Reflections on a Theme of Ulam

Ron Graham

1 Some history

The annual Southeastern International Conference on Graph Theory, Combinatorics and Computing is among the longest-running combinatorics conferences in the U.S.. Launched in 1970, it was originally held in alternate years at Louisiana State University in Baton Rouge, LA and at Florida Atlantic University (FAU) in Boca Raton, FL. However, it is now held exclusively each year at FAU under the dedicated leadership of the redoubtable Fred Hoffman. (In fact, the 48th annual Conference took place in March, 2017.) Among the many attractions of this meeting, besides the marvelous climate of Florida in March, is the laid back atmosphere that permeats the meeting environment. The beach is nearby, the tennis courts are active (and in the old days, no one could defeat Ernie Cockayne, and Steve Hedetniemi), and everyone who wants to give a talk can. (Of course, this policy can result in an interesting variety of presentations!).

Among the frequent attendees at the Southeastern Conferences were Paul Erdős, John Selfridge, Fan Chung, myself, and from time to time, Stan Ulam (along with several hundred other combinatorialists, including quite a few Canadians, whom it was rumored were instrumental in getting these meetings started in the first place). Stan (who is well known for his fundamental contributions to the Manhattan Project during World War II) had a part-time appointment at the University of Florida in Gainseville (which is reasonably close to Boca Raton), and he was also a good friend of Erdős and myself. It was during one of these meetings (the 10th Southeastern Conference, actually, in 1979) that Paul, Stan and some of us were sitting around "proving and conjecturing" (as Paul would say), when Stan mentioned an idea he had for measuring how "similar" two mathematical objects were. Namely, let X and Y be two mathematical objects (for example, graphs). Suppose you could decompose $X = X_1 \cup X_2 \cup \ldots \cup X_r$ and $Y = Y_1 \cup Y_2 \cup \ldots \cup Y_r$ into disjoint sets so that X_i and Y_i were "isomorphic"

Figure 1: (left-to-right) Paul Erdos + $\epsilon_{Mauldin}$, Dan Mauldin, Jean Larson, Stan Ulam, Francoise Ulam, Ron Graham at the University of Florida, 1979

for all i. The smaller the value of r for which this is possible, the more similar you could say X and Y were. We defined $U(X, Y)$ to be the minimum value of r for which this could be done. (Some years later, we came to call these decompositions *Ulam* decompositions. For an additional reference, see [8].)

2 Ulam decompositions

To see if this all made sense, we began by taking X and Y to be graphs, each with n vertices and the same number of edges. By "disjoint", we would mean *edge disjoint*, so that the subgraphs could share common vertices. It is clear that in this case $U(X, Y)$ was bounded above by the number of edges each graph had, since one could simply take each X_i and Y_i to be a single edge. However, since the number of edges could be as large as $\binom{n}{2}$, this decomposition only gave the bound $U(X, Y) < \frac{1}{2}n^2$. Surely, we thought, one must be able to do much better than that! What about a lower bound for $U(X, Y)$. We soon came up with the example (Figure 2) where X was a star S_{3m} with degree $3m$ and Y was m disjoint copies of a triangle K_3 together with a single isolated vertex v.

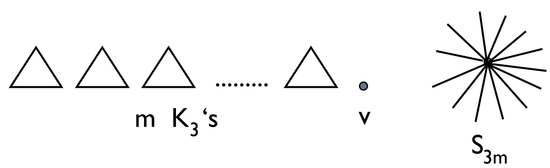

Figure 2: Two graphs with maximum Ulam value

In this case, it is not hard to see that any Ulam decomposition of X and Y must have $U(X,Y) \geq 2m = \frac{2}{3}n + o(n)$. (In general, for our bounds, we only needed that the numbers of vertices of X and Y differ by at most $o(n)$.) Was this the worst possible example?

Let us define $U_2(n)$ to be the *maximum* possible value that $U(X,Y)$ can assume when X and Y range over all possible graphs on $n + o(n)$ vertices and the same number of edges. We were eventually able to show (not at the meeting, however), that $U_2(n) = \frac{2}{3}n + o(n)$, so that the example just described is in fact as bad as it can get! This result was included in our first (joint) paper on the subject and appeared in the Proceedings of the 10th Southeastern conference [2]. The authors of this paper were Fan Chung, Paul Erdős, Stan Ulam, Frances Yao and myself. Frances Yao is a distinguished computer scientist (now living in Beijing). A natural question which came up in looking at Ulam decompositions of pairs of graphs was to determine how hard it was from a computational point of view to determine $U(X,Y)$. (In this way, Frances was drawn into the mix of people thinking about these decompositions). For example, is there an efficient algorithm to decide if $U(X,Y) = 2$? The answer is: *Probably not!* The reason for this pessimistic point of view is that Frances showed [11] that given two general graphs X and Y on n vertices with the same number of edges, it is an NP-hard problem to decide if $U(X,Y) = 2$. It is believed by most computer scientists that polynomial-time algorithms do not exist for NP-hard problems but there are some notable exceptions (see the introduction of [9]).

Given two graphs X and Y, it is perhaps even more fundamental to ask if $U(X,Y) = 1$! That is, are X and Y isomorphic? There was a major breakthrough on this classic problem recently when Babai [1] gave a quasi-polynomial algorithm for deciding this question. (The running time is $\exp((\log n)^{O(1)})$ for graphs on n vertices).

After settling the problem of Ulam decompositions for two graphs, it was inevitable that we would ask what happens for *three* graphs. That is, if X, Y and Z are three graphs on $n + o(n)$ vertices and the same number of edges, how large could $U_3(n) = \max U(X,Y,Z)$ be, where of course, $U(X,Y,Z)$ just means the least r such X, Y and Z can be decomposed into edge-disjoint graphs X_i, Y_i and Z_i such X_i, Y_i and Z_i are isomorphic. This was attacked by Fan Chung, Paul Erdős and myself a few years later (at another Southeastern Conference!), and we proved that in this case $U_3(n) = \frac{3}{4}n + o(n)$. An example of three graphs which achieve this bound is as follows: X is a star of degree n; Y is union of $\frac{n}{3}$ triangles; and Z is the union of $\frac{n+\sqrt{n}}{2}$ disjoint edges together with a complete graph on $\sqrt{n}$ vertices. Thus, this is the worst possible behavior that any three graphs can exhibit.

Well, of course, we couldn't stop there! We wondered what happens for *four* (or more) graphs? (**Spoiler alert:** $f_4(n) \neq \frac{4}{5}n + o(n)$!). Suppose for a positive integer k, we let $U_k(n)$ denote the maximum possible value of $U(G_1, G_2, \ldots, G_k)$ where the G_i range over all possible graphs on $n + o(n)$ vertices and with all having the same number of edges. We were able to prove (to our complete surprise) that $U_k(n) = \frac{3}{4}n + o(n)$ [3]. In other words, starting with four (or more) graphs, you can still find an Ulam decomposition with $\frac{3}{4}n + o(n)$ pairwise isomorphic subgraphs. The proof was quite a bit more complicated than for the case of $n = 2$. It is as if there are just three distinct "dimensions" in which graphs can differ. (Admittedly this is a somewhat vague statement!).

Finally, Fan, Paul and I went all the way (so to speak). We considered the set $\mathbf{G}(n,e)$ of *all* graphs on n vertices and e edges. We then looked at all Ulam decompositions for the set of *all* $G \in \mathbf{G}(n,e)$ and we defined $U(n,e)$ to be the maximum value that this value attains. Finally we defined $U_\infty(n)$ to be the maximum value of $U(n,e)$ over all values of e. In [4], we showed that

$$U_\infty(n) = \frac{3}{4}n + O(1).$$

Of course, this implied our earlier result on $U_k(n)$. Well, okay you say, that is well and good for graphs. But what about *hypergraphs*? Since you mentioned it, we did look at this question for hypergraphs. Of course, the situation is much more complicated. To state our results, we need another definition. For fixed k and r, we consider Ulam decompositions of k r-uniform hypergraphs $H_1, \ldots, H_k$, each having $n + o(n)$ vertices and the same (unspecified) number of hyperedges (where every hyperedge contains r vertices). Thus, $H_i = H_i(1) \cup H_i(2) \cup \ldots \cup H_i(t)$ for $1 \le i \le k$. We require that for each i, all k of the subhypergraphs $H_i(j),\ 1 \le j \le t$, be isomorphic. Let t denoted the minimum possible number of subhypergraphs in any such decomposition. Now, define $U_k(n;r)$ to be the maximum value that such a t can attain over all possible choices of k r-uniform hypergraphs (with $n + o(n)$ vertices and the same number of edges). The case of $r = 2$ corresponds to ordinary graphs. Here is what we were able to show (see [5]) (where $c_1, c_2, \ldots$ denote appropriate positive constants):

$$c_1 n^{\frac{4}{3}} \frac{\log\log n}{\log n} < U_2(n;3) < c_2 n^{\frac{4}{3}}; \tag{1}$$

$$c_3 n^{2-\frac{2}{k}-\epsilon} < U_k(n;3) < c_4 n^{2-\frac{1}{k}} \text{ for any } \epsilon > 0, \text{ provided } n > n(\epsilon); \tag{2}$$

$$c_5 n^{\frac{r}{2}} < U_2(n;r) < c_6 n^{\frac{r}{2}} \text{ for } r \text{ even}; \tag{3}$$

$$c_7 n^{\frac{(r-1)^2}{2r-1}} < U_2(n;r) < c_8 n^{\frac{r}{2}} \text{ for } r \text{ odd}; \tag{4}$$

$$n^{r-1-\frac{r}{k}} \le U_k(n;r) \le n^{r-1-\frac{1}{k}} \text{ for } r \ge 3. \tag{5}$$

Some of the techniques in proving various results concerning Ulam decompositions of graphs (and hypergraphs) involve what are called *unavoidable* graphs (and hypergraphs). We say that a graph H is *(n,e)-unavoidable* if every graph with n vertices and e edges must contain H as a (not necessarily induced) subgraph. For example, the well-known theorem of Turán [10] asserts that the complete graph K_m on m vertices is (n, e)-unavoidable if:

$$e > \frac{m-2}{2(m-1)}(n^2 - r^2) + \binom{r}{2}$$

where r satisfies $r \equiv n \mod (m-1)$ and $1 \leq r \leq m-1$.

Let $f(n, e)$ denote the largest integer m with the property that there exist an (n, e)-unavoidable graph on m edges. In the paper of Chung and Erdős [6], it is shown that:

$$f(n,e) = 1 \text{ if } e \leq \left\lfloor \frac{n}{2} \right\rfloor; \tag{6}$$

$$f(n,e) = 2 \text{ if } \left\lfloor \frac{n}{2} \right\rfloor < e \leq n; \tag{7}$$

$$f(n,e) = \left(\frac{e}{n}\right)^2 + O\left(\frac{e^3}{n^{\frac{10}{3}}}\right) \text{ if } n \leq e \leq n^{\frac{4}{3}}; \tag{8}$$

$$c_1 \frac{\sqrt{e}\log n}{\log\left(\frac{\binom{n}{2}}{e}\right)} < f(n,e) < c_2 \frac{\sqrt{e}\log n}{\log\left(\frac{\binom{n}{2}}{e}\right)}, \tag{9}$$

for some constants c_1, c_2, where $cn^{\frac{4}{3}} < e < \binom{n}{2} - n^{1+c'}$ and c and c' are between 0 and 1.

In particular, we have:

$$f(n,e) > (1 + o(1))\sqrt{(2e)} \text{ if } e \gg n^{\frac{4}{3}}, \tag{10}$$

$$f(n,e) = (1 + o(1)) \frac{\sqrt{(2e)}\log n}{\log\left(\frac{\binom{n}{2}}{e}\right)} \text{ if } n^{\frac{4}{3}} \ll e = o(n^2). \tag{11}$$

The basic technique in establishing bounds for Ulam decompositions is to sequentially remove (n, e)-unavoidable subgraphs from the G_i, depending on the current value of e. As we remove unavoidable subgraphs, the value of e keeps decreasing. These unavoidable subgraphs form the sets of isomorphic subgraphs that decompose the G_i.

In a followup paper [7], Chung and Erdős prove some analogous results for 3-uniform hypergraphs. Similar to the definition of $f(n, e)$ for graphs, we define $f_k(n,e)$ to be the largest integer m so that there is an (n,e)-unavoidable r-uniform hypergraph on m edges. In particular, it is shown that:

$$f_3(n,e) = \sqrt{\frac{e}{n}} + O(1) \ \text{ if } \ e \leq \frac{n^2}{6} - 2n; \tag{12}$$

$$c_2\sqrt{\frac{e^3}{n^5}} \leq f_3(n,e) \leq c_2\sqrt{\frac{e^3}{n^5}} \ \text{ if } \ \frac{n^2}{6} - 2n < e < n^{\frac{15}{7}}; \tag{13}$$

$$c_3\frac{e^{\frac{1}{3}}\log n}{\log\left(\frac{\binom{n}{3}}{e}\right)} < f(3(n,e) < c_3\frac{e^{\frac{1}{3}}\log n}{\log\left(\frac{\binom{n}{3}}{e}\right)} \ \text{ if } \ n^{\frac{15}{7}} < e. \tag{14}$$

As should be expected, the proofs for hypergraphs are more complex and the results are more incomplete than the corresponding results for graphs.

3 Some questions

Concerning Ulam decompositions of graphs, there are many things we still don't know. Suppose we restrict our candidate graphs G_i in some way. For example, suppose that the G_i are all bipartite. Then it turns out that the corresponding functions $U'_k(n)$ satisfy $U'_2(n) = \frac{1}{2}n + o(n)$, and $U'_k(n) = \frac{3}{4}n + o(n)$ for $k \geq 3$. Note that here, only the value of $U'_2(n)$ decreased from the value of $U_2(n) = \frac{2}{3}n + o(n)$ for general graphs.

One could ask for more precise estimates for Ulam decompositions. For example, we could define $U_k(n,e)$ to be the value of the corresponding parameter when the graphs in question all have exactly e edges. $U_k(n)$ is just the maximum value of $U_k(n,e)$ as e ranges over all possible number of edges. How does $U_k(n,e)$ change as e goes from 0 to $\binom{n}{2}$?

Of course, the analogous questions apply to the hypergraph problems. Here, we are even more in the dark. For example, does $U_k(n,3)$ differ substantially from $U_k(n,4)$?

From a broader perspective, we can follow Stan Ulam's initial instincts and look beyond graphs and hypergraphs to other mathematical (and non-mathematical) structures, such as directed graphs, posets, topological spaces, functions, chemical compounds, proteins, and in fact, almost anything! This is something I'm sure that Stan would encourage and appreciate!

References

[1] L. Babai, Graph Isomorphism in Quasipolynomial Time, (89 pp.) arXiv:1512.03547

[2] F. R. K. Chung, P. Erdős, R. L. Graham, S. M. Ulam and F. F. Yao, Minimal Decompositions of Two Graphs into Pairwise Isomorphic Subgraphs, *Proc.* 10^{th} *S-E Conf. on Comb., Graph Theory and Comp. Congressus Numerantium* **23** (1979), 3–18.

[3] F. R. K. Chung, P. Erdős and R. L. Graham, Minimal Decompositions of Graphs into Mutually Isomorphic Subgraphs, *Combinatorica*, **11** (1981), 13–24.

[4] F. R. K. Chung, P. Erdős and R. L. Graham, Minimal Decomposition of All Graphs with Equinumerous Vertices and Edges into Mutually Isomorphic Subgraphs, *Colloq. Math. Soc. Janos Bolyai*, **37**, *Finite and Infinite Sets*, Eger, Hungary, (1981), 181–222.

[5] F. R. K. Chung, P. Erdős and R. L. Graham, Minimal Decompositions of Hypergraphs into Mutually Isomorphic Subhypergraphs, *Jour. Comb. Theory (A)* **32** (1982), 241–251.

[6] F. R. K. Chung and P. Erdős, On Unavoidable Graphs, *Combinatorica* **3** (1983), 167–176.

[7] F. R. K. Chung and P. Erdős, On Unavoidable Hypergraphs, *J. Graph Theory*, **11** (1987), 251–263.

[8] R. L. Graham, A Similarity Measure for Graphs – Reflections on a Theme of Ulam, *Los Alamos Science* No. 15, (special issue) (1987), 114–121.

[9] D. E. Knuth, *The Art of Computer Programming*, Volume 4 Fascicle 6, Satisfiability (2015), xiii+310pp.

[10] P. Turán, On the theory of graphs, *Coll. Math.* **3** (1954), 19–30.

[11] F. F. Yao, Graph 2-isomorphism is NP-complete, *Inform. Process. Lett.* **9** (1979), 68–72.