SIAM J. COMPUT.
Vol. 9, No. 3, August 1980

0097-5397/80/0903-0014 $01.00/0

RANDOM GRAPH ISOMORPHISM*

LÁSZLÓ BABAI†, PAUL ERDŐS‡ AND STANLEY M. SELKOW§

Abstract. A straightforward linear time canonical labeling algorithm is shown to apply to almost all graphs (i.e. all but $o(2^{\binom{n}{2}})$ of the $2^{\binom{n}{2}}$ graphs on n vertices). Hence, for almost all graphs X, any graph Y can be easily tested for isomorphism to X by an extremely naive linear time algorithm. This result is based on the following: In almost all graphs on n vertices, the largest $n^{0.15}$ degrees are distinct. In fact, they are pairwise at least $n^{0.03}$ apart.

Key words. graph, isomorphism testing, canonical labeling, random graph, naive algorithm, average-case analysis, linear time, degree sequence of a graph

1. A straightforward algorithm.

The problem of testing graphs for isomorphism belongs to those combinatorial search problems for which no polynomial-time algorithm is available as yet. It is, however, striking, that even the most trivial isomorphism testing algorithms have a good performance if tested on randomly generated graphs. The aim of the present note is to give some theoretical background for this.

By a *canonical labeling algorithm* of the class $\mathcal{K}$ of graphs we mean an algorithm which assigns the numbers $1, \cdots, n$ to the vertices of each graph in $\mathcal{K}$, having n vertices, in such a way that two graphs in $\mathcal{K}$ are isomorphic (if and) only if the obtained labeled graphs coincide. (We assume that $\mathcal{K}$ is closed under isomorphisms.) Clearly, given a canonical labeling algorithm of $\mathcal{K}$, and an algorithm deciding whether a given graph belongs to $\mathcal{K}$ or not, we also have an algorithm, deciding whether $X \cong Y$ for any two graphs X, Y provided $X \in \mathcal{K}$. Namely, if $Y \notin \mathcal{K}$ then $X \not\cong Y$; and if $Y \in \mathcal{K}$ then we have to check whether X and Y coincide after canonical labeling.

We describe a class $\mathcal{K}$ of graphs (closed under isomorphisms) and a canonical labeling algorithm of $\mathcal{K}$. Deciding whether $X \in \mathcal{K}$ and subsequently, canonically labeling X will require *linear time* (i.e. $O(n^2)$, where n is the number of vertices) on a random access machine which operates in one step on binary words of length $O(\log n)$. We shall prove, that $\mathcal{K}$ contains *almost all* graphs on n vertices (i.e. all but $o(2^{\binom{n}{2}})$ of the graphs on a fixed vertex set of cardinality n). In particular, we prove

THEOREM 1.1. *There is an algorithm which, for almost all graphs X, tests any graph Y for isomorphism to X within linear time.*

The algorithm is as follows:

Input: a graph X having n vertices. (The graph is represented by its adjacency matrix.)

1. Compute $r = [3 \log n/\log 2]$.
2. Compute the degree of each vertex of X.
3. Qrder the vertices by degree; call them $v(1), \cdots, v(n)$. Denote by $d(i)$ the degree of $v(i): d(1) \geqq d(2) \geqq \cdots \geqq d(n)$.
4. If $d(i) = d(i+1)$ for some i, $1 \leqq i \leqq r-1$, set $X \notin \mathcal{K}$, end. Otherwise
5. Compute

$$f(v(i)) = \sum_{j=1}^{r} a(i,j)2^j \qquad (i = r+1, \cdots, n)$$

[*] Received by the editor October 9, 1977.

[†] Department of Algebra and Number Theory, Eötvös L. University, H-1445 Budapest 8, Pf. 323, Hungary.

[‡] Mathematics Institute of the Hungarian Academy of Sciences, Reáltanoda u. 13-15, H-1053 Budapest, Hungary.

[§] Department of Computer Science, Concordia University, Montréal, Quebec, Canada.

(the "code of $v(i)$ with respect to $v(1), \cdots, v(r)$"), where $a(i, j)=1$ if $v(i)$ and $v(j)$ are adjacent and $a(i, j)=0$ otherwise.

6. Order the vertices $v(r+1), \cdots, v(n)$ according to their f-value: $w(r+1), \cdots, w(n)$ where $f(w(r+1)) \geqq \cdots \geqq f(w(n))$.
7. If $f(w(i))=f(w(i+1))$ for some $i, r+1 \leqq i \leqq n-1$, set $X \notin \mathcal{K}$, end. Otherwise
8. Label $v(i)$ by i for $i=1, \cdots, r$, and $w(i)$ by i for $i=r+1, \cdots, n$. This labeling will be called canonical. Set $X \in \mathcal{K}$. End.

In other words, the first r labels will be assigned to the vertices with largest degrees, in decreasing order of the degree. If this is not unique, then $X \notin \mathcal{K}$. The rest of the labels will be assigned to the remaining $n-r$ vertices in decreasing order of their codes with respect to the first r vertices, as defined in step 5. Again, if two vertices get the same code then $X \notin \mathcal{K}$.

Obviously, this algorithm defines a canonical labeling, indeed, and $\mathcal{K}$ is closed under isomorphisms. The running time of the algorithm is $O(n^2)$, as readily verified. Our principal result is the following:

THEOREM 1.2. *The probability that a random graph on n vertices belongs to the class $\mathcal{K}$, specified by our canonical labeling algorithm, is greater than* $1-\sqrt[7]{1/n}$ *(for sufficiently large n).*

This clearly implies Theorem 1.1.

At this point we have to stress that our algorithm is not intended for practical use: more involved but still very natural heuristic algorithms are much better. Our purpose is to show that *even such an extremely naive, fast algorithm solves the problem for almost all graphs.*

The referee and the first named author share the responsibility for almost two years delay in publishing this paper. Since 1977, the paper has been circulated as a preprint essentially in its present form (except for the introduction and a simplification of the proof in § 4, suggested by the referee).

In the preprint we formulated the following two problems:

(i) Find a fast canonical labeling algorithm with exponentially small probability of rejection.

(ii) Find a canonical labeling algorithm of *all* graphs, with polynomial expected running time.

The preprint seems to have inspired further work instantaneously. Both problems have been solved shortly after submission of this paper. R. Lipton [8] gives a canonical labeling algorithm with $O(n^6 \log n)$ running time and exponentially small probability of rejection ($c^{-n}, c>1$). R. M. Karp [7] improves this, giving an $O(n^2 \log n)$ algorithm, with $O(n^{3/2} 2^{-n/2})$ probability of rejection. Babai and Kučera [1] prove that the standard vertex classification algorithm gives a canonical labeling in $O(n^2)$ time with c^{-n} probability of rejection. In addition, it is proved in [1] that the rejected graphs can be handled such as to obtain a canonical labeling algorithm of *all* graphs with *linear expected time*, i.e. the average running time over the $2^{\binom{n}{2}}$ graphs is $O(n^2)$.

This short survey tends to convince us that, despite of the long delay, the present note may merit some attention. Apart from [1], it still appears to be the only example of a *linear* time canonization of almost all graphs. [1] definitely outscores our results, but the simplicity of our algorithm can hardly be improved on, and it may be worth noting that still, such an algorithm canonizes almost all graphs.

The performance of our algorithm relies on our results on the degree sequence of a random graph. This aspect of the paper, which extends the idea of [4], may have interest on its own. The results of § 3 are stronger than what would be necessary to prove the main theorem. Recently, B. Bollobás [2] has obtained finer and more detailed results on this subject.

More about random graphs can be found in Erdős–Spencer [3].

Concerning the probabilistic analysis of some hard combinatorial problems we refer to Karp [6].

2. Preliminaries. Throughout this paper, we shall use the following notation:

$$P(m, l) = \sum_{s=l+1}^{m} \binom{m}{s}. \tag{1}$$

Clearly,

$$P(m, l) = P(m-1, l) + P(m-1, l-1). \tag{2}$$

We shall refer to the following well-known asymptotic formula:

$$\binom{m}{m/2+t} = \binom{m}{[m/2]} e^{-2t^2/m+O(\tau)} \tag{3}$$

where $\tau = t^2/m^2 + t^4m^3$ (cf. Feller [5, Chap. VII/2]). The O notation always refers to absolute constants (not depending on any of our parameters). Of course, $m/2+t$ should be an integer. This means that t is either an integer or a half-integer, depending on the parity of m. Similar restrictions on the possible values of parameters are understood throughout without explicit mention.

Random variables are denoted by block letters. A *random graph* $\mathbf{X}$ on the vertex set $V = \{1, \cdots, n\}$ assumes as its values each graph on V with probability $2^{-\binom{n}{2}}$.

We start with some elementary computation with binomial coefficients.

PROPOSITION 2.1. *If $l = m/2+t \quad (0 < t < m/2)$ and $f > r(\log 2/2)\frac{m}{t}$, then*

$$\binom{m}{l+f} < 2^{-r}\binom{m}{l}.$$

Proof.

$$\binom{m}{l+f} \Big/ \binom{m}{l} = (m-l)\cdots(m-l-f+1)/(l+f)\cdots(l+1) < \left(\frac{m-l}{l+1}\right)^f < \left(\frac{m-l}{m/2}\right)^f$$

$$= (1-2t/m)^f < \exp(-2tf/m) < \exp(-r\log 2) = 2^{-r}. \quad \square$$

COROLLARY 2.2. *If $l = m/2+t \quad (0 < t < m/2)$ then*

$$P(m, l) \Big/ \binom{m}{l} < \frac{m}{t}.$$

Proof. Let $g = [m \log 2/(2t)] + 1$. Then, by Proposition 2.1,

$$P(m, l) < \binom{m}{l} \cdot (g + g/2 + g/4 + \cdots) < 2g\binom{m}{l} < \frac{m}{t}\binom{m}{l}. \quad \square$$

On the other hand, a lower bound of the same order of magnitude also holds. To this end, we need another simple estimate:

PROPOSITION 2.3. *If $l = m/2+t \quad (0 < t < m/2)$ and $0 < f < t$, then*

$$\binom{m}{l} > \binom{m}{l-f}(1-4tf/m).$$

Proof.

$$\binom{m}{l}\Big/\binom{m}{l-f}=(m-l+f)\cdots(m-l+1)/l(l-1)\cdots(l-f+1)>((m-l)/l)^f$$
$$=\left(\frac{m/2-t}{m/2+t}\right)^f>(1-2t/m)^{2f}>1-4tf/m. \qquad \square$$

COROLLARY 2.4. *If $l=m/2+t$ $(2\sqrt{m}<t<m/30)$, then*

$$P(m,l)/\binom{m}{l}>\frac{m}{23t}.$$

Proof. For any natural number f, obviously

$$P(m,l)/\binom{m}{l}>f\binom{m}{l+f}\Big/\binom{m}{l}.$$

By Proposition 2.3, the right side exceeds $f(1-4(t+f)f/m)$. Set $f=[m/9t]+1$. So, $f>m/9t$ and $f<m/9t+1<t/27$; hence our quantity exceeds

$$\frac{m}{9t}\left(1-\frac{4\cdot 28}{9\cdot 27}-\frac{4\cdot 28}{27}\,\frac{t}{m}\right)>\frac{m}{23t}. \qquad \square$$

3. The largest degrees are distinct. Let $\mathbf{X}$ be a random graph on the vertex set $\{1,\cdots,n\}$. Let $\mathbf{d}(x)$ denote the degree of the vertex x. Let us fix a natural number d, and set $\mathbf{z}_x=0$ if $\mathbf{d}(x)\leqq d$, $\mathbf{z}_x=1$ otherwise. Let $\mathbf{z}=\sum_{x=1}^{n}\mathbf{z}_x$.

We are interested in the behavior of the expected value $\mathrm{E}(\mathbf{z})$ (depending on the choice of d).

LEMMA 3.1. *Let $m=n-1$, $d=m/2+t$ where*

$$t=t_0+\omega_m(m/\log m)^{1/2},$$

where

$$t_0=(\tfrac{1}{2}m\log m)^{1/2}-\tfrac{1}{8}(2m/\log m)^{1/2}\log\log m,$$

and

$$-\log m/\sqrt{2}<\omega_m<m^{0.7}.$$

If $\omega_m<0$ then

$$\mathrm{E}(\mathbf{z})>c_1e^{-1.4\omega_m};$$

if $\omega_m>0$ then

$$\mathrm{E}(\mathbf{z})<c_2\exp(-2.8\omega_m-2\omega_m^2/\log m).$$

If $\omega_m/\log m\to-\varepsilon/\sqrt{2}$ $(m\to\infty)$ where $O<\varepsilon<1$ (ε is fixed) then

$$\mathrm{E}(\mathbf{z})>m^{\varepsilon(2-\varepsilon+o(1))}.$$

Proof. Clearly, for any $x(1\leqq x\leqq n)$,

$$\mathrm{E}(\mathbf{z})=n\,\mathrm{E}(\mathbf{z}_x)=n2^{-n+1}P(n-1,d)$$
$$=(1+o(1))m2^{-m}P(m,d).$$

Now we apply Corollaries 2.2 and 2.4 to obtain θ, $0<\theta<1$ such that

$$P(m,d)=\frac{1}{1+22\theta}\binom{m}{d}\frac{m}{t}=\frac{1+o(1)}{1+22\theta}2^m m\,e^{-2t^2/m}/(t(\tfrac{1}{2}\pi m)^{1/2})$$

(by (3)). Hence,

$$\log \mathrm{E}(\mathbf{z}) = O(1) + 3 \log m/2 - \log t - 2t^2/m.$$

For $t = t_0$ the right side is bounded; hence in the general case,

$$\begin{aligned}\log \mathrm{E}(\mathbf{z}) &= O(1) - \log (t/t_0) - 2(t^2 - t_0^2)/m \\ &= O(1) - \log(1 + \omega_m\sqrt{2}/\log m) - 2\omega_m(\sqrt{2} - \sqrt{2} \log\log m/4 \log m + \omega_m/\log m).\end{aligned}$$

Now our assertions can be readily checked. □

COROLLARY 3.2. *With the notation of Lemma 3.1, the probability that* $\mathbf{x}$ *has a vertex of degree* $> t_0 + \omega_m(m/\log m)^{1/2}$ *is less than* $c_2 \exp(-2.8\omega_m - 2\omega_m^2/\log m)(\omega_m > 0)$.

In order to obtain the counterpart of Corollary 3.2 for $\omega_m < 0$, we have to compute the variance of $\mathbf{z}$.

LEMMA 3.3. *Let* $m = n - 1$ *and* $d = m/2 + t$, *where* $2\sqrt{m} < t < m/30$. *Then*

$$\mathrm{Var}(\mathbf{z})/\mathrm{E}(\mathbf{z})^2 < 1/\mathrm{E}(\mathbf{z}) + 67t^2/m^2.$$

Proof. Clearly,

$$\mathrm{Var}\,\mathbf{z} = \mathrm{E}(\mathbf{z}^2) - \mathrm{E}(\mathbf{z})^2 = mA + \binom{n}{2}B,$$

where

$$A = \mathrm{E}(\mathbf{z}_x)(1 - \mathrm{E}(\mathbf{z}_x) < \mathrm{E}(\mathbf{z}_x) \qquad \text{(hence } nA < \mathrm{E}(\mathbf{z})\text{)},$$

$$B = \mathrm{E}(\mathbf{z}_x\mathbf{z}_y) - \mathrm{E}(\mathbf{z}_x)^2 \qquad \text{(for any } 1 \leq x < y \leq n\text{)}.$$

Clearly, for $x \neq y$

$$\mathrm{E}(\mathbf{z}_x\mathbf{z}_y) = \mathrm{Prob}(\mathbf{d}(x) > d \text{ and } \mathbf{d}(y) > d) = (P_1 + P_2)/2,$$

where P_1, P_2 are conditional probabilities:

$$\begin{aligned}P_1 &= \mathrm{Prob}(\mathbf{d}(x) > d \text{ and } \mathbf{d}(y) > d \mid x \text{ and } y \text{ are adjacent}) \\ &= 2^{-2n+4}P(n-2, d-1)^2;\end{aligned}$$

$$\begin{aligned}P_2 &= \mathrm{Prob}(\mathbf{d}(x) > d \text{ and } \mathbf{d}(y) > d \mid x \text{ and } y \text{ are not adjacent}) \\ &= 2^{-2n+4}P(n-2, d)^2.\end{aligned}$$

It follows (using (2)), that

$$\begin{aligned}B &= 2^{-2n+2}(2P(n-2, d-1)^2 + 2P(n-2, d)^2 - P(n-1, d)^2) \\ &= 2^{-2n+2}(P(n-2, d-1) - P(n-2, d))^2 \\ &= 2^{-2n+2}\binom{n-2}{d}^2.\end{aligned}$$

Hence

$$\begin{aligned}\frac{\mathrm{Var}\,\mathbf{z}}{\mathrm{E}(\mathbf{z})^2} &< \frac{1}{\mathrm{E}(\mathbf{z})} + \binom{n}{2}B/\mathrm{E}(\mathbf{z})^2 < \frac{1}{\mathrm{E}(\mathbf{z})} + \frac{1}{2}\binom{n-2}{d}^2 / P(n-1, d)^2 \\ &< \frac{1}{\mathrm{E}(\mathbf{z})} + \frac{1}{8}\left(\binom{m}{d}/P(m, d)\right)^2 < \frac{1}{\mathrm{E}(\mathbf{z})} + \frac{1}{8}\left(\frac{m}{23t}\right)^2 \\ &< 1/\mathrm{E}(\mathbf{z}) + (23t/m)^2/8 < 1/\mathrm{E}(\mathbf{z}) + 67t^2/m^2.\end{aligned}$$

(We have used here the inequality $\binom{n-2}{d} < \frac{1}{2}\binom{n-1}{d}$ which trivially holds for $d > n/2$; and subsequently Corollary 2.4.) □

COROLLARY 3.4. *If the sequence d_n is so chosen that (setting $d = d_n$) we obtain* $\mathrm{E}(\mathbf{z}) \to \infty$ $(n \to \infty)$, *then*

$$\operatorname{Prob}(\mathbf{z} < \mathrm{E}(\mathbf{z})/2) \to 0.$$

Using the notation of Lemma 3.1, *for* $-\log m\sqrt{2} < \omega_m < 0$ *we have*

$$\operatorname{Prob}(\mathbf{z} < \mathrm{E}(\mathbf{z})/2) < c_3\, e^{1.4\omega_m}.$$

If $\omega_m/\log m \to \varepsilon/\sqrt{2}$ *where* $0 < \varepsilon < 1$ (ε *is fixed*), *then*

$$\operatorname{Prob}(\mathbf{z} < \mathrm{E}(\mathbf{z})/2) < c_4 m^{-\varepsilon(2-\varepsilon+o(1))}.$$

Proof. By Chebyshev's inequality,

$$\operatorname{Prob}(\mathbf{z} < \mathrm{E}(\mathbf{z})/2) < 4 \operatorname{Var} \mathbf{z}/\mathrm{E}(\mathbf{z})^2.$$

This implies our second statement, by Lemmas 3.3 and 3.1. Namely,

$$t^2 < m \log m + 2\omega_m^2 m/\log m = O(m \log m),$$

hence

$$1/\mathrm{E}(\mathbf{z}) + 67t^2/m^2 < e^{-1.4\omega_m} + O(\log m/m)$$

and $\log m/m = \exp(\log\log m - \log m) = o(\exp(-1.4 \log m/\sqrt{2})) = o(e^{-1.4\omega_m})$. The third statement follows similarly.

For the first statement, by Lemma 3.3 and Chebyshev's inequality we only have to prove that if $\mathrm{E}(\mathbf{z}) \to \infty$ then $t/m \to 0$. This is obvious from Lemma 3.1. □

LEMMA 3.5. *Let* $0 < k < \sqrt{n}$, $\sqrt{3}/2 < \alpha < 1$ *and* $t = \alpha(n \log n)^{1/2}$. *Then the probability of the event that* $\mathbf{X}$ *has two vertices* x, y *of degrees exceeding* $n/2 + t$ *such that* $|\mathbf{d}(x) - \mathbf{d}(y)| < k$ *is*

$$p = o(kn^{3/2-2\alpha^2}) \qquad (n \to \infty).$$

Proof. Let $n/2 < a \leqq b$. The probability that $\mathbf{d}(x) = a$ and $\mathbf{d}(y) = b$ $(x \neq y)$ is clearly

$$\frac{1}{2}\left(\binom{n-2}{a}\binom{n-2}{b} + \binom{n-2}{a-1}\binom{n-2}{b-1}\right)\Big/2^{2n-4} < 2^{-2n+4}\binom{n-2}{a-1}^2.$$

Hence, the probability that $a \leqq \mathbf{d}(x) \leqq \mathbf{d}(y) \leqq \mathbf{d}(x) + k$ is at most

$$k \cdot 2^{-2n+4} \sum_{s=a}^{n-1} \binom{n-2}{s-1}^2.$$

By Proposition 2.1, the sum here is less than

$$\frac{n}{a - n/2}\binom{n-2}{a-1}^2.$$

Setting $a = [n/2 + t]$, we obtain (by (3))

$$\begin{aligned} p &< \binom{n}{2} k \frac{n}{t} e^{-4t^2/n}/(\tfrac{1}{2}\pi n)(1 + o(1)) \\ &= \frac{\sqrt{2}(1 + o(1))}{\pi\alpha} n^2 k (n \log n)^{-1/2} n^{-2\alpha^2} \\ &< kn^{3/2-2\alpha^2}/(\log n)^{1/2} \end{aligned}$$

(for n not too small). □

Now we are in the position to prove

THEOREM 3.6. *Let $\mathbf{d}_1 \geqq \mathbf{d}_2 \geqq \cdots \geqq \mathbf{d}_n$ denote the degrees of the vertices of the random graph* $\mathbf{X}$. *Let $k = [n^{0.03}]$ and $l = [n^{0.15}]$. For n sufficiently large, the event that $\mathbf{d}_i - \mathbf{d}_j \geqq k$ for every i, j satisfying $1 \leqq i < j \leqq l$ has probability exceeding $1 - n^{-0.15}$.*

Proof. Set $\varepsilon = \frac{1}{12}$, $\alpha = 1 - \varepsilon$, $t = \alpha (n \log n)^{1/2}/\sqrt{2}$, $d = [n/2 + t]$. Then, with the notation of Lemma 3.1, $t = t_0 + \omega_m (m/\log m)^{1/2}$ where $\omega_m/\log m \to -\varepsilon/\sqrt{2}$, whence, by 3.1,

$$\mathrm{E}(\mathbf{z}) > m^{\varepsilon(2-\varepsilon+o(1))}.$$

$\varepsilon(2-\varepsilon) > 0.157$, hence $\mathrm{E}(\mathbf{z})/2 > n^{0.15} \geqq l$ for sufficiently large n. By Cor. 3.4, this implies that $\mathbf{X}$ has at least $\mathrm{E}(\mathbf{z})/2$ vertices of degree $> d$ with probability $> 1 - c_4 m^{-\varepsilon(2-\varepsilon+o(1))} > 1 - m^{-0.155}$. Finally, by Lemma 3.5, the difference between the degrees of any two of these vertices is at least k with probability $> 1 - kn^{3/2-2\alpha^2} > 1 - n^{0.03+1.5-2\alpha^2} > 1 - n^{-0.1505}$. Hence the probability that $\mathbf{X}$ does not satisfy the theorem is less than

$$n^{-0.155} + n^{-0.1505} < n^{-0.15}. \qquad \square$$

Remark. The particular corollary to Theorem 3.6, that the vertex having maximum degree is unique in almost all graphs, appears in Erdős–Wilson [4].

4. Uniqueness of the codes of the vertices. As in § 3, let $\mathbf{X}$ be a random graph having $V = \{1, \cdots, n\}$ for its vertex set. Let $\mathbf{d}_1 \geqq \mathbf{d}_2 \geqq \cdots \geqq \mathbf{d}_n$ denote the degree sequence of $\mathbf{X}$. Set $r = [3 \log n/\log 2]$, and let C denote the event that $\mathbf{d}_i \geqq \mathbf{d}_{i+1} + 3$ for $i = 1, \cdots, r+2$. We write $\bar{C}$ for the negation of C. By Theorem 3.6,

$$\mathrm{Prob}\,(\bar{C}) < n^{-0.15}.$$

For $i \neq j$, let $C(i, j)$ denote the event that in the graph $\mathbf{X}(i, j)$ obtained from $\mathbf{X}$ by deleting i and j, the largest r degrees are distinct. Clearly, C implies $C(i, j)$ for all $i, j (1 \leqq i \leqq j \leqq n)$. Let $A(i, j)$ denote the event that either $C(i, j)$ fails or i and j have identical codes with respect to the vertices having the largest r degrees in $\mathbf{X}(i, j)$.

The probability that $\mathbf{X}$ is rejected by our algorithm is less than

$$\begin{aligned} &\mathrm{Prob}\,(\bar{C}) + \mathrm{Prob}\,(C \text{ and the graph } \mathbf{X} \text{ has two vertices} \\ &\qquad \text{with identical codes}) \\ &\leqq \mathrm{Prob}\,(\bar{C}) + \sum_{i<j} \mathrm{Prob}\,(C \text{ and } A(i, j)) \\ &\leqq \mathrm{Prob}\,(\bar{C}) + \sum_{i<j} \mathrm{Prob}\,(C(i, j) \text{ and } A(i, j)) \\ &\leqq \mathrm{Prob}\,(\bar{C}) + \sum_{i<j} \mathrm{Prob}\,(A(i, j) | C(i, j)) \\ &= \mathrm{Prob}\,(\bar{C}) + \binom{n}{2} 2^{-r} < n^{-0.15} + O\left(\frac{1}{n}\right) < n^{-1/7}. \end{aligned}$$

This proves Theorem 1.2. $\square$

REFERENCES

[1] L. BABAI AND L. KUČERA, *Canonical labelling of graphs in linear average time*, 20th Annual IEEE Symp. on Foundations of Comp. Sci. (Puerto Rico) 1979, pp. 39–46.

[2] B. BOLLOBÁS, *Degree sequences of random graphs*, Aarhus University, 1978 preprint.

[3] P. ERDŐS AND J. SPENCER, *Probabilistic Methods in Combinatorics*, Akadémiai Kiadó, Budapest, 1974.

[4] P. ERDŐS AND R. J. WILSON, *On the chromatic index of almost all graphs*, J. Comb. Theory—B, 23 (1977), pp. 255–257.

[5] W. FELLER, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., John Wiley, New York, 1968.

[6] R. M. KARP, *The fast approximate solution of hard combinatorial problems*, Proc. 6th South-Eastern Conf. Combinatorics, Graph Theory and Computing (Florida Atlantic U. 1975), pp. 15–31.

[7] ———, *Probabilistic analysis of a canonical numbering algorithm for graphs*, Proc. Symposia in Pure Math. vol. 34, American Mathematical Society, Providence, RI, 1979, pp. 365–378.

[8] R. J. LIPTON, *The beacon set approach to graph isomorphism*, Yale University, 1978, preprint.