On the Complexity of Finding Sets of Edges in Graphs and Digraphs

T.I. Fenner

A.M. Frieze

ABSTRACT

Suppose we are given a graph (V,E) and a function $c:E \to Z$ where $c(e)$ is said to be the colour of edge e of E. A subset $S \subseteq E$ is said to be polychromatic if $|c(S)| = |S|$. In this paper we study the complexity of recognizing whether such an edge coloured graph has certain types of polychromatic subsets, such as spanning trees, paths, cutsets, matchings and Hamiltonian cycles.

1. Introduction

Erdős, Simonovits and Sós [6] studied the following problem: given a graph H, what is the maximum number of colours m that can be used to colour the edges of a complete graph K_n on n vertices so that no subgraph isomorphic to H has all its edges coloured differently? They considered, in particular, the cases when H is a complete graph, a path or a cycle. Hahn [9] [10] considered a similar problem when H is a star. In his doctoral thesis, Bate [1] [2] investigates, *inter alia*, related problems on the existence of polychromatic paths and cycles. (See also Galvin [7].)

In this paper we study the problem of the complexity of determining whether such a subgraph exists.

An edge colouring of a graph $G = (V,E)$ is a function $c:E \to Z$ for some set Z. A subset $X \subseteq E$ is said to be *polychromatic* if $|c(S)| = |S|$, i.e. each edge of S is coloured differently. A similar definition is assumed for digraphs.

Possible applications of this model could be to problems where

ISBN 0-12-114320-1

each colour represents an indivisible resource which can only be used once.

In the paper we study the problem of determining whether an edge coloured graph or digraph has various types of polychromatic subset, including spanning trees, spanning arborescences, $s-t$ paths, cycles, cutsets, matchings and Hamiltonian paths and cycles. Most cases turn out to be NP-Complete and in particular it is interesting to note that the case of spanning arborescences, which is a special case of a matroid-greedoid intersection problem, turns out to be NP-Complete, thus showing that this problem is unlikely to be polynomially solvable.

2. Polychromatic Spanning Trees and Arborescences

Spanning Trees

We first consider the problem of determining whether or not an edge coloured graph G contains a polychromatic spanning tree. Fortunately this is a matroid intersection problem: indeed a polychromatic spanning tree is a set of $n-1$ edges independent in both

(1) the cycle matroid of G, and

(2) the partition matroid defined by polychromatic sets of edges.

Thus the existence of a polychromatic spanning tree can be tested in polynomial time using a matroid intersection algorithm, e.g. Lawler [13] pp. 313-314. In fact, this algorithm could easily be implemented to run in time $O(|V||E|)$.

Furthermore, it follows from Edmonds' matroid intersection duality theorem [5] that G contains a polychromatic spanning tree if and only if

$$(1.1)\ r_1(S) + r_2(E-S) \geq n-1 \qquad \text{for all } S \subseteq E$$

where

(*a*) $n - r_1(S)$ = *the number of components in the graph* (V,S)

(*b*) $r_2(T)$ = *the number of distinct colours used in* T.

We can deduce from (1.1) the following simple but interesting result:

THEOREM 1.1. Let G be a complete graph on n vertices and let the edges of G be coloured with $n-1$ colours such that each colour is used at least once and no more than $\lceil n/2 \rceil + 1$ times. Then G contains a polychromatic spanning tree.

PROOF. If G does not have a polychromatic spanning tree then (1.1) fails for some S. Suppose (V,S) has p components containing $n_1, n_2, \ldots, n_p$ vertices respectively and that $r_2(E-S) = t$. We can quickly eliminate the following values for p:

(i) $p = 1$: here $r_1(S) = n-1$;

(ii) $p = 2$: here $|E-S| \geq n-1$ implies $t \geq 1$;

(iii) $p = n$: here $S = \varnothing$ implies $t = n-1$;

(iv) $p = n-1$: here $|S| = 1$ implies $r_1(S) = 1$ and $t \geq n-2$.

We can thus assume that $3 \leq p \leq n-2$ and $n \geq 5$. Since (1.1) fails we have $t \leq p-2$ and so some colour occurs at least $|E-S|/(p-2)$ times. Now

$$|E-S| \geq \binom{n}{2} - \sum_{i=1}^{p} \binom{n_i}{2}$$

$$\geq \binom{n}{2} - \binom{n-p+1}{2} = \binom{p-1}{2} + (p-1)(n-p+1).$$

Thus some colour is used at least $m = \lceil (p-1)/2 + (p-1)(n-p+1)/(p-2) \rceil$ times. It is easy to show, given our range of values for p, that $m \geq \lceil n/2 \rceil + 2$ for $n \geq 4$. The result follows.

Spanning Arborescences

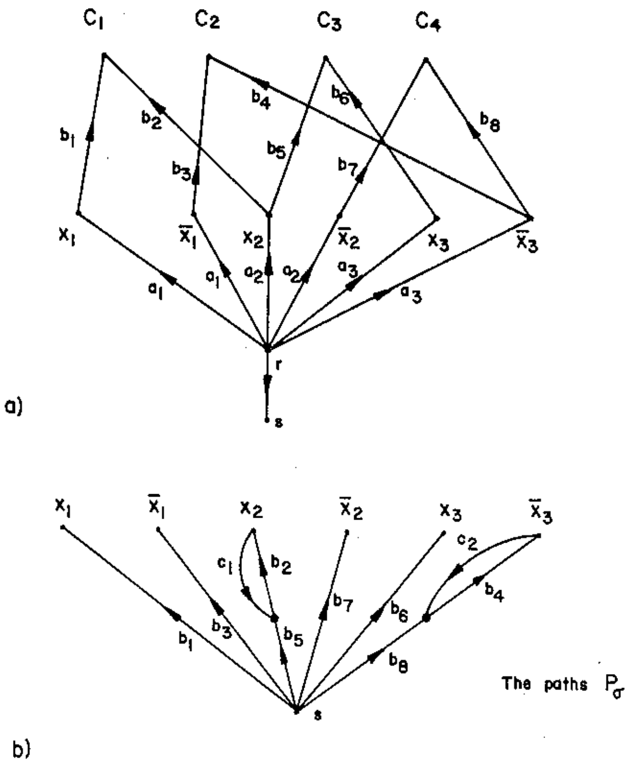
We consider next the problem (PSA) of determining whether an arc coloured digraph D contains a polychromatic spanning arborescence rooted at a given vertex r. Here an arborescence is a spanning tree in which each arc is directed away from the root and so each vertex other than r has indegree 1. The uncoloured version is solvable in time $O(|V| + |E|)$, and it appears plausible that PSA could be polynomially solvable as it involves the intersection of a greedoid, i.e. the arborescence (see Korte and Lovász [12]), and the partition matroid defined by the colouring. We show, however, that PSA is NP-Complete:

THEOREM 1.2. PSA is NP-Complete.

PROOF. We prove this by showing that SATISFIABILITY $\propto$ PSA. Suppose we are given a set $X = \{x_1, x_2, \ldots, x_n\}$ of boolean variables and a set $C = \{C_1, C_2, \ldots, C_m\}$ of clauses over X. We define an instance of PSA as follows: let $L = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$ be the set of literals, where we can assume each literal appears in at least one clause, otherwise the complementary literal can be taken as true. We start with a bipartite digraph D_0 with vertices $L \cup C$ and arcs of the form (σ, C_j) whenever literal σ occurs in clause C_j. Each arc of D_0 is differently coloured. We now add a new vertex r and join it by an arc to each member of L (see Figure 1(a)). Using n new colours

$a_1, \ldots, a_n$, we colour these arcs, so that (r,x_i) and $(r,\bar{x}_i)$ are both coloured a_i for $i = 1,2, \ldots, n$.

We next add a new vertex s and an arc (r,s) using a new colour. We then join s to each literal σ by a path P_σ (from s to σ) where the sets of interior vertices of the P_σ are disjoint. Suppose that literal σ occurs in k_σ clauses; then P_σ contains k_σ arcs. The arcs of P_σ are coloured with exactly the same k_σ colours used to colour those arcs of D_0 with initial vertex σ. Finally, each σ is joined to each interior vertex of P_σ by an arc, using a new colour each time. This completes the description of the arc coloured digraph D (see Figure 1(b)).



C = $\{\{x_1, x_2\}, \{\bar{x}_1, \bar{x}_3\}, \{x_2, x_3\}, \{\bar{x}_2, \bar{x}_3\}\}$

Figure 1

Suppose first that C is satisfied by some assignment of truth values to X. Then there exists a polychromatic spanning arborescence rooted at r with the following set of arcs:

(1) the arc (r,σ) for each true literal σ;

(2) for each $C_i, i = 1,2, \ldots, m$, an arc (σ,C_i) where σ is some true literal occurring in C_i (the existence of some such σ is guaranteed since C is satisfied);

(3) the arc (r,s);

(4) the path P_σ for each false literal σ;

(5) all arcs of the form (σ,v) where σ is a true literal and v is an interior vertex of P_σ.

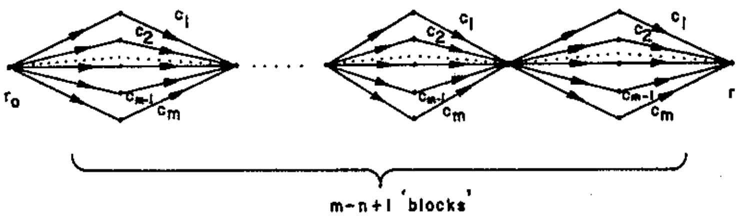
It is easy to see that (1) to (5) define a polychromatic spanning arborescence.

Now suppose, conversely, that D contains a polychromatic spanning arborescence rooted at r with arc set A. Let the true literals be those σ for which $(r,\sigma) \in A$. Note that, since (r,x_j) and $(r,\bar{x}_j)$ have the same colour, this defines a proper assignment of truth values. Any variable x_j not assigned a truth value in this way will be assigned the value *true*. We now show that this assignment of truth values satisfies C.

We note first that for each literal σ there are precisely 2 arcs in D that enter σ, i.e. (r,σ) and the terminal arc u_σ of P_σ. Now it can easily be shown, by tracing P_σ in reverse, that to avoid cycles we must have either $u_\sigma \notin A$ or $P_\sigma \subseteq A$; but if $P_\sigma \subseteq A$ then $(\sigma,C_i) \notin A$ because P_σ uses the colours of all such arcs. Thus $(\sigma,C_i) \in A$ implies that $(r,\sigma) \in A$ and hence that σ is true in our assignment. Since A spans the vertices of D this assignment satisfies C .□

We note that the problem is still hard if we do not specify the root since in the above proof r is the root of any spanning arborescence.

We further note that PSA remains NP-Complete when we restrict attention to those cases in which the number of colours used is $n-1$, where n is the number of vertices. This is easily seen by "splitting" r as in Figure 2.



(We assume the original colours of D were $c_1, c_2, \ldots, c_m$, where $m > n - 1$. Arcs without a specified colour are given distinct new colours. No. of colours = no. of vertices - 1.)

Figure 2

The problem remains NP-Complete if we further restrict attention to complete digraphs; we simply add a new vertex O and a new colour, red say, and complete the digraph with red arcs.

However, we observe that if the given digraph has no directed cycles then the problem is solvable in polynomial time. Indeed, suppose that the vertices of D are numbered $1,2,\ldots,n$ such that $i < j$ if (i,j) is an arc of D. We can obtain an arborescence of D by selecting, for each $j > 1$, an arbitrary arc entering j. Thus PSA reduces to a bipartite matching problem in which the bipartite graph G has one vertex for each $j = 2,3,\ldots,n$ and another vertex for each colour. The arcs are of the form (j,c) whenever c is the colour of some arc entering j in the digraph D. Each matching of G of cardinality $n-1$ corresponds to a polychromatic spanning arborescence of D and vice versa.

We finally note that, unless NP = Co-NP, there cannot be any *good* characterization of the set of coloured digraphs which contain a polychromatic spanning arborescence. This is in clear contradistinction to the undirected case where we have the characterization given by (1.1). These remarks will also apply to other classes of polychromatic subgraphs for which the corresponding recognition problems are NP-Complete.

3. Polychromatic Paths, Cycles and Cutsets

Paths and Cycles

If an edge coloured graph has a polychromatic spanning tree then every pair of vertices is connected by a polychromatic path. If, however, we specify a pair of vertices s and t of a graph and ask whether it contains a polychromatic path from s to t then, surprisingly perhaps, this problem is NP-Complete. We prove this by showing that 3-DIMENSIONAL MATCHING (3DM) is polynomially transformable to this problem which we call POLYCHROMATIC PATH (PP).

An instance of 3DM is described by

(a) disjoint sets X,Y,Z where $|X| = |Y| = |Z| = m$;

(b) a subset $T \subseteq X \times Y \times Z$.

The problem is to determine whether or not there exists $M \subseteq T$ such that (a) $|M| = m$, (b) each element of $X \cup Y \cup Z$ occurs in exactly one member of M. A set M satisfying (a) and (b) is called a 3-dimensional matching.

THEOREM 2.1. PP is NP-Complete.

PROOF. Given an instance of 3DM we construct the edge coloured graph G indicated in Figure 3.

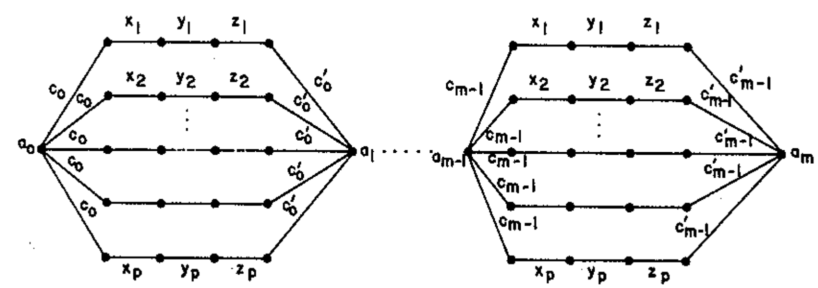


Figure 3

This can be described informally as follows. Starting with a path $a_0, a_1, \ldots, a_m$, we duplicate each edge (a_i, a_{i+1}) p times, where $p = |T|$. Each of the p edges joining a_i to a_{i+1}, $i = 0,1,\ldots,m-1$ is then replaced by a distinct path of length 5. This gives us a graph with

$m + 1 + 4mp$ vertices and $5mp$ edges as in Figure 3. Let

$$T = \{(x_i, y_i, z_i) \mid i = 1,2,\ldots,p\}.$$

For $i = 0,1,\ldots,m-1$ we order the p paths joining a_i to a_{i+1}. The 5 edges of the kth path joining a_i to a_{i+1} are coloured c_i, x_k, y_k, z_k, c_i', respectively, where c_i and c_i' , $i = 0,1, \ldots, m-1$ are $2m$ new colours.

It is now easy to see that T contains a 3-dimensional matching if and only if G contains a polychromatic path from a_0 to a_m. □

We note that the graph G constructed in Theorem 2.1 is bipartite, planar and edge series-parallel [14].

By adding a large monochromatic clique to the graph we can make the ratio of vertices to colours as large as we like; alternatively, by adding a large polychromatic clique we can make this ratio as small as we like.

By applying the transformation indicated in Figure 4 to each of the sets of edges coloured $c_0, c_0', c_1, c_1', \ldots, c_{m-1}'$, we can obtain a graph in which no vertex has degree exceeding 3.

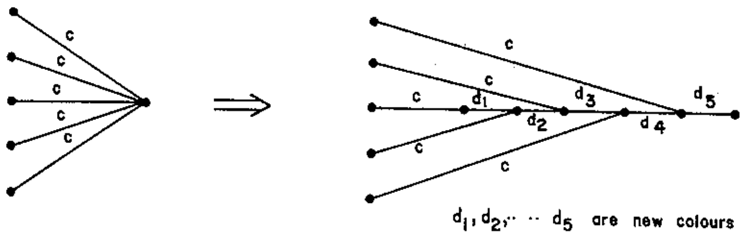


Figure 4 (for $p = 5$)

The graph is still planar, edge series-parallel and can easily be made bipartite. Thus we have

COROLLARY 2.1. PP is NP-Complete for bipartite, planar, edge series-parallel graphs with maximum degree 3. □

If we add the edge (a_o, a_m) to the graph constructed in Theorem 2.1 using a new colour (red say), we see that the problem (PC) of determining whether or not an edge coloured graph has a polychromatic cycle is also NP-Complete. Moreover, we have

COROLLARY 2.2. PC is NP-Complete for bipartite, planar, edge series-parallel graphs with maximum degree 3.

It is easy to show (Bate [1]) that an edge coloured *complete* graph contains a polychromatic cycle if and only if it contains a polychromatic triangle (if cycle C is polychromatic, any chord splits C into 2 smaller cycles one of which must be polychromatic). Clearly, checking for polychromatic triangles can be done in polynomial time.

All of the above results, except the last remark on complete graphs, carry over to digraphs without difficulty. We note that the uncoloured versions of these problems are solvable in time $O(|V| + |E|)$.

Cutsets

We now consider the problem (PCS) of whether or not an edge coloured graph has a polychromatic cutset.

THEOREM 2.2. Planar PCS is NP-Complete.

PROOF. Starting with the graph G constructed in Theorem 2.1 (see Figure 3) together with the red edge (a_0, a_m), we add "new" red edges as shown in Figure 5 to create an edge coloured plane graph H in which the common boundary of any 2 neighbouring faces is a single edge.

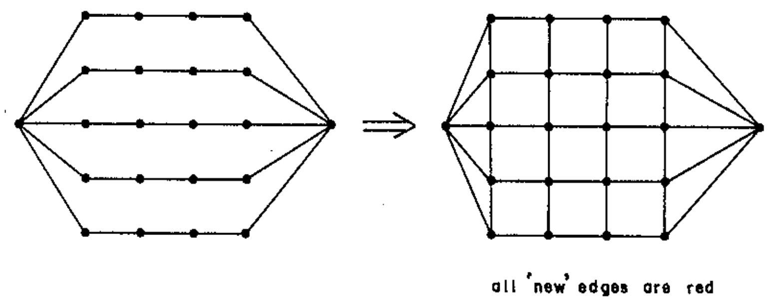


Figure 5 (for $p = 5$)

It is straightforward to show that none of these "new" red edges can occur in any polychromatic cycle of H. Now we construct the dual graph H^*, the vertices of which correspond to the faces of H. For an edge e of H, let $h(e)$ denote the edge of H^* joining the 2 faces containing e; by construction, h is a bijection. We colour $h(e)$ the same as e. Now it is well known [3] that C is a cycle of H if and only if $h(C)$ is a minimal cutset of H^*. Since C and $h(C)$ are coloured in the same way, C is polychromatic if and only if $h(C)$ is. It follows that 3DM $\propto$ Planar PCS.

For complete graphs PCS remains NP-Complete: given an edge coloured graph G, we can complete the graph, colouring each new edge with a different new colour, to create an edge coloured complete graph G' which has a polychromatic cutset if and only if G has.

4. Polychromatic Matchings

It is easy to show that the problem (PPM) of determining whether an edge coloured graph has a polychromatic perfect matching is NP-Complete, even for bipartite graphs. This contrasts with the uncoloured matching problem which is solvable in time $O(|V|^{2.5})$.

THEOREM 3.1. PPM is NP-Complete, even when the number of colours is one half of the number of vertices.

PROOF. Given an instance of 3DM, construct a bipartite graph G with vertex sets X and Y. For each $(x,y,z) \in T$ we include the edge (x,y) and colour it z. Now T contains a 3-dimensional matching if and only if G contains a polychromatic perfect matching. Note that the number of colours used is m which is half the number of vertices and hence is minimal. The above construction may produce parallel edges. We can remove parallel edges by replacing each edge by a path of length 3 as shown in Figure 6. The resulting graph is still bipartite. □

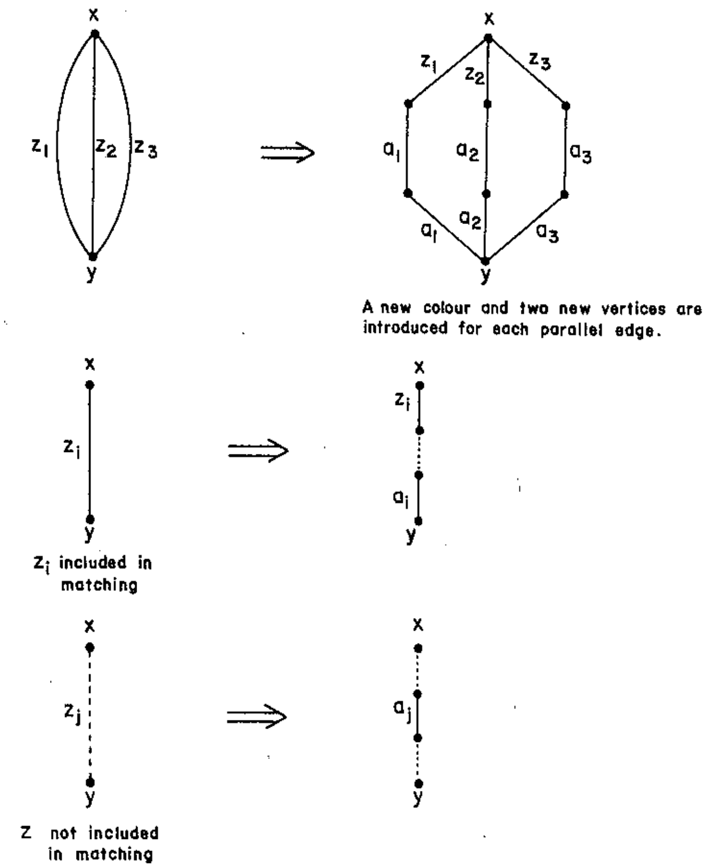


Figure 6

We show next that PPM remains NP-Complete when the graph G is a complete bipartite graph or a complete graph. Thus, even when the existence of many perfect matchings is assured, the problem remains difficult.

COROLLARY 3.2. PPM is NP-Complete for complete bipartite graphs and complete graphs.

PROOF. Consider an arbitrary edge coloured bipartite graph. Add 2 new vertices a and b, putting them into different parts of the vertex partition. Complete the graph by adding new edges all of the same new colour, red say. Any polychromatic perfect matching must contain the edge (a,b) and after deleting the red edges we are left with the previous problem. The same argument is valid whether we construct a complete graph or complete bipartite graph. □

5. Polychromatic Hamiltonian Cycles and Paths

The general problem (PHC) of whether an edge coloured graph contains a polychromatic Hamiltonian cycle is NP-Complete: we can colour each edge differently and then every Hamiltonian cycle is polychromatic. It is not obvious that the problem remains hard when we restrict it to the case (CPHC) where we have a complete graph with n vertices and only n colours are used. We give an outline proof of this result.

THEOREM 4.1. CPHC is NP-Complete.

PROOF. Suppose that we are given an arbitrary edge coloured graph G. If $m > n$ colours are used, we split vertex 1 to obtain G_1 as indicated in Figure 7, where each "block" of m parallel edges uses each colour. Then G_1 contains a polychromatic Hamiltonian cycle if and only if G does. (For the case $m=n+1$, we must add a dummy colour which is not used in G so that we have $m=n+2$.)

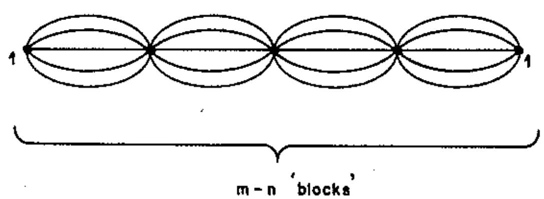
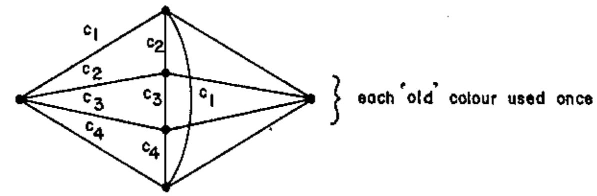


Figure 7

To remove the parallel edges, we replace each "block" of m parallel edges as indicated in Figure 8, to obtain a simple graph G_2. It is not difficult to show that G_2 has a polychromatic Hamiltonian cycle if and only if G_1 does.



(m new colours and m new vertices are added for each of the $m-n$ "blocks".)

Figure 8 (for $m = 4$)

In order to obtain a complete graph we split one vertex as indicated in Figure 9. The resulting graph has the same number of colours as vertices. □

(r, b, and g are new colours; the graph is completed by colour b.)

Figure 9

Using similar constructions it can be shown that, given a complete graph G on n vertices, edge coloured using $n-1$ colours, the problem of determining whether G has a polychromatic Hamiltonian path is NP-Complete. This problem remains NP-Complete when either one or both of the endpoints of the path is specified.

The corresponding Hamiltonian path and cycle problems for directed graphs can also be shown to be NP-Complete.

References

[1] N.G. Bate, "Circuits of Edge-Coloured Complete Graphs", Ph. D. Thesis, University of Keele, England, 1981.

[2] N.G. Bate, "Complete Graphs Without Polychromatic Circuits" to appear in Discrete Mathematics.

[3] J.A. Bondy and U.S.R. Murty, "Graph Theory with Applications", Macmillan Press, London, 1976.

[4] S.A. Cook, "The Complexity of Theorem Proving Procedures", in Proceedings 3rd Annual ACM Conference on Theory of Computing, ACM, New York (1971) 151-158.

[5] J. Edmonds, "Submodular Functions, Matroids and Certain Polyhedra" in Proceedings International Conference on Combinatorics at Calgary, Gordon and Breach, New York (1970) 69-87.

[6] P. Erdös, M. Simonovits and V.T. Sós, "Anti-Ramsey Theorems", Infinite and Finite Sets (Colloq. Keszthely, 1973) Vol. 2, Colloq. Math. Soc. János Bolyai 10, North Holland, Amsterdam (1975) 633-643.

[7] F. Galvin, Problem 6034, Amer. Math. Monthly 1975, Solution: Amer. Math. Monthly 1977, p 224.

[8] M.R. Garey and D.S. Johnson, "Computers and Intractability: a Guide to the Theory of NP-Completeness", W.H. Freeman and Co., San Francisco, 1979.

[9] G. Hahn, "Some Star Anti-Ramsey Numbers" Proceedings 8th S.E. Conference on Combinatorics, Graph Theory and Computing, Louisiana State University, Baton Rouge (1977) 303-310.

[10] G. Hahn, "More Star Sub-Ramsey Numbers", Discrete Mathematics 34 (1981) 131-139.

[11] R.M. Karp, "Reducibility among Combinatorial Problems" in Complexity of Computer Computations (eds. R.E. Miller and J.W. Thatcher), Plenum Press, New York (1972) 85-103.

[12] B. Korte and L. Lovász, "Mathematical Structures Underlying Greedy Algorithms", Fundamentals of Computation Theory, Lecture Notes in Computer Science 117, Springer-Verlag, Berlin (1981) 205-209.

[13] E.L. Lawler, "Combinatorial Optimization: Networks and Matroids", Holt, Rinehart and Winston, New York, 1976.

[14] J. Valdes, R.E. Tarjan and E.L. Lawler, "The Recognition of Series Parallel Digraphs", SIAM Journal on Computing 11 (1981) 298-313.