ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Discrete Applied Mathematics 154 (2006) 537–551

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Exact enumeration of acyclic deterministic automata☆

Valery A. Liskovets

Institute of Mathematics, National Academy of Sciences, 11 Surganov str., 220072 Minsk, Belarus

Received 28 May 2004; received in revised form 29 April 2005; accepted 29 June 2005
Available online 20 October 2005

Abstract

A linear recurrence relation is derived for the number of unlabelled initially connected acyclic automata. The coefficients of this relation are determined by another, alternating, recurrence relation. The latter determines, in particular, the number of acyclic automata with labelled states. Certain simple enumerative techniques developed by the author for counting initially connected automata and acyclic digraphs are combined and applied. Calculations show that the results obtained in this paper improve recent upper bounds for the number of minimal deterministic automata (with accepting states) recognizing finite languages. Various related questions are also discussed.

Keywords: Initially connected automaton; Quasi-acyclic automaton; Dead state; Minimal automaton recognizing a finite language; Enumerative injection method

1. Introduction

Recently, Domaratzki [3,4] (see also [5]) obtained some lower and upper bounds for the number of minimal n-state deterministic automata recognizing *finite* languages. In particular, one of the upper bounds is based upon the enumeration of initially connected acyclic automata with numbered states, where the transitions between states are compatible with the state numbers (from lesser to greater). These automata proved to be enumerated by the familiar (unsigned) Genocchi numbers [21, ex.5.8(d)] (close to the Bernoulli numbers) in the case of two input letters and by certain generalized Genocchi numbers for $k > 2$ inputs. The author noted that a better bound should follow from the enumeration of such automata as unlabelled initially connected acyclic automata. It is this problem, natural and interesting by itself, which is solved here. The idea is to combine two approaches which we developed in the 1960s for counting labelled acyclic digraphs [12] and arbitrary initially connected automata [9]. The point is that in the latter case, automata *do not* have non-trivial automorphisms; so that the problems of counting them as having labelled or unlabelled states are equivalent. As an intermediate step, we count labelled acyclic automata and, more generally, quasi-acyclic automata with a given number of absorbing states (see the precise definitions in Section 2).

Numerical calculations suggest that our formulae indeed provide a significantly better upper bound for the number of minimal n-state deterministic automata with accepting states recognizing finite languages (acceptors). This assertion remains, however, unproven since we have not extracted any asymptotics or tight estimates from the formulae obtained.

☆ Partially supported by the Belarusian Republican Fund for Fundamental Research (Grant no. F05-227).
E-mail address: liskov@im.bas-net.by.

0166-218X/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2005.06.009

Nor could we express the results in terms of generating functions. We discuss these and other related questions, including some conjectures and old results, in the second half of the paper.

Initially connected acyclic automata with a unique "pre-dead" state can also be enumerated in a similar way, and these numbers provide a somewhat better upper bound for the numbers of minimal automata.

The present research is motivated by abstract automata theory and is represented in terms of automata. However, our main results can be considered independently of automata theory as the enumeration of some rather natural types of directed graphs, which seem to have not been studied previously. The existence of the obtained simple enumerative formulae does not look a priori evident; nor is their short derivation trivial or well-known.

2. Definitions. Preliminaries

2.1. Initial automata

Generally, for background in automata theory we refer to [7]. For the reader's convenience, together with terms adopted in the present paper we point out some of their synonyms which often appear in the current literature.

Throughout the paper, we consider deterministic *initial finite completely defined* automata without outputs. Thus, an (initial) *automaton* is a quadruple $\mathscr{A} = (Q, q_0, X, \delta)$, where $Q = Q_{\mathscr{A}}$ is the set of states, $q_0 \in Q$ is the initial state, X is the input alphabet and $\delta = \delta_{\mathscr{A}} : (Q, X) \to Q$ is the transition function. The function δ extends naturally to the set X^* of all finite words over X: if $w = x_1 x_2 \ldots x_s$, then $\delta(q, w) := \delta((\delta(\delta(q, x_1), x_2) \ldots), x_s)$. By definition, $\delta(q, \varepsilon) = q$ for any state q, where ε is the *empty* word. If $\delta(q, w) = q'$, we say that the automaton $\mathscr{A}$ goes or passes from the state q to q' under the action of the input word $w \in X^*$ and that q' is *reachable* (accessible) from q in $\mathscr{A}$. The number of states $m = |Q_{\mathscr{A}}|$ is called the *size* of $\mathscr{A}$. Any input letter x determines a mapping δ^x from the set of states to itself, and δ can be identified with the set of mappings $\{\delta^x\}_{x \in X}$. Graphically, an automaton is represented by its *transition diagram* [7, 2.2] (or state transition graph), which is the digraph with Q as the set of nodes and $\{(q, \delta(q, x))\}_{q \in Q, x \in X}$ as the set of directed (X-colored) edges.

Sometimes, we admit *non-initial* automata; these are triples (Q, X, δ) in which no initial state is distinguished.

2.2. Acceptors recognizing languages

An *acceptor* means an automaton with accepting states, i.e., a pair $(\mathscr{A}, F)$, where $\mathscr{A}$ is an automaton and $F \subseteq Q_{\mathscr{A}}$ is a non-empty set of states called *accepting*, or final. The other states are called *non-accepting*. In the literature, acceptors are often called simply automata or recognizers.

Let $L(\mathscr{A}, F)$ denote the language *accepted* (or recognized) by the acceptor $(\mathscr{A}, F)$, i.e., the set of all words under which $\mathscr{A}$ goes from the initial state to an accepting state: $L(\mathscr{A}, F) := \{w \mid w \in X^*, \ \delta(q_0, w) \in F\}$. An acceptor $(\mathscr{A}, F)$ is called *minimal* recognizing a given language $L \subseteq X^*$ if $L = L(\mathscr{A}, F)$ and $\mathscr{A}$ is of minimal size (number of states) among all the acceptors recognizing L.

2.3. Recurrent and transient states

We call a state q of an automaton $\mathscr{A}$ *recurrent* if there exists a non-empty word w which returns $\mathscr{A}$ from q to itself: $\delta(q, w) = q$. Such states are also known as cyclic or looping. Non-recurrent states are called *transient*. Evidently, any (completely defined) finite automaton has recurrent states. Moreover, for any state, there is a recurrent state reachable from it. It follows that finite automata cannot be acyclic in the strict sense of this term; so we have to relax the restrictions.

2.4. Acyclic automata. Dead and pre-dead states

An automaton is called *acyclic* if it has a unique recurrent state. The recurrent state of an acyclic automaton is called its *dead state* (or sink).

It is convenient to distinguish the dead state of acyclic automata, and we will designate it separately by the letter "D" possibly with a subscript. By 2.3, the dead state D of an acyclic automaton is *absorbing* (a trap), i.e., $\delta(D, x) = D$ for any $x \in X$. The dead state is a singular element of an acyclic automaton. From now on, n will denote the number of *transient* states of acyclic automata, so that $n = m - 1$. The transient states are usually labelled by $q_0, q_1, \ldots, q_{n-1}$, where q_0 is the initial state. It is important to stress that we do *not* demand that the transition function δ be compatible with the numbers (or any other order) of the labels $q_1, q_2, \ldots$.

Call a state q of an acyclic automaton *pre-dead* if only the dead state D is reachable from q by all inputs. For $n > 0$, such states always exist: these are the sinks of the partial automaton obtained after the deletion of the dead state and all transitions to it.

2.5. Initially connected automata

A state of an automaton is referred to as a *source* (or maximal) if there are no transitions to it. It is easy to see that any non-empty acyclic automaton has at least one source.

An automaton $\mathscr{A}$ is called *initially connected* if all of its states are reachable from the initial state. An acyclic automaton is initially connected if and only if q_0 is its unique source. In the current literature, initially connected automata are sometimes referred to as accessible or start-useful automata (or automata with no start-useless state).

The transition diagram of an acyclic automaton is an acyclic (multi)digraph excluding the loops in the dead state, and in the case of initially connected acyclic automata, this is an acyclic digraph with a unique sink and a unique source.

2.6. Subautomata

Let $\mathscr{A}$ be an automaton with the set of states Q. If R is a subset of Q and if $\delta(q, x) \in R$ for any $q \in R$ and $x \in X$, then R and the restriction of δ to R form an automaton called a *subautomaton*. In other words, a subautomaton absorbs all transitions: it admits transitions to it from the outside, but all the transitions from it lead again to it. This notion is naturally extended to acceptors: $F \cap R$ serves as the set of accepting states. The following lemma is obvious.

Lemma 2.1. *For any state $q \in Q$, all states reachable from it in an automaton $\mathscr{A}$ form an initially connected subautomaton $\mathscr{A}^{(q)}$ with the initial state q.*

By what is said above it is obvious that $\mathscr{A}^{(q)}$ is the minimal subautomaton containing q. This subautomaton is said to be *generated* by the state q.

The subautomaton $\mathscr{A}^{(q_0)}$ is called the *initially connected component* of $\mathscr{A}$.

By definition, subautomata generated by states satisfy the following heredity property: if q' is reachable from q, then $\mathscr{A}^{(q')}$ is a subautomaton of $\mathscr{A}^{(q)}$ and $\mathscr{A}^{(q)(q')} = \mathscr{A}^{(q')}$.

2.7. Isomorphism

Two automata $\mathscr{A} = (Q, q_0, X, \delta)$ and $\mathscr{A}' = (Q', q_0', X, \delta')$ with the same input alphabet X are called *isomorphic* (by states) if there is a one-to-one correspondence (isomorphism) between their sets of states $\rho : Q' \to Q$ such that $\rho(q_0') = q_0$ and $\rho(\delta'(q', x)) = \delta(\rho(q), x)$ for all states $q' \in Q'$ and all $x \in X$. An isomorphism of acceptors must additionally preserve the property of states to be or not to be accepting.

Isomorphisms from $\mathscr{A}$ to $\mathscr{A}$ are called *automorphisms*. All automorphisms of $\mathscr{A}$ form a group.

Two states q and q' of an automaton $\mathscr{A}$ are called *similar* if the subautomata $\mathscr{A}^{(q)}$ and $\mathscr{A}^{(q')}$ are isomorphic. An automaton $\mathscr{A}$ is referred to as a *primitive* automaton if all its subautomata generated by a single state are pairwise non-similar.

The following assertion is well-known (see, e.g., [9]) and easily provable since any automorphism preserves the initial state and all paths from it:

Lemma 2.2. *The group of automorphisms of an initially connected automaton is trivial.*

2.8. Finite languages and minimal acceptors

Consider an acceptor $(\mathscr{A}, F)$. If there is a recurrent state q in it reachable from the initial state q_0 and an accepting state q' reachable from q, then it is evident that the language $L = L(\mathscr{A}, F)$ accepted by $(\mathscr{A}, F)$ is infinite. Conversely, if $\mathscr{A}$ is acyclic and $D \notin F$, then $L(\mathscr{A}, F)$ is finite. These facts explain a particular interest of researchers to acyclic automata and acceptors, which prove to be efficient tools for representing and processing formal and natural languages; see, in particular, [18,2].

The following important claim is valid (see, e.g., [17]):

Proposition 2.1.

1. *For any finite language L, there exists a unique (up to isomorphism) minimal acceptor $(\mathscr{A}, F) = (\mathscr{A}_L, F_L)$ recognizing it. Moreover*:
2. *$\mathscr{A}_L$ is an initially connected acyclic automaton.*
3. *For any state different from the dead state, there is an accepting state reachable from it.*

The first assertion is a direct corollary of the famous Myhill–Nerode theorem [7]; the second and third assertions are evident. In the literature, automata satisfying properties 2 and 3 are sometimes called stripped or trim, and automata satisfying property 3 are called co-accessible (or having no final-useless state).

In fact, the minimal acceptors are known to be completely characterized by one more property. Call $(\mathscr{A}, F)$ *reduced* if $L_{q'} \neq L_q$ for any two different states q' and q, where L_q denotes the set of all words recognizable by the subautomaton $\mathscr{A}^{(q)}$ (more exactly, by the corresponding acceptor): $L_q := \{w \mid \delta(q, w) \in F\}$. In particular, $L_{q_0} = L$. If $L_{q'} = L_q$, the states q' and q are said to be *equivalent*, and if such $q' \neq q$ exist, the acceptor $(\mathscr{A}, F)$ is called *reducible*.

Lemma 2.3. *An acceptor $(\mathscr{A}, F)$ with the finite language $L = L(\mathscr{A}, F)$ is minimal (for L) if and only if $(\mathscr{A}, F)$ satisfies assertions* 2 *and* 3 *of Proposition* 2.1 *and is reduced.*

Two elementary facts concerning acceptors should also be mentioned: if $L = L(\mathscr{A}, F)$ is finite, then $D \notin F$; $\varepsilon \in L$ if and only if $q_0 \in F$.

2.9. Enumerators

Now, we can obtain some upper bounds for the number $M_k(n)$ of minimal $(n+1)$-state acceptors recognizing finite languages over a k-letter alphabet. Denote by $C_k(n)$ the number of initially connected acyclic automata, counted up to isomorphism (that is, *unlabelled*), with n transient states and k inputs. It is clear from assertion 3 of Proposition 2.1 that in any minimal acceptor $(\mathscr{A}, F)$ recognizing a finite language, F must contain all the pre-dead states. Consequently, in any automaton $\mathscr{A}$ there are no more than 2^{n-1} ways to choose F. Therefore,

$$M_k(n) \leqslant 2^{n-1} C_k(n). \tag{1}$$

Moreover, we can strengthen this bound. As we have just seen, if a minimal acceptor had two or more pre-dead states, then all of them would be accepting. But then they would be equivalent, which is impossible for minimal acceptors by Lemma 2.3. Thus, we obtain the following (cf. [15]).

Corollary 2.1. *Any minimal acceptor recognizing a finite non-empty language has only one pre-dead state q_*. The state q_* is accepting, and it is reachable from any transient state.*

Therefore, to estimate the number of minimal acceptors, we may restrict ourselves to initially connected acyclic automata with a unique pre-dead state. Denoting by $C_k^{(1)}(n)$ their number, we obtain instead of (1) a tighter upper bound:

$$M_k(n) \leqslant 2^{n-1} C_k^{(1)}(n). \tag{2}$$

This inequality, however, does not strengthen (1) very significantly; see Table 5, conjectured formula (15) in Section 7.2 and the discussion therein.

Our main aim is to obtain a formula for the number of unlabelled initially connected acyclic automata $C_k(n)$. To derive it we first count labelled acyclic automata; let $a_k(n)$ denote the number of them with $n+1$ states including D.

2.10. Quasi-acyclic automata

We also need a generalization of acyclic automata called *quasi-acyclic*: these are automata in which all recurrent states are absorbing. This auxiliary class of automata is not very popular in automata theory since an acceptor with more than one absorbing states cannot be minimal. Just as in the case of acyclic automata (see 2.4) we will refer to absorbing states of quasi-acyclic automata as *dead* states.

By $a_k(n,r)$ we will denote the number of quasi-acyclic automata with $r \geqslant 1$ dead states $D_1, D_2, \ldots, D_r$ and n transient labelled states. Thus, $a_k(n,1) = a_k(n)$. It is important that instead of being the dead states, $D_1, D_2, \ldots, D_r$ may form an arbitrary subautomaton: $a_k(n,r)$ counts also the number of all automata with such a fixed absorbing subautomaton (*"black hole"*) and n other, transient, states. Later on, we will make use of this fruitful treatment, in particular in the formula for the number of labelled initially connected acyclic automata $c_k(n)$.

3. Main results

We begin with quasi-acyclic automata, not necessarily initially connected.

Theorem 3.1. $a_k(0,r) = 1$, *and for* $n \geqslant 1$ *the quantity* $a_k(n,r)$ *is determined by the following recursion*:

$$a_k(n,r) = \sum_{t=0}^{n-1} \binom{n}{t} (-1)^{n-t-1} (t+r)^{k(n-t)} a_k(t,r), \quad r \geqslant 1. \tag{3}$$

Proof. We reason as in the case of acyclic digraphs [12]. Consider arbitrary quasi-acyclic automata with k inputs, n (labelled) transient states and r dead states. Let $Y \subseteq Q$ be a set of $n-t$ transient states $(0 \leqslant t \leqslant n)$. Introduce the property Π_Y of an automaton to have Y as a subset of its sources. There are $(t+r)^{k(n-t)} a_k(t,r)$ such automata: we take an arbitrary quasi-acyclic automaton with the set $Q \backslash Y$ of transient states, add Y to it and define the $k(n-t)$ transitions from Y to $(Q \backslash Y) \cup Z$ in an arbitrary way, where $Z = Z_r$ denotes the set of dead states. Now, by the inclusion–exclusion method we can count the number of automata possessing none of these properties, and it should be equated to 0, since any non-empty acyclic automaton possesses a source. Thus, we obtain the formula

$$\sum_{t=0}^{n} \binom{n}{t} (-1)^{n-t} (t+r)^{k(n-t)} a_k(t,r) = 0, \quad n \geqslant 1, \quad r \geqslant 1, \tag{4}$$

which is equivalent to (3). □

Theorem 3.2. $c_k(1) = 1$, *and for* $n > 1$, *the number of labelled initially connected acyclic automata* $c_k(n)$ *is determined by the following recursion*:

$$\sum_{t=1}^{n} \binom{n-1}{t-1} a_k(n-t, t+1) c_k(t) = a_k(n), \tag{5}$$

where $a_k(n) = a_k(n,1)$.

Proof. In [9] (see also [10,11]) we used a simple enumerative method, which we call the "injection method", in order to count arbitrary labelled initially connected automata (see formula (11) below). This method generalizes the well-known method of counting connected graphs of various types and related objects ("exponential structures" by Stanley [21, 5.5]). In practice, it is applicable fruitfully to digraphs possessing a generalized connectivity. Briefly, the idea is to "inject" the (connected) digraphs under consideration $\mathfrak{C}$ into an appropriate class of digraphs $\mathfrak{A}$ in such a way that any graph $\Delta \in \mathfrak{A}$ contains a uniquely determined subgraph Γ (its "connected" component) belonging to $\mathfrak{C}$. And, conversely, we require that the number $\alpha(n,t)$ of graphs $\Delta \in \mathfrak{A}$ with a given component $\Gamma \in \mathfrak{C}$ depend only on the sizes t of Γ and n of Δ (see [13] for a more general and abstract description of this method, which covers Theorem 3.1

as well). If these properties hold, we obtain immediately a linear recurrence relation of form

$$\sum_{t=1}^{n} \binom{n}{t} \alpha(n,t)c(t) = a(n), \tag{6}$$

where $a(n)$ and $c(n)$ stand for the number of graphs with n nodes, resp., in $\mathfrak{A}$ and $\mathfrak{C}$. The factor $\binom{n}{t}$ corresponds to the case when the component can contain any t-element set of nodes; for graphs with a distinguished root this factor is replaced by $\binom{n-1}{t-1}$, and so on. We called $\alpha(n,t)$ the *kernel* of Eq. (6).

In the problem under consideration, $\mathfrak{C}$ is the class of initially connected acyclic automata, and we take the set of acyclic automata as $\mathfrak{A}$. In any acyclic automaton (or, equivalently, its transition diagram) $\varDelta \in \mathfrak{A}$, we select its initially connected component $\varGamma = \varDelta^{(q_0)}$. Now, given an initially connected acyclic component $\varGamma$ with t labelled transient states, we consider the possible acyclic automata $\varDelta$ with n transient states over it. Following the idea formulated in Section 2.10 we may interpret these automata as the quasi-acyclic ones with $t+1$ dead and $n-t$ transient states. Therefore, regardless of a particular choice of $\varGamma$, there are $\alpha(n,t) = a_k(n-t, t+1)$ such $\varDelta$, and the injection method is applicable here. To complete the proof of (5) we need only to add that t states of the component $\varGamma$ including q_0 can be chosen in $\binom{n-1}{t-1}$ ways. $\square$

Now, according to Lemma 2.2, for unlabelled initially connected acyclic automata we have the formula

$$C_k(n) = \frac{c_k(n)}{(n-1)!}. \tag{7}$$

It is interesting to note that we do not know formal, purely *analytical* reasons which would explain why the solution of Eq. (5) is divisible by $(n-1)!$ for any n. The same remark applies also to formulae (7′) and (11).

3.1. Automata with one pre-dead state

Similar arguments can be applied to acyclic automata with a unique pre-dead state.

Consider labelled automata which have q_1 as the pre-dead state. Let $b_k(n,r)$ denote the number of quasi-acyclic automata which have n transient states *different* from q_1, r dead states including D and the property that q_1 is the unique (pre-dead) state such that all transitions from it go to D. Reasoning as in the proof of Theorem 3.1 with $Y \subseteq Q\backslash\{q_1\}$ we obtain the recurrence relation

$$b_k(n,r) = \sum_{t=0}^{n-1} \binom{n}{t} (-1)^{n-t-1}[(t+r+1)^k - 1]^{n-t} b_k(t,r), \quad r \geqslant 1, \tag{3′}$$

which together with the initial conditions $b_k(0,r) = 1$ determines the function $b_k(n,r)$ for all $r \geqslant 1$. In particular, $b_k(n,1) = b_k(n)$ is the number of acyclic automata with q_1 as the unique pre-dead state and n other transient states. The factor $[(t+r+1)^k - 1]^{n-t}$ in Eq. (3′) is the number of admissible transitions from Y, where $|Y| = n-t$, to the other $t+r+1$ states including q_1: for every state $q \in Y$, there is only one inadmissible set of transitions, all to D.

Let $\bar{c}_k(n)$ denote the number of corresponding initially connected automata. Take an acyclic automaton $\varDelta$ with q_1 as the unique pre-dead state. It is initially connected component $\varGamma$ contains q_1, for otherwise a pre-dead state of $\varGamma$ would be another pre-dead state of $\varDelta$. Let $\varGamma$ contain $t \geqslant 0$ other transient states. Then reasoning just as in the proof of Theorem 3.2 we obtain

$$\sum_{t=1}^{n} \binom{n-1}{t-1} b_k(n-t, t+1)\bar{c}_k(t) = b_k(n), \quad n \geqslant 1. \tag{5′}$$

Finally, due to Lemma 2.2 we have the following (cf. (7)).

Theorem 3.3. *The number of unlabelled initially connected acyclic automata with a unique pre-dead state satisfies the following equation*:

$$C_k^{(1)}(n+1) = \frac{\bar{c}_k(n)}{(n-1)!}, \quad n \geqslant 1 \quad \textit{and} \quad C_k^{(1)}(1) = 1, \tag{7'}$$

where $\bar{c}_k(n)$ is determined by formulae (5′) and (3′), and $n+1$ is the number of transient states including *the pre-dead state*.

Remark. There are some reasons to rescale formulae (3′) and (5′) replacing $b_k(n,r)$ and $\bar{c}_k(n)$ by new quantities which are closer to $a_k(n,r)$ and $c_k(n)$, namely, $a_k^{(1)}(n,r) := nb_k(n-1,r)$, the number of labelled quasi-acyclic automata with r dead states, and $c_k^{(1)}(n) := (n-1)\bar{c}_k(n-1)$, $n > 1$, the number of initially connected acyclic automata. In both cases, n is the total number of transient states, one of them is distinguished, and the distinguished state is the only state which has all transitions going to D. The distinguished state is an *arbitrary* state, not necessarily q_1 (but it is clearly different from q_0 in the case of $c_k^{(1)}(n)$).

4. Autonomous automata

Consider the particular case of automata with one input: $k = 1$. Such automata are usually called *autonomous* (or unary). It is evident that autonomous acyclic n-state automata are equinumerous with labelled trees on $n+1$ nodes. So, $a_1(n) = (n+1)^{n-1}$. More generally, quasi-acyclic automata are in one-to-one correspondence with forests of rooted labelled trees, and there are

$$r(n+r)^{n-1} = a_1(n,r) \tag{8}$$

of them with $n+r$ nodes and $r \geqslant 1$ trees, where every dead state serves as the distinguished root of a tree [21, Proposition 5.3.2].

An autonomous acyclic automaton is initially connected if and only if it is a chain starting at q_0 and finishing at D. There are $c_1(n) = (n-1)!$ such labelled chains (hence $C_1(n) = 1$ for all n). Therefore, formula (5) for $k = 1$ turns into the following simple identity:

$$\sum_{t=1}^{n} \binom{n-1}{t-1} (n+1)^{n-t}(t+1)(t-1)! = (n+1)^n, \tag{9}$$

which is equivalent to the familiar Riordan identity [19] (cf. also [10]).

5. Minimal acceptors

The exact enumeration of minimal acceptors recognizing finite languages remains an open problem (cf. [5]). Here, we are interested in the relationship between initially connected acyclic automata and minimal acceptors corresponding to them. We begin with several new (for this paper) definitions.

5.1. Rank and diameter

By the *rank* of a state q of an acyclic automaton we understand the number equal to 1 less than the maximal length of (simple) paths from q to the dead state. For automata with a unique pre-dead state q_*, this is the maximal length of paths (words) leading from q to q_*. In particular, the rank of q_* is 0. States of rank 1 are the states becoming sinks after the deletion of the dead and pre-dead states (“pre-pre-dead”). In the literature, rank is also known under other names such as height or layer.

The maximal rank of states is called the *diameter* of an acyclic automaton. The diameter of an initially connected acyclic automaton is equal to the rank of the initial state, and for the minimal acceptor recognizing a finite language L it is equal to the maximal length of words in L.

5.2. "Useless" automata

There exist initially connected acyclic automata with a unique pre-dead state which cannot become minimal acceptors for *any* choice of the set of accepting states; for instance, such are automata with 3 or more states of rank 1 in which all transitions from them lead to the pre-dead state q_*. Indeed, for any choice of F, at least two states of rank 1 are both accepting or both non-accepting. Consequently, they are equivalent and may be merged together.

More generally, minimal acceptors recognizing finite languages can have no more than $2(2^k - 1)$ states of rank 1. Indeed, all transitions from a state of rank 1 lead to the dead or pre-dead states. Hence there are $2^k - 1$ possible sets of transitions (we must exclude the only case where all transitions lead to D: it would create one more pre-dead state). Now, any such set of transitions may be implemented no more than twice, once in an accepting state and once in a non-accepting state, and the estimate follows by Lemma 2.3.

There are similar constraints, though less restrictive, concerning states of rank 2 or more.

5.3. Primitive automata

At the opposite extreme, there are initially connected acyclic automata with a unique pre-dead state for which any F containing the pre-dead state gives rise to minimal acceptors. Such automata can be easily characterized.

Proposition 5.1. *Let $\mathscr{A}$ be an initially connected acyclic automaton with a unique pre-dead state q_*. Any F containing q_* gives rise to minimal acceptors $(\mathscr{A}, F)$ if and only if $\mathscr{A}$ is primitive.*

Proof. If $\mathscr{A}$ contains two similar states q' and q (see the definition in Section 2.7) then we can easily find a subset F such that the acceptor $(\mathscr{A}, F)$ is reducible. For example, if F contains all transient states both of $\mathscr{A}^{(q')}$ and $\mathscr{A}^{(q)}$, then $L_{q'} = L_q$. By Lemma 2.3, $(\mathscr{A}, F)$ is not minimal.

On the contrary, suppose that $(\mathscr{A}, F)$ is reducible. This means that there are different equivalent states q' and q, i.e., states such that $L_{q'} = L_q$. It is evident that for any F containing q_*, the rank of q is equal to the maximal length of words in the language L_q: the longest path from q to F terminates in q_*. The same is valid for q', therefore q' and q are of the same rank. Take now equivalent q' and q of minimal rank. We have $L_q = L(\mathscr{A}^{(q)}, \bar{F})$ and $L_{q'} = L(\mathscr{A}^{(q')}, \bar{F}')$, where $\bar{F}$ and $\bar{F}'$ denote the sets of all accepting states reachable from q and q', resp. Now, if the acceptors $(\mathscr{A}^{(q)}, \bar{F})$ and $(\mathscr{A}^{(q')}, \bar{F}')$ are reduced (minimal), then by Proposition 2.1(1), they coincide up to isomorphism; so that $\mathscr{A}$ is not primitive.

Suppose conversely that the acceptor $(\mathscr{A}^{(q)}, \bar{F})$ is reducible. Hence there are different states q'' and q''' in it such that $L_{q''} = L_{q'''}$. Now, $L_{q''} = L(\mathscr{A}^{(q)(q'')}, \bar{F}'') = L(\mathscr{A}^{(q'')}, \bar{F}'')$ and similarly for $L_{q'''}$, what means that q'' and q''' are equivalent in $(\mathscr{A}, F)$. But q'' and q''' are of rank smaller than q, which contradicts the choice of q and q'. □

6. Calculations and estimates

6.1. Tables

We restrict our calculations mainly to automata with two inputs. We used Maple in all computations. Tables 1 and 2 contain data for quasi-acyclic, acyclic and initially connected acyclic automata with labelled states.

In Table 3, we give numerical values for unlabelled initially connected acyclic automata and compare them with known lower and upper bounds. Inequality (1) together with a lower bound for $M_2(n)$ obtained in [5] give rise to the inequality

$$(2n-1)!! = 1 \cdot 3 \cdot 5 \cdots (2n-1) \leqslant C_2(n). \tag{10}$$

These numbers, as well as the ratios $C_2(n)/(2n-1)!!$, are also contained in Table 3.

Table 1
The number of labelled quasi-acyclic automata $a_2(n, r)$ with n transient and r dead states

n/r	1	2	3	4	5
0	1	1	1	1	1
1	1	4	9	16	25
2	7	56	207	544	1175
3	142	1780	9342	32848	91150
4	5941	103392	709893	3142528	10682325
5	428856	9649124	82305144	440535696	1775027000
6	47885899	1329514816	13598786979	85529171200	398824865275
7	7685040448	254821480596	3046304952000	22041805076944	116816612731200

$$
\begin{aligned}
a_2(0,r) &= 1.\\
a_2(1,r) &= r^2.\\
a_2(2,r) &= 2r^2+4r^3+r^4.\\
a_2(3,r) &= 21r^2+60r^3+48r^4+12r^5+r^6.\\
a_2(4,r) &= 568r^2+1920r^3+2160r^4+1040r^5+228r^6+24r^7+r^8.\\
a_2(5,r) &= 29705r^2+111400r^3+150400r^4+97160r^5+33190r^6+6280r^7\\
&\quad +680r^8+40r^9+r^{10}.\\
a_2(6,r) &= 2573136r^2+10379520r^3+15778080r^4+12160800r^5+5330520r^6\\
&\quad +1406592r^7+231360r^8+24240r^9+1590r^{10}+60r^{11}+r^{12}.
\end{aligned}
$$

Table 2
The number of labelled acyclic and initially connected acyclic automata

n	$a_2(n)$	$c_2(n)$	$a_2(n)/c_2(n)$
1	1	1	1.000
2	7	3	2.333
3	142	32	4.438
4	5941	762	7.797
5	428856	32712	13.110
6	47885899	2235360	21.422
7	7685040448	224100000	34.293
8	1681740027657	31115906640	54.048
9	482368131521920	5733129144960	84.137
10	175856855224091311	1356239286057600	129.665
11	79512800815739448576	401263604225164800	198.156
12	43701970591391787395197	145349590736723788800	300.668
13	28714779850695689959247872	63331019483788869120000	453.408
14	22239820866807621347245261875	32702367239716877602099200	680.068
15	200060586399267989706814051311616	9760224335684945097034649600	1015.200

$C_2(n)$ are compared with the Genocchi numbers G_{2n} which count, by [4], initially connected acyclic automata in which states are properly ordered. Accordingly, the last column of Table 3 represents

> the average number of numberings (orderings) compatible with the transition functions in initially connected acyclic automata.

Table 4 contains intermediate data for quasi-acyclic automata with a distinguished pre-dead state (formula (3′)). Numerical data for $C_2^{(1)}(n)$ and $C_3^{(1)}(n)$, and their ratios with $C_2(n)$ and $C_3(n)$ are contained in Table 5.

The upper bounds by inequalities (1) and (2) for the number of minimal automata are provided in Table 6; these data are compared with the exact values and bounds published in [5,3].

Table 3
The number of unlabelled initially connected acyclic automata $C_2(n)$

n	I (2n-1)!!	II $C_2(n) = \frac{c_2(n)}{(n-1)!}$	III G_{2n}	II/I	III/II
1	1	1	1	1.000	1.000
2	3	3	3	1.000	1.000
3	15	16	17	1.067	1.063
4	105	127	155	1.210	1.220
5	945	1363	2073	1.442	1.521
6	10395	18628	38227	1.792	2.052
7	135135	311250	929569	2.303	2.987
8	2027025	6173791	28820619	3.046	4.668
9	34459425	142190703	1109652905	4.126	7.804
10	654729075	3737431895	51943281731	5.708	13.898
11	13749310575	110577492346	2905151042481	8.042	26.273
12	316234143225	3641313700916	191329672483963	11.515	52.544
13	7905853580625	132214630355700	14655626154768697	16.724	110.847
14	213458046676875	5251687490704524	1291885088448017715	24.603	245.994
15	6190283353629375	226664506308709858	129848163681107301953	36.616	572.865

Table 4
The number of labelled quasi-acyclic automata $b_2(n, r)$ with a distinguished pre-dead state, $n + 1$ transient and r dead states

n/r	1	2	3	4	5
0	1	1	1	1	1
1	3	8	15	24	35
2	39	176	495	1104	2135
3	1206	7784	29430	84600	204470
4	69189	585408	2791125	9841728	28569765
5	6416568	67481928	389244600	1627740504	5518006200
6	881032059	11111547520	75325337235	364616173440	1413735254155
7	168514815360	2483829653544	19371055651200	106576788695352	465181963908480

6.2. "Cyclic" automata

For comparison and completeness, we also calculate all initially connected automata, not necessarily acyclic. $h_k(n)$ denotes the number of such labelled automata with n states and k inputs. Then $h_k(1) = 1$ and by [9]

$$h_k(n) = n^{kn} - \sum_{t=1}^{n-1} \binom{n-1}{t-1} n^{k(n-t)} h_k(t). \tag{11}$$

Now, $H_k(n) = h_k(n)/(n-1)!$ is the number of unlabelled initially connected automata. Numerical data for them with $k = 2, 3$ are given in Table 7.

As follows easily from formula (11), $H_k(n)$ is divisible by n^k, see [9]. Note, incidentally, that some similar observations are valid for $C_k^{(1)}(n)$; in particular, $(2^k - 1)$ divides $C_k^{(1)}(n)$.

7. Further discussion

7.1. Possible generalizations

Instead of completely defined automata we could consider partial deterministic automata, that is automata for which the transition function is defined not necessarily for all pairs (q, x). In this case, we could exclude the dead state and

Table 5
The number of unlabelled initially connected acyclic automata with a unique pre-dead state $C_k^{(1)}(n)$, $k = 2, 3$

n	I $C_2^{(1)}(n)$	II $C_2(n)$	I/II	III $C_3^{(1)}(n)$	IV $C_3(n)$	III/IV
1	1	1	1.000000	1	1	1.000000
2	3	3	1.000000	7	7	1.000000
3	15	16	0.937500	133	139	0.956835
4	114	127	0.897638	5362	5711	0.938890
5	1191	1363	0.873808	380093	408354	0.930793
6	15993	18628	0.858546	42258384	45605881	0.926599
7	263976	311250	0.848116	6830081860	7390305396	0.924195
8	5189778	6173791	0.840614	1520132414241	1647470410551	0.922707
9	118729335	142190703	0.835001	447309239576913	485292763088275	0.921731
10	3104549229	3737431895	0.830664	…	…	0.921060
20			0.813154			0.919137
40			0.805872			0.918746
60			0.803707			0.918682
80			0.802679			0.918661
100			0.802082			0.918652
150			0.801310			
200			0.800935			
250			0.800715			

Table 6
Upper bounds for the number of minimal acceptors $M_2(n)$

n	I $M_2(n)$: [5]	II $2^{n-1}C_2^{(1)}(n)$: formula (2)	III $2^{n-1}C_2(n)$: formula (1)	IV Upper bound: [3]	II/I	III/I	IV/I
1	1	1	1	1	1.000	1.000	1.000
2	6	6	6	6	1.000	1.000	1.000
3	60	60	64	64	1.000	1.067	1.067
4	900	912	1016	1120	1.013	1.129	1.244
5	18480	19056	21808	26432	1.031	1.180	1.430
6	487560	511776	596096	889216	1.050	1.223	1.824
7		16894464	19920000				
8		664291584	790245248				
9		30394709760	36400819968				

consider genuine acyclic automata. This class does not introduce anything substantially new, since we can transform it bijectively into the class of completely defined automata considered above by adding a new dead state and all undefined transitions as leading to it. If necessary, we could enumerate partial acyclic automata specified additionally by the number of transitions between states (or, equivalently, complete acyclic automata specified by the number of transitions to the dead states).

There is a less trivial generalization of automata under consideration which often appears in the literature; the class of *multi-initial* automata, that is deterministic automata with a distinguished set of initial states. By a slight modification of the proofs given in Section 3, the formulae for *labelled* initial acyclic automata can be generalized to multi-initial as well as to *multi-initially connected* automata (automata in which every state is reachable from an initial state). Note, however, that multi-initially connected automata can have non-trivial automorphisms (preserving the property of states to be initial); so that the enumeration of such unlabelled automata is an additional non-trivial problem.

Table 7
The number of unlabelled initially connected automata $H_k(n),\ k=2,3$

n	$H_2(n)=h_2(n)/(n-1)!$	$H_2(n)/n^2$	$H_3(n)=h_3(n)/(n-1)!$
1	1	1	1
2	12	3	56
3	216	24	7965
4	5248	328	2128064
5	160675	6427	914929500
6	5931540	164765	576689214816
7	256182290	5228210	500750172337212
8	12665445248	197897582	572879126392178688
9	705068085303	8704544263	835007874759393878655
10	43631250229700	436312502297	1510492370204314777345000
11	2970581345516818	24550259053858	3320470273536658970739763334
12	220642839342906336	1532241939881294	8718034433102107344888781813632

7.2. Asymptotics

Asymptotics of $a_k(n)$, $C_k(n)$ and $C_k^{(1)}(n)$ remain open problems. As Table 1 suggests (and as is typical for deterministic automata), only a small fraction of acyclic automata are initially connected. The data in two last columns of Table 3 suggest that $C_2(n)$ is closer to the lower bound. Note that the Genocchi numbers grow much faster than $(2n-1)!!=(2n)!/n!2^n$: asymptotically as $n\to\infty$, $G_{2n}\sim 4(2n)!/\pi^{2n}$.

For arbitrary k, more generally (cf. (10)),

$$\prod_{i=1}^{n}(i^k-(i-1)^k)\leqslant C_k(n), \tag{12}$$

which follows easily from the enumeration of chain-like initially connected acyclic automata. These are $(n+1)$-state automata of diameter $n-1$. It follows that all of them are reduced since there is only one state of each possible rank, whereas similar states of an initially connected acyclic automaton are necessarily of the same rank. Hence by Proposition 5.1,

$$2^{n-1}\prod_{i=1}^{n}(i^k-(i-1)^k)\leqslant M_k(n), \tag{13}$$

a result of [3].

We assume that a significant (i.e., not tending to 0 as n tends to infinity) fraction of initially connected acyclic automata with a unique pre-dead state are primitive. This fraction increases with k but, presumably, it does not tend to 1 as n tends to infinity taking into account the arguments given in Section 5.2: there is a significant fraction of initially connected acyclic automata with three or more states of rank 1, and in a significant fraction of them at least two such states are similar. Thus, these automata give rise to no acceptors (note, incidentally, that pre-dead states are similar; so that an initially connected acyclic automaton with several pre-dead states is not primitive). If this assumption is valid, by Proposition 5.1 we get the following hypothetical relationship (cf. (2)): $M_k(n)=\Theta(2^{n-1}C_k^{(1)}(n))$, $n\to\infty$. Moreover, we conjecture the validity of the following asymptotic formula:

$$M_k(n)\sim\gamma_k\,2^{n-1}C_k^{(1)}(n),\quad n\to\infty, \tag{14}$$

where γ_k is a constant depending on k, $0<\gamma_k<1$ for $k>1$, and $\gamma_k\to 1$ as $k\to\infty$.

The similarity of formulae (3′), (5′) and (7′) to, respectively, (3), (5) and (7) suggests that the numbers $C_k^{(1)}(n)$ should be close to $C_k(n)$ for large n. As extensive calculations show, this is apparently the case; moreover, the fraction of automata with a unique pre-dead state among all initially connected automata decreases monotonically and tends to a

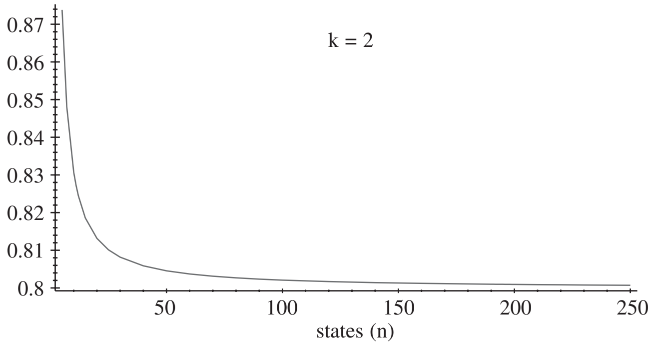

Fig. 1. Fraction of initially connected automata with a unique pre-dead state.

positive limit as n grows. So we conjecture that

$$C_k^{(1)}(n) \sim \beta_k\, C_k(n), \quad n \to \infty, \;\; 0 < \beta_k < 1, \;\; k > 1. \tag{15}$$

From our calculations, we conclude that if (15) is valid, then $\beta_2 \approx 0.800$, $\beta_3 \approx 0.918$, $\beta_4 \approx 0.963$, $\beta_5 \approx 0.982$ and $\beta_6 \approx 0.991$. The corresponding data for $k = 2$ and 3 are represented in Table 5; for $k = 2$ see also Fig. 1.

If both conjectures are valid, then

$$M_k(n) \sim \beta_k\, \gamma_k\, 2^{n-1} C_k(n), \quad n \to \infty. \tag{16}$$

There are other intriguing questions, in particular the distribution of the diameter and the number of pre-dead states in acyclic and initially connected acyclic automata. For comparison, according to McKay [16], the diameter of a random acyclic (labelled) digraph has an asymptotically normal distribution with mean μn, where $\mu \approx 0.764$. In a random acyclic digraph, the mean number of sinks tends to $1.488\ldots$ and the mean number of pre-sink nodes tends to $1.326\ldots$ (see [12,14]). Almost all acyclic digraphs are connected [1].

7.3. Splittable kernels

We return to the general linear recurrence relation of form (6). There is a simple sufficient condition that allows to represent $c(n)$ in a more convenient form. The kernel $\alpha(n, t)$ of (6) is said to be *splittable* if it can be written as the product of single-variable functions of n, t and $n - t$:

$$\alpha(n, t) = f(n)g(t)h(n - t) \tag{17}$$

for all non-negative n, t and $n - t$ (we might consider $\binom{n}{t}$ as a part of the kernel as well, and this factor is clearly splittable). If (17) is valid, then (6) turns into the convolution

$$\sum_{t=1}^{n} \frac{h(n-t)}{(n-t)!} \frac{c(t)g(t)}{t!} = \frac{a(n)}{f(n)n!}, \tag{18}$$

which can be easily represented in terms of appropriate generating functions. Such formulae facilitate extracting asymptotics (see, e.g., [20,12] for the case of acyclic digraphs).

The kernels of recurrence relations for (initially connected) automata are typically unsplittable (unlike the case of general (di)graphs). There is an elementary necessary condition:

Lemma 7.1 (*Liskovets [13]*). *If $\alpha(n, t)$ is splittable, then there exist numbers U and V, not both equal to* 0, *such that for all $n > 2$,*

$$\begin{vmatrix} \alpha(n, n-1) & U\alpha(n-1, n-1) \\ \alpha(n, n-2) & V\alpha(n-1, n-2) \end{vmatrix} = 0. \tag{19}$$

By (19) it is easy to conclude rigorously that the kernel of formula (3) is unsplittable.

7.4. Asymptotics of general initially connected automata

The kernel of formula (11) for all initially connected automata is also unsplittable, and this simple recurrent formula is not very suitable for obtaining asymptotics (numerical experiments show, however, that it is not so bad for *approximate* calculations, contrary to what we expected formerly). For fixed $k > 1$, we managed only to extract the asymptotics $h_k(n) = y_k^{-n} n^{kn+\mathrm{O}(\sqrt{n \log n})}$, where

$$y_k = z_k \mathrm{e}^k (1 - z_k)^{k-1} \tag{20}$$

and z_k is the real root of the equation

$$z\mathrm{e}^{k(1-z)} = 1 \tag{21}$$

different from 1 (thus, $y_2 \approx 1.196$); see [11]. Later on, Korshunov [8] developed a strong technique which enabled him to prove that

$$h_k(n) \sim v_k y_k^{-n} n^{kn+1}, \quad n \to \infty, \tag{22}$$

(where v_k is a complicated constant) and which has nothing to do with the exact enumeration. Hopefully his technique can be modified so as to cover the case of acyclic automata.

Acknowledgements

I am grateful to J. Shallit and M. Ito for attracting my attention to the paper [3]. My thanks go also to Timothy Walsh and the anonymous referees whose valuable remarks helped me to improve the exposition of this article in general, its language and style.

Added in Proof. Recently, Hanna [6] obtained convincing numerical evidences of the validity of various simple formulae for $C_2(n)$. In particular, presumably, $\sum_{n=0}^{\infty} C_2(n) x^n \prod_{i=1}^{n+1} (1 - \mathrm{i}x) = 1$ in formal power series (where $C_2(0) = 1$) and $C_2(n) = T(n, n)$, where $T(n, m) = T(n, m-1) + (m+1)T(n-1, m)$ for $n > m > 0$, $T(n, n) = T(n, n-1)$ and $T(n, 0) = 1$ for $n \geqslant 1$. There are also formulae in terms of Stirling numbers and Dyck paths, and some similar formulae for $C_2^{(1)}(n)$ and $C_k(n)$, $k > 2$.

References

[1] E.A. Bender, R.W. Robinson, The asymptotic number of acyclic digraphs. II, J. Combin. Theory Ser. B 44 (3) (1988) 363–369 MR0941444 (90a:05098).

[2] J. Daciuk, S. Mihov, B.W. Watson, R.E. Watson, Incremental construction of minimal acyclic finite-state automata, Comput. Linguistics 26 (1) (2000) 3–16 MR1783750 (2001d:68123).

[3] M. Domaratzki, Improved bounds on the number of automata accepting finite languages, Internat. J. Found. Comput. Sci. 15 (1) (2004) 143–161 see also in: M. Ito, M. Toyama (Eds.), Developments in Language Theory (DLT 2002), Proceedings of the Sixth International Conference Kyoto, Lecture Notes in Computer Science, vol. 2450, Springer, Berlin, 2003, pp. 209–219, MR2056728 (2005c:68135).

[4] M. Domaratzki, Combinatorial interpretations of a generalization of the Genocchi numbers, J. Integer Seq. 7(3) (2004) Article 04.3.6, 11p, electronic; see also: M. Domaratzki, A generalization of the Genocchi numbers with applications to enumeration of finite automata, School of Computing Technical Report No. 2001-449, Queen's University Kingston, Canada, September 2001, MR2110777 (2005h:11031).

[5] M. Domaratzki, D. Kisman, J. Shallit, On the number of distinct languages accepted by finite automata with n states, J. Autom., Lang. Combin. 7 (4) (2002) 469–486 see also in: J. Dassow, D. Wotschke (Eds.), Descriptional Complexity of Automata, Grammars and Related Structures, 2001, Proceedings of the Third International Workshop, Vienna, 2001, pp. 66–78, MR1990452 (2004i:68113).

[6] P.D. Hanna, personal communication (2005). See also sequences A082161, A082162, A082163 and A102316, in: N.J.A. Sloane (Ed.), The on-line encyclopedia of integer sequences, 2005, published electronically at http://www.research.att.com/~njas/sequences/.

[7] J.E. Hopcroft, J.D. Ullman, Introduction to Automata Theory Languages, and Computation, Addison-Wesley, Reading, MA, 1979 MR0645539 (83j:68002).

[8] A.D. Korshunov, Enumeration of finite automata, Problemy Kibernet. (34) (1978) 5–82, 272 (in Russian), MR0517814 (80a:68050).

[9] V.A. Liskovets, The number of connected initial automata, Kibernetika (Kiev) (3) (1969) 16–19 (in Russian; English translation: Cybernetics 4 (1969) 259–262), MR0302356 (46#1509).

[10] V.A. Liskovets, On a recurrence method of counting graphs with labelled vertices, Dokl. Akad. Nauk SSSR 184 (6) (1969) 1284–1287 (in Russian; English translation: Soviet Math. Dokl. 10(1) (1969) 242–246), MR0244109 (39#5426).

[11] V.A. Liskovets, Some enumerative problems of graph theory, Ph.D. Thesis, Institute of Mathematics, Academy of Sciences of BSSR, Minsk, 1969, 160pp. (in Russian).
[12] V.A. Liskovets, On the number of maximal vertices of a random acyclic digraph, Teor. Verojatnost. i Primenen. 20 (2) (1975) 412–421 (in Russian; English translation: Theory Probab. Appl. 20(2) (1975) 401–409), MR0380925 (52#1822).
[13] V.A. Liskovets, On a general enumeration scheme for labelled graphs, Dokl. Akad. Nauk BSSR 21 (6) (1977) 496–499 (in Russian), MR0505801 (58#21797).
[14] V.A. Liskovets, Unsolved problem 17, in: Combinational and Asymptotic Analysis, No. 2 (coll. papers; Krasn. State University, Krasnoyarsk, 1977), p. 176 (in Russian).
[15] D. Maurel, Pseudo-minimal transducer, Theoret. Comput. Sci. 231 (1) (2000) 129–139 (Implementing automata, London, Ont., Canada, 1997) MR1732692 (2000j:68107).
[16] B.D. McKay, On the shape of a random acyclic digraph, Math. Proc. Cambridge Philos. Soc. 106 (3) (1989) 459–465 MR1010370 (90f:05067).
[17] S. Mihov, Direct construction of minimal acyclic finite states automata, Annuaire Univ. Sofia Fac. Math. Inform. 92 (1–2) (1998) 87–104 MR1781170 (2001j:68072).
[18] D. Revuz, Minimisation of acyclic deterministic automata in linear time, Theoret. Comput. Sci. 92 (1) (1992) 181–189 (Combinatorial Pattern Matching School, Paris, 1990), MR1143138 (92m:68046).
[19] J. Riordan, Enumeration of linear graphs for mappings of finite sets, Ann. Math. Statist. 33 (1962) 178–185 MR0133250 (24#A3084).
[20] R.W. Robinson, Counting labeled acyclic digraphs, in: New directions in the theory of graphs, Proceedings of the Third Ann Arbor Conference, University of Michigan, Ann Arbor, MI, 1971, Academic Press, New York, 1973, pp. 239–273, MR0363994 (51#249).
[21] R.P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, Cambridge, 1999 MR1676282 (2000k:05026).