COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

17,4 (1976)

PROBABILISTIC RECONSTRUCTION FROM SUBGRAPHS

Vladimír MÜLLER, Praha

Abstract: In particular, it is proved that Ulam conjecture is true with probability 1.

Key words: Finite undirected graphs, automorphisms of graphs, Ulam conjecture.

AMS: 05C05 Ref. Ž.: 8.83

Introduction: It is proved that, given $\varepsilon > 0$, asymptotically the most graphs with n vertices have all its subgraphs with at least $\frac{n}{2}(1 + \varepsilon)$ vertices asymmetric (see [1]) and mutually non-isomorphic. Particularly, from this follows that the Ulam's conjecture [4] is true with probability 1. The line analog of this result was proved in [2]. Moreover, the following stronger result holds: For every $\varepsilon > 0$ there exists n_0 such that for every $n > n_0$ the most graphs with n vertices can be uniquely reconstructed from its $\frac{n}{2}(1 + \varepsilon)$-vertex subgraphs. On the other hand, V. Nýdl (Prague, Charles University) exhibitěd in his thesis an example of two non-isomorphic graphs G, H with 2n vertices with the same collection of (n - 1)-vertex subgraphs.

We consider finite undirected graphs without loops and multiple edges. The set of vertices and the set of edges of a graph G are denoted $V(G)$ and $E(G)$, respectively.

A bijection $f\colon V(G) \longrightarrow V(H)$ is called isomorphism from graph G to graph H if $\{x,y\} \in E(G) \Longleftrightarrow \{f(x),f(y)\} \in E(H)$.

An isomorphism $f\colon G \longrightarrow G$ is called automorphism of G. In the usual sense, the term type of an automorphism is used.

A graph with n vertices will be shortly denoted n-graph. For natural numbers p,k,n, $p \geq 2$, $kp \leq n$, we shall denote $S_{k,p}(n)$ the number of all n-graphs having some automorphism of the type $(\underbrace{p,p,\ldots,p}_{k\text{-times}},1,1,\ldots,1)$.

A graph having a non-trivial automorphism is called symmetric, a graph which is not symmetric is asymmetric. Further denote $S(n)$ the number of all symmetric n-graphs and $G(n) = 2^{\binom{n}{2}}$ the number of all n-graphs.

Two statements are obvious:

1) $S(n) \leq \sum_{\substack{k\geq 1,\, p\geq 2\\ kp\leq n}} S_{k,p}(n)$

2) $S_{k,p}(n) \leq \binom{n}{p}\binom{n-p}{p}\cdots\binom{n-kp+p}{p}\frac{1}{k!}\cdot[(p-1)!]^{k}\cdot 2^{\binom{n-kp}{2}}\cdot 2^{(n-kp)k}\cdot 2^{\frac{k}{2}p}\cdot 2^{\binom{k}{2}p} = R_{k,p}(n)$ for every $k \geq 1$, $p \geq 2$, $kp \leq n$.

Let $p \geq 2$, $(l + 1)p \leq n$. It is

$$\frac{R_{k+1,p}(n)}{R_{k,p}(n)} = \frac{n-kp}{p(k+1)}\cdot(p-1)!\cdot 2^{-A} \quad , \qquad \text{where}$$

$A = np - kp^2 - n - \frac{p^2}{2} + kp.$

Lemma 1: Let $p \in N$, $p \geq 1$. Then $p! \leq 2^{\frac{p^2-1}{2}}$.

Proof: Lemma 1 can be easily proved by induction on p.

Lemma 2: Let $p \geq 2$, $k \geq 1$, $n = (k+1) \cdot p$. Then

$$\frac{R_{k+1,p}(n)}{R_{k,p}(n)} \leq 1 .$$

Proof: It is $\frac{R_{k+1,p}(n)}{R_{k,p}(n)} = \frac{1}{k+1} \cdot \frac{(p-1)!}{2^{\frac{p^2}{2}-p}} \leq \frac{1}{k+1} \leq 1$.

Remark: It holds for $p = 2$, $k \geq 1$, $n = 2(k+1) + 1$

$$\frac{R_{k+1,p}(n)}{R_{k,p}(n)} = \frac{3}{2(k+1)} \leq p$$

Lemma 3: Let either $p \geq 3$, $n \geq (k+1).p$ or $p = 2$, $n \geq 2k + 3$.

Then $$\frac{R_{k+1,p}(n+1)}{R_{k,p}(n+1)} \leq \frac{R_{k+1,p}(n)}{R_{k,p}(n)} .$$

Proof: It is $\frac{R_{k+1,p}(n+1)}{R_{k,p}(n+1)} \cdot \frac{R_{k,p}(n)}{R_{k+1,p}(n)} = \frac{n+1-kp}{n-kp-p+1} \cdot \frac{1}{2^{p-1}}$

If $p \geq 3$, $n \geq (k+1) \cdot p$ then $\frac{n+1-kp}{n-kp-p+1} \cdot \frac{1}{2^{p-1}} \leq \frac{p+1}{2^{p-1}} \leq 1$.

If $p = 2$, $n \geq 2k + 3$ then $\frac{n+1-kp}{n-kp-p+1} \cdot \frac{1}{2^{p-1}} \leq \frac{p+2}{2 \cdot 2^{p-1}} = 1$.

Corollary: Let $p \geq 2$, $k \geq 1$, $n \geq (k+1).p$. Then

$$\frac{R_{k+1,p}(n)}{R_{k,p}(n)} \leq 1 .$$

Proof: Follows immediately from the previous lemmas.

Proposition 1: Let p,k,s,n be natural numbers, $p \geq 2$, $k \geq s \geq 1$, $n \geq kp$. Then $R_{k,p}(n) \leq R_{s,p}(n)$.

Putting $k = 1$ in the definition of $R_{k,p}(n)$, we get

$$R_{1,p}(n) = \binom{n}{p}(p-1)! \cdot 2^{\binom{n-p}{2}} \cdot 2^{n-p} \cdot 2^{\frac{p}{2}} \text{ and}$$

$$\frac{R_{1,p+1}(n)}{R_{1,p}(n)} = \frac{n-p}{p+1} \cdot p \cdot 2^{-n+p+\frac{1}{2}} \qquad \text{for } n \geq p+1.$$

<u>Lemma 4:</u> Let p 2, $n = p + 1$. Then $\frac{R_{1,p+1}(n)}{R_{1,p}(n)} \leq 1$.

Proof: It is $\frac{R_{1,p+1}(p+1)}{R_{1,p}(p+1)} = \frac{p}{p+1} \cdot \frac{1}{\sqrt{2}} \leq 1$.

<u>Lemma 5:</u> Let p 2, n $p + 1$. Then

$$\frac{R_{1,p+1}(n+1)}{R_{1,p}(n+1)} \leq \frac{R_{1,p+1}(n)}{R_{1,p}(n)}.$$

Proof: It is $\frac{R_{1,p+1}(n+1)}{R_{1,p}(n+1)} \cdot \frac{R_{1,p}(n)}{R_{1,p+1}(n)} = \frac{n+1-p}{n-p} \cdot \frac{1}{2} \leq 1$.

<u>Proposition 2:</u> Let $p \geq q \geq 2$, $n \geq p$. Then $R_{1,p}(n) \leq$ $\leq R_{1,q}(n)$.

Proof: Follows easily from the lemmas 4, 5.

Using the propositions 1, 2, we get the following bound:

$$S(n) \leq \sum_{\substack{p \geq 2 \\ kp \leq n}} S_{k,p}(n) \leq \sum_{\substack{p \geq 2 \\ kp \leq n}} R_{k,p}(n) \leq R_{1,2}(n) +$$

$$+ \sum_{k=2}^{n} R_{k,2}(n) + \sum_{\substack{p \geq 3 \\ kp \leq n}} R_{k,p}(n) \leq R_{1,2}(n) + n \cdot R_{2,2}(n) +$$

$$+ n^2 \cdot R_{1,3}(n) = \binom{n}{2} 2^{\binom{n-2}{2}} \cdot 2^{n-2} \cdot 2 + \frac{n}{2}\binom{n}{2}\binom{n-2}{2} \cdot$$

$$2^{\binom{n-4}{2}} \, 2^{2n-8} \cdot 4 \cdot 4 + n^2 \binom{n}{3} \cdot 2 \cdot 2^{\binom{n-3}{2}} \, 2^{n-3} \, 2^{\frac{3}{2}} \leq$$

$$2n^2 \cdot 2^{\frac{n^2-3n}{2}} + 12 \cdot n^5 \, 2^{\frac{n^2-5n}{2}}.$$

Remark: It is clear that the number of graphs with an automorphism of the type $(2,1,\ldots,1)$ is bounded by the first term, the second term bounds the number of all other symmetric graphs. Obviously the first term is greater than the second one for n sufficiently large.

Lemma 6: Let $n \in N$, $a < \frac{1}{n}$. Then $(1 - a)^n \geq 1 - na$.

Proof: This is a well-known inequality. (It is also easy to prove by binomic development of $(1 - a)^n$.)

Lemma 7: Let $k \in N$, $k \geq 2$. Then $\frac{(k+1)^{k+1}(k-1)^{k-1}}{k^{2k}} > 1$.

Proof: It is $\frac{(k+1)^{k+1}(k-1)^{k-1}}{k^{2k}} = \left(1+\frac{1}{k}\right)^{k+1}\left(1-\frac{1}{k}\right)^{k-1} =$

$$= \left(1-\frac{1}{k^2}\right)^{k-1} \cdot \left(1+\frac{1}{k}\right)^2 \geq \left(1-\frac{k-1}{k^2}\right) \cdot \left(1+\frac{2}{k}\right) \geq \left(1-\frac{1}{k}\right)\left(1+\frac{2}{k}\right) \geq 1 .$$

Lemma 8: Let $\varepsilon > 0$, $r \in N$. Then

$$\lim_{n \to \infty} n^r \binom{n}{\left[\frac{n}{2}(1-\varepsilon)\right]} \cdot 2^{-n} = 0 .$$

Proof: It is enough to take $\varepsilon = \frac{1}{k}$ and to prove

$$\lim \frac{n^r}{2^n} \cdot \binom{2kn'+z}{\left[\frac{n}{2}\ \frac{k-1}{k}\right]} = 0$$ for every $z = 0,1,\dots$

$\dots, 2k - 1$, $n = 2kn' + z$. It is

$$\left[\frac{n}{2} \cdot \frac{k-1}{k}\right] = n'(k-1) + \left[\frac{z(k-1)}{2k}\right] = n'(k-1) + z' .$$

Denote $A_{n'} = (2kn' + z)^r \cdot \binom{2kn'+z}{n'(k-1)+z'} \cdot 2^{-(2kn'+z)}$

Let us count the limit $\lim_{n' \to \infty} \frac{A_{n'+1}}{A_{n'}} = \lim_{n' \to \infty} \frac{1}{2^{2k}} \cdot$

$$\cdot \frac{(2n'k+2k+z)\dots(2n'k+z+1)}{(n'(k-1)+z')\dots(n'(k-1)+z'+k-1)(n'(k+1)+z-z'+1)\dots(n'(k+1)+z-z'+k+1)} =$$

$$= \frac{1}{2^{2k}} \frac{(2k)^{2k}}{(k-1)^{k-1}(k+1)^{k+1}} < 1 ,$$

by the lemma 7 and by the d'Alambert's convergence criterion there is $\lim_{n'\to\infty} A_{n'} = 0$.
This proves the lemma 8.

<u>Notation</u>: Let $G = \langle V(G), E(G) \rangle$ be a graph. Denote $s(G) = \min \{ |M|, M \subset V(G)$ and $G|_{V(G)-M}$ is symmetric $\}$. (I.e. $s(G)$ is the minimal number of vertices of G, the deleting of which makes the graph symmetric.)
For $r \le n$ denote further $S^r(n)$ the number of all n-graphs G satisfying $s(G) = r$.

<u>Theorem 1</u>: Let $\varepsilon > 0$. Then

$$\lim \frac{1}{G(n)} \cdot \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S^r(n) = 0 .$$

(i.e. the most of graphs have all its subgraphs with at least $\frac{n}{2}(1 + \varepsilon)$ vertices asymmetric).

Proof: Denote $S'^r(n)$ the number of all n-graphs $G = \langle V(G), E(G) \rangle$ which satisfies $s(G) = r$ and there exists a set $M \subset V(G)$, $|M| = r$ such that the graph $G|_{V(G)-M}$ has an automorphism of the type $(2,1,\dots,1)$. Denote $S''^r(n) = S^r(n) - S'^r(n)$. It holds

$$S'^r(n) \le \binom{n}{r} 2 (n-r)^2 \, 2^{\frac{(n-r)^2 - 3(n-r)}{2}} \, 2^{\binom{n}{2}} \, 2^{r(n-r-1)} .$$

The first two terms bound the number of $(n - r)$-graphs having an automorphism of the type $(2,1,\dots,1)$ and the last two terms bound the number of all possible completions to an n-graph. In the last exponent we use the fact that $s(G)$ is exactly equal to r and not $s(G) < r$.
Further it holds

$$S''^r(n) \le \binom{n}{r} 12 (n-r)^5 \, 2^{\frac{(n-r)^2 - 5(n-r)}{2}} \, 2^{\binom{r}{2}} \, 2^{r(n-r)} .$$

and $\frac{1}{G(n)} \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S^{r}(n) = \frac{1}{G(n)} \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S'^{r}(n) + \frac{1}{G(n)} \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S''^{r}(n).$

We have $\frac{1}{G(n)} \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S''^{r}(n) \leq 12\, n^5 \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} 2^{-n+2r} \leq 12\, n^6\, 2^{-n\varepsilon}$.

The last formula is o(1). At the same time we have

$$\frac{1}{G(n)} \cdot \sum_{r=0}^{[\frac{n}{2}(1-\varepsilon)]} S'^{r}(n) \leq 2\, n^2 \sum^{[\frac{n}{2}(1-\varepsilon)]} \binom{r}{n} \cdot 2^{-n} \leq 2n^3 \binom{n}{[\frac{n}{2}(1-\varepsilon)]} \cdot 2^{-n} \longrightarrow 0$$

for $n \longrightarrow \infty$ (see the previous lemma 8).

This proves the theorem 1.

Theorem 1 cannot be improved as follows by the following proposition.

<u>Proposition 3:</u> Let $G = \langle V(G), E(G) \rangle$ be a graph, $|V(G)| = n = 2k + 1$ ($k \in N$). Then there exists a symmetric subgraph of G with at least $k + 1$ vertices.

If $n = 2k$ then there exists a symmetric subgraph of G with at least $k + 1$ vertices.

Proof: Let $G = \langle V(G), E(G) \rangle$ be a graph, $|V(G)| = n = 2k + 1$. For $x, y \in V(G)$ let us denote $d_G(x) = |\{z \in V(G), \{x,z\} \in E(G)\}|$ the degree of x in G, $d_G(x,y) = |\{z \in V(G), \{x,z\} \in E(G)$ and $\{y,z\} \in E(G)\}|$, $\bar{d}_G(x) = d_{K_n - G}(x)$, $\bar{d}_G(x,y) = d_{K_n - G}(x,y)$. It holds

$$\sum_{x \in V(G)} \binom{d_G(x)}{2} + \sum_{x \in V(G)} \binom{\bar{d}_G(x)}{2} = \sum_{\substack{x,y \in V(G) \\ x \neq y}} d_G(x,y) + \sum_{\substack{x,y \in V(G) \\ x \neq y}} \bar{d}_G(x,y) .$$

As there is $d_G(x) + \bar{d}_G(x) = n - 1$ for every $x \in V(G)$, it must be $\sum_{x \in V(G)} \binom{d_G(x)}{2} + \sum_{x \in V(G)} \binom{\bar{d}_G(x)}{2} \geq 2n \binom{\frac{n-1}{2}}{2}$ and there exist two points $x, y \in V(G)$, $x \neq y$ such that $d_G(x,y) + \bar{d}_G(x,y) \geq$

$$\geq \frac{2n \binom{\frac{n-1}{2}}{2}}{\binom{n}{2}} = \frac{n-3}{2} .$$

It means that graph G has the symmetric subgraph with $\frac{n+1}{2}$ vertices induced by the set $\{x,y\} \cup \{z \in V(G), \{z,x\} \in E(G)$ and $\{z,y\} \in E(G)\} \cup \{z \in V(G), \{z,x\} \in E(G)$ and $\{z,y\} \in E(G)\}$.

This subgraph has the non-trivial automorphism exchanging the points x and y.

Analogously, for n even there can be proved the existence of a symmetric subgraph with $\frac{n}{2} + 1$ vertices.

Let $\varepsilon > 0$, $r, n \in N$, $k = \left[\frac{n}{2}(1+\varepsilon)\right]$, $2k - n \le r \le k - 1$. Denote by $K_r(n)$ the number of n-graphs G satisfying

1) there exist two different isomorphic k-subgraphs of G having precisely r common vertices

2) all subgraphs of G with at least $\frac{n}{2}(1+\varepsilon)$ vertices are asymmetric.

<u>Theorem 2:</u> $$\lim_{n \to \infty} \frac{\sum_{r=2k-n}^{k-1} K_r(n)}{G(n)} = 0 .$$

Proof: Put $K(n) = \sum_{r=2k-n}^{k-1} \frac{K_r(n)}{G(n)}$ and

$k' = \left[\frac{n}{2}\left(1+\frac{\varepsilon}{2}\right)\right]$ (we write shortly k' instead of $k'(n)$ as well as k instead of $k(n)$). Obviously it is $K(n) = K'(n) + K''(n) + K_{k-1}(n)$, where $K'(n) =$

$$= \sum_{r=2k-n}^{k'} \frac{K_r(n)}{G(n)} \quad \text{and } K''(n) = \sum_{r=k'+1}^{k-2} \frac{K_r(n)}{G(n)} .$$

We divide the proof into three cases:

I. Let $2k - n \le r \le k'$. It holds (even for every r)

$$K_r(n) \le \binom{n}{k}\binom{k}{r}\binom{n-k}{k-r} \cdot 2^{\binom{k}{2}} \cdot k! \cdot 2^{\binom{n-2k+r}{2}} 2^{(2k-r)(n-2k+r)} 2^{(k-r)(k-r)}$$

and $\dfrac{K_r(n)}{G(n)} \leq \dfrac{\binom{n}{k}\binom{k}{r}\binom{n-k}{k-r}}{2^{\binom{k}{2}-\binom{r}{2}}} = \overline{K}_r(n)$. Further it holds

$$\frac{\overline{K}_{r+1}(n)}{\overline{K}_r(n)} = \frac{(k-r)^2}{(r+1)(n-2k+r+1)} \cdot 2^r \geq \frac{1}{n^2} \cdot 2^{2k-n} \geq \frac{1}{n^2}\, 2^{n\varepsilon-2} ,$$

hence $\overline{K}_n(n) \leq \overline{K}_{k'}(n)$ for every sufficiently large n and for every r satisfying the conditions of the case I. Hence for sufficiently large n there is

$$K'(n) \leq n\,\overline{K}_{k'}(n) = n \cdot \frac{\binom{n}{k}\binom{k}{k'}\binom{n-k}{k-k'}}{2^{\binom{k}{2}-\binom{k'}{2}}} \leq$$

$$\leq \frac{n \cdot n!\; k!}{k'!\,(k-k')!\,(n-2k+k')!} \cdot \frac{1}{2^{n^2\left(\frac{\varepsilon}{8}+\frac{3\varepsilon}{32}\right)-n\left(1+\frac{3\varepsilon}{8}\right)+1}} .$$

Obviously for sufficiently large n it is

$$K'(n) \leq \frac{n \cdot n! \cdot k!}{2^{\frac{n^2\varepsilon}{16}}} \longrightarrow 0 \qquad \text{for } n \to \infty .$$

II. Let $k' + 1 \leq r \leq k - 2$. We suppose that all subgraphs with at least $\frac{n}{2}\left(1 + \frac{\varepsilon}{2}\right)$ vertices are asymmetric. Hence

$$K_r(n) \leq \binom{n}{k}\binom{k}{r}\binom{n-k}{k-r} \cdot 2^{\binom{k}{2}} \cdot (k-r)!\binom{k}{r} \cdot$$
$$\cdot\, 2^{\binom{n-2k+r}{2}}\, 2^{(n-2k+r)(2k-r)}\, 2^{(k-r)(k-r)}$$

and $\dfrac{K_r(n)}{G(n)} \leq \dfrac{\binom{n}{k}\binom{k}{r}^2\binom{n-k}{k-r}(k-r)!}{2^{\binom{k}{2}-\binom{r}{2}}} = \overline{K}'_r(n)$.

Analogously as in case I we can derive

$$\frac{\overline{K}'_{r+1}(n)}{\overline{K}'_r(n)} = \frac{(k-r)}{(r+1)(n-2k+r+1)} \cdot 2^r \geq \frac{1}{n^2} \cdot 2^{\frac{n}{2}(1+\varepsilon)}$$

The last number is greater than 1 for every sufficiently large n and for every r satisfying the conditions of the case II. Thus for sufficiently large n it is $K_r(n) \leq$ $\leq K_{k-2}(n)$ and

$$K''(n) \leq n \cdot \overline{K}'_{k-2}(n) = n \cdot \frac{\binom{n}{k}\binom{k}{2}^2\binom{n-k}{2} \cdot 2}{2^{\binom{k}{2}-\binom{k-2}{2}}} \leq \frac{n^7}{2^{2k-n}} \leq \frac{n^7}{2^{n\varepsilon-2}}.$$

The last term tends to 0 for $n \longrightarrow \infty$.

III. Let us notice that the number of all k-graphs G satisfying

(i) all subgraphs of G with at least $k - 2$ points are asymmetric,

(ii) there exist two different isomorphic $(k - 1)$-subgraphs of G

by $k \cdot 2^{\binom{k-1}{2}} \cdot (k - 1) \cdot (k - 1) \cdot 2 \leq 2 \cdot k^3 2^{\binom{k-1}{2}}$.

Further let us notice that there is no asymmetric $(k + 1)$-graph G which satisfies:

(i) there exist two different copies of some k-graph G_1 as subgraphs of G,

(ii) all $(k - 1)$-subgraphs of G_1 are asymmetric and non-isomorphic to one another.

From these facts it follows

$$K_{k-1}(n) \leq \binom{n}{k} \cdot k \cdot (n - k) \cdot 2 \cdot k^3 \; 2^{\binom{k-1}{2}} \cdot \; 2k \cdot 2^{\binom{n-k-1}{2}} \cdot$$

$$\cdot \; 2^{(n-k-1)(k+1)} \qquad \text{and} \quad \frac{K_{k-1}(n)}{G(n)} \leq \frac{\binom{n}{k} \cdot n^6}{2^{2k}} .$$

However, the last term tends to 0 for $n \longrightarrow \infty$ (see Lemma 8). This proves the theorem 2.

From Theorems 1, 2 it follows easily:

Corollary 1: Let $\varepsilon > 0$. Then the most of n-graphs (in the sense of limit) have all its subgraphs with at least $\frac{n}{2}(1 + \varepsilon)$ vertices asymmetric and non-isomorphic to one another.

Corollary 2: Let $\varepsilon > 0$. Then the most of n-graphs (in the sense of limit $n \longrightarrow \infty$) are uniquely determined (up to isomorphism) by the family of its subgraphs with $\left[\frac{n}{2}(1+\varepsilon)\right]$ vertices.

Proof: Every graph which has all its subgraphs on $\left[\frac{n}{2}(1+\varepsilon)\right]$ vertices asymmetric and non-isomorphic to one another, has the described property.

From Corollary 2 it easily follows that the Ulam's hypothesis is true with probability 1.

References

[1] F. HARARY: Graph theory, Addison Wesley, Reading, (1969).

[2] V. MULLER: The edge reconstruction hypothesis is true for graphs with more than n log n edges (to appear in Journal of Comb. Theory (B)).

[3] J. NEŠETŘIL: On approximative isomorphisms and Ulam-Kelly conjecture, Berichte der XVIII. IWK, TH Ilmenau(1973), 17-18.

[4] S.M. ULAM: A collection of mathematical problems, Wiley (Interscience, New York, 1960).

Matematický ústav ČSAV
Žitná 25, 11000 Praha 1
Československo

(Oblatum 6.7. 1976)