Electronic Colloquium on Computational Complexity, Report No. 13 (2012)

Packing Edge-Disjoint Triangles in Given Graphs

Tomás Feder* Carlos Subi†

Abstract

Given a graph G, we consider the problem of finding the largest set of edge-disjoint triangles contained in G. We show that even the simpler case of decomposing the edges of a sparse split graph G into edge-disjoint triangles is NP-complete. We show next that the case of a general G can be approximated within a factor of $\frac{3}{5}$ in polynomial time, and is NP-hard to approximate within some constant $1-\epsilon$, $0<\epsilon<1$. We generalize this to the question of finding the largest set of edge-disjoint copies of a fixed graph H in G, which we approximate within $\frac{2}{|E(H)|+1}$ in polynomial time, and show hard to approximate when $H=K_r$ or $H=C_r$ is a fixed clique or cycle on at least three vertices. This relates to a problem solved by Kirkpatrick and Hell.

We finally determine optimal solutions for the case where G is a clique of any size by examining a family of dense split graphs. The motivation for the case of cliques is the theory of block designs, in particular the case of Steiner triple systems, and a conjecture of Erdös, Faber, and Lovasz which states that the union of n edge-disjoint cliques of size n has vertex chromatic number n. The connection to this conjecture is obtained through a dualization argument.

1 Introduction

We begin with a known result from the theory of block designs, namely a characterization of the existence of Steiner triple systems. This result was first shown by Kirkman (1847) (see e.g. [3], p. 113), and our proof will lead from the known Theorem 1 to the more general approximation result, Theorem 2.

Theorem 1 *A clique K_n with $n\geq 1$ vertices can have its edges decomposed into edge-disjoint triangles if and only if $n=6x+1$ or $n=6x+3$, $x\geq 0$ integer.*

We call numbers n *valid* if K_n can have its edges decomposed into edge-disjoint triangles. We call numbers n *good* if $n=6x+1$ or $n=6x+3$, $x\geq 0$ integer. Thus Theorem 1 says that n is valid if and only if n is good.

The motivation for this result is a conjecture of Erdös, Faber, and Lovasz, that states that if a graph G is the union of n edge-disjoint cliques K^i, $1\leq i\leq n$, where each K^i has at most n vertices, then G has vertex chromatic number at most n.

We shall not study this conjecture, but just examine the structure of some the graphs involved in the conjecture. In the conjecture, we may assume that each vertex in G belongs to at least two K^i; otherwise remove vertices in G that belong to only one K^i, and if the vertices of the resulting graph can be colored with n colors then the removed vertices can be colored as well, since each K^i has at most n vertices. We may as well assume next that any two K^i, K^j share a vertex; otherwise

*268 Waverley Street, Palo Alto, CA 94301. Email: tomas@theory.stanford.edu

†Email: carlos.subi@hotmail.com

ISSN 1433-8092

add a new vertex v_{ij} shared by these two cliques, and the sizes of the cliques will never exceed $n-1$.

Consider now the dual clique graph $H = K_n$ whose vertices correspond to cliques in G. If a vertex v in G belongs to $r_v \geq 2$ cliques K^i, then v corresponds to a clique K_v of size r_v in H. The cliques K_v partition the edges of the clique H. Furthermore v, w are independent in G if and only if K_v, K_w are vertex-disjoint in H. If $r_v = 2$ for all v, then the conjecture is easy to prove (just partition the dual clique graph H into at most n matchings, see the Walecki construction [1, 6], see Lemma 1), so one can assume at least one $r_v \geq 3$.

We ask the following question. For which values of n is it possible that each v belongs to exactly $r_v = 3$ of the n cliques in G. This is equivalent to partitioning the edges of $H = K_n$ into triangles K_v. Theorem 1 answers this question by showing that the values of n are those for which n is good, while Theorem 2 below implicitly characterizes which combinations of numbers of $r_v = 2$ and $r_v = 3$ are possible, with no $r_v \geq 4$.

We consider next a class of graphs larger than that of cliques, namely the class of split graphs, which is a special case of chordal graphs. A *split graph* is a graph G whose vertices can be partitioned into two disjoint parts R and S, such that R induces an independent set in G and S induces a clique in G. We say that a split graph G is *sparse* if each vertex in R has degree two in G, and we say that a split graph G is *dense* if each vertex in R is adjacent to all of S.

We show that the problem of determining if a sparse split graph can have its edges partitioned into edge-disjoint triangles is NP-complete, and attempt to characterize the values of $r = |R|$ and $s = |S|$ for which the corresponding dense split graph can have its edges partitioned into edge-disjoint triangles. This leads to the following result.

Theorem 2 *The maximum number of edge-disjoint triangles contained in a clique K_n, $n = 6x + i, 0 \leq i \leq 5$ is $(\binom{n}{2} - k)/3$, where the number k of edges of K_n not covered by the chosen triangles is $k = 0$ for $i = 1$ and $i = 3$, $k = 4$ for $i = 5$, $k = \frac{n}{2}$ for $i = 0$ and $i = 2$, and $k = \frac{n}{2} + 1$ for $i = 4$.*

We also show that the NP-complete question of finding the largest set of edge-disjoint triangles contained in a given graph G can be approximated in polynomial time within a factor of $\frac{1}{3}$ by a greedy algorithm, and within a factor of $\frac{1}{2}$ by an augmentation algorithm. The approximation factor with a more complex augmentation improves to $\frac{3}{5}$. We next show that this question is NP-hard to approximate within some constant $1 - \epsilon$, $0 < \epsilon < 1$, even in the case where the edges of G can be partitioned into edge-disjoint triangles.

We then generalize the problem to the question of finding the largest set of edge-disjoint copies of a fixed graph H with $e = |E(H)|$ edges contained in a given graph G. We approximate this problem within $\frac{1}{e}$ by a greedy algorithm, and within $\frac{2}{e+1}$ by an augmenting algorithm, in polynomial time. We finally show that this question is NP-hard to approximate within some constant $1 - \epsilon$, $0 < \epsilon < 1$, if $H = K_r$ or $H = C_r$ is a clique or cycle with $r \geq 3$ vertices.

This relates to a problem shown NP-complete by Kirpatrick and Hell [5]. Given graphs G and H with $|V(G)| = q|V(H)|$, can the vertices of G be partitioned into q disjoint sets V_i, $1 \leq i \leq q$, such that the subgraph of G induced by each V_i is isomorphic to H? This question remains NP-complete for any fixed H that contains at least 3 vertices. The analogous problem in which the subgraphs induced by V_i need only have the same number of vertices as H and contain a subgraph isomorphic to H is also NP-complete for any fixed H that contains a connected component of three or more vertices. Both problems can be solved in polynomial time (by matching) for any H not meeting the stated restrictions.

2 Proof of Theorem 1

The following is known as the Walecki construction [1, 6].

Lemma 1 *The edges of a clique K_n can be decomposed into $n-1$ edge-disjoint perfect matchings if $n = 2k$ is even, and can be decomposed into n edge disjoint matchings that each match all but one vertex if $n = 2k-1$ is odd.*

Proof. Label the vertices $0, 1, \ldots, 2k-2$ modulo $2k-1$, plus an additional vertex v if n is even. The $2k-1$ matchings M_i, $0 \leq i < 2k-1$, have edges $(i+j, i-j)$ modulo $2k-1$ for $1 \leq j \leq k-1$, plus an additional edge (i, v) if n is even. □

We now proceed to the proof of Theorem 1.

Lemma 2 *If $n \geq 1$ is valid, then n is good.*

Proof. Suppose n is valid. If n is even, then each vertex v of K_n has an odd number $n-1$ edges coming out of v. But these $n-1$ edges are paired by triangles, so $n-1$ is even, a contradiction. So n is odd.

If $n = 3k+2$, then the number $\binom{n}{2} = \frac{n(n-1)}{2}$ of edges in K_n is not divisible by 3, but the partition of the edges of K_n into triangles implies that the number of edges in K_n is divisible by 3, a contradiction. So $n = 3k$ or $n = 3k+1$. Furthermore $n = 3k$ is odd only if $k = 2x+1$ is odd, in which case $n = 6x+3$, and $n = 3k+1$ is odd only if $k = 2x$ is even, in which case $n = 6x+1$. Therefore n must be good. □

Lemma 3 *If $n = a$ is valid, then $2a+1$ is also valid.*

Proof. Decompose the vertices of K_{2a+1} into a set S of size a and a set T of size $a+1$. Since a is good because it is valid, we have that a is odd and $a+1$ is even. We first decompose the edges of the clique induced by S into edge-disjoint triangles, since $a = |S|$ is valid. We then decompose the edges of the clique induced by T with $a+1 = |T|$ even into a perfect matchings M_i, $1 \leq i \leq a$ (see the Walecki construction [1, 6], Lemma 1). We finally label the vertices of S by v_i, $1 \leq i \leq a$, and join each v_i with each edge of the corresponding M_i to form $\frac{a+1}{2}$ extra triangles for each v_i. □

Lemma 4 *If $n = b$ is valid, then $3b$ is also valid.*

Proof. Decompose the vertices of K_{3b} into three sets S, T, U of size $|S| = |T| = |U| = b$. The cliques induced by S, T, U are first decomposed into edge-disjoint triangles, since b is valid. Next label the vertices of S, T, U as s_i, t_i, u_i respectively for $0 \leq i < b$, and add triangles (s_i, t_j, u_k), where k is $i+j$ modulo b. □

Lemma 5 *If $n = c$ is valid, then $3c-2$ is also valid.*

Proof. Let $c = b+1$, and decompose the vertices of $K_{3c-2} = K_{3b+1}$ into three sets S, T, U of size $|S| = |T| = |U| = b$ plus a vertex v. The cliques induced by $S \cup \{v\}, T \cup \{v\}, U \cup \{v\}$ are first decomposed into edge-disjoint triangles, since $c = b+1$ is valid. Next label the vertices of S, T, U as s_i, t_i, u_i respectively for $0 \leq i < b$, and add triangles $(s_i, t_j, u_k$, where k is $i+j$ modulo b. □

Lemma 6 *$n = 13$ is valid.*

Proof. Let the vertices of K_{13} be $0, i, i', i''$ for $1 \le i \le 4$. Then the triangles are (1) $11'1''; 01'2'', 01''3'', 014''$; (2) $022', 12'2'', 1''22''; 033', 1'3'3'', 133''; 044', 1''4'4'', 1'44''$; (3) $1'23, 1''34, 124; 1''2'3', 13'4', 1'2'4'; 2''3''4''$; (4) $23'4'', 2'3''4, 2''34'; 23''4', 2'34'', 2''3'4$. □

Lemma 7 *If $n = d$ is valid, then $6d - 5$ is also valid.*

Proof. Let $d = f + 1$, and decompose the vertices of $K_{6d-5} = K_{6f+1}$ into six sets S^i, $1 \le i \le 6$, of size $|S^i| = f$, plus a vertex v. The cliques induced by $S^i \cup \{v\}$ are first decomposed into edge-disjoint triangles, since $d = f + 1$ is valid. Next since d is good and odd, f is even, and we may decompose each S^i into two sets of size $g = f/2$. Collapsing each such set of size g to a single super-vertex, we have that each S^i consists of two super-vertices, plus v, for a total of 13 super-vertices. By Lemma 6, we can decompose K_{13} into edge-disjoint triangles of super-vertices, including the triangles joining v with the two super-vertices of each S^i. The remaining triangles are of the form ABC where each super vertex has size $|A| = |B| = |C| = g$. We expand these three super-vertices for a triangle to sets of size g, and the super-triangle is replaced by g^2 triangles with one vertex from each of A, B, C as in Lemma 4. □

Lemma 8 *If $n = e = 6x + 3$ is valid, then $2e + 3$ is also valid.*

Proof. Decompose K_{2e+3} into a set S of size e and a set T of size $e + 3$. Decompose S into $f = e/3 = 2x + 1$ triples S_i, $1 \le i \le f$, and Decompose T into $f + 1$ triples T_j, $0 \le i \le f$. The clique induced by S is first decomposed into edge-disjoint triangles, since e is valid. Next each triangle T_j is chosen. Finally decompose the clique of size $f + 1$ even whose super-vertices are the T_j into f perfect matchings (by the Walecki construction [1, 6], Lemma 1), and if the ith such matching contains the super-edge T_j, T_k, choose the super-triangle S_i, T_j, T_k. When each such super-vertex is replaced by three vertices, the super-triangles S_i, T_j, T_k can be replaced by 9 triangles as in Lemma 4. □

Lemma 9 *If $n \ge 1$ is good, then n is valid.*

Proof. We proceed by induction on n. The base case is $n = 1$, without edges.

Assume first $n = 6x + 1$. If $x = 2r + 1$ is odd, then $a = 3x = 6r + 3$ is valid and the result for $n = 2a + 1$ follows from Lemma 3. If $x = 2r$ is even, then $n = 12r + 1$. Then $n = 3c - 2$, for $c = 4r + 1$. If $r = 3s$, then $c = 12s + 1$ is valid and the result follows for $n = 3c - 2$ from Lemma 5. If $r = 3s + 2$ then $c = 12s + 9$ is valid and the result follows for $n = 3c - 2$ from Lemma 5. Finally if $r = 3s + 1$ then $n = 36s + 13 = 6d - 5$ for $d = 6s + 3$ valid and the result follows for $n = 6d - 5$ from Lemma 7.

Assume next $n = 6x + 3$. If $x = 2r$ is even, then $a = 3x + 1 = 6r + 1$ is valid and the result for $n = 2a+1$ follows from Lemma 3 If $x = 2r+1$ is odd, then $n = 12r+9$. Then $n = 3b$, for $b = 4r+3$. If $r = 3s$, then $b = 12s + 3$ is valid. and the result follows for $n = 3b$ from Lemma 4. If $r = 3s + 1$ then $b = 12s + 7$ is good and the result follows for $n = 3b$ from Lemma 4. Finally if $r = 3s + 2$ then $n = 36s+33 = 2e+3$ for $e = 18s+15$ valid and the result follows for $n = 2e+3$ from Lemma 8. □

The two directions of Theorem 1 then follow from Lemma 2 and Lemma 9.

3 Split Graphs

We say that a graph G is *valid* if G can have its edges partitioned into edge-disjoint triangles.

Lemma 10 *The problem of deciding whether a graph G is valid is NP-complete.*

Proof. The problem of deciding whether a cubic graph G is edge-3-colorable is NP-complete [4]. This question is equivalent to partitioning the edges of a cubic graph G into three perfect matchings. Given such a graph G, let G' be obtained from G by letting $V(G') = V(G) \cup \{v_1, v_2, v_3\}$ where the v_i are three new vertices not in $V(G)$, and letting $E(G') = E(G) \cup \{(u_j, v_i), u_j \in V(G), 1 \leq i \leq 3\}$. Then the triangles involving each of the three vertices v_i must form a perfect matching in G, so G' is valid if and only if G is edge-3-colorable. □

Theorem 3 *The question of whether a sparse split graph G is valid is NP-complete.*

Proof. Given an arbitrary graph G, let G' be obtained from G by adding for each edge $e = (u, v) \notin E(G')$ a triangle (u, v, w_e), where each w_e is a new vertex of degree two in G' for each such e. Thus the w_e added form an independent set R with vertices of degree two in G', while the original set of vertices $S = V(G)$ form a clique in G', since all missing edges of such a clique were added. Therefore G' is a sparse split graph. Furthermore, the added triangles must be chosen since each w_e has degree two, so G' is valid if and only if G is valid, and the NP-completeness follows from Lemma 10. □

Consider now a dense split graph $G = (R \cup S, E)$ having $R \cap S = \emptyset$ and $r = |R|$ independent vertices in R adjacent to all $s = |S|$ vertices in the clique S. We say that r, s is *good* if r is odd, $s = r + 2t + 1$ is even with $t \geq 0$, and no additional conditions are imposed if $t = 3u, u \geq 0$, while if $t = 3u + 1, u \geq 0$ then $r = 6x + 3$, $x \geq 0$, and if $t = 3u + 2, u \geq 0$ then $r = 6x + 1$, $x \geq 0$.

Theorem 4 *If r, s define a valid dense split graph then r, s is good.*

Proof. The r vertices of R give r perfect matchings in S to form triangles, so $s \geq r + 1$ and s is even. Furthermore after removing these r perfect matchings, the vertices of S must each have even degree $s - 1 - r = 2t$ so that they can each belong to t triangles in S. Therefore $s = r + 2t + 1$ is even with $t \geq 0$, and so r is odd. If $t = 3u + 1$ then $s - 1 - r = 2t = 6u + 2$ and the number of edges remaining for S is ts, so to decompose S into triangles we must have s even divisible by 3, so $6u + 2 = s - 1 - r = 6v - 1 - r$, and $r = 6(v - u) - 3 = 6x + 3$, $x \geq 0$. If $t = 3u + 2$ then $s - 1 - r = 2t = 6u + 4$ and the number of edges remaining for S is ts, so to decompose S into triangles we must have s even divisible by 3, so $6u + 4 = s - 1 - r = 6v - 1 - r$, and $r = 6(v - u) - 5 = 6x + 1$, $x \geq 0$. Therefore r, s is good. □

We prove a partial converse to Theorem 4.

Theorem 5 *If $r, s = r + 2t + 1$ is good then the dense split graph G is valid, provided one of the seven following cases holds: $r = 1$; $r = 3$; $r = 5$; $t = 0$; $t = 1$; $t = 2$; or $t = 3$ and $r = 6x + 5$.*

Proof. If $r = 1$ then the graph G is a clique on $s + 1 = 2t + 3$ vertices, so either $t = 3u$ and $s + 1 = 6u + 3$ is valid, or $t = 3u + 2$ and $s + 1 = 6u + 7$ is valid.

If $r = 3$ then the graph G is a clique on $s + 3 = 2t + 7$ vertices minus a triangle which may be assumed chosen, so either $t = 3u$ and $s + 3 = 6u + 7$ is valid, or $t = 3u + 1$ and $s + 3 = 6u + 9$ is valid.

If $t = 0$ then $s = r + 1$, $r + s = 2r + 1$ and the result follows as in Lemma 3.

If $t = 1$ then $s = r + 3$, $r + s = 2r + 3$, $r = 6x + 3$ and the result follows as in Lemma 8.

If $t = 2$ then $s = r + 5$, $r + s = 2r + 5$, $r = 6x + 1$ and $s = 6x + 6$. We must then decompose the edges of the clique with vertex set S into r perfect matchings (to form triangles with the r vertices of R) plus triangles. Decompose the $s = 6x + 6$ vertices of S into $2y = 2x + 2$ vertex disjoint triangles, which we shrink to obtain $2y$ super-vertices. Decompose the shrunk graph into $y - 1$ perfect matchings on $2y$ super-vertices (by the Walecki construction [1, 6], Lemma 1). When the $2y$ triangles are unshrunk, each of the $2y - 1$ perfect super-matchings becomes three perfect matchings. In particular, if two super-vertices $S_1 = \{u_1, v_1, w_1\}$ and $S_2 = \{u_2, v_2, w_2\}$ were joined by a super-edge in one of super-matchings, then the three perfect matchings will have edges u_1u_2, v_1v_2, w_1w_2 for M_1, edges u_1v_2, v_1w_2, w_1u_2 for M_2, and u_1w_2, v_1u_2, w_1v_2 for M_3. This gives $3(2y - 1)$ perfect matchings. However, if for just one of the super-matchings, we include just M_1 (and not M_2, M_3), then we only have $3(2y-1)-2 = 3(2x+1)-2 = 6x+1 = r$ perfect matchings as desired, and the super-edge that joins S_1 and S_2 with the edges of M_2 and M_3 can be decomposed into four triangles $u_1v_2w_1, v_1w_2u_1, w_1u_2v_1, u_2v_2w_2$.

If $t = 3$ and $r = 6x + 5$, then $s = r + 7 = 6x + 12$. Decompose the $s = 6x + 12$ vertices of S into $2y = 2x + 4$ vertex disjoint triangles, which we shrink to obtain $2y$ super-vertices. Decompose the shrunk graph into $y - 1$ perfect matchings on $2y$ super-vertices (by the Walecki construction [1, 6], Lemma 1). When the $2y$ triangles are unshrunk, each of the $2y-1$ perfect super-matchings becomes three perfect matchings. In particular, if two super-vertices $S_1 = \{u_1, v_1, w_1\}$ and $S_2 = \{u_2, v_2, w_2\}$ were joined by a super-edge in one of super-matchings, then the three perfect matchings will have edges u_1u_2, v_1v_2, w_1w_2 for M_1, edges u_1v_2, v_1w_2, w_1u_2 for M_2, and u_1w_2, v_1u_2, w_1v_2 for M_2. This gives $3(2y - 1)$ perfect matchings. However, if for just two of the super-matchings, we include just M_1 (and not M_2, M_3), then we only have $3(2y - 1) - 4 = 3(2x + 3) - 4 = 6x + 5 = r$ perfect matchings as desired. The two super-matchings that were treated differently decompose into even cycles $S_1, S_2, \ldots, S_{2k}, S_1$, $k \geq 2$. The super-edge that joins S_1 and S_2 with the edges of M_2 and M_3 can be decomposed into three triangles $u_1v_2w_1, v_1w_2u_1, w_1u_2v_1$. The super-edges joining S_i and S_{i+1}, or S_{2k} and S_1 are treated similarly, so each vertex will participate in exactly three triangles (two as for S_1 and one as for S_2 along the cycle of length $2k$).

If $r = 5$, then $2t + 1 = 6u + 1$ for $u = 3t$, and $s = 5 + 2t + 1 = 6u + 6 = 6x$. We must thus decompose the edges of the clique S with $s = |S| = 6x$ into triangles plus five perfect matchings to be joined with the five vertices of R respectively, to form triangles involving R. We write $x = 6z + i$ for $0 \leq i \leq 5$.

If $i = 1$ or $i = 3$, then we partition the $s = 6x$ vertices of S into x sets of size six S_j, $1 \leq j \leq x$, and then shrink each set S_j to a super-vertex so that S becomes a clique of size $x = 6z + i$ that can be decomposed into triangles by Theorem 1. After unshrinking the sets S_j, each triangle S_j, S_k, S_ℓ can be replaced by 36 triangles as in Lemma 4. The edges within each S_j decompose into five perfect matchings (by the Walecki construction [1, 6], Lemma 1) to be joined with the five vertices of R respectively, to form triangles involving R.

If $i = 0$ or $i = 2$, then we again partition the $s = 6x$ vertices of S into x sets of size six S_j, $1 \leq j \leq x$, and then shrink each set S_j to a super-vertex so that S becomes a clique of size $x = 6z + i$ that can be decomposed into triangles plus a perfect matching (since a clique of size $x + 1 = 6z + i + 1$ can be decomposed into triangles by Theorem 1). After unshrinking the sets S_j, each triangle S_j, S_k, S_ℓ can be replaced by 36 triangles as in Lemma 4. We are left with sets of 12 vertices in S, namely pairs S_j, S_k that are matched. The edges of these sets of 12 vertices decompose into five perfect matchings plus triangles by the case $t = 3$ and $r = 5$ above, since $r + 2t + 1 = 12$.

If $i = 4$, then we partition the $s = 6x$ vertices of S into $2x$ sets of size three S_j, $1 \le j \le 2x$, and then shrink each set S_j to a super-vertex so that S becomes a clique of size $2x = 2(6z+i) = 12z+8$ that can be decomposed into triangles plus a perfect matching (since a clique of size $2x+1 = 6z+9$ can be decomposed into triangles by Theorem 1). After unshrinking the sets S_j, each triangle S_j, S_k, S_ℓ can be replaced by 9 triangles as in Lemma 4. We are left with sets of 6 vertices in S, namely pairs S_j, S_k that are matched. The edges of these sets of 6 vertices decompose into five perfect matchings plus triangles by the Walecki construction [1, 6], Lemma 1.

Finally if $i = 5$, then we partition the $s = 6x$ vertices of S into $2x$ sets of size three S_j, $1 \le j \le 2x$, and then shrink each set S_j to a super-vertex so that S becomes a clique of size $2x = 2(6z + i) = 12z + 10$ that can be decomposed into triangles plus three perfect matchings (since a clique of size $2x + 3 = 6z + 13$ can be decomposed into triangles by Theorem 1). After unshrinking the sets S_j, each triangle S_j, S_k, S_ℓ can be replaced by 9 triangles as in Lemma 4. We are left with three perfect matchings M_1, M_2, M_3 of the sets S_j. The matching M_3 gives three perfect matchings joining pairs S_j, S_k when these sets are unshrunk, so it remains to find two more perfect matchings. The union of M_1 and M_2 decomposes into a union of even length cycles $S_1, S_2, \ldots, S_{2k}, S_1$. We may choose one of the three matchings joining pairs S_j, S_k for M_1 and for M_2, completing the five matchings. The remaining edges in the cycles described above decompose into triangles, say if $S_i = \{u_1, v_1, w_1\}$ and $S_{i+1} = \{u_2, v_2, w_2\}$ then the chosen matching from M_1 or M_2 is u_1u_2, v_1v_2, w_1w_2, and the chosen triangles are $u_1v_2w_1, v_1w_2u_1, w_1u_2v_1$. We thus obtain five perfect matchings of S plus triangles to cover all the edges of S. $\square$

4 Proof of Theorem 2

We now prove Theorem 2.

If $i = 1$ or $i = 3$, then all edges joining the $n = 6x+i$ vertices of K_n can be covered by triangles, so $k = 0$ is optimum.

If $i = 5$, then we decompose $n = 6x + 5$ into $r = 5$ and $s = 6x$, and apply the case $r = 5$ of Theorem 5. The only edges not covered are the 10 edges of R, which can be decomposed into two triangles and one four-cycle. We select the two triangles, and the four-cycle gives $k = 4$. To show that $k = 4$ is optimum, note that $k = 0$ is not possible by Theorem 1, and since each vertex of K_n has even degree $6x+4$, the number of edges not covered by triangles incident to each vertex is also even. A graph with at least one edge and all vertices of even degree must contain a cycle, and this cycle cannot be a triangle else the triangle would have been chosen, so it must be a cycle of length at least four, so $k = 4$ is indeed optimum.

If $i = 0$ or $i = 2$, then we may add one vertex to obtain K_{n+1} decomposable into triangles by Theorem 1. Removing the added vertex leaves a matching in K_n not covered by triangles, giving $k = \frac{n}{2}$. To show that this is optimum, note that $n = 6x + i$ is even and so the vertices of K_n have odd degree and thus at least one edge not covered by triangles per vertex, which requires at least $k = \frac{n}{2}$ such edges in total.

Finally if $i = 4$, then we may add one vertex to obtain K_{n+1} decomposable into triangles plus an uncovered K_5 as in $i = 5$. When we remove one of the vertices of this K_5, we are left with an unchosen K_4 plus an unchosen matching of the remaining $n - 4$ vertices. We then choose one triangle from the K_4, leaving 3 edges unchosen in the K_4 and $\frac{n}{2} - 2$ unchosen edges in the remaining matching, for a total of $k = \frac{n}{2} + 1$ unchosen edges. To show that this is optimum, note again that $n = 6x + 4$ is even and so the vertices of K_n have odd degree and thus at least one edge not covered by triangles per vertex, which requires at least $k \geq \frac{n}{2}$ such edges in total. If this total were achievable, then it would be achieved with an uncovered matching of K_n, which could be chosen to form triangles with an additional vertex in K_{n+1}. But $n + 1 = 6x + 5$, so K_{n+1} cannot be so decomposed into triangles by Theorem 1, a contradiction. So $k = \frac{n}{2} + 1$ is optimum.

5 Approximation Results

The optimization problem can be approximated in polynomial time within a constant factor.

Theorem 6 *The NP-complete question of finding the largest set of edge-disjoint triangles contained in a given graph G can be approximated within a factor of $\frac{1}{3}$ in polynomial time by a greedy algorithm.*

Proof. The NP-completeness of this question was shown in Theorem 10. The greedy polynomial time approximation algorithm repeatedly selects triangles that are edge-disjoint from previously chosen triangles until no more such triangles can be found. Each triangle chosen by the algorithm can share edges with at most three triangles of an optimal solution O. Therefore at least one more triangle of O will be edge-disjoint from the triangles chosen until the optimal solution O has at most three times the number of triangles found greedily. This shows that the greedy algorithm finds at least $\frac{1}{3}|O|$ triangles. □

Theorem 7 *The NP-complete question of finding the largest set of edge-disjoint triangles contained in a given graph G can be approximated within a factor of $\frac{1}{2}$ in polynomial time by an augmenting algorithm.*

Proof. The algorithm begins as before by greedily selecting a maximal set T of edge-disjoint triangles. Then the algorithm repeatedly performs the following augmenting step: Determine for each $t_0 \in T$ if after removing t_0 from T to obtain T' with $|T'| = |T| - 1$, we can add to T' two triangles t_1, t_2 that are edge-disjoint from each other and from the triangles in T', to obtain T'' with $|T''| = |T'| + 2 = |T| + 1$, and then proceed to add a further edge-disjoint triangle t_3 to T'' if possible. The algorithm terminates when such an augmentation is not possible.

Notice that t_1, t_2, and possibly t_3 must each share an edge with the removed triangle t_0. Let O be an optimum solution, that is, O is a largest set of edge-disjoint triangles. For each $t \in T$ and each $o \in O$ such that t, o share at least one edge, let w_{to} be the number of triangles in T that share an edge with o. Then

$$\begin{aligned} |O| &= \sum_{o \in O} (\sum_t \frac{1}{w_{to}}) \\ &= \sum_{t \in T} (\sum_o \frac{1}{w_{to}}) \\ &\leq \sum_{t \in T} 2 = 2|T|, \end{aligned}$$

proving $|T| \geq \frac{1}{2}|O|$ as desired. The reason is that $\sum_o \frac{1}{w_{to}} > 2$ only if there are three such o sharing an edge with t, and at least two of them, say o_1, o_2 have $w_{to} = 1$. But then we could perform an augmenting step removing t and adding $t_1 = o_1$, $t_2 = o_2$, contrary to the assumption that the algorithm had ended. The bound $|T| \geq \frac{1}{2}|O|$ gives the desired $\frac{1}{2}$ approximation. $\square$

Theorem 8 *The NP-complete question of finding the largest set of edge-disjoint triangles contained in a given graph G can be approximated within a factor of $\frac{3}{5}$ in polynomial time by an augmenting algorithm.*

Proof. Let T be a set of edge-disjoint triangles in G. Initially T is empty. A *k-augmentation* replaces T with $T' = T \setminus T_{k-1}$ and then replaces T' with $T'' = T' \cup T_k$. Here T_{k-1} is a set of $k-1 = |T_{k-1}|$ triangles in T and T_k is a set of $k = |T_k|$ triangles that are edge-disjoint from one another and from the triangles in T'.

The algorithm performs a series of k-augmentations with possible values $k = 1, 2, 3, 4$, until no such augmentation is possible. Note that the $\frac{1}{3}$ approximation algorithm above is the case where $k = 1$, and the $\frac{1}{2}$ approximation algorithm above is the case where $k = 1, 2$.

When the algorithm ends, we will have as in Theorem 7

$$|O| = \sum_{t \in T} (\sum_{o} \frac{1}{w_{to}})$$

with $w_t = \sum_o \frac{1}{w_{to}} \leq 2$, since we cannot have $w_{to_1} = w_{to_2} = 1$ (otherwise we could perform a 2-augmentation as before).

We split T into three sets T_1, T_2, T_3. Let T_1 be the triangles with $w_{to_1} = 1$, $w_{to_2} = w_{to_3} = \frac{1}{2}$, Let T_2 be the triangles not in T_1 with $w_{to_1} = 1$, $w_{to_2} = \frac{1}{2}$. Let T_3 be the triangles not in T_1 or T_2. Note that $w_t \leq \frac{11}{6}$ for $t \in T_2$ and $w_t \leq \frac{5}{3}$ for $t \in T_3$.

If $t \in T_1 \cup T_2$ with $w_{to_1} = 1$, $w_{to_2} = \frac{1}{2}$, we say that *t attributes to $t' \in T$ via o_2* if t' is the other triangle with $w_{to_2} = \frac{1}{2}$. Note that if t' is so attributed, then we cannot have $w_{t'o_3} = 1$, otherwise we would have a 3-augmentation removing t, t' and adding o_1, o_2, o_3. So an attributed t' must be in T_3 and must have $w_{t'} \leq \frac{3}{2}$. Note that t' cannot be attributed by another t, say t'' with $w_{t''o'_1} = 1$, $w_{t''o'_2} = \frac{1}{2}$, $w_{t'o'_2} = \frac{1}{2}$, otherwise we would have a 4-augmentation removing t, t', t'' and adding o_1, o_2, o'_1, o'_2. Similarly, for $t \in T_1$, we cannot have that t attributes to t' via both o_2, o_3, otherwise we would have a 3-augmentation removing t, t' and adding o_1, o_2, o_3.

Therefore each $t' \in T_3$ is attributed at most once, each $t \in T_1$ attributes to two $t' \in T_3$ (call T_4 the set of such attributed t'), each $t \in T_2$ attributes to one $t' \in T_3$ (call T_5 the set of such attributed t', and let $T_6 = T_3 \setminus (T_4 \cup T_5)$). Therefore $|T_4| = 2|T_1|$ and $|T_5| = |T_2|$. Then

$$\sum_{t \in T_1 \cup T_4} w_t \leq 2|T_1| + \frac{3}{2}|T_4|$$

$$= 5|T_1| = \frac{5}{3}(|T_1| + |T_4|),$$

and

$$\sum_{t \in T_2 \cup T_5} w_t \leq \frac{11}{6}|T_2| + \frac{3}{2}|T_5|$$

$$= \frac{10}{3}|T_2| = \frac{5}{3}(|T_2| + |T_5|).$$

Therefore

$$|O| = \sum_{t \in T} w_t$$

$$= \sum_{t \in T_1 \cup T_4} w_t + \sum_{t \in T_2 \cup T_5} w_t + \sum_{t \in T_6} w_t$$
$$\le \frac{5}{3}(|T_1| + |T_4| + |T_2| + |T_5| + T_6|) = \frac{5}{3}|T|.$$

This gives $|T| \ge \frac{3}{5}|O|$, proving the $\frac{3}{5}$ approximation bound. □

We prove a hardness of approximation counterpart to these approximation algorithms. Let S be the collection of 3SAT instances with some upper bound c on the number of occurrences per variable. It is known that there exists a constant $0 < \gamma < 1$ such that it is NP-hard to distinguish instances in S that are satisfiable from instances in S that cannot have a fraction $1 - \gamma$ of the clauses satisfied [2, 7]. We refer to this approximation problem as MAX-3SAT(c).

Theorem 9 *Let U be the collection of cubic graphs without triangles. There exists a constant $0 < \delta < 1$ such that it is NP-hard to distinguish instances in U whose edges can be covered by three matchings from instances in U that cannot have a fraction $1 - \delta$ of the edges covered by three matchings.*

Proof. We reduce 3SAT instances S with c occurrences per variable to instances U of edge-3-colorability (partition the edges into three perfect matchings) by the reduction in [4]. The reduction has a constant M such that each clause is represented by a gadget with at most M edges, and c occurrences of each variable are represented by a gadget with at most M edges. If we fail to assign one of the three matchings to a fraction only δ of the edges, we fail to assign a variable properly or a clause properly to a fraction only δM of variables plus clauses, therefore fail a fraction only $\delta c M < \gamma$ of the clauses by choosing $\delta < \frac{\gamma}{cM}$. □

Theorem 10 *There exists a constant $0 < \epsilon < 1$ such that it is NP-hard to distinguish graphs whose edges can be covered by edge-disjoint triangles from graphs that cannot have a fraction $1 - \epsilon$ of the edges covered by edge-disjoint triangles. Furthermore, we may take $\epsilon = \delta$ from Theorem 9.*

Proof. It suffices to represent a cubic graph without triangles G from Theorem 9 by constructing G', adding three independent vertices to G that are set adjacent to all of G. As in Theorem 10, the chosen triangles in G' will involve the three added vertices (since G alone has no triangles), and the triangles chosen for each of the three added vertices will give the corresponding three matchings in G. If a fraction ϵ of the edges is not covered by edge-disjoint triangles in G', then a fraction $\delta = \epsilon$ of the edges is not covered by the three matchings in G. The reason is that if $n = |V(G)|$, then $|E(G)| = \frac{3}{2}n$ and $|E(G')| = |E(G)| + 3n = 3|E(G)|$; and if the three matchings select k edges of G, then the k triangles chosen for G' have $3k$ edges, so $1 - \delta = \frac{k}{|E(G)|} = \frac{3k}{|E(G')|} = 1 - \epsilon$. □

6 The Case of General H

We now generalize the problem form triangles to arbitrary fixed graphs H.

Theorem 11 *Let H be a fixed graph with $e = |E(H)| \ge 1$ edges. The question of finding the largest number of edge-disjoint copies of H (induced, or not necessarily induced) that can be found in a given graph G can be approximated within $\frac{1}{e}$ by a greedy algorithm in polynomial time.*

Proof. As in Theorem 6, we repeatedly select copies of H in G, until no more can be found. Clearly each H found will overlap in some edge with at most e copies of H in an optimal solution, and the optimal solution cannot have any more copies of H. Therefore the opitmal solution has at most e times the number of copies of H in the solution found, proving the $\frac{1}{e}$ approximation bound. $\square$

Theorem 12 *Let H be a fixed graph with $e = |E(H)| \geq 1$ edges. The question of finding the largest number of edge-disjoint copies of H (induced, or not necessarily induced) that can be found in a given graph G can be approximated within $\frac{2}{e+1}$ by an augmenting algorithm in polynomial time.*

Proof. As in Theorem 7, we first select copies of H greedily, and then continue by attempting to remove one H from the solution found and add two copies of H that overlap the removed H instead, adding extra copies of H if more appear after the removal as well.

If T is the solution found, and O is the optimal solution, then for each $t \in T$ and $o \in O$ that overlap in some edge e, we let w_{toe} be the number of edges of o that appear in some element of T. As before,

$$\begin{aligned} |O| &= \sum_{o \in O} (\sum_{t,e} \frac{1}{w_{toe}}) \\ &= \sum_{t \in T} (\sum_{o,e} \frac{1}{w_{toe}}) \\ &\leq \sum_{t \in T} \frac{e+1}{2} = \frac{e+1}{2}|T|, \end{aligned}$$

proving $|T| \geq \frac{2}{e+1}|O|$ as desired. The reason is that $\sum_{o,e} \frac{1}{w_{toe}} > \frac{e+1}{2}$ only if there are at least two o sharing an edge e, with t, say o_1, e_1 and o_2, e_2, with $w_{toe} = 1$. But then we could perform an augmenting step removing t and adding $t_1 = o_1$, $t_2 = o_2$, contrary to the assumption that the algorithm had ended. The bound $|T| \geq \frac{2}{e+1}|O|$ gives the desired $\frac{2}{e+1}$ approximation. $\square$

The approximation ratio can again be improved to slightly above $\frac{e}{2}$ as in Theorem 8, but we omit this result. We prove a partial hardness of approximation result for this problem.

Theorem 13 *Let $H = K_r$ or $H = C_r$ be a fixed clique or cycle with $r \geq 3$ vertices. The question of finding the largest number of edge-disjoint copies of H (induced, or not necessarily induced) that can be found in a given graph G is NP-hard to approximate within some constant $1 - \epsilon$, $0 < \epsilon < 1$, which can be taken as the constant of Theorem 10 except in the case of C_4. In the case of $H = C_r$, the hardness is with regards to a solution that covers all the edges in G.*

Proof. Suppose first $H = K_r$, $r \geq 3$. The cubic graphs in Theorem 9 were triangle-free, so the graphs in Theorem 10 are K_4-free. Given G as in Theorem 10, for each triangle t in G, we add $r - 3$ vertices G to form a clique of size r containing t in G'. Clearly the cliques of size r in G' are precisely those constructed for triangles t in G, so the hardness of approximation follow from Theorem 10.

Suppose next $H = C_r$, $r \geq 3$. Write $r = 2a + b$ for $b = a, a+1, a+2$, and in the construction of Theorem 10, replace the edges of the cubic graph by paths of length b, and replace the edges joining to the three additional independent vertices by paths of length a. Clearly all the edges will be covered in a solution of cycles C_{2a+b} if and only if all paths of length b from the original cubic graph are joined by two paths of length a to an added vertex to form C_{2a+b}, while otherwise no cycles not going through the path of length b will be found unless $r = 4$. $\square$

References

[1] Brian Alspach, The wonderful Walecki construction, Bull. Inst. Combin. Appl., 52 (2008), 7–20.

[2] Sanjeev Arora, Probabilistic checking of proofs and hardness of approximation probles, revised version of dissertation submitted at CS Division, U C Berkeley, in August 1994. CS-TR-476-94.

[3] Béla Bollobás, Combinatorics, Cambridge University Press, 177 pp.

[4] I. Holyer, The NP-completeness of edge-coloring, SIAM J. Comput. 10 (1981), 718–720.

[5] D.G. Kirpatrick and P. Hell, On the complexity of a generalized matching problem, Proc. 10th Ann. ACM Symp. on Theory of Computing (1978), 240–245.

[6] E. Lucas, Récréations Mathématiques, four volumes, Gautheir-Villars, New York, 1968.

[7] Rudich et al. Computational Complexity Theory, IAS/Park City Mathematics Series, 2004, page 108. ISBN 0-8218-2872-X.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de