CROATICA CHEMICA ACTA CCACAA **74** (2) 207–223 (2001)

ISSN-0011-1643
CCA-2726 *Original Scientific Paper*

Two Models for Random Graphs with Bounded Degree

Krystyna T. Balińska,[a] *Michael L. Gargano,*[b] *and Louis V. Quintas*[c,*]

[a] *Computer Science Center, The Technical University of Poznań, pl. M. Sklodowskiej-Curie 5, 60-965 Poznań, Poland*

[b] *Computer Science Department, Pace University, New York, NY 10038, USA*

[c] *Mathematics Department, Pace University, New York, NY 10038, USA*

Received September 16, 1999; revised August 23, 2000; accepted August 24, 2000

Two digraphs both of whose nodes consist of the set of unlabeled graphs of order n having bounded vertex degree equal to f are studied. Adjacency in these digraphs is defined *via* one-edge transformations of the node graphs. Probabilities on the arcs are introduced so that one digraph is a strictly evolving absorbing Markov chain and the other an ergodic Markov chain. Probabilistic and deterministic results and problems concerning these Markov chains are presented. An example of physical interest in these chains is in models where the nodes of the digraphs are identified with chemical species.

Key words: random graph, Markov chain, bounded degree.

1. INTRODUCTION

A graph on n vertices that has no vertex of degree greater than a given positive integer f is called an *f-graph*. The *Random f-Graph Process* (R*f*GP)[1] is a probability model for f-graphs that generalizes the classical random graph process ($f = n - 1$) introduced by P. Erdős and A. Rényi. The R*f*GP is defined as follows.

* Author to whom correspondence should be addressed. (E-mail: lquintas@pace.edu)

Starting with the empty graph G_1 on n labeled vertices obtain a sequence of f-graphs G_1, G_2,... by uniformly selecting an edge in the complement of G_t for each $t \geq 1$ such that its insertion in G_t results in an f-graph G_{t+1}. The process stops when the complement of G_t contains no edges whose insertion in G_t will produce an f-graph.

If the RfGP is modified to a process which starts with an f-graph G_1 on n labeled vertices which is not necessarily the empty graph and a sequence G_1, G_2,... of f-graphs is generated such that at each step $t \geq 1$ an edge is uniformly either inserted in G_t or deleted from G_t to obtain G_{t+1} in the sequence, then a model called the *Reversible Random f-Graph Process* (RRfGP) is obtained.[2] The edges that can be inserted or deleted from G_t are called *admissible edges*.

In Ref. 3 a survey is given of results on random walks on graphs with the transition probability at each vertex v to each neighbor of v uniformly equal to $1/\deg(v)$. We note that our models can be studied as random walks on graphs each of whose vertices is a labeled f-graph. In this case, the degree of a vertex depends on the structure of its associated f-graph, namely, it is equal to the number of admissible edges.

Although the RfGP and the RRfGP are defined in terms of labeled graphs, our interest is in the *transition digraphs* $D(n, f)$ and $R(n, f)$ both of which have as their node set the isomorphism classes of f-graphs on n vertices and each state of the RfGP and RRfGP is represented by an unlabeled f-graph of order n. Let this common node set be denoted $\mathcal{G}(n, f)$ and where no ambiguity is possible, simply as $\mathcal{G}$. The *number of elements* in $\mathcal{G}$ is denoted $N(n, f)$. There exists an arc (G, H) in the transition digraph $R(n, f)$ if and only if an edge when either inserted in or deleted from a labeled copy of G results in a labeled f-graph isomorphic to H. The arc (G, H) is labeled with the transition probability of G going to H. For $D(n, f)$ only insertions are considered. It is easily seen that both $D(n, f)$ and $R(n, f)$ are bipartite.

We shall consider probabilistic and deterministic properties and problems in these two models. In earlier work on the RfGP we obtained results concerning *state probabilities*, namely, the probability of arriving at a particular state after starting at the empty graph, the structure of terminal graphs, the cycle structure of graphs with $f = 2$, the asymptotic degree distribution of evolving f-graphs, node degrees and distance properties.[1,4–6]

Algorithmic methods that were used in the study of the RfGP can be applied to the RRfGP. Since the transition probabilities in the two models are different the results for analogous random variables, in general, are different. However, the distance properties of the digraph $R(n, f)$ are precisely those of the underlying graph of $D(n, f)$. Determining the properties of the node degrees of $D(n, f)$ and $R(n, f)$ is a study in itself.[4]

If the nodes of D(n, f) and R(n, f) are thought of as chemical species, for example, as molecular graphs, then the arcs can be thought of as one-bond transformations between these species. Furthermore, a shortest sequence of arcs between species determines a distance between these which in turn can be used as a measure of similarity.

2. THE MARKOV CHAIN VIEW

From the definitions it is clear that the RfGP and the RRfGP are, respectively, absorbing and ergodic Markov chains[7] and the sequences G_1, G_2,... are random walks in D(n, f) and R(n, f).

Let $\boldsymbol{D}$ denote the transition matrix for D(n, f). The ij-th entry of $\boldsymbol{D}^k = [D_{ij}^{(k)}]$, with $1 \le i, j \le N(n, f)$, is the probability that a random walk starting at state S_i, that is, at a particular node of D(n, f), will be at state S_j after k steps.

In general, for an absorbing chain two basic things to determine are:

(1) The expected number of steps s_i for which a random walk starting at nonabsorbing state S_i will be absorbed, and

(2) The probability B_{ij} that a random walk starting at a nonabsorbing state S_i will end at the absorbing state S_j.

To do this, using linear algebra, we order the states of a Markov chain so that the nonabsorbing states come first and absorbing states (these are the *terminal nodes* of D(n, f)) come last. Then, if v and u denote the number of nonabsorbing and absorbing states, respectively, then the transition matrix has the form

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{Q}_v & \boldsymbol{P} \\ \boldsymbol{0} & \boldsymbol{I}_u \end{bmatrix}$$

where $\boldsymbol{I}_u$ is the identity matrix of order u and $\boldsymbol{0}$ is the zero matrix of size $u \times v$, with $v = N(n, f) - u$.

The matrix $\boldsymbol{M} = (\boldsymbol{I}_v - \boldsymbol{Q}_v)^{-1}$ is called the *fundamental matrix* for the Markov chain. The entries M_{ij} of $\boldsymbol{M}$ are interpreted as the expected number of times that a random walk starting at a nonabsorbing state S_i will be at nonabsorbing state S_j before absorption.

Let $\boldsymbol{s} = [s_i]$, where s_i is the sum of the entries in the i-th row of $\boldsymbol{M}$. Then, s_i is interpreted as the expected number of steps before absorption for a random walk that starts at state S_i (see (1) above). The *absorption probabilities* B_{ij}, by definition, the probability that a random walk starting at

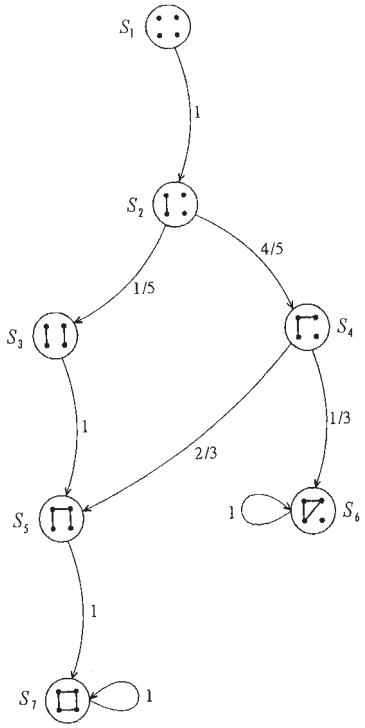

Figure 1. The transition digraph D(4, 2).

nonabsorbing state S_i will terminate at absorbing state S_j, are the entries of $\boldsymbol{B} = \boldsymbol{MP}$ (see (2) above).

For the R*f*GP the initial state S_1 is unique and corresponds to the empty graph. Thus, we focus our attention on s_1 and B_{1j}.

For example, the transition matrix for D(4, 2) (*cf.* Figure 1) is

$$\boldsymbol{D} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & \frac{4}{5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

From this we obtain

$$\boldsymbol{M} = \begin{bmatrix} 1 & 1 & \frac{1}{5} & \frac{4}{5} & \frac{11}{15} \\ 0 & 1 & \frac{1}{5} & \frac{4}{5} & \frac{11}{15} \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{s} = \begin{bmatrix} \frac{56}{15} \\ \frac{41}{15} \\ 2 \\ \frac{5}{3} \\ 1 \end{bmatrix}, \text{ and } \boldsymbol{B} = \begin{bmatrix} \frac{4}{15} & \frac{11}{15} \\ \frac{4}{15} & \frac{11}{15} \\ 0 & 1 \\ \frac{1}{3} & \frac{2}{3} \\ 0 & 1 \end{bmatrix}.$$

Therefore, $s_1 = \frac{56}{15}$ is the expected number of steps for the R2GP of order 4 to reach a terminal state and $B_{1,6} = \frac{4}{15}$, $B_{1,7} = \frac{11}{15}$ respectively, are the state probabilities of the terminal states S_6 and S_7.

The transition digraph D(4, 3) is shown in Figure 2.

Observe that for D(n, $n-1$) there is always exactly one absorbing state and in this case, since every random walk will end up in it with probability 1, the matrix $\boldsymbol{B}$ is always the $(N(n, n-1)-1) \times 1$ matrix with all 1's. Consequently, the interest in the RfGP in this context is in D(n, f) with $f < n-1$, where there is always more than one absorbing state. In Ref. 1 we presented a survey of the results concerning the RfGP known up to 1993 and a list of open problems, some of which are still unsolved. In particular, we asserted that problems concerning the RfGP for both finite n and n going to infinity are of interest. In Ref. 5 algorithms were specifically developed or extended for the study of the RfGP. Some details concerning the content of[1,5] are given here in Section 3.

For the RRfGP let $\boldsymbol{R}$ denote the transition matrix for R(n, f). Note that in D(n, f) all arcs that have nonzero probability D_{ij} go from an f-graph G_i to a supergraph G_j of G_i. The transition probabilities D_{ij} are related to the transition probabilities R_{ij} of R(n, f) by

$$D_{ij} = R_{ij} \Big/ \sum_k R_{ik}$$

where the summation is taken over all k such that G_k is a one-edge extended supergraph of G_i. For an ergodic chain two basic things to determine are the equilibrium vector $\boldsymbol{w}$ and whether the chain is reversible.

Given any ergodic Markov chain, there is a unique probability vector $\boldsymbol{w} = [w_i]$ (sum of the entries w_i is 1) that is fixed by the transition matrix $\boldsymbol{R}$ of the chain. Namely, $\boldsymbol{wR} = \boldsymbol{w}$, or equivalently, $\boldsymbol{R}'\boldsymbol{w}' = \boldsymbol{w}'$, where the superscript $'$ denotes the *transpose* of a matrix. $\boldsymbol{w}$ is called the *equilibrium vector* for the chain. For R(n, f) the i-th component w_i of $\boldsymbol{w} = [w_1\, w_2 \ldots w_{N(n,f)}]$ is the proportion of time in the long run that a random walk will be at state S_i (see

Ref. 7). If a walk has just visited state S_i, then the expected number of steps before it returns to S_i is $1/w_i$ (*e.g.,* see Ref. 3).

An ergodic chain is said to be *reversible*, $w_i R_{ij} = w_j R_{ji}$ for all i, j. If we start a reversible Markov chain with initial probabilities w_i (that is, in equilibrium) and traverse a sequence of states, for example, S_a, S_b, S_c, S_d, then the probability that this sequence of states occurs in this order is $w_a R_{ab} R_{bc} R_{cd}$. The probability that the sequence S_d, S_c, S_b, S_a occurs is $w_d R_{dc} R_{cb} R_{ba}$. For a reversible Markov chain, these two probabilities are equal. This means that when a reversible Markov chain is started in equi-

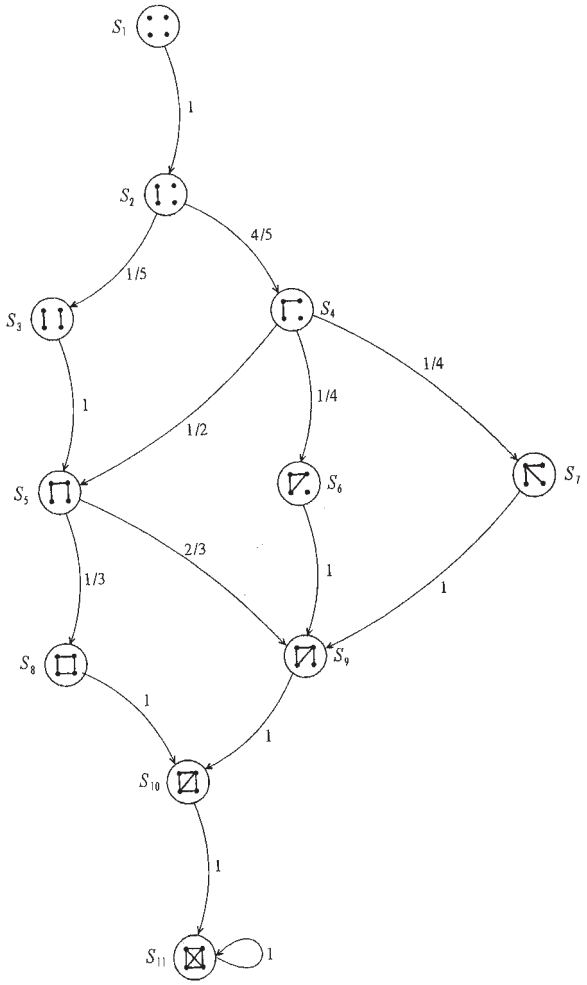

Figure 2. The transition digraph D(4, 3).

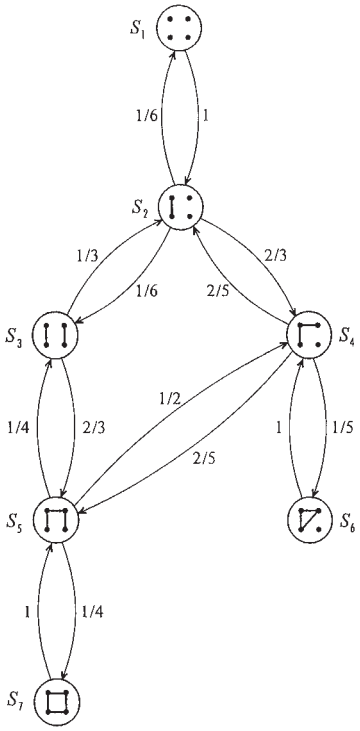

Figure 3. The transition digraph R(4, 2).

librium, probabilities for sequences of states in one order are the same as those with the order reversed. Thus, the RR*f*GP is reversible, if $\boldsymbol{WR}$ is symmetric, where $\boldsymbol{W} = [W_{ij}]$, the $N(n, f) \times N(n, f)$ *m*atrix, with $W_{ij} = w_i$ when $i = j$ and 0 otherwise.

For example, the transition matrix for R(4, 2) (*cf.* Figure 3) is

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{6} & 0 & \frac{1}{6} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 & 0 \\ 0 & \frac{2}{5} & 0 & 0 & \frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The solution to $\boldsymbol{wR} = \boldsymbol{w}$ is $[a\ 6a\ 3a\ 10a\ 8a\ 2a\ 2a]$.

We set $a = 1$ and then normalize to get the desired unique probability vector $\boldsymbol{w}$. That is, since $1 + 6 + 3 + 10 + 8 + 2 + 2 = 32$, we have

$$\boldsymbol{w} = \frac{1}{32}[1\ 6\ 3\ 10\ 8\ 2\ 2] = \left[\tfrac{1}{32}\ \tfrac{6}{32}\ \tfrac{3}{32}\ \tfrac{10}{32}\ \tfrac{8}{32}\ \tfrac{2}{32}\ \tfrac{2}{32}\right].$$

If $\boldsymbol{R}$ is the transition matrix for R(4, 3) (*cf.* Figure 4), then

$$\boldsymbol{w} = \left[\tfrac{1}{64}\ \tfrac{6}{64}\ \tfrac{3}{64}\ \tfrac{12}{64}\ \tfrac{12}{64}\ \tfrac{4}{64}\ \tfrac{4}{64}\ \tfrac{3}{64}\ \tfrac{12}{64}\ \tfrac{6}{64}\ \tfrac{1}{64}\right].$$

It is straightfoward to check that both of the above processes are reversible as ergodic chains.

A combinatorial approach to the study of the RR*f*GP using the transition digraph with labeled *f*-graphs as its nodes yields the following.

THEOREM 2.1. If w_{G} is the component of $\boldsymbol{w}$ corresponding to the unlabeled *f*-graph G in the RR*f*GP of order n, then

$$w_{\mathrm{G}} = \frac{F(\mathrm{G})}{\sum_{\mathrm{H}\in\mathcal{G}} F(\mathrm{H})}$$

where $F(\mathrm{G}) = c_{\mathrm{G}}\alpha_{\mathrm{G}}$, with c_{G} the number of labeled copies of G, α_{G} the number of admissible edges for G, and $\mathcal{G}$ the set of all unlabeled *f*-graphs of order n.

PROOF. A random walk in the RR*f*GP can be thought of as a random walk in $\tilde{\mathrm{R}}(n, f)$, the digraph whose node set is the set of labeled *f*-graphs of order n with adjacency defined as in $\mathrm{R}(n, f)$. In $\tilde{\mathrm{R}}(n, f)$ let a random walk start at node G_j (the j-th labeled copy of an unlabeled *f*-graph G) and then let an admissible edge $\{a, b\}$ of G_j be uniformly chosen and either inserted or deleted from G_j to obtain the next labeled *f*-graph H_k in the random walk. Thus, the walk evolves as follows:

$$\mathrm{G}_j \to (\mathrm{G}_j, \{a, b\}) = \mathrm{H}_k \to (\mathrm{H}_k, \{c, d\}) = \ldots$$

Let

$\mathrm{X} = \{(\mathrm{L}_i, \{a, b\})$: L_i is a labeled *f*-graph of order n and $\{a, b\}$ is an admissible edge of $\mathrm{L}_i\}$.

Then, $|\mathrm{X}| = \sum_{\mathrm{H}\in\mathcal{G}} c_{\mathrm{H}}\alpha_{\mathrm{H}}$, with $c_{\mathrm{H}} = n!/|Aut(\mathrm{H})|$ the number of labeled copies of H, α_{H} the number of admissible edges for H, and $\mathcal{G}$ the set of all unlabeled *f*-graphs of order n.

If a random walk starting at a uniformly chosen labeled graph G_j proceeds through $\tilde{\mathrm{R}}(n, f)$ as described above for a sufficiently large number of

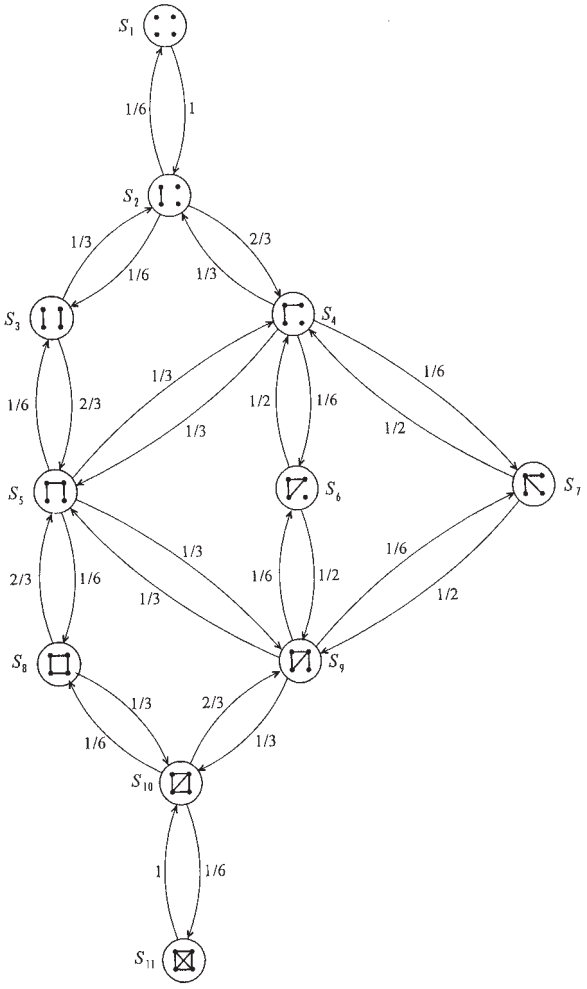

Figure 4. The transition digraph R(4, 3).

steps tending to infinity, then the proportion of time in the long run that the walk passes through a graph isomorphic to a given graph G will be $P(G) = c_G\alpha_G / \sum_{H\in\mathcal{G}} c_H\alpha_H$. But, $P(G)$ is precisely the definition of w_G. ■

COROLLARY 2.2. If $f = n - 1$, then $w_G = \dfrac{c_G}{2^{\binom{n}{2}}}$.

PROOF. $\alpha_H = \dbinom{n}{2}$ for all H and $\sum_{H\in\mathcal{G}} c_H = 2^{\binom{n}{2}}$. ■

COROLLARY 2.3. If $f = n - 1$, then $w_G = w_{G^c}$, where G^c is the complement of G.

PROOF. Since, for any G, $c_G = n!/|Aut(G)|$, we have $c_G = n!/|Aut(G)| = n!/|Aut(G^c)| = c_{G^c}$, so that $w_G = w_{G^c}$. ■

A graph G such that $|Aut(G)| = 1$, is called an *identity graph*.

COROLLARY 2.4. If $f = n - 1$ and $n \geq 6$, the maximum value of a component of $\boldsymbol{w}$ is $n!/2^{\binom{n}{2}}$ and this is realized for each identity graph G. ■

COROLLARY 2.5. If $f = n - 1$, the minimum value of a component of $\boldsymbol{w}$ is $1/2^{\binom{n}{2}}$ and this value is uniquely realized by K_n^c and K_n. ■

THEOREM 2.6. The RR*f*GP of order n is reversible as an ergodic Markov chain for all n and f with $2 \leq f \leq n - 1$.

PROOF. We shall show that $w_G R_{GH} = w_H R_{HG}$ for all graphs G and H in the RR*f*GP of order n. Let G and H be fixed and $\tilde{G}$ and $\tilde{H}$ be the subsets of X (defined in proof of Theorem 2.1) where $\tilde{G}$ is the Cartesian product of the set of labeled copies of G and the set of admissible edges for G and $\tilde{H}$ is defined analogously. Then, $|\tilde{G}| = c_G \alpha_G$ and $|\tilde{H}| = c_H \alpha_H$.

The proportion of elements of $\tilde{G}$ that correspond to elements of $\tilde{H}$ is $(\alpha_{GH}/\alpha_G) c_G \alpha_G$, where α_{GH} is the number of admissible edges for G each of which when combined with a labeled graph isomorphic with G yield a graph isomorphic to H. This proportion is equal to the proportion of elements of $\tilde{H}$ that correspond in the reverse manner to elements of $\tilde{G}$, namely, $(\alpha_{HG}/\alpha_H) c_H \alpha_H$. If we now divide both sides of $(\alpha_{GH}/\alpha_G) c_G \alpha_G = (\alpha_{HG}/\alpha_H) c_H \alpha_H$, we obtain

$$\frac{\frac{\alpha_{GH}}{\alpha_G} c_G \alpha_G}{\sum_{H \in \mathcal{G}} c_H \alpha_H} = \frac{\alpha_{GH}}{\alpha_G} w_G = \frac{\frac{\alpha_{HG}}{\alpha_H} c_H \alpha_H}{\sum_{H \in \mathcal{G}} c_H \alpha_H} = \frac{\alpha_{HG}}{\alpha_H} w_H .$$

Noting that $\frac{\alpha_{GH}}{\alpha_G} = R_{GH}$ and $\frac{\alpha_{HG}}{\alpha_H} = R_{HG}$ concludes the proof. ■

Among the parameters of interest in Markov chains is that of *cover time*, this being the expected number of steps that it will take a random walk starting from a given distribution to reach every node in the chain.We shall use the following two theorems of U. Feige as lemmas for the proofs of Theorem 2.7 and Corollary 2.8.

LEMMA F1 (Theorem 2.1b in Ref. 3) The cover time for any starting node in a graph with n nodes is at least $(1 - o(1))\ n \log n$ and at most $(4/27 + o(1))n^3$. ■

LEMMA F2 (Theorem 2.1c in Ref. 3) The cover time for a regular graph on n nodes is at most $2n^2$ ■

THEOREM 2.7. Let $\tilde{R}(n, f)$ be the labeled graph transition digraph for the RRfGP, then $\tilde{R}(n, f)$ has

(1) order $\sum_{G \in \mathcal{G}} c_G = \tilde{N}(n, f)$, where $\mathcal{G}$ is the set of unlabeled f-graphs of order n,

(2) outdegree sequence $(...\alpha_i...)\ i = 1,2,...,\tilde{N}(n, f)$ and this also can be written $(...\alpha_G^{c_G}...)\ G \in \mathcal{G}$, where $\alpha_G^{c_G}$ means α_G appears c_G times in the degree sequence,

(3) size $\sum_{i=1}^{\tilde{N}(n,f)} \alpha_i = \sum_{G \in \mathcal{G}} c_G \alpha_G$, and

(4) the cover time from any starting node is at most $(\frac{4}{27} + o(1))\ (\tilde{N}(n, f))^3$.

PROOF. (1) The order is the number of labeled copies of f-graphs of order n, namely, $\sum_{G \in \mathcal{G}} c_G$.

(2) The outdegree of each node G_j (the j-th labeled copy of an unlabeled f-graph G) in $\tilde{R}(n, f)$ is equal to the number of admissible edges α_G at G_j. Since there are c_j copies of G, the number α_G appears c_j times in the degree sequence. Let the index i range over all states of $\tilde{R}(n, f)$ to obtain the degree sequence to be $(...\alpha_i...)\ i = 1,2,...\tilde{N}(n, f)$. Since each graph isomorphic to a given G appears c_G times, α_G will appear c_G times in the outdegree sequence. Thus, the outdegree sequence also can be written $(...\alpha_G^{c_G}...)\ G \in \mathcal{G}$.

(3) The size of $\tilde{R}(n, f)$ is the sum of the outdegrees. Thus, the size is $\sum_{i=1}^{\tilde{N}(n,f)} \alpha_i = \sum_{G \in \mathcal{G}} c_G \alpha_G$.

(4) The outdegree at each node G of $\tilde{R}(n, f)$ is the number of admissible edges α_G. The underlying digraph of $\tilde{R}(n, f)$ is a symmetric digraph and the underlying simple graph of $\tilde{R}(n, f)$ is obtained by replacing each 2-cycle between adjacent nodes with an undirected edge. A random walk on $\tilde{R}(n, f)$ is equivalent to a random walk on its underlying graph where a walk at vertex G moves to a neighbor of G uniformly with probability $1/\alpha_G$. Thus, applying Lemma F1 yields the assertion. ■

COROLLARY 2.8. If $f = n - 1$, the cover time for $\tilde{R}(n, n - 1)$ is at most $2^{n(n-1)+2}$.

PROOF. Each node of $\tilde{\mathrm{R}}(n, n-1)$ has outdegree $\alpha_i = \binom{n}{2}$, that is, the underlying graph of $\tilde{\mathrm{R}}(n, n-1)$ is an $\binom{n}{2}$-regular graph of order $\tilde{N}(n, n-1) = 2^{\binom{n}{2}}$. Thus, from Lemma F2 we obtain the cover time to be $2\binom{2^{\binom{n}{2}}}{} = 2^{n(n-1)+2}$. ■

Still to be resolved is the following problem.

PROBLEM 1. What is the cover time for $\mathrm{R}(n, f)$?

3. ALGORITHMS

An algorithmic approach has been applied to obtain results in $\mathrm{D}(n, f)$ and $\mathrm{R}(n, f)$. Essentially, two types of algorithms have been used.

These are *exact algorithms* that generate all nodes of $\mathrm{D}(n, f)$ (and $\mathrm{R}(n, f)$) and calculate transition and state probabilities. For those values of n for which these algorithms can be implemented, the (numerical) distribution of any random variable on f-graphs in the RfGP can be obtained and enumerative/structural problems for f-graphs in this process can be solved.

The main algorithm, IMAGEf, uses the *Breadth First Search* method for generating and visiting nodes. Its time complexity depends on the number of nodes and on the complexity of the algorithm used for testing the isomorphism of two graphs. Using the IMAGEf algorithm we have computed the transition digraphs for the RRfGP for up to $n = 7$. The exponential growth of the number of f-graphs limits the application of this algorithm. IMAGEf has been extended for constructing the underlying (undirected) graph $\mathrm{U}(n, f)$ of $\mathrm{D}(n, f)$. This extension has been applied to studying purely graph theoretical properties of $\mathrm{U}(n, f)$ such as its order (see Table I), size, degrees of its vertices and diameter.[5]

Sampling based algorithms are used for generating and identifying f-graphs with large order to study probabilistic problems in RfGP and RRfGP. This type of algorithm has also been used for f-graphs having specific properties, such as graphs with integral spectra. Searching in the RRfGP has been found to be more effective than using a strictly evolutionary search. Algorithms in this class are polynomial in n. In general, this method is used to generate samples of processes and collect data on chosen properties. This information has been combined with approximation functions for studying the properties of the RfGP as functions of n and f (see Ref. 5).

3.1. Some Comments on the RfGP

The common order of D(n, f) and R(n, f) is denoted $N(n, f)$. Using the algorithm IMAGE*f* values of $N(n, f)$ have been computed for $n \leq 10$ and are given in Table I.

For the R*f*GP we have studied the structure of the terminal states and the probability of the process terminating at an *f*-graph with a given structure.[1,5] Since the number of unlabeled *f*-graphs together with the number of terminal *f*-graphs both increase quite rapidly with n, it is natural to partition the terminal graphs into equivalence classes according to some structural property. For example, let $NT(n, f)$ denote the *number of terminal nodes (absorbing states)* in D(n, f). As noted above, for some values of n and f, values of $N(n, f)$ (see Table I) and $NT(n, f)$ have been determined using algorithms specifically designed for the study of the R*f*GP (see Ref. 5). For $f = 2$, formulas for $N(n, 2)$ and $NT(n, 2)$ have been derived.[1,5] Namely,

THEOREM 3.1. Let $N(t, n, 2)$ denote the number of nodes of R(n, 2) associated with 2-graphs of size t, $1 \leq t \leq n - 1$. Then,

$$N(t, n, 2) = N(t, n-1, 2) + N(2t-n, t, 2)$$

where $N(x, y, 2) = 0$, if $x < 0$ or $x > y$, $N(0, n, 2) = 1$, if $n \geq 1$, and $N(n, n, 2) = r(n, 3)$, the number of partitions of n having no part less than 3,

$$N(n, 2) = \sum_{t=0}^{n} N(t, n, 2)$$

and

$$NT(n, 2) = p(n) - p(n-2) - p(n-3) + p(n-5)$$

where $p(k)$ is the number of unrestricted partitions of k. ■

TABLE I

The number of nodes $N(n, f)$

N	$f = 2$	3	4	5	6	7	8	9
4	7	11						
5	11	23	34					
6	19	62	122	156				
7	29	150	510	888	1044			
8	46	424	2590	7459	11302	12346		
9	70	1165	15471	84245	197867	262322	274668	
10	106	3547	108376	1249973	5203135	9808968	11730500	12005168

A graph that is a terminal node in D(n, f) is an f-graph to which no edge can be added without introducing a vertex of degree greater than f. If m is the number of vertices of degree less than f in a terminal node, then $m \leq \min \{n(n–f–1) / (2(n–f)–1), f\}$ (see Theorem 2.3 in Ref. 5).

Let the terminal nodes of D(n, f) be classified in accordance with their number m of vertices of degree less than f. A problem in random graph theory, verbally posed by Paul Erdős in 1985, and formally stated on p. 359 of Ref. 8 (see also Refs. 1,5,9), is that of determining the probability distribution of these classes of terminal nodes. Algorithmic methods[5] have been applied to study the Erdős problem. Let $P(m; n; f)$ denote the probability that the RfGP will terminate in a graph with m vertices of degree less than f. Using an exact algorithm DOV, $P(m; n; 2)$ has been calculated for n from 4 to 1532. It is known that as n goes to infinity, the probability of RfGP terminating at an f-regular graph tends to 1, when nf is even, and to a terminal graph with exactly one vertex of degree $f - 1$, when nf is odd. From the exact calculation noted above for $f = 2$ we have $P(0; n; 2) = 0.8948$ when $n = 1532$. In general, it is of interest to determine the probability $P(0; n; f)$ that the RfGP will terminate at an f-regular graph, when nf is even and n is fixed (a similar question can be considered when nf is odd and n is fixed). The distribution $P(m; n; f)$ has been studied using both exact (IMAGEf) and simulation algorithms. We have obtained an approximation of this distribution as a function of both n and f within the interval for n and f up to 240 (see Ref. 5).

These methods have been applied to determine distributions of a number of other random variables and properties of the RfGP. For example, the probability of a big cycle $P(M > n/2; 2)$ in R2GP has been determined using an exact algorithm PBC for n from 4 to 63 (see Ref. 5).

3.2. Some Comments on the RRfGP

The *distance* between two graphs G and H in R(n, f) is defined as the least number of deletions and insertions of edges in G needed to obtain H and is denoted d(G, H). This is also the distance between the nodes in R(n, f). The *diameter* of the digraph R(n, f) is defined as max{d(G, H): G, H nodes of R(n, f)} and is denoted diamR(n, f). Since R(n, f) is a symmetric digraph, its diameter is the same as that of the underlying graph of D(n, f). The diameters of R(4, 2) and R(4, 3) are obviously equal to 4 and 6, respectively. We have proven the following result (see Theorems 1.1 and 2.9 in Ref. 6).

THEOREM 3.2. If $f = 2$ or $f \geq (n - 1) / 2$ then diamR(n, f) = $\lfloor nf / 2 \rfloor$. ■

PROBLEM 2. What is the diameter of R(n, f) when $3 \leq f \leq (n - 1) / 2$?

PROBLEM 3. Obtain eccentricity results in general such as radius, center, and periphery properties of $R(n, f)$.

In the theory of ergodic Markov chains it is of interest to know $\lim_{k\to\infty} R_{ij}^{(k)}$. Since $R(n, f)$ is bipartite, it follows that $\lim_{k\to\infty} R_{ij}^{(k)}$ does not exist. However, $\lim_{k\to\infty} R_{ij}^{(2k)}$ and $\lim_{k\to\infty} R_{ij}^{(2k+1)}$ exist and are disjoint on their nonzero entries. Thus $\lim_{k\to\infty}(R_{ij}^{(2k)} + R_{ij}^{(2k+1)}) / 2$ exists. In particular, $\lim_{k\to\infty}(R_{ij}^{(2k)} + R_{ij}^{(2k+1)}) / 2 = w_j$ and $\lim_{k\to\infty}(\boldsymbol{R}^{(2k)} + \boldsymbol{R}^{(2k+1)}) / 2$ is the $N(n, f) \times N(n, f)$ matrix with each row equal to $\boldsymbol{w}$, the unique probability vector of $\boldsymbol{R}$.

If n and f are such that the transition matrix $\boldsymbol{R}$ for the RRfGP of order n is not available, an approximation for $\boldsymbol{w}$ can be obtained from the following computational experiment.

If a sufficiently large number of random walks of lengths $L_0 = 2k$ and $L_1 = 2k + 1$, respectively, that start at K_n^c are generated, the relative frequency distribution of the ending states of these random walks will yield an approximation to the probability distributions $R_{1j}^{(L_0)}$ and $R_{1j}^{(L_1)}$ and consequently $\boldsymbol{w}$ can be approximated. Note that for large n it is not expected that the set of generated ending states of these random walks will coincide with the entire node set of $R(n, f)$. First, those states with low probability may not appear as outcomes of this experiment and second, if n is large it may be computationally unfeasible to generate enough random walks of length sufficiently greater than $\text{diam}R(n, f)$ to make the results statistically meaningful.

For $R(n, f)$ and $x = N(n, f)$, we have compared outcomes $\boldsymbol{z} = [z_1\ z_2\ \ldots\ z_x]$ of the above experiment with the theoretical results $\boldsymbol{w} = [w_1\ w_2\ \ldots\ w_x]$ that we have obtained. This was done using $y = \sqrt{(1/x)(\sum_{i=1}^{x}(w_i - z_i)^2)}$ as a measure of comparison. The experimental data was obtained using samples of 10,000 random walks with $k = c\lfloor nf/2 \rfloor$, $c = 1,2,\ldots,100$. The approximations for $n = 4$ and 5 yielded average values of y between 0.001 and 0.002.

Sampling algorithms have an important place in these studies, however more theoretical results are still needed. For example, of particular interest for search problems it is useful to know the answer to Problem 1 and to have solutions to hitting time questions such as the following.

PROBLEM 4. In the RRfGP, what is the expected number of steps for a random walk that starts at K_n^c to reach its first f-graph having a specified property?

A variety of other random variables in the RRfGP are being studied using an algorithm that generates f-graphs with specified properties.

In addition to the exact and sampling algorithmic approaches we have presented, it might be productive in the study of the RRfGP to apply the as-

ymptotic method that was used in Refs. 10–12. Here the order of the node graphs of the RR*f*GP would go to infinity. Although what questions might be answered is not clear at this time, the questions posed would relate to asymptotic structural properties of the node graphs.

Acknowledgments. – Support of this work was provided in part by grants from The Technical University of Poznań and The Dyson College of Arts and Sciences, Pace University. Some computations were done in The Poznań Supercomputing and Networking Center.

REFERENCES

1. K. T. Balińska and L. V. Quintas, *The Random f-Graph Process*, in: J. Gimbel, J. W. Kennedy, and L. V. Quintas (Eds.), *Quo Vadis, Graph Theory?* Fairbanks, Alaska, 1990; *Ann. Discrete Math*. **55** (1993) 333–340.
2. K. T. Balińska and L. V. Quintas, *The Reversible Random f-Graph Process,* CSIS Pace University Technical Report Series, Report No. 132, 1998, pp. 1–11.
3. L. Lovász, *Random Walks on Graphs,* in: D. Miklós, V. T. Sós, and T. Szőnyi (Eds.), *Combinatorics, Paul Erdős is Eighty* (Volume 2), Keszthely, Hungary, 1993; Bolyai Society Mathematical Studies, 1996, 353–398.
4. K. T. Balińska and L. V. Quintas, *Maximum Outdegrees in a Digraph whose Nodes are Graphs with Bounded Degree,* in: A. Sali, M. Simonovits, and V. T. Sós (Eds.), *Research Communications of the Conference held in the Memory of Paul Erdős*, J. Bolyai Mathematical Society, Budapest, Hungary, 1999, 23–26.
5. K. T. Balińska, *Algorithms for Random Graphs with Bounded Degree,* Politechnika Poznańska, Rozprawy, 314, Poznań, 1996.
6. K. T. Balińska, G. R. Brightwell, and L. V. Quintas, *J. Math. Chem.* **24** (1998) 109–121.
7. P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks*, Carus Mathematical Monographs, No. 22, Mathematical Association of America, 1984.
8. K. T. Balińska and L. V. Quintas, *Random Graph Models for Physical Systems,* in: R. B. King and D. H. Rouvray (Eds.), *Graph Theory and Topology in Chemistry,* Studies in Physical and Theoretical Chemistry **51**, Elsevier, Amsterdam, 1987, pp. 349–361.
9. K. T. Balińska and L. V. Quintas, *On Generating Random f-Graphs,* Graph Theory Notes of New York **XII**, New York Acad. Sci., 1986, pp. 22–24.
10. K. T. Balińska, H. Galina, and L. V. Quintas, *A Kinetic Approach to the Random f-Graph Process,* Graph Theory Notes of New York **XXIII**, New York Acad. Sci. 1992, pp. 35–40.
11. K. T. Balińska, H. Galina, L. V. Quintas, and J. Szymański, *Discrete Applied Math.* **67** (1996) 13–25.
12. K. T. Balińska, H. Galina, L. V. Quintas, and J. Szymański, *J. Chem. Inf. Comput. Sci.* **36** (1996) 347–350.

SAŽETAK

Dva modela za slučajne grafove s ograničenim stupnjem čvorova

Krystyna T. Balińska, Michael L. Gargano i Louis V. Quintas

Izučavan je par usmjerenih grafova čiji su čvorovi $\{v\}$ neoznačeni grafovi reda n koji sadrže vrhove ograničenog stupnja jednakog f. Bliskost u ovim usmjerenim grafovima definirana je preko jednobridnih transformacija nad $\{v\}$. Vjerojatnosti nad oba grafa su definirane tako, da je jedan usmjereni graf apsorbirajući a drugi ergodički Markovljev lanac. Razmotreni su probabilistički i deterministički rezultati koji vrijede za te Markovljeve lance. Jedan od primjera fizikalnog interesa za takve lance jesu modeli u kojima se čvorovi $\{v\}$ identificiraju s različitim kemijskim spojevima.