procedure and the savings to be gained by using the inclusion-exclusion principle. Then in Section 3 the full apparatus of cycle indices is introduced and applied to the computation of A_p. As in the labeled case, the inclusion-exclusion principle is applied in order to eliminate the necessity for keeping a separate accounting of the out-points. This allows considerable reduction of the amount of storage and the number of arithmetic operations required to calculate A_p. The previous method, from [7, §2], is recalled for purposes of comparison in Section 4. Also presented in that section is an alternative approach to counting unlabeled acyclic digraphs by number of out-points, based on the ideas of the previous section. In the closing section are indications of how to include the number of arcs as an enumeration parameter in the counting of unlabeled acyclic digraphs and how to count those which are weakly connected. Some related topics to be developed elsewhere are also mentioned.

2. LABELED ACYCLIC DIGRAPHS

Since every acyclic digraph on $p \geq 1$ points has at least one out-point, all such digraphs can be represented uniquely by a set of k out-points, $1 \leq k \leq p$, together with arcs from the out-points to the acyclic digraph induced by the remaining p-k points. If the number k of out-points and the labeled acyclic digraph δ on the p-k non-out-points are fixed, there are $\binom{p}{k}$ ways to intermingle the labels from these point sets. There are $2^{k(p-k)}$ different sets of arcs from the out-points to δ, but some of these are not admissible because they leave other out-points in addition to the k given ones. If δ has s out-points, each of these must have at least one arc adjacent to it; thus $(2^k-1)^s 2^{k(p-k-s)}$ is the total number of suitable sets of arcs from the out-points to δ.

In Figure 2, one way to extend a particular labeled acyclic digraph on three points by two new out-points is illustrated. There are a total of $\binom{5}{2}(2^2-1)^2 2^2 = 360$ such extensions possible.

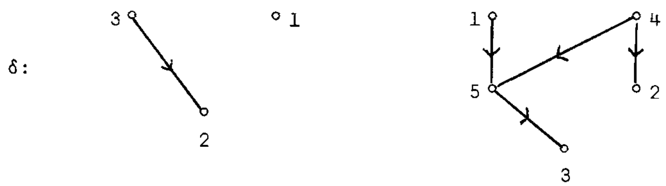

FIGURE 2. A labeled acyclic digraph δ and an extension of δ by two out-points.

Now let $a_{k,m}$ denote the number of different labeled acyclic digraphs on k+m points, exactly k of them out-points. The number of ways of extending with k new out-points a labeled acyclic digraph on m points of which exactly s are out-points is

$$(3) \qquad \binom{k+m}{k}(2^k-1)^s 2^{k(m-s)} a_{s,m-s}$$

by the reasoning above. Summing over s gives $a_{k,m}$ in the form

$$(4) \qquad a_{k,m} = \sum_{0<s\leq m} \binom{k+m}{k}(2^k-1)^s 2^{k(m-s)} a_{s,m-s},$$

which is valid for $k \geq 1$, $m \geq 1$. With the initial conditions $a_{s,0} = 1$ for all $s > 0$, this yields a straightforward method of computing $a_{k,m}$ in terms of the numbers $a_{s,n}$ for $s+n \leq m$. One can find a_p for $p > 0$ by the sum

$$(5) \qquad a_p = \sum_{0<k\leq p} a_{k,p-k},$$

which in conjunction with (4) will require that all the numbers $a_{s,n}$ for $s+n \leq p$ be determined in the process of computing a_p. This rather inefficient method of computing a_p was first discussed in [7, §1].

To obtain a recurrence relation for a_p in terms of the numbers a_k for $k<p$, we make use of the standard inclusion-exclusion principle, as in [3, Section 2 of Chapter 4]. For $1 \leq i \leq p$ we denote by P_i the set of labeled acyclic digraphs on point set $\{1,2,\ldots,p\}$ in which i is an out-point. As $p > 0$, the number of acyclic digraphs on this point set which is in none of the P_i's is 0. By the inclusion-exclusion principle we can express this as the sum

$$(6) \qquad 0 = a_p - \sum_{1\leq i\leq p} |P_i| + \sum_{1\leq i<j\leq p} |P_i \cap P_j| \mp \cdots .$$

In general, if $s \geq 1$ different numbers $\sigma(1),\sigma(2),\ldots,\sigma(s)$ from $\{1,2,\ldots,p\}$ are specified, the cardinality of $P_{\sigma(1)} \cap \cdots \cap P_{\sigma(s)}$ is just $2^{s(p-s)}a_{p-s}$. For we have s points which must be out-points. These s points can only be adjacent to arcs directed toward the remaining p-s points, making a total of $2^{s(p-s)}$ possible configurations of arcs. Independently, the subgraph induced by the p-s other points may be any of the a_{p-s} possible labeled acyclic digraphs. Summing over the $\binom{p}{s}$ different choices of s different points, we obtain

$$(7) \qquad \binom{p}{s} 2^{s(p-s)} a_{p-s}$$

for the s^{th} term in equation (6). This is analogous to (3), but owes its simpler form to the fact that our s points no longer have to constitute all of the out-points.

Combining equations (6) and (7), we find

$$a_p = \sum_{0<s\leq p} (-1)^{s+1}\binom{p}{s}2^{s(p-s)}a_{p-s} \tag{8}$$

for all $p>0$ with the initial condition $a_0 = 1$. This was shown originally in [8, equation (13)] by a somewhat different method. It is clear that in (8) one only needs to know $a_0,a_1,\ldots,a_{p-1}$ to be able to compute a_p. By using a generalized version of the inclusion-exclusion principle, as in [3, Section 3 of Chapter 4], one can obtain an expression for $a_{k,m}$ in terms of $a_0,a_1,\ldots,a_m$. Such an expression was derived in [8, equation (14)].

3. UNLABELED ACYCLIC DIGRAPHS BY NUMBER OF POINTS

The number A_p of unlabeled acyclic digraphs is determined by means of the corollary to Burnside's Lemma, equation (2). This requires that for any $g \in S_p$ we find the number $N(g)$ of labeled acyclic digraphs α on the set $\{1,2,\ldots,p\}$ such that $g^*(\alpha) = \alpha$. In this section the inclusion-exclusion principle will be used to derive a recurrence relation for $N(g)$. As in the labeled case, this is a much more efficient relation than the one obtained in the straightforward way by keeping explicit account of the out-points. The latter will be presented for comparison purposes in the next section.

We may suppose that g is the product of disjoint cycles $\gamma_1,\gamma_2,\ldots,\gamma_m$ of lengths $\ell(1),\ell(2),\ldots,\ell(m)$ respectively. If $g^*(\alpha) = \alpha$, the points in any cycle γ_i for $1\leq i\leq m$ must either all be out-points of α or else all be non-out-points of α. Thus if $p>0$ there is at least one cycle of g which consists of out-points of α. Let Q_i for $1\leq i\leq m$ be the set of labeled acyclic digraphs on $\{1,2,\ldots,p\}$ as point set such that $g^*(\alpha)=\alpha$ and the points of γ_i are all out-points of α. Then the number of labeled acyclic digraphs fixed by g and not in any of the sets $Q_1,Q_2,\ldots,Q_m$ is 0. By the inclusion-exclusion principle, in parallel with (6), we can express this as the sum

$$0 = N(g) - \sum_{1\leq i\leq m} |Q_i| + \sum_{1\leq i<j\leq m} |Q_i \cap Q_j| \mp \cdots .$$

Solving for $N(g)$, one has

$$N(g) = -\sum_{\emptyset\neq I\subseteq\{1,\ldots,m\}} (-1)^{|I|}\Big|\bigcap_{k\in I} Q_k\Big|. \tag{9}$$

Thus for any non-empty subset I of $\{1,\ldots,m\}$ we need to find the number $|\bigcap_{k\in I} Q_k|$ of labeled acyclic digraphs left fixed by g^* for which all the points in cycles γ_k for $k\in I$ are out-points. To start, consider cycles γ_i and γ_j with $i\in I$ and $j\notin I$.

For points x in γ_i and y in γ_j, if the arc $\langle x,y\rangle$ from x to y is in some digraph α left fixed by g^* then α also contains the arcs $\langle g(x),g(y)\rangle, \langle g^2(x),g^2(y)\rangle, \ldots$. In this case these arcs are just $\langle \gamma_i(x),\gamma_j(y)\rangle, \langle \gamma_i^2(x),\gamma_j^2(y)\rangle, \ldots$, and we arrive at $\langle \gamma_i^n(x),\gamma_j^n(y)\rangle = \langle x,y\rangle$ only when the cycle lengths $\ell(i)$ and $\ell(j)$ both divide n. Thus we see that any cycle of arcs from γ_i to γ_j has as its length the least common multiple $[\ell(i),\ell(j)]$ of the point cycle lengths. As there are a total of $\ell(i)\ell(j)$ arcs from γ_i to γ_j, they fall into exactly $(\ell(i),\ell(j))$ different cycles, the number being the greatest common divisor of the point cycle lengths. As each cycle of arcs must either all be contained in α or else all not be in α, this gives a total of $2^{(\ell(i),\ell(j))}$ possibilities for the arcs between γ_i and γ_j. If we multiply together all the independent possibilities for arcs from point cycles which are specified to be out-points to the remaining point cycles of g, we have in all

$$2^{\sum_{i\in I}\sum_{j\notin I}(\ell(i),\ell(j))}$$

different configurations of such arcs which are possible for digraphs fixed by g^*. In addition one can specify independently the subgraph induced by the points of the cycles γ_j for $j\notin I$, which could be any acyclic digraph fixed by $\left(\prod_{j\notin I}\gamma_j\right)^*$. Thus in all there are

$$2^{\sum_{i\in I}\sum_{j\notin I}(\ell(i),\ell(j))}\, N\left(\prod_{j\notin I}\gamma_j\right)$$

different labeled acyclic digraphs fixed by g^* for which all points in the cycles γ_k for $k\in I$ are out-points.

This allows (9) to be put in the more explicit form

$$\text{(10)}\qquad N(g) = -\sum_{\emptyset\neq I\subseteq\{1,\ldots,m\}} (-1)^{|I|}\, 2^{\sum_{i\in I}\sum_{j\notin I}(\ell(i),\ell(j))}\, N\left(\prod_{j\notin I}\gamma_j\right).$$

It is clear by induction on the total number p of points in (10) that N(g) depends only on the lengths of the cycles in the disjoint cycle decomposition of g. Let g have σ_1 cycles of length 1, σ_2 cycles of length 2, and so on up to σ_p cycles of length p. The *cycle type* Z(g) of g is the monomial in the commuting variables $a_1,a_2,a_3,\ldots$ defined by

$$\text{(11)}\qquad Z(g) = \prod_{1\leq i\leq p} a_i^{\sigma_i}.$$

Since the cycle type specifies the number of cycles of the various lengths in the disjoint cycle decomposition of g, we may consider N as a function of cycle types. Thus $N\left(\prod_i a_i^{\sigma_i}\right)$ is the common value of N(g) for all $g\in S_p$ with exactly σ_i cycles of length i for each i. Of course $p=\Sigma i\sigma_i$ is also determined by the cycle type.

To put (10) in terms of cycle types, we group the subsets I according to the numbers τ_i of cycles of length i contained in I. For fixed values of τ_i, there are $\prod_i \binom{\sigma_i}{\tau_i}$ such sets I. In each case $|I| = \sum_i \tau_i$, and one has a total of

$$2^{\Sigma_{i,j}\tau_i(\sigma_j-\tau_j)(i,j)}\quad N\left(\prod_j a_j^{\sigma_j-\tau_j}\right)$$

for the absolute value of the summand. Thus (10) becomes

$$(12)\qquad N\left(\prod_j a_j^{\sigma_j}\right) = -\sum_{0\leq\tau_i\leq\sigma_i}{}' (-1)^{\sum_i\tau_i} N\left(\prod_j a_j^{\sigma_j-\tau_j}\right) 2^{\Sigma_{i,j}\tau_i(\sigma_j-\tau_j)(i,j)} \prod_i \binom{\sigma_i}{\tau_i}$$

where $\sum'$ is a reminder that the term corresponding to $\tau_1 = \tau_2 = \ldots = \tau_p = 0$ is omitted from the summation. We must have the initial value $N(1) = 1$ to deal properly with the case in which all of the points are specified to be out-points, that is when $\tau_i = \sigma_i$ for all i. For there is just one acyclic digraph to be counted in this case, the one with $p = \sum_i i\,\sigma_i$ points and no arcs.

From (2) we have

$$A_p = \frac{1}{p!}\sum_{g\in S_p} N(g)$$

for the number of unlabeled acyclic digraphs on p points. In terms of cycle types, this can be written

$$(13)\qquad A_p = \sum N\left(\prod_i a_i^{\sigma_i}\right) \Big/ \prod_i \sigma_i!\,i^{\sigma_i}$$

where the sum is over all sequences $\sigma_1,\sigma_2,\ldots,\sigma_p$ of non-negative integers such that $p = \sum_i i\sigma_i$. To verify this, we show that there are precisely $p!/\prod_i \sigma_i!\,i^{\sigma_i}$ elements g of S_p with $Z(g) = \prod_i a_i^{\sigma_i}$, for any such sequence. Every g with the given cycle type can be obtained from some assignment of the numbers $\{1,2,\ldots,p\}$ to a fixed array of cycles with σ_i of length i for all $1 \leq i \leq p$. There are p! such assignments, but we must consider two of them equivalent if they represent the same permutation. All equivalent assignments are generated by swapping the entire cycles of length i among themselves, which can be done in $\sigma_i!$ different ways for each i, and for each cycle of length i rotating the assignment by any of the i possible amounts keeping the cyclic order fixed. The latter gives i^{σ_i} different possibilities just from the cycles of length i, so in all we have $\prod_i \sigma_i!\,i^{\sigma_i}$ for the size of each equivalence class, as required.

Since we must determine the numbers $N\left(\prod_i a_i^{\sigma_i}\right)$ recursively, (13) suggests using the generating function Z(A) in $a_1,a_2,a_3,\ldots$ which is given by

$$(14)\qquad Z(A) = \sum_{\sigma_i\geq 0} N\Big(\prod_i a_i^{\sigma_i}\Big)\prod_i a_i^{\sigma_i}\Big/\prod_i \sigma_i!\,i^{\sigma_i}.$$

Then A_p can be found from $Z(A)$ by simply summing the coefficients of all the monomials $\prod_i a_i^{\sigma_i}$ with *weight* p, that is, with $p = \sum_i i\sigma_i$.

If (12) is multiplied by $\prod_j a_j^{\sigma_j}/\prod_j \sigma_j!\,j^{\sigma_j}$ and summed over all monomials of positive weight, one obtains $Z(A) - 1$ on the left. Setting $\nu_j = \sigma_j - \tau_j$ for all $j \geq 0$ on the right, the resulting equation is

$$(15)\qquad Z(A)-1 = -\sum_{\tau_i\geq 0}{}' \sum_{\nu_j\geq 0} \frac{(-1)^{\sum_i \tau_i}\prod_i a_i^{\tau_i}\, N\big(\prod_j a_j^{\nu_j}\big)\prod_j a_j^{\nu_j}\, 2^{\sum_{i,j}\tau_i\nu_j(i,j)}}{\prod_i \tau_i!\,i^{\tau_i}\prod_j \nu_j!\,j^{\nu_j}}.$$

To allow a more compact representation of this relation, we define a product $*$ for monomials by setting

$$\Big(\prod_i a_i^{\tau_i}\Big) * \Big(\prod_j a_j^{\nu_j}\Big) = 2^{\sum_{i,j}\tau_i\nu_j(i,j)}\prod_i a_i^{\tau_i}\prod_j a_j^{\nu_j}.$$

The product is then extended by specifying that it be a bilinear operation on the ring $Q[[a_1,a_2,a_3,\ldots]]$ of formal generating functions. Then the double sum can be separated into a product, giving

$$Z(A)-1 = \left(-\sum_{\tau_i\geq 0}{}' \prod_i \frac{(-a_i/i)^{\tau_i}}{\tau_i!}\right) * \left(\sum_{\nu_j\geq 0} N\Big(\prod_j {}^{\nu_j}\Big)\prod_j \frac{(a_j/j)^{\nu_j}}{\nu_j!}\right).$$

The first factor can be put in the exponential form $1 - e^{-\sum_{i\geq 1} a_i/i}$, and the second is just $Z(A)$ again. So we have the relation

$$(16)\qquad Z(A)-1 = \Big(1 - e^{-\sum_{i\geq 1} a_i/i}\Big) * Z(A).$$

Observing that $1 * Z(A) = Z(A)$, this can be put in the neater form

$$(17)\qquad e^{-\sum_{i\geq 1} a_i/i} * Z(A) = 1.$$

The recursive solution of this relation is more direct from the form (16). For it is apparent that the polynomial $Z_{\leq p}(A)$ consisting of the terms of weight $\leq p$ in $Z(A)$ is all that contributes in the product on the right to the terms $Z_{p+1}(A)$ of weight p+1 on the left. Thus we can start with $Z_0(A) = 1$ and find $Z_p(A)$ for successive values of p. One need only keep $1 - e^{-\sum_{i\geq 1} a_i/i}$ and the running total $(1 - e^{-\sum_{i\geq 1} a_i/i}) * Z_{\leq p}(A)$. From the latter one picks out the terms of weight p+1, which form $Z_{p+1}(A)$. Then advantage is taken of the distributive property of $*$ over + to update the running total by simply adding $(1 - e^{-\sum_{i\geq 1} a_i/i}) * Z_{p+1}(A)$. In practice one settles first on the maximum weight $p^{\#}$ to which it is intended to compute the terms of $Z(A)$. At each stage one keeps only the terms of weight $\leq p^{\#}$. This procedure was followed on a digital computer for $p^{\#} = 18$; the running time was about 7 hours and 1597 terms were required. $Z_{\leq 6}(A)$ is displayed in Table 1, and the coefficient sums A_p are shown for all p up to 18 in Table 2.

TABLE 1. Terms of weight up to 6 in Z(A).

$$
\begin{aligned}
&1 + a_1 + (\tfrac{3}{2}a_1^2 + \tfrac{1}{2}a_2) + (\tfrac{25}{6}a_1^3 + \tfrac{3}{2}a_1a_2 + \tfrac{1}{3}a_3) \\
&+ (\tfrac{181}{8}a_1^4 + \tfrac{25}{4}a_1^2a_2 + \tfrac{7}{8}a_2^2 + a_1a_3 + \tfrac{1}{4}a_4) \\
&+ (\tfrac{29281}{120}a_1^5 + \tfrac{181}{4}a_1^3a_2 + \tfrac{57}{8}a_1a_2^2 + \tfrac{25}{6}a_1^2a_3 + \tfrac{1}{2}a_2a_3 + \tfrac{3}{4}a_1a_4 + \tfrac{1}{5}a_5) \\
&+ (\tfrac{420167}{80}a_1^6 + \tfrac{29281}{48}a_1^4a_2 + \tfrac{1215}{16}a_1^2a_2^2 + \tfrac{289}{48}a_2^3 + \tfrac{181}{6}a_1^3a_3 \\
&\qquad + \tfrac{25}{6}a_1a_2a_3 + \tfrac{5}{6}a_3^2 + \tfrac{25}{8}a_1^2a_4 + \tfrac{7}{8}a_2a_4 + \tfrac{3}{5}a_1a_5 + \tfrac{1}{6}a_6)
\end{aligned}
$$

TABLE 2. The numbers of unlabeled acyclic digraphs on up to 18 points.

A_p	p
1	0
1	1
2	2
6	3
31	4
302	5
5 984	6
243 668	7
20 286 025	8
3 424 938 010	9
1 165 948 612 902	10
797 561 675 349 580	11
1 094 026 876 269 892 596	12
3 005 847 365 735 456 265 830	13
16 530 851 611 091 131 512 031 070	14
181 908 117 707 763 484 218 885 361 402	15
4 004 495 398 476 191 849 391 903 634 065 582	16
176 332 675 845 335 018 307 024 273 011 267 894 000	17
15 530 301 094 140 830 400 618 221 389 766 986 731 287 870	18

We conclude this section by interpreting the factors in equation (17) in terms of cycle indices. If H is any subgroup of S_p, the *cycle index* Z(H) of H is the polynomial given by the average of Z(g) over H, that is

$$Z(H) = \frac{1}{|H|}\sum_{g\in H} Z(g).$$

The terminology is derived from the name *Zyklenzeiger* which Pólya [4] gave to these polynomials. Cycle indices play a central role in the formulation of Pólya's now famous counting theorem. Redfield [5] had earlier made extensive use of cycle indices for enumeration, but he called them *group reduction functions* and his contributions were not recognized until much later.

In justifying (13) we argued that for any given monomial $\prod_i a_i^{\sigma_i}$ with $p = \Sigma i\sigma_i$, there are exactly $p!/\prod_i \sigma_i! i^{\sigma_i}$ elements of Z_p with that cycle type. Thus we can write the cycle index $Z(S_p)$ of the symmetric group S_p as $Z(S_p) = \sum \prod_i a_i^{\sigma_i}/\prod_i \sigma_i! i^{\sigma_i}$, which is a well-known expression for it. The sum is over all sequences $\sigma_1, \sigma_2, \ldots, \sigma_p$ of non-negative integers such that $p = \Sigma i\sigma_i$. Summing over all p, one obtains the exponential form:

$$\sum_{p\geq 0} Z(S_p) = e^{\sum_{i\geq 1} a_i/i} .$$

Thus the first factor in (17) is obtained from the sum of the cycle indices of the symmetric groups by substituting $-a_1$ for a_1, $-a_2$ for a_2, and so on. The usual notation for this is

$$\sum_{p\geq 0} Z(S_p; -a_1, -a_2, \ldots) = e^{-\sum_{i\geq 1} a_i/i}$$

Using a weighted version of Burnside's Lemma, it can be shown that Z(A) is the sum of the cycle indices of the automorphism groups of the unlabeled acyclic digraphs. The idea first appears in Redfield's Decomposition Theorem [5, p.445]. Alternatively, the fact can be drawn as a corollary of a more general result of the author [6, equation (2)]. In Figure 3 the six unlabeled acyclic digraphs on $p = 3$ points are illustrated. Shown with each is the cycle index of its automorphism group. These sum to $\frac{25}{6}a_1^3 + \frac{3}{2}a_1a_2 + \frac{1}{3}a_3$, in agreement with the terms of weight 3 in Z(A) as given in Table 1.

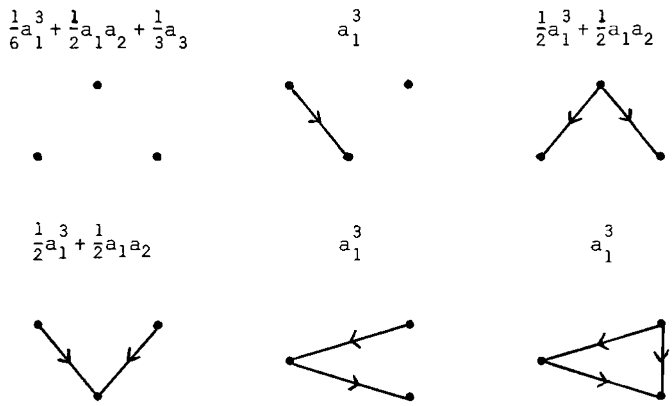

FIGURE 3. The six unlabeled digraphs on three points.

4. UNLABELED ACYCLIC DIGRAPHS BY NUMBER OF OUT-POINTS

First it is indicated how to find the number $A_p^{(k)}$ of unlabeled digraphs on p points with exactly k out-points using the inclusion-exclusion method of the previous section. For purposes of comparison an earlier method [7, §2] is recalled which relies on distinguishing explicitly the contributions of the out-points. The two methods are then contrasted in the light of computational experience.

Just as shown in the previous section for N(g), the number of labeled acyclic digraphs left fixed by g* for some $g \in S_p$ and having exactly k out-points depends only on the cycle type $Z(g) = \prod_i a_i^{\sigma_i}$. Denoting this common number by $N^{(k)} \prod_i a_i^{\sigma_i}$, we have in the same way as (13) from Burnside's lemma

$$A_p^{(k)} = \sum N^{(k)}\Big(\prod_i a_i^{\sigma_i}\Big) \Big/ \prod_i \sigma_i!\, i^{\sigma_i}, \tag{18}$$

the sum being over all monomials of weight p. The analogue of Z(A) for acyclic digraphs with exactly k out-points is then $Z(A^{(k)})$, given by

$$Z(A^{(k)}) = \sum_{\sigma_i \geq 0} N^{(k)}\Big(\prod_i a_i^{\sigma_i}\Big) \prod_i a_i^{\sigma_i} \Big/ \prod_i \sigma_i!\, i^{\sigma_i} \tag{19}$$

In this case the relation which is deduced from the inclusion-exclusion principle is

$$Z(A^{(k)}) = \Big(Z(S_k)\, e^{-\sum_{i\geq 1} a_i/i}\Big) * Z(A). \tag{20}$$

Using the method of the previous section to calculate Z(A), one can thus find the terms of $Z(A^{(k)})$ through any given weight, and hence $A_p^{(k)}$ by summing the coefficients of the terms of weight p. The analogue of (20) for labeled acyclic digraphs appeared in [8, equation (14)].

An approach to finding $A_p^{(k)}$ based on explicitly distinguishing the contributions of out-points from the others in the cycle type was developed in [7, §2] and later received an exposition in [2, Section 8 of Chapter 8]. One way to base this on Burnside's Lemma is to consider S_k as the symmetric group on $\{1,2,\ldots,k\}$ and S_{p-k} as the symmetric group on $\{k+1,k+2,\ldots,p\}$. Then the *product* S_kS_{p-k} is the permutation group which acts on $\{1,2,\ldots,p\}$ in the obvious way and is isomorphic to the direct product of S_k and S_{p-k}; see [2, p.37] for a formal definition. For $gh \in S_kS_{p-k}$ we consider $(gh)^*$ as acting, by relabeling, on the set of labeled acyclic digraphs on $\{1,2,\ldots,p\}$ for which the out-points are precisely $\{1,2,\ldots,k\}$. Then $A_p^{(k)}$ is the number of orbits of this representation $(S_kS_{p-k})^*$ of S_kS_{p-k}. The number of labeled acyclic digraphs left fixed by $(gh)^*$ depends as usual only on the cycle types $Z(g)$ and $Z(h)$. In order to keep these separate, we write $Z(h)$ in terms of the variables $b_1,b_2,b_3,\ldots$ in place of $a_1,a_2,a_3,\ldots$.

One can therefore let $N(\prod_i a_i^{\tau_i} \prod_j b_j^{\sigma_j})$ denote the number of labeled acyclic digraphs left fixed by $(gh)^*$ for any $g \in S_k$ and $h \in S_{p-k}$ such that $Z(g) = \prod_i a_i^{\tau_i}$ and $Z(h) = \prod_j b_j^{\sigma_j}$. Then, much as in (13), Burnside's lemma takes the form

$$A_p^{(k)} = \sum N(\prod_i a_i^{\tau_i} \prod_j b_j^{\sigma_j})/\prod_i \tau_i!\,i^{\tau_i}\prod_j \sigma_j!\,j^{\sigma_j}, \tag{21}$$

where the summation is over all monomials in which $k = \sum_i i\tau_i$ and $p-k = \sum_j j\sigma_j$. The analogue to $Z(A)$ in which cycles of out-points are represented by $a_1,a_2,a_3,\ldots$ and non-out-points are represented by $b_1,b_2,b_3,\ldots$, we denote by $Z_{S,N}(A)$. It is defined by

$$Z_{S,N}(A) = \sum_{\tau_i,\sigma_j \geq 0} N(\prod_i a_i^{\tau_i}\prod_j b_j^{\sigma_j})\prod_i a_i^{\tau_i}\prod_j b_j^{\sigma_j}/\prod_i \tau_i!\,i^{\tau_i}\prod_j \sigma_j!\,j^{\sigma_j}. \tag{22}$$

As before the numbers $N(\prod_i a_i^{\tau_i}\prod_j b_j^{\sigma_j})$ satisfy a recurrence relation which is equivalent to a functional relation satisfied by $Z_{S,N}(A)$. In order to express this compactly, it is again necessary to define a product. For a monomial $\prod_m a_m^{\nu_m}$ in the a_i's and a second monomial $\prod_i a_i^{\tau_i}\prod_j b_j^{\sigma_j}$ in the a_i's and b_i's together, we let

$$(\prod_m a_m^{\nu_m}) \circledast (\prod_i a_i^{\tau_i}\prod_j b_j^{\sigma_j}) = 2^{\sum_{m,j}\nu_m\sigma_j(m,j)} \prod_i (2^{\sum_m \nu_m(m,i)} - 1)^{\tau_i} \prod_m a_m^{\nu_m}\prod_i b_i^{\tau_i}\prod_j b_j^{\sigma_j}.$$

Essentially the integer factor is the number of ways to select arcs from points being permuted according to some g' with $Z(g') = \prod_m a_m^{\nu_m}$ to the points of some acyclic digraph left fixed by (gh) with $Z(gh) = \prod_i a_i^{\tau_i}\prod_i b_i^{\sigma_j}$, in such a way that the points acted on by g' become the only outpoints and the enlarged acyclic digraph is left fixed by (g'gh).

The product $\circledast$ is extended to generating functions by imposing the condition that it be bilinear. Then the relation satisfied by $Z_{S,N}(A)$ is

$$Z_{S,N}(A) - 1 = (-1 + e^{\sum_{i\geq 1} a_i/i}) \circledast Z_{S,N}(A). \tag{23}$$

This should be compared with the relation (16) satisfied by $Z(A)$. It is clear that in (23) the terms of total weight $\leq p$ in $Z_{S,N}(A)$ are the only ones which contribute in the $\circledast$-product to the terms of total weight $p+1$ in $Z_{S,N}(A)$. Thus, starting with 1 for weight 0 one can calculate the terms of successively higher total weights in $Z_{S,N}(A)$. In view of (21) and (22), $A_p^{(k)}$ is found by summing the coefficients of all terms in which the weight of the factors in $a_1, a_2, a_3, \ldots$ is k while the total weight is p.

The disadvantage of (23) compared to (16) for computing A_p is obvious. There will be many more terms of total weight p in $Z_{S,N}(A)$ then in $Z(A)$, due to the distinction made between cycles of out-points and other point cycles. Thus in solving for $Z_{S,N}(A)$ compared to solving for $Z(A)$ there are correspondingly more arithmetic operations to be performed, and hence greater requirements for both storage and time. (For $p \leq 15$, the method based on (23) required 10,840 terms and took about 13 hours while the method based on (16) required only 684 terms and took about 2 hours.) In computing $A_p^{(k)}$ for fixed k, similar advantages in using (16) to determine $Z(A)$ then (20) to find $Z(A^{(k)})$ can be expected, compared to computing all of the relevant terms in $Z_{S,N}(A)$ based on (23).

The situation is reversed if it is desired to compute $A_p^{(k)}$ for all $1 \leq k \leq p$. In terms of storage, there are still savings in the combination of (16) and (20). However the number of arithmetic operations required is now about the same as for (23), and the directness of the latter makes for quicker implementation. (For $p \leq 15$, the method based on (16) and (20) took about 22 hours, compared to about 13 hours for the method based on (23).)

5. EXTENSIONS AND RELATED RESULTS

An obvious extension of the results of the previous sections is to include the number of arcs as an additional enumeration parameter. For each $m \geq 0$ let $N_m(\prod_i a_i^{\sigma_i})$ be the number of labeled acyclic digraphs with exactly m arcs left fixed by g^* for any permutation g such that $Z(g) = \prod_i a_i^{\sigma_i}$. Then the ordinary generating function $N'(\prod_i a_i^{\sigma_i})$ given by

$$N'(\prod_i a_i^{\sigma_i}) = \sum_{m\geq 0} y^m N_m(\prod_i a_i^{\sigma_i})$$

is used to define an extension $Z'(A)$ of $Z(A)$ in which a count is kept of the number of arcs, namely

$$Z'(A) = \sum_{\sigma_i \geq 0} N'(\prod_i a_i^{\sigma_i}) \prod_i a_i^{\sigma_i} \Big/ \prod_i \sigma_i!\, i^{\sigma_i}.$$

The product $*'$ which extends $*$, again by keeping track of the number of arcs, is defined for monomials by

$$(\prod_i a_i^{\tau_i}) *' (y^m \prod_j a_j^{\sigma_j}) = y^m \prod_{i,j} (1+y^{[i,j]})^{\tau_i \sigma_j (i,j)} \prod_i a_i^{\tau_i} \prod_j a_j^{\sigma_j}.$$

The functional relation satisfied by $Z'(A)$ now takes the same form as (17), that is

$$e^{-\sum_{i\geq 1} a_i/i} *' Z'(A) = 1. \tag{24}$$

Similarly one can define $Z'(A^{(k)})$ in the obvious way. It will satisfy

$$Z'(A^{(k)}) = (Z(S_k) e^{-\sum_{i\geq 1} a_i/i}) *' Z'(A) \tag{25}$$

which generalises (20) directly.

To obtain a version of (23) which keeps track of the number of arcs, let $Z'_{S,N}(A)$ denote the generating function in $y, a_1, b_1, a_2, b_2, \ldots$ which extends $Z_{S,N}(A)$ by grouping as the factors of y^m for each m the terms corresponding to acyclic digraphs with m arcs. Likewise the $\circledast$-product must be extended to a product which we denote by $\circledast'$ and which for monomials is defined by

$$(\prod_m a_m^{\nu_m}) \circledast' (y^k \prod_i a_i^{\tau_i} \prod_j a_j^{\sigma_j}) =$$

$$y^k \prod_{m,j} (1+y^{[m,j]})^{\nu_m \sigma_j (m,j)} \prod_i (-1 + \prod_m (1+y^{[m,i]})^{\nu_m (m,i)})^{\tau_i} \prod_m a_m^{\nu_m} \prod_i b_i^{\tau_i} \prod_j b_j^{\sigma_j}.$$

In terms of this product the functional relation satisfied by $Z'_{S,N}(A)$ is

$$Z'_{S,N}(A) - 1 = (-1 + e^{\sum_{i\geq 1} a_i/i}) \circledast' Z'_{S,N}(A). \tag{26}$$

In spite of the similarity between (24), (25), (26) and (17), (20), (23) there will be approximately $\binom{p}{2}$ times as many terms of weight p in the generating functions which distinguish the numbers of arcs as in those that do not incorporate this distinction. The inevitable effect is a dramatic increase in the amount of storage and the number of arithmetic operations required.

A research report containing extensive tables of numerical results on unlabeled acyclic digraphs is in preparation. The combinations of enumeration parameters considered will include the following: points only; points and arcs; points and out-points; points, arcs and out-points. The report will be available on request from the author of the present paper.

A second obvious extension is to count weakly connected unlabeled acyclic digraphs, that is, those which are connected in the ordinary sense after each arc is replaced by an undirected line. This can be done with relative ease, for it does not require a generating function in the point cycle variables $a_1, a_2, a_3, \ldots$ but only in a single variable, say x, to keep track of the total number of points. Thus we can work with the ordinary generating function A(x) given by

$$A(x) \quad = \quad \sum_{p\geq 0} A_p x^p .$$

Let C_p denote the number of different weakly connected unlabeled acyclic digraphs on p points. The ordinary generating function C(x) defined by

$$C(x) \quad = \quad \sum_{p\geq 1} C_p x^p$$

is related to A(x) by the equation

$$A(x) \quad = \quad e^{\sum_{i\geq 1} C(x^i)/i} \tag{27}$$

This is a standard application of Pólya's *Hauptsatz*; see for example [2, Section 2 of Chapter 4]. Given A_p for $p \leq m$, say from (16), it is then straightforward to solve (27) for the numbers C_p in the same range $p \leq m$. One can keep track of the number of arcs or the number of out-points in addition to the total number of points by including a separate variable for the purpose and treating it just like x in generalising (27).

Another extension of the ideas of Section 3 is the enumeration of unlabeled digraphs with given strong components. This will appear in a separate paper. It will be applied to counting unlabeled digraphs which are strongly connected, those which are unilaterally connected, and those which contain a source. The labeled versions of these results are discussed in [8, especially Section 4].

The results of Sections 3 and 4 can also be used to obtain asymptotic expressions for A_p, and $A_p^{(k)}$ for fixed k, as $p \to \infty$. The methods developed in [8, Section 3] for labeled acyclic digraphs can be applied with little difficulty. It is found that for fixed k, $A_p^{(k)}/A_p \to r_k$ where r_k is the same positive constant depending on k as for the labeled case.

REFERENCES

[1] W. Burnside, *Theory of Groups of Finite Order*. 2nd ed., Cambridge Univ. Press, London, 1911. Reprinted by Dover, New York, 1955.

[2] F. Harary and E.M. Palmer, *Graphical Enumeration*. Academic Press, New York, 1973.

[3] C.L. Liu, *Introduction to Combinatorial Mathematics*. McGraw-Hill, New York, 1968.

[4] G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen, *Acta Math.* 68 (1937) 145-254.

[5] J.H. Redfield, The theory of group reduced distributions, *Amer. J. Math.* 49 (1927) 433-455.

[6] R.W. Robinson, Enumeration of nonseparable graphs, *J. Combinatorial Theory* 9 (1970) 327-356.

[7] R.W. Robinson, Enumeration of acyclic digraphs, *Combinatorial Mathematics and Its Applications*. (R.C. Bose *et al.*, eds) Univ. of North Carolina, Chapel Hill (1970) 391-399.

[8] R.W. Robinson, Counting labeled acyclic digraphs, *New Directions in the Theory of Graphs*. (Frank Harary, ed.) Academic Press, New York (1973) 239-273.

Department of Mathematics,
University of Newcastle.