Reflections on a Problem of Erdős and Hajnal

András Gyárfás*

Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, Hungary

Summary. We consider some problems suggested by special cases of a conjecture of Erdős and Hajnal.

1. Epsilons

The problem I am going to comment reached me in 1987 at Memphis in a letter of Uncle Paul. He wrote: 'We have the following problem with Hajnal. If $G(n)$ has n points and does not contain induced C_4, is it true that it has either a clique or an independent set with n^{ϵ} points? Kind regards to your boss + colleagues, kisses to the ϵ -s: E.P.' After noting that ϵ have been used in different contexts I realized soon that $\frac{1}{3}$ is a good ϵ (in both senses since I have three daughters). About a month later Paul arrived and said he meant C_5 for C_4. And this minor change of subscript gave a problem still unsolved. And this is just a special case of the general problem formulated in the next paragraph.

2. The Erdős-Hajnal problem (from [7])

Call a graph H-free if it does not contain induced subgraphs isomorphic to H. Complete graphs and their complements are called *homogeneous sets.* As usual, $\omega(G)$ and $\alpha(G)$ denotes the order of a maximum clique and the order of a maximum independent set of G. It will be convenient to define $hom(G)$ as the size of the largest homogeneous set of G, i.e. $hom(G) = max\{\alpha(G), \omega(G)\}$ and

$$hom(n, H_1, H_2, ...) = min\{hom(G) : |V(G)| = n, G \text{ is } H_i\text{-free}\}$$

A well-known result of Paul Erdős ([5]) says that there are graphs of n vertices with $hom(G) \leq 2\log n$ (log is of base 2 here). The following problem of Erdős and Hajnal suggests that in case of forbidden subgraphs $hom(G)$ is much larger: Is it true, that for every graph H there exists a positive ϵ and n_0 such that every H-free graph on $n \geq n_0$ vertices contains a homogeneous set of n^{ϵ} vertices? If such ϵ exists for a particular H, one can define the 'best' exponent, $\epsilon(H)$ for H as

$$\epsilon(H) = sup\{\epsilon > 0 : hom(n, H) \geq n^{\epsilon} \text{ for } n \geq n_0\}$$

The existence of $\epsilon(H)$ is proved in [7] for P_4-free graphs (usually called *cographs* but in [7] the term *very simple graphs* have been used). In fact, a stronger statement is proved in [7]: if $\epsilon(H_i)$ exists for $i = 1, 2$ and H is a graph formed by putting all or no edges between vertex disjoint copies of H_1 and H_2 then $\epsilon(H)$ is also exists. Combining this with the well known fact that P_4-free graphs are perfect ([14]), it follows that $\epsilon(H)$ exists for those graphs H which can be generated from the one-vertex graph and from P_4, using the above operations. In the spirit of [7], call this

* Supported by OTKA grant 2570

class SVS (still very simple). In terms of graph replacements (see 4. below), SVS is generated by replacements into two-vertex graphs starting from K_1 and P_4. As far as I know, the existence of $\epsilon(H)$ is not known for any graph outside SVS.

3. Large perfect subgraphs

A possible approach to find a large homogeneous set in a graph is to find a large perfect subgraph. It was shown in [7] that any H-free graph of n vertices has an induced cograph of at least $e^{c(H)\sqrt{\log n}}$ vertices for sufficiently large n. This shows that the size of the largest homogeneous set makes a huge jump in case of any forbidden subgraph. In a certain sense it is not so far from $n^{c(H)}$... What happens if cographs are replaced by other perfect graphs? A deep result of Prömel and Steger ([13]) says that almost all C_5-free graphs are perfect (generalized split graphs). This suggest the possibility to find a large (n^{ϵ}) generalized split graph in a C_5-free graph of n vertices and prove the existence of $\epsilon(C_5)$ this way.

4. Replacements

A well-known important concept in the theory of perfect graphs is the replacement of a vertex by a graph. The replacement of vertex x of a graph G by a graph H is the graph obtained from G by replacing x with a copy of H and joining all vertices of this copy to all neighbors of x in G. According to a key lemma (Replacement Lemma) of Lovász (see for example in [12]), perfectness is preserved by replacements. The property of being H-free is obviously preserved by replacements if (and in some sense only if) H can not be obtained from a smaller graph by a nontrivial (at least two-vertex) replacement. For such an H, replacements can be applied to get an upper bound on $\epsilon(H)$. Analogues of the Replacement Lemma can be also useful to find large homogeneous sets (an example is Lemma 7.1 below).

5. Partitions into homogeneous sets

For certain graphs H, the existence of $\epsilon(H)$ follows from stronger properties. An H-free graph G may satisfy $\chi(G) \leq p(\omega(G))$ or $\theta(G) \leq p(\alpha(G))$ or more generally $cc(G) \leq p(\alpha(G), \omega(G))$ where p is a polynomial of constant degree and $\chi, \theta, \alpha, \omega, cc$ denote the chromatic number, clique cover number, independence number, clique number and cochromatic number of graphs. Using terminology from [10], p is a *polynomial binding function* (for χ, θ, cc, respectively). It is clear that if H-free graphs have a polynomial binding function of degree k then $\epsilon(H) \geq \frac{1}{k+1}$. Binding functions for χ (for θ) in H-free graphs may exist only if H ($\overline{H}$) is acyclic. However, the existence of a polynomial binding function for cc in H-free graphs is equivalent with the existence of $\epsilon(H)$.

6. Small forbidden subgraphs

The existence of $\epsilon(H)$ follows if H has at most four vertices since these graphs are all in SVS. But, as will be shown below, to find $\epsilon(H)$ for these small graphs is not always that simple...

Since $hom(n,H) = hom(n,\overline{H})$ from the definition, it is enough to consider one graph from each complementary pair. For $H = K_m$, finding $hom(n,H)$ is the classical Ramsey problem. In case of $m = 2,3,4$, $\epsilon(K_2) = 1$ (trivial), $\epsilon(K_3) = \frac{1}{2}$ (from Uncle Paul's lower bound on $R(3,m)$ in [4]), $\frac{1}{3} \le \epsilon(K_4) \le 0.4$ (the upper bound is due to Spencer [15]). If $H = P_3$ or $H = P_4$ then an H-free graph is perfect and thus $\epsilon(H) = \frac{1}{2}$.

There are four more graphs with four vertices to look at. Let H_1 be $K_{1,3}$, the *claw*, and let H_2 be K_3 with a pendant edge. It is not difficult to see that $\epsilon(H_i) = \frac{1}{3}$ in this case. The construction is simple: let G be a graph on m vertices with no independent set of three vertices and with no complete subgraph of much more than $\sqrt{m}$ vertices ([4]). Take about $\frac{\sqrt{m}}{2}$ disjoint copies of G. This graph is H_1-free, has $\frac{m^{\frac{3}{2}}}{2}$ vertices and has no homogeneous subset with much more than $m^{\frac{1}{2}}$ vertices. The complement of this graph is good for H_2. On the the other hand, let G be an H_i-free graph with n vertices (i is 1 or 2). If the degree of a vertex v is at least $n^{\frac{2}{3}}$ then $\Gamma(v)$ (the set of vertices adjacent to v) contains a homogeneous set of at least $n^{\frac{1}{3}}$ vertices (in case of H_1 by Ramsey's theorem, in case of H_2 by perfectness). Otherwise G has an independent set of at least $n^{\frac{1}{3}}$ vertices. This gives

Proposition 6.1. *If H is the claw or K_3 with a pendant edge then $\epsilon(H) = \frac{1}{3}$.*

The remaining two H-s are the C_4 and K_4 minus an edge (the *diamond*). The following argument is clearly discovered by many of us, could be heard from Uncle Paul too. It was used for example in [9], [16]. Let $S = \{v_1, \ldots, v_\alpha\}$ be a maximum independent set of a C_4-free or diamond-free graph G. Then $V(G)$ is covered by the following $\binom{\alpha+1}{2}$ sets: $A(i)$, $B(i,j)$, $1 \le i < j \le \alpha$ where

$$A(i) = \{v \in V(G) - S : \Gamma(v) \cap S = \{v_i\}\} \cup \{v_i\}$$

and

$$B(i,j) = \{v \in V(G) - S : \Gamma(v) \cap S \supseteq \{v_i, v_j\}\}$$

The sets $A(i)$ induce complete subgraphs by the maximality of S and the sets $B(i,j)$ induce homogeneous sets (complete if G is C_4-free, independent if G is diamond-free). This gives

Proposition 6.2. *If G is C_4-free or diamond-free then $cc \le \binom{\alpha+1}{2}$.*

Corollary 6.1. *If H is either C_4 or the diamond then $hom(n,H) \ge (2n)^{\frac{1}{3}}$. Therefore $\epsilon(H) \ge \frac{1}{3}$.*

Vertex disjoint union of complete graphs shows that $\epsilon(H) \le \frac{1}{2}$ for any connected graph H. In case of $H = C_4$ this upper bound can be improved as follows. Let $R(C_4,m)$ be the smallest integer k such that any graph on k vertices either contains a C_4 (not necessarily induced C_4!) or contains an independent set on m vertices. F.K.Chung gives a graph G_m in [3] which shows that $R(C_4,m) \ge m^{\frac{4}{3}}$ for infinitely many m. Replacing each vertex of G_m by a clique of size $\frac{m}{3}$ we have a graph with no induced C_4 and with no homogeneous subset larger than m. This gives

Proposition 6.3. $\epsilon(C_4) \le \frac{3}{7}$

Notice that if $R(C_4,m) \ge m^{2-\epsilon}$ with every $\epsilon > 0$ as asked by Uncle Paul then the replacement described above would show that $\epsilon(C_4) = \frac{1}{3}$.

Perhaps the next construction has a chance to improve the upper bound on $\epsilon(H)$ if H is the diamond. The vertices of G_q are the points of a *linear complex*

([11]) of a 3-dimensional projective space of order q. Two points are adjacent if and only if they are on a line of the linear complex. The graph G_q has q^3+q^2+q+1 vertices, $\omega(G_q)=q+1$ and G_q is diamond-free. But is it true that $\alpha(G_q)<q^{\frac{3}{2}-\epsilon}$ for some positive ϵ and for infinitely many q? Thanks for the conversations to T.Szőnyi who thinks this is not known.

Problem 6.1. Improve the exponents in the above estimates of $hom(n,H)$ if H is either C_4 or the diamond.

What happens if G is C_4-free *and* diamond-free? In these graphs each four cycle induces a K_4. The sets $B(i,j)$ collapse implying

Corollary 6.2. $hom(n, C_4, Diamond) \geq \sqrt{\frac{2}{3}n}-1$

Problem 6.2. Is it true that $hom(n, C_4, Diamond) = \sqrt{n}+o(\sqrt{n})$?

During the years between the submission and publication of this paper, Problem 6.2 had been answered affirmatively, in fact $hom(n, C_4, Diamond) = \lceil\sqrt{n}\rceil$ ([6]).

There are eight five-vertex graphs outside SVS. Keeping one from each complementary pair reduces the eight to five: $K_{1,3}$ with a subdivided edge, the *bull* (the self-complementary graph different from C_5), C_4 with a pendant edge, P_5, C_5. The existence of $\epsilon(H)$ is open for all of them, perhaps the list is about in the order of increasing difficulty. The construction in Proposition 1 shows that $\epsilon(H)$ is at most $\frac{1}{3}$ for all but C_5. In case of C_5 repeated replacements of $\overline{C_7}\cup K_3$ into itself shows $\epsilon(C_5)\leq\frac{\log 3}{\log 10}$. (Any C_5-free graph G with $hom(G)=3$ and with at least 11 vertices would improve this.)

7. Forbidden complementary pairs

Perhaps an interesting subproblem is to find bounds on $hom(n,H,\overline{H})$. In case of four-vertex H, the structure of graphs which are both H-free and $\overline{H}$-free is well understood and values of $hom(n,H,\overline{H})$ can be determined as follows: $n^{\frac{1}{2}}$ if $H=P_4$ (from perfectness); $n^{\frac{1}{2}}-1$ if $H=P_3+K_1$ (from structure, [10]); $\frac{n-1}{2}$ if $H=C_4$ (from structure, [2]); $n-4$ if H is a diamond (from structure, [10]); $\frac{2n}{5}$ if H is a claw (from structure, [10]).

The rest of this section is devoted to the case $H=P_5$. The upper bound $hom(n,P_5,\overline{P_5})\leq n^{\frac{1}{\log 5}}$ is shown by replacing repeatedly C_5 into itself. The lower bound $n^{\frac{1}{3}}$ will follow from Corollary 7.1 which is the consequence of the following result.

Theorem 7.1. *If G is P_5-free and $\overline{P_5}$ -free then G satisfies the following property SP^*: there is an induced perfect subgraph of G whose vertices intersect all maximal cliques of G.*

Notice that property SP^* is a generalization of *strong perfectness* introduced by Berge and Duchet in [1]. (Maximal clique is a clique which is not properly contained in any other clique.) By Theorem 7.1, if G is both P_5-free and $\overline{P_5}$-free then G can be partitioned into at most $\omega(G)$ vertex disjoint perfect subgraphs. Each of these perfect graphs has clique number at most $\omega(G)$ thus each has chromatic number at most $\omega(G)$. This gives the next corollary.

Corollary 7.1. *If a graph is both P_5-free and $\overline{P_5}$-free then $\chi \le \omega^2$.*

The proof of Theorem 7.1 is combining a result of Fouquet [8] with the following analogue of Lovász replacement lemma.

Lemma 7.1. *Property SP^* is preserved by replacements.*

Proof. The proof of the lemma is along the same line as the replacement lemma of Lovász. Assume that G and H have property SP^* and R is the graph obtained by replacing $v \in V(G)$ by H. Let G_1 and H_1 be perfect subgraphs of G and H such that $V(G_1)$ intersects all maximal cliques of G and $V(H_1)$ intersects all maximal cliques of H.

Case 1. $v \notin V(G_1)$. We claim that $V(G_1)$ intersects all maximal cliques of R. Let K be a maximal clique of R. If $V(K) \cap V(H)$ is empty then the claim follows from the definition of G_1. Otherwise $\{v\} \cup (K \cap V(G))$ is a clique of G which can be extended in G to a maximal clique K' intersecting $V(G_1)$. Since K is obtained by replacing $v \in K'$ by $K \cap V(H)$, K intersects $V(G_1)$.

Case 2. $v \in V(G_1)$. By the Lovász replacement lemma, the subgraph Z of R induced by $(V(G_1) \cup V(H_1)) - \{v\}$ is perfect. If a maximal clique K of R intersects $V(H)$, it intersects it in a maximal clique of H which (by the definition of H_1) intersects $V(H_1)$. If K does not intersect H then it does not contain v so it intersects $V(G_1) - \{v\}$ by the definition of G_1. Therefore K intersects Z.

Theorem 7.2 ((Fouquet, [8])). *Each graph from the family of P_5-free and $\overline{P_5}$-free graphs is either perfect or isomorphic to C_5 or can be obtained by a nontrivial replacement from the family.*

Now Theorem 7.1 follows by induction from Lemma 7.1 and Theorem 7.2.

8. Berge graphs

These are graphs which do not contain induced subgraphs isomorphic to C_{2k+1} or to $\overline{C_{2k+1}}$ for $k \ge 2$. According to the Strong Perfect Graph Conjecture (of Berge), Berge graphs are perfect. The following weaker form of this conjecture is attributed to Lovász in [7] (illustrating the difficulty of proving the existence of $\epsilon(C_5)$).

Problem 8.1. There exists a positive constant ϵ such that Berge graphs with n vertices contain homogeneous subsets of n^ϵ vertices.

Similar problems can be asked for subfamilies of Berge graphs for which the validity of SPGC is not known. One of them is the following.

Problem 8.2. Show that C_4-free Berge graphs with n vertices contain homogeneous subsets of $n^{\frac{1}{2}}$ (or at least $cn^{\frac{1}{2}}$) vertices.

Thanks to an unknown referee for valuable remarks.

References

1. C.Berge, D.Duchet, Strongly perfect graphs, in Topics on Perfect graphs, Annals of Discrete Math. Vol 21 (1984) 57-61
2. Z.Blázsik, M.Hujter, A.Pluhár, Zs.Tuza, Graphs with no induced C_4 and $2K_2$, Discrete Math. 115 (1993) 51-55.
3. F.R.K.Chung, On the covering of graphs, Discrete Math. 30 (1980) 89-93.
4. P.Erdős, Graph Theory and Probability II., Canadian J.Math. 13, (1961) 346-352.
5. P.Erdős, Some Remarks on the Theory of Graphs, Bulletin of the American Mathematics Society, 53 (1947), 292-294.
6. P.Erdős, A.Gyárfás, T.Luczak, Graphs in which each C_4 spans K_4, submitted.
7. P.Erdős, A.Hajnal, Ramsey Type Theorems, Discrete Applied Math. 25 (1989) 37-52.
8. J.L.Fouquet, A decomposition for a class of $(P_5, \overline{P_5})$-free graphs, preprint.
9. A.Gyárfás, A Ramsey type theorem and its applications to relatives of Helly's theorem, Periodica Math. Hung. 3 (1973) 299-304.
10. A.Gyárfás, Problems from the world surrounding perfect graphs, Zastowania Matematyki, Applicationes Mathematicae, XIX 3-4 (1987) 413-441.
11. Hirschfeld, Finite Projective Spaces of three dimensions, Clarendon Press, Oxford, 1985.
12. L.Lovász, Perfect Graphs, in Selected Topics in Graph Theory 2. Academic Press (1983) 55-87.
13. H.J.Prömel, A.Steger, Almost all Berge graphs are perfect, Report 91715, Forschungsinstitut fur Diskrete Mathematik, Bonn.
14. D.Seinsche, On a property of the class of n-colorable graphs, Journal of Combinatorial Theory B. 16 (1974) 191-193.
15. J.Spencer, Ten lectures on the Probabilistic method, CBMS-NSF Conference Series, 52.
16. S.Wagon, A bound on the chromatic number of graphs without certain induced subgraphs, J.Combinatorial Theory B. 29 (1980) 345-346.