THE UNIVERSITY OF MANITOBA

ON REDUCIBLE CONFIGURATIONS FOR THE FOUR COLOUR PROBLEM

by

FRANCIS RONALD ALLAIRE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

WINNIPEG, MANITOBA

May 1977

"ON REDUCIBLE CONFIGURATIONS FOR THE FOUR COLOUR PROBLEM"

by

FRANCIS RONALD ALLAIRE

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

ABSTRACT

The theory surrounding reducible configurations and open sets of schemes is developed. This leads to two heuristics to distinguish configurations whose reducibility can be shown by the simple examination of Kempe chains. E-reducibility, an extension of D-reducibility, is described together with a means of checking for it. These developments are implemented in a computer programme that determines the reducibility of configurations. Techniques to minimise the time and storage requirements are described. Results of this programme are summarised in appendices. A last chapter indicates how reducible configurations are used to obtain further information about irreducible graphs.

Contents

Outline

This thesis describes the reducible configurations approach to the Four Colour Problem. The preface and first four chapters are introductory. The former details the source of the problem and its early history. Using the graph theoretical terms described in Chapter 1, the problem is formalized as the Four Colour Conjecture. The attack on this conjecture is indirect, i.e., we assume the existence of a counter-example. Such a graph with a minimum number of vertices is called irreducible, and Chapter 2 describes several easy to prove properties of such a graph.

Schemes and Kempe chains are described in Chapter 3, and from these are derived the Kempe constraints for realisable sets of schemes. Using these constraints, several more properties of irreducible graphs are demonstrated.

Finally, reducible configurations are defined in Chapter 4. The classical approach to proving the reducibility of a configuration involves the use of a reducer and perhaps the Kempe chain constraints. The problems with automating this approach are outlined and an alternate method is described. This involves the examination of the entire set of schemes and the use of all the Kempe constraints in the so-called Heesch algorithm. If these constraints are sufficient to demonstrate reducibility, then the configuration is called D-reducible. Otherwise, a suitable reducer must be found and its suitability verified in a final step.

In Chapter 5, we show that the Heesch algorithm determines a closure system for sets of schemes. Using the properties of this system, the first heuristic for distinguishing reducible configurations is developed and expressed as Conjectures 1 and 2.

Chapter 6 describes how the closure system can be used to help determine a suitable reducer for a D-irreducible configuration.

Chapter 7 develops the theory of anti-sets. This theory ties together most of the computed results. To begin, it explains why some configurations are not D-reducible. Further, if a D-irreducible configuration has a suitable reducer, the anti-set theory explains much, if not all, of the structure of the suitable reducer. If no suitable reducer exists, then this too can usually be explained by anti-sets through what is called a reduction obstacle. This leads to the second heuristic: only configurations with reduction obstacles are not reducible.

The phenomenon of E-reducibility is described and explained by anti-sets. Theorem 7.2 describes a property of maximal reducers which is used in the determination of E-reducibility. Numerous examples of anti-sets are given as well as an attempt to generalize these examples.

The computer programme which determines the reducibility of configurations is described in Chapter 8. The isotope form of the Kempe constraints is developed and an efficient way of implementing it is described. Other techniques to minimize time and storage requirements are outlined. Results of this programme are summarized in appendices 2 to 9.

The final chapter develops discharging schemes and shows how they are used with reducible configurations to obtain further information about irreducible graphs. Two face discharging schemes and two vertex discharging schemes are outlined. Appendix 1 describes fully a fifth scheme and how it is used to prove that an irreducible graph must have a 6-valent vertex.

Appendix 10 describes the quantitative form of the Kempe constraints and how this form is stronger than the logic form.

Preface

Around 1850 Francis Guthrie, a graduate student at University College, London, noticed that four colours were sufficient to colour the counties on a map of England. He wondered whether four colours would suffice for any map. The earliest written source indicating this problem is in a letter of Augustus De Morgan to Sir William Rowan Hamilton, dated October 23, 1852:

"A student of mine [Fredrick Guthrie, younger brother to Francis] asked me today to give him a reason for a fact which I did not know was a fact, and do not yet. He says, that if a figure be anyhow divided, and the compartments differently coloured, so that figures with any portion of common boundary line are differently coloured--four colours may be wanted, but no more. Query--cannot a necessity for five or more be invented?" May[20]

De Morgan communicated the problem to his students and to other mathematicians without arousing substantial interest. The first printed references are by Cayley [13], who in 1878 wrote to the London Mathematical Society and to the Royal Geographical Society asking whether the conjecture had been proved. Since then, numerous articles and even a few books have been written on this problem. Despite the simplicity of its statement, this conjecture is notorious for the several proofs that have been accepted and published, only to be discredited later.

While investigating a related three colour theorem, I was introduced to Ore's book The Four Color Problem [22]. My interest was spurred by his presentation of reducible configurations. The extensive bibliography in this book is complemented by that contained in a paper by Saaty 'Thirteen colorful variations on Guthrie's four color conjecture'

[24]. Saaty includes a section containing historical highlights of this problem. Although it has been investigated by numerous capable mathematicians, only recently has significant progress towards a solution been made. In their endeavours, several authors have introduced concepts that in turn have developed into fruitful areas of research. The offshoots of the problem enable Saaty to reformulate or strengthen this conjecture in over thirty ways. More than any other problem, I feel that the four-colour conjecture has been instrumental in the development of the branch of mathematics called Graph Theory.

The text of this thesis is formatted by a computer and this places restrictions on the character set. Since subscripts are not available, the full-size digits 0,1,2,... ,9 are used, as well as the letters i,j,k,m, and n. If a subscript is an expression, it is enclosed within round brackets.

The characters '+' and '*' are used for several operations depending on the operands. If the operands are numeric, then the operations are 'PLUS' and 'TIMES' respectively. For logic-valued operands, they mean 'OR' and 'AND'. For sets, they mean 'UNION' and 'INTERSECTION'. If this introduces no ambiguity in an expression, ')*(' may be replaced by ')('. Other logic symbols are '->' for 'IMPLIES' and '<->' for 'IFF'. Otherwise, '=' is used for EQUIVALENCE. For sets, '≤' and '≥' indicate 'SUBSET' and 'SUPERSET' respectively. Set 'DIFFERENCE' is indicated by '-'.

CHAPTER 1

Introduction

Guthrie's problem can be reworded as the following conjecture:

C0: Four colours are sufficient to colour the regions of any planar map.

Many early papers on this problem were phrased in terms of maps and colouring countries. Given any map, mark a capital city inside every country. Now for every two countries that have a common border, draw a railroad from capital to capital that does not pass through any other country. Each railroad must therefore cross the common border. If two countries have several border sections in common, then draw separate railroads from capital to capital, one across each section. On the other hand, if four or more countries meet at a common point, then some pairs of these countries may have only a point of their borders in common. The capitals of these countries will not be connected by a railroad. With this restriction, the railroads can be drawn on the map in such a way that they do not cross and they meet only at the capital cities. This resulting arrangement of capital cities and railroads joining them forms a network that is dual to the original map. If the countries of the map are coloured, then we can assign to each capital city the colour of the country. Now any two capital cities joined by a railroad are assigned different colours. Conversely, if the capital cities are coloured such that each city receives a colour different from all those cities to which it has a direct rail connection, then the map of countries can also be coloured. Thus, the problem of colouring the countries of a map is equivalent to the problem of colouring the dual network of cities joined by railroads.

The terms map and country are formalized in Graph Theory as map and face. Unfortunately, faces are a derived rather than an intrinsic part of a graph. On the other hand, the railroad-capital city problem translates directly to a problem in Graph Theory. The capital cities are represented by vertices and the railroads by edges joining the vertices. Possibly because this formulation can be generalised in several directions, it has been more common in recent publications. In this presentation, only the dual formulation will be used.

A graph G is a triple (V;E;I), where V is a finite non-empty set called the set of vertices, E is a (possibly empty) finite set disjoint from V called the set of edges, and I:E->V&V is a function called the incidence mapping. Here, V&V is the unordered product of V with itself. If I(e)=(v&w), we say that vertices v and w are incident with the edge e and that they are adjacent to each other. They are called the end points of the edge. An edge is a loop if both its end points are the same vertex. Two edges are parallel if they have the same end points. A graph is simple if it has no loops or parallel edges.

A graph G1=(V1;E1;I1) is a subgraph of G=(V;E;I) if V1≤V, E1≤E, I1=I, and I1:E1->V1&V1, i.e., for every edge of E1, both end points are in V1. We can also express G1 as G[V1;E1]. A vertex-generated subgraph G[V1] is the maximal subgraph of G on the vertex set V1, i.e., every edge of G with both ends in V1 is in G[V1].

A sequence of n edges e1,e2,...,en in a graph G is called an edge progression of length n if there exists an appropriate sequence of n+1 (not necessarily distinct) vertices v0,v1,,...,vn such that ei is incident with v(i-1) and vi, i=1,...,n. The edge progression is *closed* (*open*) if v0=vn (v0≠vn). If ei≠ej for all i≠j, the edge progression is called a *chain progression*. The set of edges is said to form a *chain* joining v0 and vn. The chain is closed if v0=vn. If only v0=vn and all the other vertices are distinct, the edges are said to form a *circuit*. A circuit of a graph is *proper* if only consecutive vertices are adjacent.

A graph is *connected* if each pair of vertices can be joined by a chain. A *connected component* (component) of a graph is a maximal connected subgraph. A graph is *p-connected* if each pair of vertices v and w is connected by at least p chains which have no vertices in common other than v and w. A graph is p-connected iff it is not disconnected or made trivial by the removal of p-1 or fewer vertices.

The *degree* or *valence* of a vertex is the number of edges incident with that vertex.

A graph can be represented by a drawing in which each vertex is represented by a point and each edge by a line segment joining the points corresponding to its end vertices. These points and lines are called vertices and edges respectively. A graph is *planar* if it can be embedded (drawn) in a plane (or on a 2-sphere) in such a way that no two edges meet except at a vertex.

A *k-colouring* of a graph is an assignment of k colours, one to each vertex of the graph, in such a way that no two adjacent vertices receive the same colour. A graph is *k-colourable* if it has a k-colouring. A *full* k-colouring of a graph is one that uses exactly k colours. A 3-colouring is a special type of 4-colouring, but it is not a full

4-colouring. A loopless graph is <u>k-chromatic</u> if it has a k-colouring but no (k-1)-colouring.

We can now state the Four Colour Conjecture:

C1: Every planar graph is 4-colourable.

While it is quite easy to show that every planar graph is 5-colourable, the open question is whether there exist planar graphs that <u>need</u> five colours or alternately whether four colours will always suffice. Note that loops and multiple edges do not affect the colourability of a graph. Further, the problem of colouring a graph reduces to colouring its connected components. Therefore we may restrict our attention to simple planar connected graphs.

In an embedding of a connected planar graph, the edges cut the plane into regions that are called the <u>faces</u> of G. One of these is the <u>outside</u> face, and for every face of G there exists an embedding with that face as the outside face.

A <u>face-boundary</u> of G is the minimal closed edge progression that forms the boundary of a face. A face is a triangle if its face boundary consists of three edges. If Q is a face boundary of G, then (G;Q) represents an embedding of G with Q the boundary of an inside face and (Q;G) an embedding with Q the boundary of the outside face. If Q is a circuit, the remainder of G lies on one side of Q, the inside or the outside, unless G consists solely of Q. A circuit Q made up of edges of G is <u>vertex-separating</u> in G if G has a vertex inside Q and a vertex outside Q. Similarly a circuit can be <u>edge-separating</u>.

If G is simple, then the vertex sequence Qn=(v1,v2,v3,... ,vn) corresponding to a face boundary can also be used to describe the face boundary. Note v0 has been dropped since it is the same as vn. Any sequence of n successively adjacent vertices such that the first and last vertices are adjacent and such that it is a face boundary in some subgraph of G is called an _n-ring_. The face boundary condition excludes sequences of vertices corresponding to some closed edge progressions, for example, those that cross. In a 4-colouring of G, the sequence of colours assigned to the vertices of Qn is such that consecutive colours are distinct. (The first and last colours are also considered to be consecutive.) Any such sequence of n colours is called an _n-scheme_. If Qn is the boundary of a face of a graph G and G can be 4-coloured with scheme s on Qn, then we say scheme s is _extensible_ to G. This assumes that Qn and s are aligned so that the kth colour in s is assigned to the kth vertex of Qn.

A graph is a _triangulation_ if it has a planar representation with all faces triangles. It is a _near-triangulation_ if every face but one is a triangle. A _k-wheel_ is the graph Wk formed from a circuit of k edges by adding a new vertex, the 'hub' vertex, and k new edges called 'spokes', with each vertex of the circuit joined to the hub by a spoke. In a simple triangulation with at least four vertices, each vertex is the hub of a wheel.

CHAPTER 2

Irreducible Graphs

Several steps in the attack on the Four Colour Problem use the indirect method of proof, i.e., a theorem is proved by assuming the opposite and deriving a contradiction from this assumption. For example, to prove that every planar graph is 4-colourable, we assume the opposite, the existence of planar non-4-colourable graphs. Since loops or multiple edges do not affect the colourability of a graph, each of these graphs has an underlying non-4-colourable simple planar subgraph. Since every planar graph is 5-colourable, these underlying graphs are 5-chromatic. Further, there will be 5-chromatic planar graphs with a minimum number of vertices, and these graphs are called irreducible. The size of their vertex set is called the Birkhoff number, designated B#.

Properties of irreducible graphs are investigated with the aim of showing that they cannot in fact exist. For example, it is obvious that an irreducible graph (if one exists) is connected and has no vertex of degree less than four. Further, if it has a cut vertex v whose removal disconnects G into non-null graphs G1=(V1;E1;I) and G2=(V2;E2;I), then G[V1+{v}] and G[V2+{v}] are 4-colourable. Any 4-colourings of these graphs can be made to have the same colour on vertex v and thus lead to a 4-colouring of G. Similarly, if G is 2-connected but has a cut pair {x,y}, two vertices whose removal disconnects G into G1 and G2, then G[V1+{x,y};E1+{(x&y)}] and a similar graph derived from G2 are 4-colourable and lead to a 4-colouring of G. The edge (x&y) was added to ensure that x and y are assigned different colours in both derived graphs so that again at most a permutation of the colours is necessary to assign the same colour to x in both graphs and the same second colour to y. This argument can be used to show that G cannot have a set of three vertices whose removal disconnects the graph, i.e., irreducible graphs are at least 4-connected.

Whitney's Theorem states that a 3-connected planar graph has an essentially unique representation in the plane [22]. Therefore we can refer to an abstract irreducible graph G by its planar embedding.

Theorem 2.1 An irreducible graph is a triangulation.

Proof If an irreducible graph G has a face that is not a triangle, this face will have four consecutive vertices w, x, y, and z in its boundary circuit Q. We can assume this face to be an inside face, and hence there are no edges inside Q. Since G is planar, at most one of the edges (w&y) or (x&z) can occur outside Q. Therefore we can assume that w and y are not joined by an edge and further that we can pull w and y together inside Q and merge them into a new vertex v, deriving a new planar graph G'. Now G' has one fewer vertex than G; so it must have a 4-colouring. This colouring can be applied to G by assigning to both w and y the colour on v to produce a 4-colouring of G. This contradiction invalidates the assumption that some face was not a triangle.∎

Corollary An irreducible graph is edge-critical, i.e., the removal of any edge results in a 4-colourable graph.

Theorem 2.2 An irreducible graph has no vertices of degree four.

Proof If an irreducible graph G has a vertex v of degree four, then its neighbours form a 4-circuit. As before, this circuit can be labelled wxyz in such a way that w and y are not adjacent. Now delete v and its incident edges to form

U, and from U produce G' by merging w and y. Now G' is planar and has fewer vertices than G; so it is 4-colourable. Any 4-colouring of G' leads to a 4-colouring of U with w and y assigned the same colour. Using the 4-colouring of U as a partial colouring on G, since only three colours are used among w,x,y, and z, the fourth colour can be assigned to v, thus producing a 4-colouring of G. This contradicts the irreducibility of G.∎

CHAPTER 3

Kempe Chains

In the preceding theorem, we derived from G a subgraph U and from U a reduced graph G' such that every 4-colouring of G' leads to a 4-colouring of U which can then be extended to a 4-colouring of G. If one or more 4-colourings of U cannot be extended directly to G, it may be possible to transform the 4-colouring on U to another 4-colouring of U that can be extended to G. These transformations are accomplished by Kempe chain interchanges and the proof that such an interchange is always possible is by an exhaustive examination of cases.

Consider a 4-colouring of a planar graph G with colours from the set {a,b,c,d}. There are three colour partitions of this set into two pairs of colours, namely (ab/cd), (ac/bd) and (ad/bc). For each colour pair, say {a,b}, let G(a,b) denote the subgraph of G consisting of those vertices coloured a or b and the edges joining them. The connected components of G(a,b) are called Kempe Chains belonging to {a,b}. A. B. Kempe used (misused?) these chains in his "proof" of the Four Colour Conjecture [18]. For a colour partition, G(ab/cd) denotes the subgraph of G whose edges are from G(a,b) or G(c,d). A Kempe interchange with respect to the colour partition (ab/cd) is an interchange of the colours a and b in one of the (ab)-chains or of the colours c and d in one of the (cd)-chains. If there are m Kempe chains in G(ab/cd), then 2^{m-2} possibly distinct 4-colourings of G can be obtained by fixing the colours on one chain of each colour pair, and performing or not performing the interchange on each of the remaining m-2 chains.

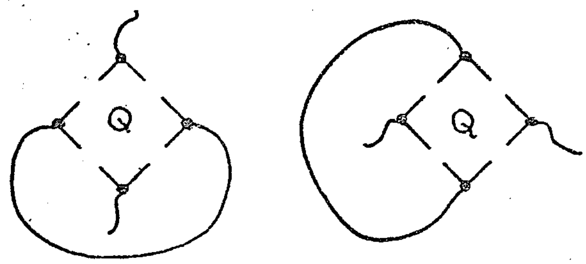


Figure 3.1 Q(ab/cd) has four components

Let U be a near-triangulation with the non-triangular (inside) face bounded by the circuit Q, and let U be 4-coloured. For any pair of colours {a,b}, the components of Q(a,b) may be connected in U by a Kempe chain of U(a,b). Since U is planar, this chain restricts the existence of Kempe chains joining the Q-components of the complementary pair of colours {c,d}. For example, if there are exactly four components of Q(ab/cd) then the Kempe chains of U(ab/cd) must conform to one of two dispositions with respect to Q. As in Figure 3.1, either the (ab)-components that meet Q are joined in U by a Kempe chain outside Q and this chain separates the (cd)-chains that meet Q, or the (cd)-chains that meet Q are joined and the (ab)-chains that meet Q are distinct.

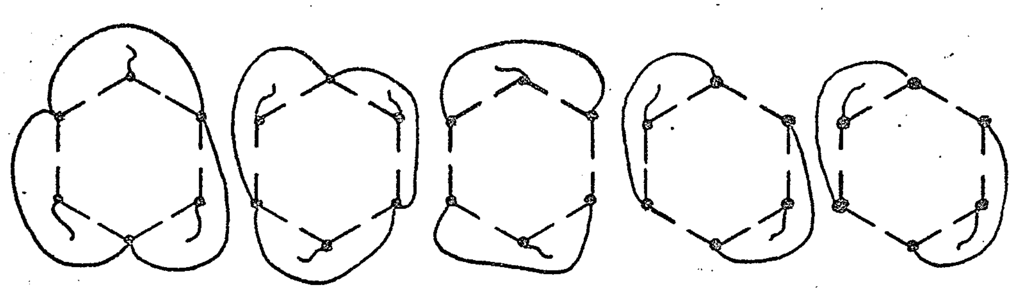


Figure 3.2 Q(ab/cd) has six components

Figure 3.2 shows the five dispositions consistent with Q(ab/cd) having six components. Note that the first and third generate all five by rotation. If Q(ab/cd) has eight components, the fourteen dispositions can be generated by the three in Figure 3.3.

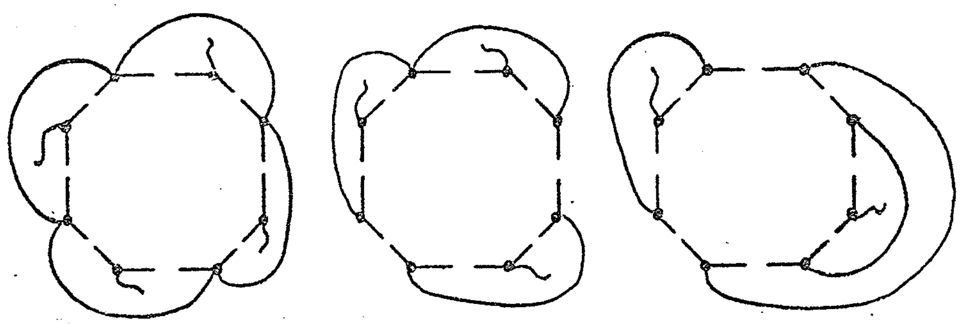


Figure 3.3 Q(ab/cd) has eight components

These limitations lead to constraints on the set of schemes on Q that are extensible to U. If Q is a 4-circuit, there are four distinct colour schemes that can appear on Q, namely, #1 abab, #2 abac, #3 abcb, and #4 abcd. If some 4-colouring of U assigns #1 to Q, then U(ac/bd) must have one of the two dispositions of Figure 3.1. In one, an (ac)-interchange is allowed that leads to a 4-colouring of U with #3 on Q; in the other, a (bd)-interchange is allowed that leads to a 4-colouring of U with abad on Q, which is essentially the same scheme (with respect to Q) as #2. Other interchanges using the Kempe chains of U(ac/bd) that meet Q do not lead to further distinct schemes on Q.

If #i also denotes a logic variable set to the true/false value of the statement:

#i on Q is extensible to U,

then we can describe the limitation by the constraint:

#1 -> #2 + #3.

Similarly we can obtain:

#2 -> #1 + #4

#3 -> #1 + #4

#4 -> #2 + #3.

All four constraints can be indicated by the expression:

#1 + #4 <-> #2 + #3.

In general, for a given scheme on Q, Q(ab/cd) will have 2k components. Denote the scheme on Q by (ø) and let (i) denote the scheme on Q obtained from (ø) by performing the Kempe interchange on the ith component (numbered consecutively from some starting point) of Q(ab/cd). The scheme on Q obtained by several Kempe interchanges with respect to (ab/cd) is similarly denoted (ijk...). Note that (ij)=(ji), and (ii)=(ø). Also, (2k-1)=(1357...2k-3) and (2k)=(246...2k-2).

From any scheme (ø) such that Q(ab/cd) has four components, we can derive four <u>4-component</u> or <u>4-break</u> constraints. As before, a single expresion can be used to indicate all four constraints.

(ø) -> (1) + (2)

(1) -> (ø) + (12)

(2) -> (12) + (ø)

(12) -> (2) + (1)

(ø) + (12) <-> (1) + (2)

For 2k=6 there are five dispositions, and so five summands in the constraint. For each disposition, three new schemes can be obtained on Q by one or more Kempe interchanges with respect to (ab/cd). For example, if the first arrangement of Figure 3.2 occurs and the dangling components are numbered 1,3, and 5, then (ø) can be transformed into any of (1), (3), or (5)=(13). This is expressed as the <u>product</u> (1)(3)(13). The full 6-break constraint is:

(ø) -> (1)(3)(13) +(2)(4)(24) +(1)(4)(14) +(2)(13)(123) +(3)(24)(234)

In general, Q(ab/cd) has 2k components and the constraint is a sum of products. There are C(2k,k)/(k+1) summands, one for each possible disposition, and each summand is a product of $2^{k-1}-1$ schemes. Here, C(n,r) is the number of ways of selecting an r-subset from a set of n elements, i.e., C(n,r)=n!/(r!•(n-r)!).

In a triangulation G with edge e=(x&z), the vertices adjacent to both x and z are called the apices of the edge e. Label them w and y. If e is replaced by the edge f=(w&y), we call f the conjugate of e. In effect, the edge (x&z) is twisted until it becomes (y&w) forming a graph Tw(G,e).

Theorem 3.1([12], 421; also [7]) If G is irreducible, then Tw(G,e) is not.

Proof Let G-e be the subgraph of G obtained by deleting e. Schemes #2 and #4 on wxyz are not extensible to G-e because these colourings are valid for G. If Tw(G,e) is also irreducible, then #3 and #4 must be forbidden as well. From #1 -> #2 + #3, if #1 on wxyz is extensible to G-e, then so is #2 or #3. Therefore #1 cannot be extensible to G-e. Thus, G-e must be non-4-colourable contradicting the fact that G is edge-critical.▪

Theorem 3.2 An irreducible graph is 5-connected.

Proof Let G be an irreducible graph and {w,x,y,z} a separating set. Since G is at least 4-connected, this set is minimal; and since G is a triangulation, it can be shown to form a circuit Q=(wxyz). Let U1 be the subgraph of G consisting of Q and the vertices and edges inside Q, and U2 the subgraph consisting of Q and the outside. U1 and U2 are 4-colourable, and the sets of schemes on Q that are extensible to U1 and to U2 must satisfy the scheme constraints. Further, no colour scheme on Q can be extensible to both U1 and U2, for that would lead to a 4-colouring of

G. There are only two disjoint pairs of non-empty solutions to the constraints on #1, #2, #3, and #4. In the first, one side, say U1, admits 4-colourings with either #1 or #2 on Q but no other schemes, and U2 admits 4-colourings with only #3 or #4 on Q. In this case, U1 with the face bounded by Q triangulated by the edge (w&y) would be non-4-colourable. Since G has at least one vertex outside Q, this contradicts the minimality of G. In the second case, only #1 and #3 are extensible to U1, and U1 with the edge (x&z) would be non-4-colourable.■

Corollary In an irreducible graph, every 4-circuit has a diagonal.

In summary, for every 4-circuit of an irreducible graph, one side of this circuit contains no vertices and is triangulated by a single edge. Further, if that edge is replaced by its conjugate, the resulting graph is 4-colourable. Before proving similar results for 5-circuits, we introduce reducing options.

Consider a near-triangulation U with the special face a 4-gon Q=(wxyz). If U with the edge (w&y) is 4-colourable, then either #3 or #4 appears on Q. We call this a reducing option for U and write it #3 + #4. Similarly, we get the option #2 + #4 if U with the edge (x&z) is 4-colourable. If vertices w and y are pulled together inside the special face and merged into one vertex and the resulting graph is 4-colourable, then we have #1 + #2 for U. Similarly #1 + #3 if U with x and z merged together is 4-colourable.

Suppose an irreducible graph G has a 5-circuit Q=(vwxyz). If there exists an edge (v&x), then (xyzv) is a 4-circuit and the side not containing w is triangulated by an edge, i.e., Q is triangulated by two edges. Further, no edge joining two vertices of Q can exist on the other side of Q;

for then G would have exactly 5 vertices and 9 edges, and this graph is 4-colourable. Otherwise every 5-circuit of G is a proper 5-circuit.

For Q any 5-circuit of an irreducible graph G, if the vertices and edges on one side of Q are deleted forming U, a 4-colouring of U assigns one of ten schemes to Q. Exploiting the rotational symmetry, they can be labelled as follows:

#11 ababc #12 abacb #13 abcab #14 abcac #15 abcbc
#21 abcad #22 abcdc #23 abcbd #24 abacd #25 abcdb

The set of schemes on Q that are extensible to U satisfies the following constraints:

#11 + #22 <-> #15 + #24
#12 + #23 <-> #11 + #25
#13 + #24 <-> #12 + #21
#14 + #25 <-> #13 + #22
#15 + #21 <-> #14 + #23

Unless they are joined by an edge, merging v with y produces the reducing option #13 +#14 +#21. If at least one vertex was removed to produce U, then Q triangulated by (z&w) and (z&x) produces the reducing option #11 +#21 +#23 +#24. Other reducing options are obtained by rotating these.

<u>Theorem 3.3</u> ([12], 430; also [1]) Let G be an irreducible graph with a 5-circuit Q=(vwxyz) whose inside consists of two diagonals e1=(w&z) and e2=(x&z). Then replacing these two diagonals by any other pair results in a 4-colourable graph.

<u>Proof</u> Two of the cases are Tw(G,e1) and Tw(G,e2). The remaining cases are Tw(Tw(G,e1),e2) and Tw(Tw(G,e2,)e1). Since G is irreducible, schemes #11, #21, #23 and #24 cannot appear on Q for any 4-colouring of Q and the outside. If

Tw(Tw(G,e1),e2) is also non-4-colourable, then #15, #25, #22, and #23 are excluded as well. But this violates the reducing option #15 +#11 +#23 which must hold for Q and the outside, and this supplies a contradiction. By symmetry, Tw(Tw(G,e2),e1) is also 4-colourable.∎

<u>Theorem 3.4</u> ([11]) If Q is a vertex-separating 5-circuit of an irreducible graph G, then either the inside or the outside of G consists of a single vertex.

<u>Proof</u> In the following table, #i in the left column indicates that #i on Q is extensible to the inside of Q. Similarly the right column indicates which schemes on Q are extensible to the outside of Q. Since G is irreducible, #i can appear in at most one column, and if it appears in one, we place ¬#i in the other.

Assume the outside has at least two vertices. Then by replacing the outside with a single vertex joined to all the vertices of Q, we get another reducing option #11 +#12 +#13 +#14 +#15 for the inside. By rotational symmetry we need examine only two cases. Either consecutively numbered (#11 follows #15) schemes admit a 4-colouring of the inside, or no two consective schemes from #11 to #15 admit a 4-colouring of the inside.

Inside	Outside	
#11,¬#12,¬#15	¬#11	Case 1
#23	¬#23	#11->#12+#23
#24	¬#24	#11->#15+#24
#21	¬#21	#24->#12+#21

But #11 +#21 +#23 +#24 is a reducing option for the outside.

#13,#14	¬#13,¬#14	Case 2
¬#21	#21	#13+#14+#21
#12	¬#12	#13->#12+#21

#15	¬#15	#14->#15+#21
¬#24	#24	#21->#13+#24
¬#23	#23	#21->#14+#23
#11	¬#11	#11+#21+#23+#24
¬#22	#22	#24->#22+#11
¬#25	#25	#23->#25+#11

Thus the outside can be 4-coloured only if #21, #22, #23, #24, or #25 appears on Q. If the inside of Q also has more than one vertex, then replacing the inside with a single vertex joined to all the vertices of Q produces a 5-chromatic graph on fewer vertices than G. This contradiction ensures that the inside of Q must be a single 5-valent vertex.∎

In summary, for every 5-circuit of an irreducible graph, one side of the circuit is triangulated by either two diagonals or a single hub vertex and five spokes.

Proper 6-circuits in irreducible graphs were investigated by A. Bernhart [7] who showed that there are only six pairs of disjoint sets satisfying the scheme constraints and several reducing options. For three of these solutions, the structure of one side of the circuit was detemined.

CHAPTER 4

Reducible Configurations

A graph is a <u>reducible configuration</u> if it cannot occur as a subgraph of an irreducible graph. It is minimally reducible if it has no reducible subconfiguration. For example, the 4-wheel is a reducible configuration. If this graph occurs as a subgraph of a planar graph, the hub vertex will still be 4-valent; so the graph cannot be irreducible. Similarly, any graph containing a 4-circuit with at least one vertex on each side or a 5-circuit with at least two vertices on each side is a reducible configuration.

If the four-colour conjecture is proved by some means, then irreducible graphs do not exist, and every graph is a reducible configuration. We restrict the definition of reducible to those graphs that can be shown to be reducible without first proving the four-colour conjecture. Using this restriction allows us to describe some configurations as not reducible.

Another type of reducible configuration resembles a separating circuit but may be thicker, and is called a <u>belt</u> [9]. This class of reducible configurations has been of limited use in deriving any further information about irreducible graphs.

A more common type of reducible configuration is called a <u>cluster</u> in [9]. It is composed of a boundary circuit Q of vertices of unspecified degrees whose inside is triangulated by a specified structure of vertices and edges, forming a near-triangulation $T=(Q;T)$. No inside edge is a diagonal of Q.

Since a cluster is a near-triangulation without diagonals, its structure can be described by indicating the degrees and adjacencies of the interior vertices. We use the following notation. The string 5-6[5665]-9 indicates the cluster configuration with 5 inner vertices: a central 6-valent vertex adjacent in order to vertices of degrees 5,6,6, and 5. These vertices are called the first neighbours of the centre as are the remaining two neighbours whose degrees are not specified. As a place-holder, we use an asterisk (*) to indicate a vertex whose degree is not specified. The inner vertices triangulate a boundary circuit of nine vertices of unspecified degrees. The configuration 6-8[56(5)75]-12 is similar except that it has a "cap", a vertex adjacent to two specified vertices. In this case the cap is 5-valent and adjacent to the vertices of degrees 6 and 7 on the side opposite to that of the central 8-valent vertex.

To reduce the number and levels of bracketting, a cap on two first neighbours of the centre may be abbreviated 0 for (5), 1 for (6), 2 for (7), and 3 for (8).

For large configurations, it may be desirable to use two central vertices with the vertex adjacent to both central vertices delimited by blanks. For example, 8-5[6078(5)05]-12 is also 8-78[56 5 555]-12. Since the cluster is a near-triangulation, the prefix and suffix can be derived from the remainder of the configuration; so they are usually omitted. In a later section, one configuration is derived from another by deleting a vertex or an edge. This configuration is indicated by placing a caret (^) over the digit corresponding to the deleted vertex, or a caret between the digits corresponding to the end vertices of the deleted edge.

In figures, the degree of a vertex that is the hub of a wheel may be indicated by the following key:

5-valent	•		
6-valent	△	at least 6-valent	✳
7-valent	o	at least 7-valent	(
8-valent	□	at least 8-valent	⊏
9-valent	▽		

n-valent, $n \geq 10$ ⓝ

If its degree is unknown, or if it is not the hub of a wheel, then a vertex is indicated by the point common to its incident edges, without any special symbol. Using this key, it is usually possible and convenient to represent a cluster configuration by drawing only the interior vertices, and the edges joining these. If necessary to avoid ambiguity, some of the remaining edges can be indicated.

In splice diagrams (Chapter 7), the degrees of the inside vertices are not important. So that the two edges incident with a 2-valent vertex are not misinterpreted as a single edge, the vertex is indicated by a full circle (•), or a stroke (+).

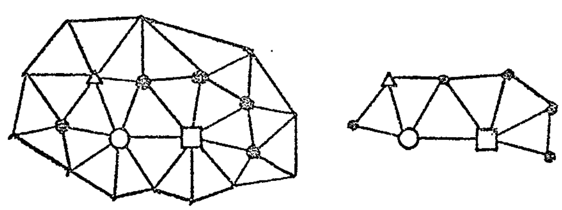


Figure 4.1 78[56 5 555] with/without boundary circuit

If (Q;T) appears as a subgraph of a planar graph G, then the circuit Q decomposes G into two near-triangulations, (U;Q) and (Q;T). With respect to their vertex and edge sets, the graphs T and U have G as their union and Q as their intersection. Conversely, U and T can be fitted together, filling the face of U bounded by Q with T. We write this as G=(U;Q)*(Q;T) or G=U*T. By describing Q as a sequence of vertices rather than as set of edges, the way T is oriented inside U can be well defined: qi of Q in (U;Q) corresponds to qi of Q in (Q;T).

In a simple planar connected graph U, any inside face bounded by an n-ring can be filled with a simple planar connected graph T whose outside face is bounded by an n-ring to produce a new graph. If these n-rings are not circuits, then it is convenient to represent U and T in such a way that the boundary Q is a circuit. This is accomplished through the following operation called <u>vertex splitting</u>. A vertex v is replaced by two images v1 and v2 joined by a double edge. Each vertex of the original graph that is adjacent to v is joined to one or both of these images. If the original graph is simple, it can be retrieved by deleting the double edge, merging its end vertices into one vertex, and removing any duplicate edges. If the original graph is planar and v is incident with the edges (u&v) and (v&w), then by splitting v along these two edges so that they form a 4-circuit (uv1wv2), we can join v1 and v2 by a double edge and this graph-like diagram is planar. Further, not both of (v1&u) and (v2&u) are required to ensure that the original graph is retrieved by merging v1 with v2. Two graph-like diagrams are <u>equivalent</u> if the same graph is produced from each by merging the ends of every double edge and removing duplicate edges.

By (Q;T) we will mean a planar graph-like diagram (which will be a graph if there are no double edges) equivalent to T in which the outside face is bounded by a circuit Q of normal edges. We can restrict the double edges to be diagonals of Q. If T has no loops, then any 4-colouring of T leads to a 4-colouring of any equivalent graph-like diagram, for example (Q;T), by assigning the colour on each vertex of T to all its images in (Q;T). The ends of every double edge are assigned the same colour, and the ends of every normal edge are assigned different colours. Conversely, any such 4-colouring of (Q;T) leads to a 4-colouring of T. The schemes that are extensible to T are extensible to (Q;T). Therefore a graph-like diagram is called a <u>colouring diagram</u>. A graph with a loop can be k-colourable, but when considered as a colouring diagram, it cannot be k-coloured for any k.

Now for any simple planar connected graphs U=(U;Qn) and T=(Qn;T), we can produce (U;Qn)*(Qn;T) by filling the face of (U;Qn) bounded by Q with (Qn;T). By U*T or G=(U;Qn)*(Qn:T), we mean the graph equivalent to the colouring diagram (U;Qn)*(Qn;T). If U and T are simple, planar, and connected, then U*T is planar and connected, but not necessarily simple, since it may have a loop. The 4-colourings of U*T lead to 4-colourings of (U;Q) and (Q;T) unless U*T has a loop, in which case some normal edge of (U;Q)*(Q;T) joins two vertices that are also joined by a chain of double edges. If this edge is in (U;Q), then the scheme assigned to Q in U*T is not extensible to (U;Q). This edge is a loop in U*T and its end vertex is assigned only one colour; but both images in (U;Q) cannot be assigned this colour since they are joined by an edge. Conversely, only those 4-colourings of U that assign to Q a scheme that is extensible to T can be used to produce a 4-colouring of (U;Q)*(Q;T). If no scheme is extensible to both T and U, then T*U either has a loop or is 5-chromatic.

The double edges were introduced only as an intermediate step to obtain U*T from U and T. This helps overcome problems associated with U and T being related to U*T as colouring diagrams, but not necessarily being subgraphs of U*T.

The classical method for determining the reducibility of a cluster configuration was outlined by Birkhoff [11], who phrased it in terms of deleting edges from a map. This restriction was removed by A. Bernhart [7]. In the dual formulation of colouring vertices, the method proceeds as follows. Assume the configuration (Qn;T) is a subgraph of an irreducible graph G=(U;Qn)*(Qn;T), and also that the circuit Qn is proper in G.

1. Replace (Qn;T) in G by a <u>reducer</u> configuration (Qn;R), where R is such that G'=(U;Qn)*(Qn;R) is 4-colourable. For example, if R has fewer vertices than T, then G' will be 4-colourable. Alternately, (Qn;R) could be obtained from (Qn;T) by replacing any edge not on Qn by its conjugate. The circuit Qn need not remain a proper circuit in G', for (Qn;R) may contain some arrangement of diagonals. These diagonals may be normal or double edges, i.e., two or more vertices of Qn in (U;Qn) may be images of the same vertex in G'.

2. Any 4-colouring of G' leads directly to a 4-colouring of (U;Qn) such that the scheme on Qn is extensible to R. All the R-extensible n-schemes are enumerated.

3. If every scheme extensible to (Qn;R) is also extensible to (Qn;T), then G can be 4-coloured and T is a reducible configuration. In Heesch's notation ([16]), T is <u>A-reducible</u> and others may use the term <u>directly reducible</u>.

On the other hand, if every such scheme can be shown to lead to a T-extensible scheme either immediately or by some sequence of Kempe chain interchanges in U, regardless of the structure of U or of the way the scheme on Q is extended to U, then G will again be 4-colourable, and T is called an <u>indirectly reducible</u> configuration. Heesch calls such configurations B- or C-reducible depending of the complexity of the argument required to prove that for every scheme extensible to R, there exists some sequence of Kempe interchanges that achieves the required result.

4. We now examine the possibility that Qn is not a proper circuit in G. The same argument applies except if the 4-colouring of G' is not valid on (U;Qn). This will occur if R merges two vertices of Qn that are joined by an edge of U. Therefore, for every pair of vertices of Qn that are joined by a double edge in (Qn;R), we must consider the possibility that these vertices are joined by an edge of U outside Qn in the original graph G=(U;Qn)*(Qn;T). If this implies the existence of an alternate reducible configuration, then the edge cannot exist.

While reducible configurations are defined in terms of being a subgraph of another graph, the way reducible configurations will be used in Chapter 9 does not exclude the possibility that two vertices of Qn in (Qn;T) could be images of the same vertex of an irreducible graph G=(U;Qn)*(Qn;T), i.e., in (U;Qn), they could be joined by a double edge outside Qn. Now if these images are adjacent in R, the scheme on Qn in G' is not extensible to R since a loop in G' is a normal edge of (Qn;R). For all pairs of vertices of Qn that are adjacent in R but not consecutive, we must consider the possibility that they are images of the same vertex in G, and supply an alternate reducible configuration.

In general, the definition of a reducible configuration must be extended to any graph (Qn;T) that cannot occur in (U;Qn)*(Qn;T), if the graph equivalent to this diagram is irreducible. Now to qualify a reducer R, we must show that U*R cannot have a loop, i.e., we must show that in (U;Qn)*(Qn;R), no normal edge joins two vertices that are also joined by a chain of double edges.

Consider the following example.

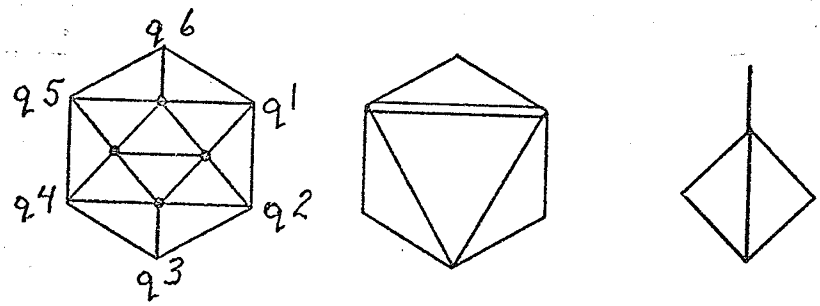


Figure 4.2 5[555] and a reducer

Theorem 4.1 The cluster 5[555] is a reducible configuration.

Proof This cluster is drawn in Figure 4.2 with its reducer. Call it T and the reducer R. If G=(U;Q6)*(Q6;T) is irreducible, then U*R is 4-colourable. If Q6 is a proper circuit in G, then this 4-colouring leads to a 4-colouring of (U;Q6) with Q6 assigned one of the six schemes extensible to R. Five of these are also extensible to T.

q1	q2	q3	q4	q5	q6	t1	t2	t3	t4
a	b	c	b	a	b	see below			
a	b	c	b	a	c	b	c	a	d
a	b	c	b	a	d	b	c	a	d
a	b	c	d	a	b	c	d	a	b
a	b	c	d	a	c	b	d	a	c
a	b	c	d	a	d	b	d	a	c

For the first scheme, we consider the six-break constraint that arises from the ac/bd chains.

(ø) -> (1)(3)(13) +(2)(4)(24) +(1)(4)(14) +(2)(13)(123) +(3)(24)(234)

Regardless of which of the five dispositions U(ac/bd) has, if (ø) appears on Q, then some Kempe interchange is allowed that transforms (ø) into (2) or (3) or (4). For ø=abcbab, the components of Q(ac/bd) are each a single vertex; if we number them 561234, then (2) is abcdab, (3) is abcbcb, and (4) is abcbad. Now (2) and (4) have already been considered, and (3)=abcbcb admits the colouring t1t2t3t4=dcda. Thus if Q6 is a proper circuit, then G cannot be irreducible.

If there is an edge of (U;Q6) joining q1 and q5 outside Q6, then q1t1q5 is a separating 3-circuit of G; so G cannot be irreducible. Similarly if q1 and q3 are images of the same vertex of G, then q1t2t3q3 is a separating 3-circuit of G. In general, if Q6 has a normal or double diagonal edge in (U;Q6), then U*T contains a reducible configuration. Thus, if Q6 is not a proper circuit, G cannot be irreducible. Therefore, 5[555] is a reducible configuration.∎

This classical procedure for determining reducibility does not lend itself to effective automation on a computer. First, a judicious choice of reducer must be made. If the choice is poor, it may not be possible to demonstrate reducibility, and the efforts expended may be wasted.

Secondly, the choice of scheme constraint to consider is important. The preceding example worked well, but suppose the (ad/bc) colour partition were considered. The constraint obtained would be:

abcbab -> abcbdb + abcbac

The scheme abcbac is extensible to 5[555], but the other scheme is not. To continue along this branch and investigate the scheme abcbdb requires a choice between the colour partitions (ab/cd) and (ac/bd), and again the choice may not be the best.

Returning to the example, suppose only (2) and (4) were extensible to the configuration. At least one scheme from each product (1)(3)(13) and (3)(24)(234) must be transformable into an extensible scheme. It may be advantageous to investigate the scheme (3) since it appears in both products, but there is no guarantee that the required sequence of Kempe chain interchanges exists. To eliminate one scheme from each product may in turn require the examination of several further schemes, and the number of schemes that must be examined may increase geometrically.

An alternate procedure was first adumbrated by Heesch[16] and more details can be found in [27] by Tutte and Whitney. Essentially, the steps of the classical procedure are reversed. In particular, the choice of a reducer is left until a later step, and in fact is often not necessary. Without a reducer to start with, we must investigate the entire set of n-ring schemes.

Assume that a configuration (Qn;T) is a subgraph of a colouring diagram (U;Qn)*(Qn;T) equivalent to an irreducible graph U*T=G.

1. (U;Qn) is 4-colourable but no colouring of U can assign to Q a scheme that is extensible to T. These schemes constitute the <u>initial forbidden set</u>.

2. The set of schemes that are extensible to U must nevertheless satisfy the scheme constraints. For each remaining scheme, we consider the constraints for which it is the left hand side. If, for any constraint, at least one scheme in each product of the right side is already forbidden, then that scheme cannot be extensible to U since that would imply that some forbidden scheme is extensible to U. In this way, the forbidden set is augmented until a maximal forbidden set is obtained.

3. Often, this <u>final set</u> is the entire set of n-ring schemes. This contradicts the fact that (U;Qn) is 4-colourable; so T must be a reducible configuration. Heesch calls these configurations <u>D-reducible</u>, and this is the case in which no reducer is required. Other authors may use

freely reducible, since any reducer can be used to reduce this configuration by the classical method.

4. Otherwise the configuration is called D-irreducible, but it may yet be reducible. We now examine the final forbidden set with the aim of finding a subset that is the extensible set for a configuration (Qn;R) that can serve as a reducer for T. If such a configuration exists, then since (U;Qn)*(Qn;R) is 4-colourable, it follows that (U;Qn) can be 4-coloured with a forbidden scheme on Qn. If the reducer has any single or double edges as diagonals of Qn, then the reducer must be qualified as before. Alternately, two or more reducers could be supplied such that at least one reducer is valid for every possible arrangement of single or double edges of (U;Qn) joining the vertices of Qn. ∎

The final set contains all schemes which can be forbidden by the simple examination of Kempe chains. If the reducibility of a configuration can be shown by this method, and if a reducer is required, then its extensible set must be contained in this final set of forbidden schemes. The only step that is not well-defined is how to find the reducer.

An extended form of the constraints exists, and this form occasionally forbids more schemes, and thus allows more freedom in the search for a reducer. As described in Appendix 10, this extension is orders of magnitude more difficult and lengthy to implement. Further, there are theoretical reasons why this extension cannot lead to a reducer for most configurations that are not reducible by the Heesch algorithm. The smallest configuration which may reduce by this method is 7[5665]-10, and the next is 7[56705]-11. All others are bounded by a circuit of at least 12 vertices. An attempt to reduce 7[5665] by this method has been made by E. R. Swart, who claims that the configuration is reducible. If this result is verified, this would be the first case that the reducibility of a configuration depends on the extended form of the Kempe constraints.

Chronology 1

Chapters 1 to 4 describe most of the published material on reducible configurations, that was available when I started my investigations in May, 1973. In his book [22], Ore uses the classical method to reduce several configurations. Without any knowledge of Heesch's work, I too realised the advantages of reversing the steps of the classical procedure. Using programmes written in APL, I verified the reducibility of the small configurations, those bounded by an 8-circuit or smaller.

For configurations bounded by a 9-circuit or larger, I switched to batch processing. Using a Fortran mainline to facilitate I/O operations, most of the work was performed by subroutines written in Assembler/360. This programme confirmed the reducibility of 5[56666], 7[55506], 7[550516], 7[560506], and 8[55555], which were listed by Ore and Stemple [23]. Further, I obtained four reducible configurations for which I could find no reference, namely, 6[5665], 7[5565], 6[505*605], and 7[555*505*]. While visiting Waterloo in December 1973, I was mentioning these results to Prof. W. T. Tutte when we were interrupted by a knock on the door to his office. Frank Bernhart was the caller, and after we were introduced, he confirmed my findings. These results, he said, were contained in his dissertation. Encouraged, I returned to Winnipeg to continue with the determination of reducible configurations bounded by a 10-circuit.

(to be continued)

CHAPTER 5

Sets of Schemes

Using the Heesch approach, we must investigate the entire set of n-circuit schemes. Sets of schemes have an underlying mathematical structure, namely, they form a partial order under the subset relation. Further, the finally forbidden sets can be described as closed. Investigating closure leads to several results that indicate the existence or non-existence of a reducer for a given D-irreducible configuration.

A set of n-schemes is realizable if it is the extensible set for a planar graph T bounded by an n-ring Qn. Relating this set to the graph, we indicate this set of schemes by (T). For example, the entire set of n-ring schemes is denoted by (Qn). In Chapter 7, a realisable set (T) will also be represented by a drawing of T or (Qn;T).

If (Qn;T) has a diagonal joining two non-consecutive vertices of Qn, then every scheme of (T) assigns to the ends of this edge the same colour if the edge is double, or different colours if the edge is normal. Such a configuration (Qn;T) is called degenerate, as is any set of schemes for which every element assigns different colours (or the same colour) to two non-consecutive boundary vertices.

Let (S) be a set of n-schemes and t any n-scheme. If there exists a constraint t->p1+p2+...+pk such that every product pi contains a factor from the set (S), or if t is in (S), then we say that t is simply immersible in (S). Let f(S) be the set of schemes that are simply immersible in (S). Let $f^0(S)=(S)$ and $f^k(S)=f(f^{k-1}(S))$, $k\geq 1$. A scheme is crudely immersible in (S) if it belongs to $f^k(S)$ for some $k\geq 0$. Since (Qn) is finite, $f^{k+1}(S)=f^k(S)$ for some k. This set $f^k(S)$ is the set of all schemes that are crudely immersible in (S) and is denoted Cl(S).

Subsets of (Qn) are partially ordered by set inclusion. For any set of schemes (S), Cl(S) is uniquely determined. For any sets (S) and (T) we have:

1. $(S) \leq Cl(S)$
2. $(S) \leq (T) \Rightarrow Cl(S) \leq Cl(T)$
3. $Cl(Cl(S)) = Cl(S)$

Therefore Cl is a closure relation ([19], 289), and Cl(S) is called the <u>closure</u> of (S). A set of schemes is <u>closed</u> if its closure is itself.

<u>Lemma 5.1</u> The intersection of closed sets is closed.

A set of schemes is <u>open</u> if it is the complement of a closed set. We now show that what we referred to as scheme constraints can more correctly be called <u>open set constraints</u>.

<u>Theorem 5.1</u> A set of schemes (S) is open iff:

for every s in (S)
and for every constraint of the form s->...
there exists a product (f1)(f2)...(fk) such that
fi is in (S) for all i=1,2,...,k.

<u>Proof</u> (S) is open
<=> (Qn)-(S) is closed
<=> s in (S) is not simply immersible in (Qn)-(S)
<=> the conditions of the theorem.∎

<u>Lemma 5.1'</u> The union of open sets is open.

While this definition of simply immersible in terms of open set constraints is well-defined, the theoretical results can be obtained and understood more easily in terms of the Kempe chains from which the constraints are derived.

<u>Lemma 5.2</u> A scheme t is simply immersible in a set of schemes (S) not containing t iff

there exists a colour partition (ab/cd), such that
for every configuration (Qn;T) such that t on Qn is extensible to T, then
for every extension of t to T, the Kempe chains of Q(ab/cd) are disposed in such a way that
there exists a sequence of Kempe interchanges with respect to (ab/cd), that transforms the scheme on Qn into some member of (S).

Lemma 5.3 A set of n-schemes (S) is open iff
for every s in (S), and
for every colour partition (ab/cd),
there exists a Kempe chain disposition such that
if s appears on Qn, then every scheme on Qn obtained by Kempe chain interchanges with respect to (ab/cd), and consistent with the disposition, is in (S).

Proof The negation of simply immersible contains the phrase:
there exists a configuration (Qn;T) such that s is extensible to T, and
there exists an extension such that the Kempe chains are disposed in such a way that
every Kempe interchange with respect to (ab/cd) transforms the scheme on Qn into some member of (S).

In particular, given a Kempe chain disposition with the required property, we can derive a configuration by joining with a single or double edge those components of Q(ab/cd) that are connected in the disposition, and this configuration will have the required property. ∎

Corollary A realisable set is open.

Theorem 5.2 If a set of schemes (T) is open and disjoint from (S), then (T) is disjoint from Cl(S).

Proof (T) and (S) are disjoint
$\Leftrightarrow$ (S) $\leq$ (Qn)-(T)

=> $Cl(S) \le Cl((Qn)-(T)) = (Qn)-(T)$, since (T) is open;

=> (T) and Cl(S) are disjoint.∎

Int(S), the <u>interior</u> of (S) is defined as the union of all open sets contained in (S). By Lemma 5.1', Int(S) is the largest open set contained in (S). Ext(S), the <u>exterior</u> of (S), is the interior of the complement of (S). $Ext(S) = Int((Qn)-(S))$, and is the largest open set disjoint from (S). By Theorem 5.2, Ext(S) is also disjoint from Cl(S) and, since the complement of Cl(S) is open, we have

$Ext(S)=(Qn)-Cl(S)$. Further,

$Int(S) = Ext((Qn)-(S)) = (Qn)-Cl((Qn)-(S))$.

To determine the reducibility of a configuration (Qn;T), we determine (T) and its closure Cl(T). The exterior of (T) can be obtained by complementing Cl(T). If $Ext(T)=\emptyset$, then the configuration realized by (T) is D-reducible. Tutte and Whitney call the set (T) <u>dominant</u>. Kurosh calls a set with a null exterior <u>dense</u>. Otherwise $Ext(T)\neq\emptyset$, and we can take its closure, and the complement of the closure is $Ext(Ext(T)) = Int(Cl(T))$. The next two results justify our interest in this set.

<u>Lemma 5.4</u> If (S) is open, then $(S) \le Int(Cl(S))$.

<u>Theorem 5.3</u> For a reducible configuration T with reducer R, we have $(R) \le Int(Cl(T))$.

<u>Proof</u> (R) is realisable; so it is open. R reduces T; so $(R)\le Cl(T)$. Thus, $(R)\le Int(Cl(T))$.∎

<u>Corollary</u> For a configuration (T;Qn), if $Int(Cl(T))=(T)$, then T is either directly reducible or has no reducer.

<u>Proof</u> If there is a reducer R, then $(R)\le Int(Cl(T))=(T)$; so T is directly reducible.∎

For a configuration T, if $Int(Cl(T)) = (T)$, call T <u>symmetrically D-irreducible</u>. The only configuration of this type that is known to be reducible is 6[666666]-12. The direct reducer for this configuration (the results of my programmes indicate that it is unique) was found by Franklin. By generalising this result, he proved the reducibility of a class of belt configurations [14]. Since this is the only exception in over a hundred examples I have found, it seems likely that the only symmetrically D-irreducible configurations that are reducible are of the form 2m[66...6]-4m, m≥3.

For a non-dominant realisable set (T), if Int(Cl(T))≠(T), then we call the configuration realised by (T) asymmetrically D-irreducible. Of the more than two hundred examples I have found of this type, none has failed to yield a reducer.

Conjecture 1 All symmetrically D-irreducible configurations, except possibly those of the form 2m[66...6]-4m, are not reducible by the simple examination of Kempe chains.

Conjecture 2 All asymmetrically D-irreducible configurations are nevertheless indirectly reducible.

Subsets of (Qn) do not form a topology. Two more conditions would be required ([19], 291):

4. Cl(S+T) = Cl(S) + Cl(T)

5. for all schemes s in (Qn), Cl({s}) = {s}.

Condition 5 holds, but not condition 4. Since the terms exterior and interior usually refer to sets in a topological space ([26], 53-62), care must be taken to ensure that topological results are not assumed for sets of schemes. For example, the open sets of a topological space form a distributive lattice ([26], 88-89). The lattice formed by the open sets of (Qn) is not distributive.

CHAPTER 6

Reducer Finding Routine

Not only does the use of closure indicate the existence of a reducer, it can also be used to help determine the structure of a reducer for a given configuration. For a D-irreducible configuration T, if a reducer R exists, then $(R)\leq Cl(T)$. Since (R) is open, we can refine this subset relation to $(R)\leq Int(Cl(T))$. For any pair of non-consecutive boundary vertices qi and qj, we can partition Int(Cl(T)) into (S^0) and (S^1) according as the schemes assign the same or different colours to qi and qj. If qi and qj are the same vertex in R, then $(R)\leq(S^0)$; if qi and qj are adjacent in R, then $(R)\leq(S^1)$. Again, these relations can be refined to $(R)\leq Int(S^0)$ and $(R)\leq Int(S^1)$. If $Int(S^0)$ is the empty set, then no reducer of T has qi, and qj merged together, and a similar conclusion follows if $Int(S^1)$ is the null set. Any restriction of Qn can be applied to Int(Cl(T)), and all extensible sets of reducers with this restriction will be subsets of the interior of this restricted set of schemes.

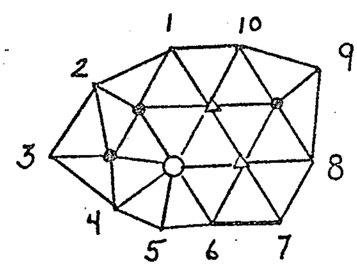


Figure 6.1 7[55606]

While these interior sets may be a union of open sets, they are sometimes realisable and hence form the extensible set for a possible reducer. For the configuration T= 7[55606], with the boundary vertices labelled as in Figure 6.1, Cl(T) contains 2096 schemes and Int(Cl(T)) has 1916 elements, a small improvement. If vertex q1 is restricted to have the same colour as vertex q6, then (S^0) is a set of 600 schemes and is open, i.e., $\mathrm{Int}(S^0)=(S^0)$. Since there are exactly 600 10-ring schemes with q1 and q6 assigned the same colour, (S^0) is precisely this set and a reducer is found. If q1 and q6 are joined by an edge outside T, then the graph will contain a vertex-separating 4-circuit. Therefore the reducer is valid, and 7[55606] is a reducible configuration.

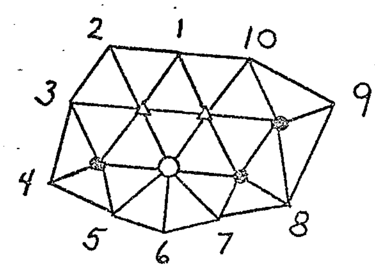


Figure 6.2 7[56605]

Secondly consider T=7[56605], drawn in Figure 6.2. The subset (S^1) of Int(Cl(T)) with vertex q1 a colour different from q6 contains 1033 of the 1173 schemes of Int(Cl(T)), and $\mathrm{Int}(S^1)$ has only 800 elements. There are 2461-600=1861 schemes with these two vertices assigned different colours; so if $\mathrm{Int}(S^1)$ is realisable, then this is only part of the structure. The remaining structure is found by noting that the same set of 800 schemes is the interior set of two other restrictions of Int(Cl(T)), namely, q1≠q5 and q1≠q7. Since only 800 10-ring schemes have q1 different in colour from q5, q6, and q7, the realisation of $\mathrm{Int}(S^1)$ is complete, and a reducer is found. If q1 is joined to any of q5, q6, or q7 by a double edge outside T, then the graph will contain a vertex-separating 3-circuit. Therefore the reducer is valid, and 7[56605] is a reducible configuration.

Occasionally the realisation of Int(S) will contain a non-degenerate part, for example, a pair or a triad of 5-valent vertices inside a 6-ring. More often, the only restriction on a sub-circuit of a reducer is that consecutive vertices receive different colours, for example, the sub-circuit q1q2q3q4q5 in both above reducers. These reducers each define a class of reducers, since any restriction on this sub-circuit will be realised by a subset of the extensible set for the unrestricted reducer. If we wish to verify the reducibility of a configuration by the classical method, then we can choose a reducer that comes closest to directly reducing the configuration.

This technique can be used in a brute force way by considering all possible diagonals. For an n-circuit there are n•(n-3)/2 non-consecutive pairs of boundary vertices and so n•(n-3) sets to consider. Hopefully, one of the non-empty derived interior sets can be realised. By using such a minimal restriction, these interior sets often are a union of unrelated open sets, and hence are not realisable. Then it is necessary to further partition or restrict one of these sets, and determine its interior before a reducer is obtained. The amount of computing required to determine the closure of a set of schemes increases markedly as the size of the boundary circuit. The cost of using this brute force method is high for 11-circuit configurations, and prohibitive for configurations bounded by larger circuits. Instead of starting with this method, it is advantageous to obtain, by other means, the partial structure of a likely reducer, and use this structure as an initial restriction. If it is not a reducer, and the interior set is non-null, then, hopefully, this set will be realisable. The remaining structure of the reducer can then be determined by the brute force method of considering all diagonals.

A disadvantage of using diagonals is that a reducer with diagonals must be shown to be compatible with all allowable outside diagonals. If a reducer with no diagonals can be found, then this check is not necessary. This method does not lend itself to finding reducers with no diagonals.

Chronology 2

Before investigating 10-ring clusters, I made several changes to my programmes. On June 4,1974, I sent to Frank Bernhart a list of approximately forty reducible configurations bounded by a 10-circuit. This list was restricted to configurations with no inside vertex adjacent to four consecutive or three non-consecutive boundary vertices. At this time, I was using a fixed set of approximately 500 reducer candidates. I also listed seven configurations for which no candidate from this set was a reducer. These configurations were:

7[5665]

7[56605]

7[506605]

8[55655]

7[56565]

7[555*606*]

*[5(5)06060(5)5].

Of these, I knew 7[506605] was reducible, but its reducer was 8[55655]. Later, I noticed that two candidates from my list reduced 7[555*606*], and that I had merely overlooked them in scanning my output.

Undaunted, I proceeded to examine 11-ring clusters. To make the transition from 9-ring to 10-ring clusters, I had to write a new subroutine to analyse the 10-break constraints. Without a similar obstacle, it was merely a matter of increasing the array sizes to get the 11-ring programmes operable.

In July, I recieved a letter from Prof. E. R. Swart of the University of Rhodesia, stating that he too was investigating reducible configurations bounded by a 10-circuit. In particular, he had found a reducer for 8[55655]. We agreed to present these 10-ring results in a joint paper [3].

At this time, little theory was available on open sets. Several terms were used for Exterior set, for example, 'bad' schemes, 'recalcitrant' set, and non-immersible set. It was Prof. Swart who suggested that the Exterior of the Exterior would be an interesting set to examine. He called this set the inversely non-immersible set. Theorems 5.2 and 5.3 and the two conjectures in Chapter 5 are his. Since 7[5665] and *[5(5)C6060(5)5] were symmetrically D-irreducible, these were eliminated from further consideration. This left only two undecided configurations, 7[56605], and 7[56565]. The word from the grapevine was that H. Heesch in Hannover had been able to reduce these configurations. However, he was keeping his results to himself.

In my investigations of 11-circuit configurations, I found that 7[56665] was reducible, indicating that 7[56565] should reduce. The exterior set for 7[56665] was the same as the exterior set for 8[565565] and 7[565075]. Eventually, I noticed that these three configurations all have the same substructure, namely, $7[56\hat{6}65] = 8[565\hat{5}65] = 7[56\hat{5}\hat{0}75]$. When I ran my programme with this substructure as the inside configuration, the exterior set was unchanged! Thus, this substructure was another reducer for these configurations. The same trick showed that a reducer for 7[56565] could be obtained by deleting the central 5-valent vertex. Several 11-ring configurations yielded reducers in this way, leaving only a few undecided 11-ring clusters.

In October, 1974, I uncovered the reducer-finding technique described in Chapter 6. Using it, I found a reducer for 7[56605] on Oct. 11. This completed the analysis of the 10-ring. On Oct. 16, I completed the analysis of the 11-ring.

I sent these results to Swart for verification, and also to F. Bernhart. Frank distributed these results to several people, including J. Mayer who used them to raise the Birkhoff number to 72.

(to be continued)

CHAPTER 7

7.1 Anti-Sets

The theory of anti-sets explains the D-irreducibility of many configurations. Further, it aids in the determination of a good initial restriction for the reducer-finding routine. Two non-null open subsets of (Qn) are called anti-sets if they are disjoint. For example, any non-dominant open set and its exterior are anti-sets. In general, this exterior set is a union of open sets, and each of these is an anti-set to the original open set.

For any edge-separating circuit Q of a non-4-colourable planar graph G=(T;Q)*(Q;U), we have (T)≤Ext(U) and (U)≤Ext(T). If G is irreducible, then (T) and (U) are non-null; so they are anti-sets. The 5-wheel is not a reducible configuration; so suppose that there is an irreducible graph G=W5*(Q5;U). Now Ext(W5)={#21,#22,#23,#24,#25}; and, since this set is minimally open, (U)=Ext(W5). Just as the set of schemes (W5) is represented by the configuration W5, we represent Ext(W5) by the pseudo-configuration of Figure 7.1, a 5-circuit with a circle inside it. Bernhart calls it the anti-pent [10], and we say the full 4-colourings of (Q5) are extensible to this anti-pent. Similarly, the full 4-colourings of (Qn) are represented by an n-circuit with a circle inside. For n≥6, these sets are not minimally open (Figure 7.2), and are distinguished from the anti-pent by inserting the digit n inside the circle.

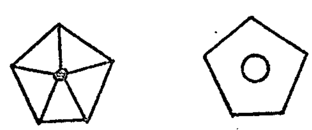


Figure 7.1 W5 and the anti-pent

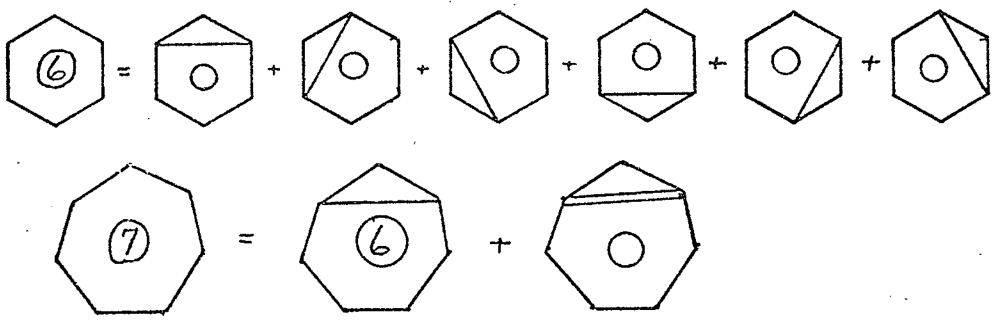


Figure 7.2 Ext(W6) and Ext(W7) are not minimally open

Two other non-degenerate 6-ring configurations are not reducible, namely, [55] and 5[55]. It is easily verified that Ext([55]) can be represented by the diagram of Example 7.1 below. This diagram represents all the 6-ring schemes that are extensible to its inside structure, i.e., for which a colour can be assigned to u such that the schemes on (q1q2q3q4u) and (q4q5q6q1u) are full.

For the other configuration, Ext(5[55]) cannot be represented in a similar fashion by anti-pents. This set is minimally open, and is represented by the pseudo-configuration of Figure 7.3, called the anti-triad [10].

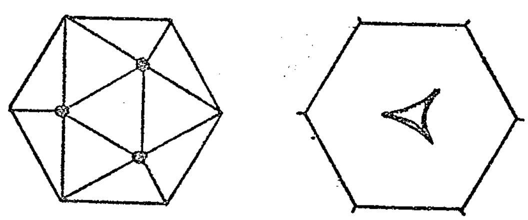


Figure 7.3 5[55] and the anti-triad

While a fourth is described in Appendix 3, one more pseudo-configuration is sufficient for this discussion. It represents Ext(6[565]). The configuration 6[565] is reducible but not D-reducible. Ext(6[565]) is not minimally open, but only 76 of the its 154 schemes can be represented by diagrams involving the anti-pent or anti-triad (see Appendix 3). This new pseudo-configuration is called the anti-diamond.

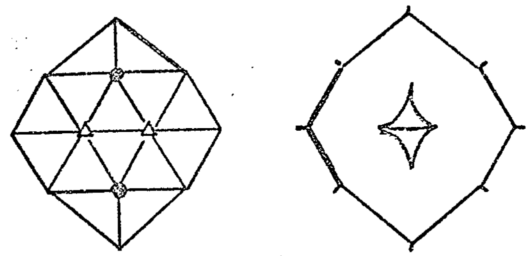


Figure 7.4 6[565] and the anti-diamond

Occasionally, the exterior set for a configuration will be represented as in Figure 7.5.

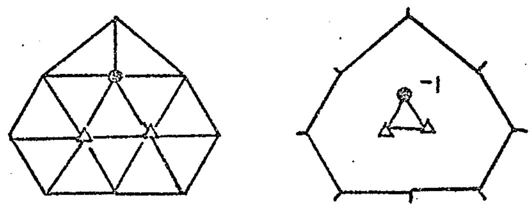


Figure 7.5 6[56] and its exterior set

If a colouring diagram is equivalent to a simple planar graph, then we call the colouring diagram consistent. For a consistent colouring diagram, a face is a region that corresponds to a face in the equivalent graph. A circuit corresponds to an n-ring in the corresponding graph. A colouring diagram is proper if every circuit bounding an inside face is a proper circuit.

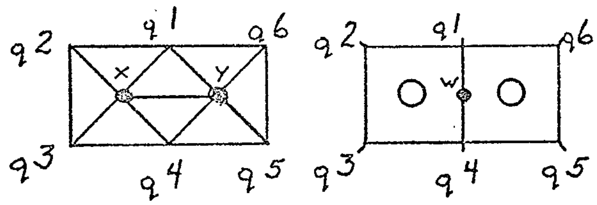
In a colouring of a proper consistent colouring diagram, any n-scheme can appear on an n-circuit bounding an inside face. A splice diagram is a generalisation of a proper consistent colouring diagram in which to each face is assigned a non-degenerate set of schemes. The splice diagram represents the set of boundary schemes that are consistent with the diagonals and extensible to the inside vertices in such a way that every face circuit is colored with a scheme from its assigned set. We also say that the schemes are extensible to the splice diagram.

Theorem 7.1 If the sets of schemes assigned to every face of a splice diagram are open, then the splice diagram represents an open set.

Proof Consider any scheme extensible to the splice diagram, an extension, and any colour partition. For every inside face, there exists a Kempe chain disposition for this colour partition such that all recolourings allowed by this disposition are in the open set assigned to this face. Assume such a disposition to occur inside every face. This assigns to the outside boundary circuit a disposition of Kempe chains such that every recolouring allowed is extensible to the splice diagram. Therefore the set of schemes represented by the splice diagram is open.■

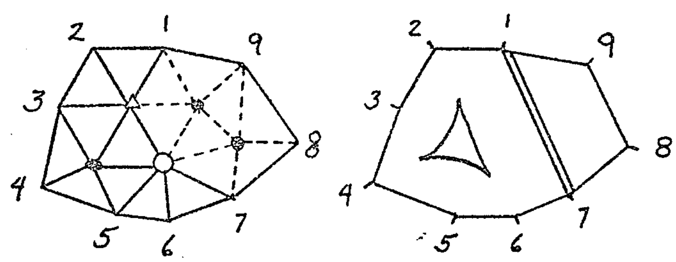
Splice diagrams conveniently express anti-sets to configurations. Since both sets are open, it remains to show that they are disjoint. To do this, we assume they are not, i.e., some scheme is extensible to both the configuration and the splice diagram. By imposing both extension conditions on such a scheme, if a contradiction is forced, then the scheme cannot exist and the sets are disjoint.

The way the contradiction is obtained is described by Bernhart as "chasing definitions". To begin this chase, there are four frequent starting points, described in Examples 7.1, 7.2, 7.3, and 7.4. The diagrams describe a pair of open sets, and this is followed by a proof that the open sets are disjoint. In the second and third examples, a counting argument shows that the splice diagram represents every scheme in the exterior set for the configuration. These exterior sets are determined by a computer programme.



Example 7.1 [55] and its exterior

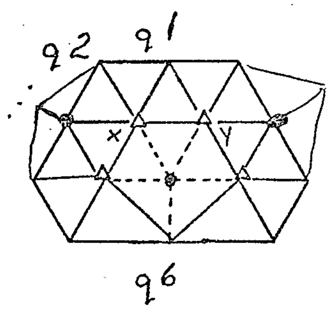
Proof Assume there is a 6-circuit scheme that is extensible to both the configuration and the splice diagram. Since x is adjacent to q1q2q3q4, at most three colours can appear among these boundary vertices. Since q1q2q3q4 and w comprise a full 4-colouring, there are exactly three colours among the boundary vertices, and the fourth is assigned to w. This fourth colour is also assigned to x. Similarly, w and y are assigned the same colour. But x and y are adjacent and cannot be assigned the same colour. Thus, the two open sets are disjoint.∎

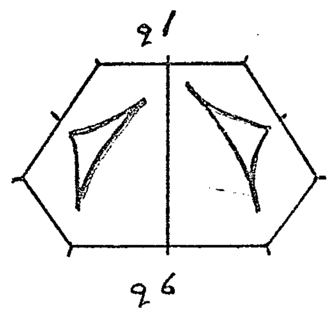


Example 7.2 7[5655] and its exterior

Proof Assume a scheme is extensible to both. For this scheme, q1=q7 and the 6-ring scheme on q1q2q3q4q5q6 is extensible to the anti-triad. This 6-ring scheme must be extensible to 5[55], since the colours assigned to the central vertex and first two first neighbours could be assigned to a triad of 5-valent vertices. Since no scheme is extensible to both 5[55] and the anti-triad, the comfiguration and splice diagram are anti-sets.

The anti-triad represents fourteen schemes, of which one is the 2-colouring. In the splice diagram, if the scheme on q1q2q3q4q5q6 is not the 2-colouring, then q8 can be coloured with any colour different from q7, i.e., a choice from three colours, and q9 any colour different from q8 and q1=q7≠q8, i.e., a choice from two colours. If the 2-colouring is used, then only three distinct schemes result. Therefore, the splice diagram represents 13•3•2+1•3=81 schemes. The exterior of 7[5655] has precisely 81 schemes; so the anti-set describes it completely.▪





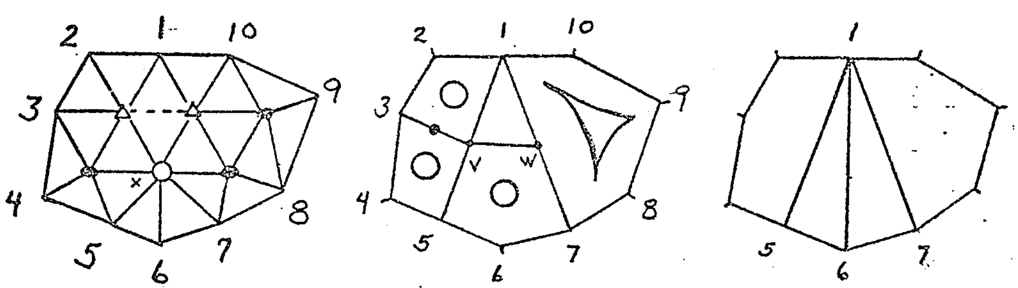
Example 7.3 5[606606] and its exterior set

Proof For any scheme in common, q1 and q6 are assigned different colours. If x and q6 are assigned different colours, then the scheme on q1q2q3q4q5q6 will be extensible to 5[55]. To see this, it is convenient to imagine q6 joined to q1 and to x in the configuration. But from the anti-set, the scheme on q1q2q3q4q5q6 is extensible to the anti-triad. Therefore x and q6 must be assigned the same colour. Similarly, y and q6 must be the same colour, and we have a contradiction.

The schemes of the anti-triad combine in this splice diagram to form 13•(13•2+1)+1•14=365 10-ring schemes. The exterior set has 365 schemes; so it is described entirely by the splice diagram.▪

In Example 7.2, the anti-set consists of a 'splice' of

Ext(5[55]) with (Q3) at the vertex q1=q7. In Example 7.3, Ext(5[55]) is spliced with itself along the edge (q1,q6), and in Example 7.1, Ext(W5) is spliced with itself along two edges (q1,u) and (u,q4). A splice along n edges is called an n-splice. The above anti-sets are examples of a 0-splice, a 1-splice, and a 2-splice. As above, one way to determine these sets is to combine every scheme of one set with every allowable scheme of the other set, and consider all permutations that produce distinct boundary schemes. Using this, it is easy to determine the number of schemes extensible to a 0-splice or a 1-splice if the sizes of the spliced sets are known. This is not as easy if the splice diagram is composed only of 2-splices because the resulting boundary schemes are not always distinct.



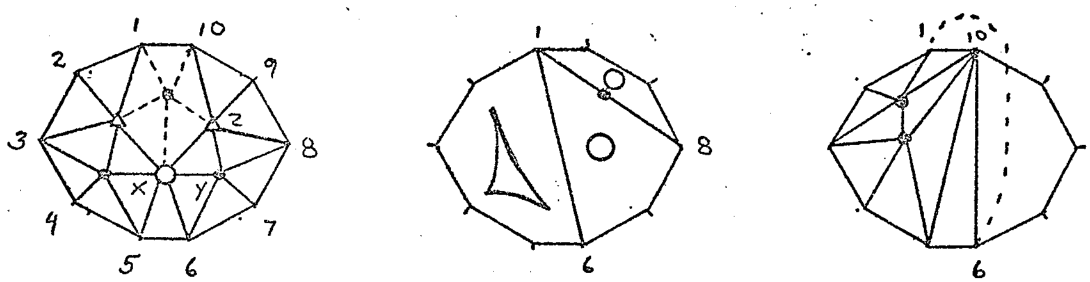
Example 7.4 7[56605], an anti-set, and a reducer

<u>Proof</u> Suppose the configuration and the splice diagram have a scheme in common. The colour assigned to x must appear at v or w since vwq7q6q5 is full. If v=x, then xq1q2q3q4q5 is extensible to [55], and the same scheme on vq1q2q3q4q5 is in the anti-set for [55]. If w=x, then xq7q8q9q10q1 is extensible to 5[55] and in its anti-set.■

<u>7.2 Reducers</u>

A reducer for a configuration must be disjoint from all the anti-sets of the configuration. If an anti-set has a diagonal that is a double edge, the reducer can ensure disjointness by a single edge and vice-versa. If an anti-set has the sub-configuration formed by q1wq7q6q5v of Example 7.4, then it can be avoided by the three edges (q1,q5), (q1,q6), and (q1,q7). For any scheme common to the anti-set and the reducer, the colour on q1 cannot appear at v, w, q5, q6, or q7; so the scheme on vwq5q6q7 cannot be full.

Sometimes the fact that an anti-set and a reducer are disjoint is less obvious.

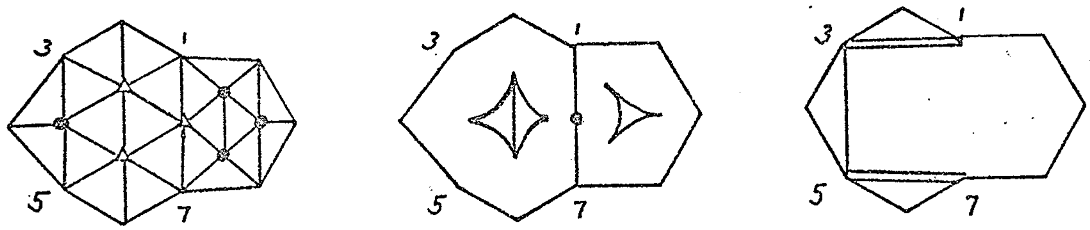


Example 7.5 7[56565], an anti-set, and a reducer

Proof that the reducer and anti-set are disjoint: For any scheme in common, q1≠q6; so in the reducer imagine an edge from q1 to q6 <u>around the outside of q10</u>. Now q1q2q3q4q5q6 is extensible to 5[55] by the reducer and to the anti-triad by the anti-set.∎

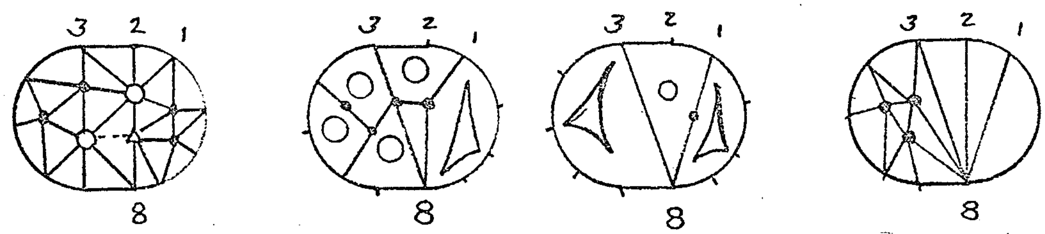
By symmetry, another anti-set is the mirror image of this one, but it is obvious that this reducer avoids it.

The only reducer for 6[565]-8 that can be demonstrated by simple Kempe chaining is q1=q3≠q5=q7. Therefore, if the anti-diamond is in a face of a splice diagram, the anti-set can be avoided by this reducer.



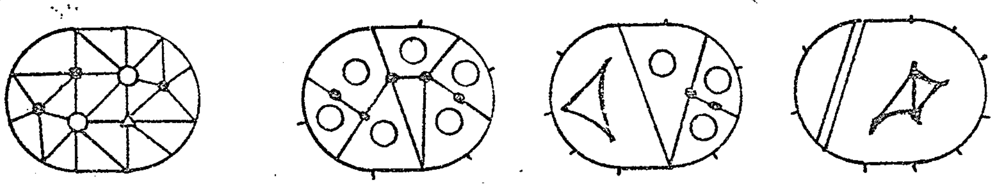
Example 7.6 6[606*505*], an anti-set, and a reducer

It seems that this idea also works in reverse. The anti-pent and anti-triad represent open sets that cannot be avoided by any valid candidate reducer for their respective configurations. Therefore, these open sets are <u>reduction obstacles</u> for these configurations. In Example 7.1, the 2-splice of two anti-pents is a reduction obstacle for [55]. If this anti-set can be avoided by some configuration that is a reducer for [55], then it seems reasonable that a reducer for W5 should be obtainable by isolating this reducer to one or the other anti-pent of the 2-splice.



Example 7.7 7[5570(5)6]-11, two anti-sets and a reducer

To avoid the first anti-set we may use q8≠q1q2q3. Since this precludes q3=q8, to avoid the second anti-set we fill the circuit q3q4q5q6q7q8 with 5[55].



Example 7.8 7[55706] and three anti-sets

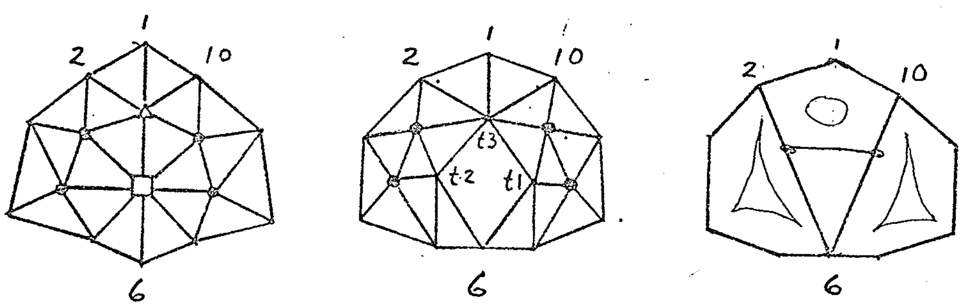
As before, to avoid the first two anti-sets, we could use q8≠q1q2q3 and 5[55] in q3q4q5q6q7q8; but then no consistent condition can avoid the third anti-set. Together, these three anti-sets form a reduction obstacle for 7[55706].

For all but a few symmetrically D-irreducible configurations, enough anti-sets can be found to provide a reduction obstacle.

7.3 Relaxations and Restrictions

Often, the proof that a configuration has a specified anti-set does not depend on every edge or incidence of the configuration. In particular, a configuration (Qn;T) may contain an m-ring Qm surrounding a subconfiguration (Qm;H), such that the anti-set remains an anti-set if the vertices and edges (normal and double) inside Qm are deleted, producing a configuration U. U has an outside face bounded by Qn and an inside face bounded by Qm, and is denoted by (Qn;U;Qm). (Qn;T)=(Qn;U;Qm)*(Qm;H). Since T=U*H, we use U=T/H to indicate that the configuration H has been deleted from T. Any configuration T/H is called a <u>relaxation</u> of T because any n-scheme on Qn extensible to T is extensible to T/H, i.e., (T)≤(T/H). To make the relation transitive, $T/H^1/H^2$ is also called a relaxation of T. Conversely, for any U=(Qn;U;Qm) and H=(Qm;H), U*H is a <u>restriction</u> of U. Only those schemes on Qn that are extensible to U, in such a way that the scheme assigned to Qm is in (H), can be extended to a colouring of U*H. If some relaxation of a configuration maintains all the anti-sets, then the exterior set will be maintained and, since the extensible set for the relaxed configuration is disjoint from this exterior set, this relaxation is a possible reducer for the configuration. If it is a reducer, then we call the configuration <u>E-reducible</u> or relaxation reducible.

For example, 7[5655] can be relaxed to 7[5655] without decreasing the exterior set. Since this relaxation has fewer vertices than the original, it can serve as a reducer. Since this reducer has no diagonals, no qualification is required, and 7[5655] is E-reducible.



Example 7.9 8[55655], a relaxation, and an anti-set

A qualification is required, however, if the relaxation has more vertices than the original configuration. By splitting the 8-valent vertex, the configuration T=8[55655] can be relaxed to T'=7[506^605] without altering the exterior set. To be a reducer, we must show that if G=(U;Q10)*(Q10;T) is irreducible, then G'=(U;Q10)*(Q10;T') is 4-colourable. From G', obtain G" by merging q6 with t2. Now G" has the same number of vertices as G, but t1 and t3 are 4-valent. Therefore G" cannot be irreducible, i.e., it must be 4-colourable. Since any 4-colouring of G" leads to a 4-colouring of G', G' is 4-colourable. Therefore, T' is a reducer for T and 8[55655] is E-reducible.

It should be noted that q6 can be identified with t2 without introducing any inconsistency. In particular, q6 cannot be joined to q2, q1, or q10 by a double edge since this would lead to a separating 3-circuit in G.

In general, a relaxation obtained by splitting a vertex of degree 9 or less is a reducer if it has the same exterior set as the original, since the graph corresponding to G" will have a 4-valent vertex. If a vertex of degree 10 or greater is split, then we must show that T" and hence G" contains a reducible configuration. This subconfiguration of T" may be composed of more than one vertex.

In Examples 7.1 to 7.8, those configurations drawn with one or more dashed edges are E-reducible. A maximal E-reducer can be obtained by removing all the broken edges and any resulting isolated vertices. In Example 7.9, a double edge from t1 to t3 is removed to form the E-reducer.

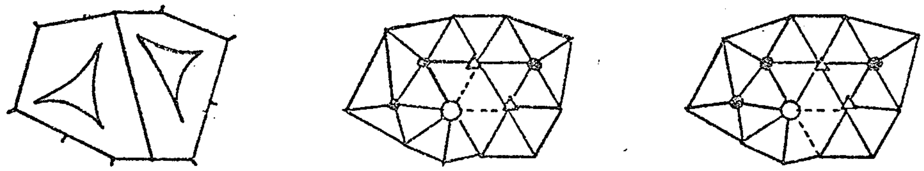
Not all relaxations of a configuration need be considered to discover E-reducibility. In particular, if an E-reducer contains a 4-valent vertex, then another E-reducer can be formed by deleting this vertex. This result can be proved in a way similar to the following.

Theorem 7.2 If a configuration (Qn;T) contains a D-reducible subconfiguration (Qm;H), i.e., T=(Qn;T/H;Qm;H), then Cl(Qn;T)=Cl(Qn;T/H)

Proof Let (U;Qn) be any configuration, and consider the graph (U;Qn)*(Qn;T/H;Qm)*(Qm;H). For any t^0 in Cl(T/H) such that t^0 on Qn is extensible to U, t^0 can be chained into some scheme t^1, where t^1 is extensible to U and in (T/H). An extension of t^1 to T/H assigns to Qm a scheme h^0. Now h^0 on Qm is extensible to U*T/H. Since H is D-reducible, h^0 is in Cl(H). Therefore the scheme h^0 can be chained into h^1, where h^1 is extensible to U*T/H and in (H). This transforms t^1 on Qn into t^2, and t^2 is in (T). Thus t^0 is in Cl(T) so $Cl(T/H) \leq Cl(T)$. Since T/H is a relaxation of T, $(T) \leq (T/H)$; so $Cl(T) \leq Cl(T/H)$, and equality follows.▪

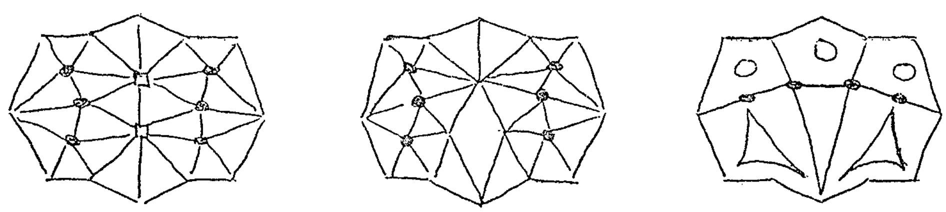
Now consider the configuration T=5[606606] of Example 7.3, and let T/e be the relaxation of T formed by deleting the dashed edge for which q6 and y are the apices. Let T/W5 be the relaxation of T formed by deleting all the dashed edges. Any scheme t^0 in Cl(T/W5) can be chained into a scheme t^1 in (T/W5). If an extension of t^1 to T/W5 assigns the same colour to y and q6, then t1 is in (T/e). If y and q6 are assigned different colours, then imagine them joined by an edge in T/e. Now the central vertex is the hub of a 4-wheel, a D-reducible configuration, and the scheme on the rim can be chained into a scheme that is extensible to the hub vertex. This transforms t^1 into t^2, a scheme in (T/e). Thus $Cl(T/W5) \leq Cl(T/e)$. Since T/W5 is a relaxation of T/e, $(T/W5) \geq (T/e)$; so $Cl(T/W5) \geq Cl(T/e)$, and equality follows. If T/e is an E-reducer, then (T/e) is disjoint from Ext(T) and, by Theorem 4.3, Cl(T/e) and Ext(T) are disjoint. Since $(T/W5) \leq Cl(T/W5) = Cl(T/e)$, T/W5 is also an E-reducer.

Note that the above discussion also holds for any reducer with a 4-valent vertex. Thus, every inside vertex in a maximal reducer is at least 5-valent, where reducers are ordered by their extensible sets using the subset relation. Note that a configuration may have more than one maximal E-reducer (Example 7.10).



Example 7.10 Ext(7[55606]) and two maximal E-reducers

Anti-sets that do not depend on every incidence of a configuration are more easily determined if those incidences are relaxed. This is especially true for anti-sets that remain anti-sets when a vertex is split into two images. Example 7.11 shows 8[5508055], a relaxation, and an anti-set for both. By symmetry there is another anti-set. Both anti-sets can be avoided by the configuration in Figure 7.6. While this configuration is not planar, it can be restricted to the second diagram in Figure 7.6. This restricted configuration is planar and is a reducer for 8[5508055].



Example 7.11 8[5508055], a relaxation, and an anti-set

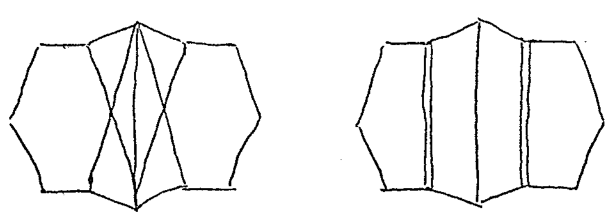
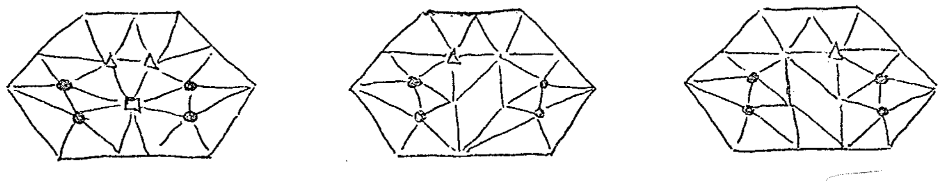


Figure 7.6 Anti-set avoider and a reducer for 8[5508055]



Example 7.12   8[556655] and two relaxations

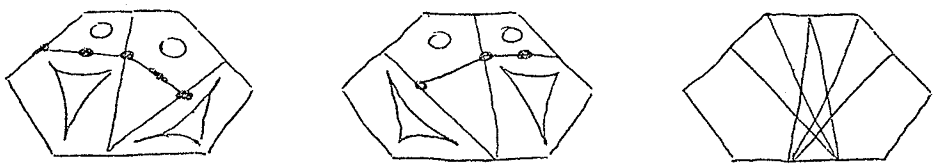


Figure 7.7 Two anti-sets for 8[556655] and an avoider

Example 7.12 shows 8[556655] and two relaxations. These relaxations have the anti-sets drawn in Figure 7.7, and these anti-sets can be avoided by the third drawing in that figure. This anti-set avoider has no restriction that transforms it into a planar configuration of the form (Q;R). Since there seems to be no other way to avoid both anti-sets, these anti-sets form a reduction obstacle for 8[556655]. Next to the 2-splice, a pair of anti-sets of this form is the most common reduction obstacle. Alone, each anti-set can be avoided by an inside configuration consisting of three diagonals. However, these diagonals intersect in such a way that both sets cannot be incorporated into any planar configuration of the form (Q;R). Note, however, the anti-sets and the reducer for 8[5566606]-13, listed in Appendix 9.

7.4 Determining Anti-Sets

If a configuration has no E-reducer, then we must resort to the seed technique for finding a reducer. Obviously, the more the exterior set can be accounted for by anti-sets, the better the chances that the seed will be a reducer, or at least be very close to a reducer, with the brute force method determining the remaining structure. The determination of anti-sets is largely by trial and error, but the following rules are helpful.

Rule #1 (0-splice) The following pairs of splice diagrams represent anti-sets.

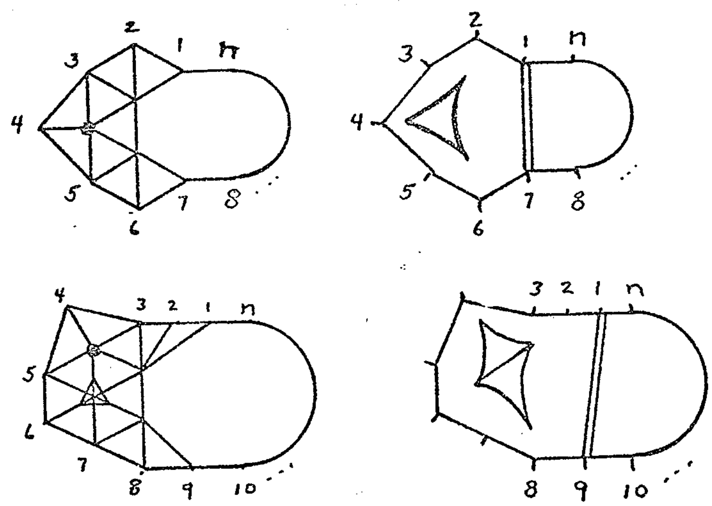


Figure 7.8 0-splice anti-sets

This is easily seen by superimposing one diagram of the pair on the other. If there is a scheme extensible to both diagrams, then some scheme is extensible to 5[55] and in its anti-set, or is extensible to 6[565] and in its anti-set.

Rule #2 (2-splice) If (A) and (A^{-1}) are anti-sets, then the following pair of splice diagrams represent anti-sets.

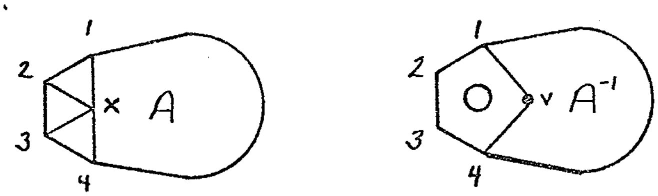


Figure 7.9 2-splice of anti-pent

Proof For a common extensible scheme, exactly three colours appear among q1, q2, q3, and q4; so the fourth cclour appears at both v and x. This implies a scheme is common to both (A) and (A^{-1}).∎

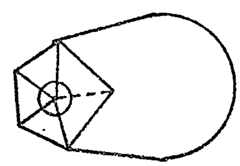


Figure 7.10 The missing edge

If cne diagram is superimposed upon the other, then it is the missing (dashed) edge that forces the same colour at both v and x to avoid an early contradiction. This idea of missing edges can be applied to other configurations.

If (A) and (A^{-1}) are anti-sets then so are the leftmost pairs of the following triples.

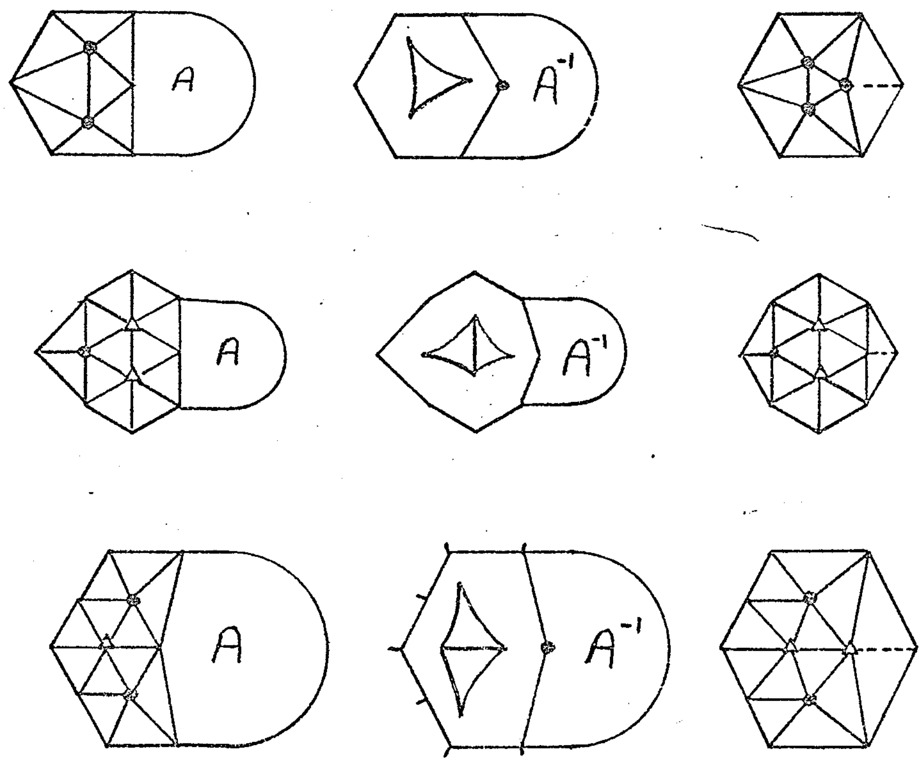


Figure 7.11 Other 2-splice anti-sets

If both open sets in a 2-splice are reduction obstacles for their associated configurations, then the 2-splice is a reduction obstacle for its associated configuration. If this configuration is also symmetrically D-irreducible, then it is almost certainly not reducible.

The first example of Rule #2 shows that any configuration with an inside vertex that meets four consecutive boundary vertices is not likely minimally reducible. If B and B^{-1}, and C and C^{-1} are also anti-sets, then so are these pairs.

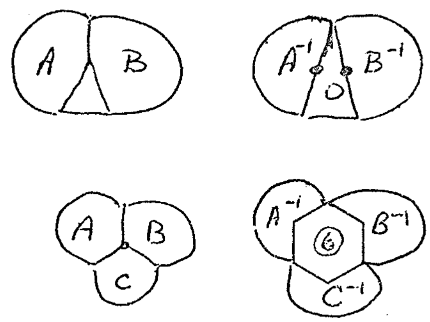


Figure 7.12 Reduction obstacles

Therefore, excluding the 4-valent vertex, every inside vertex of a minimally reducible configuration meets at most three boundary vertices, and if it meets three, they must be consecutive.

If one part of the 2-splice is the anti-triad, then usually the associated configuration has a pair of 5-valent vertices that are adjacent only to each other and to one other vertex in the configuration. An example is 6[505*55*]-8. This is not forced, however, as shown by the other two examples in Figure 7.13.

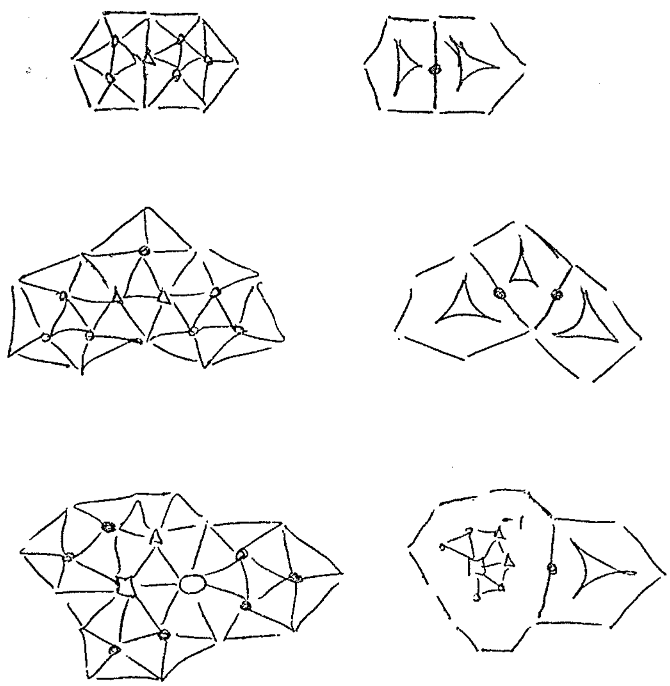


Figure 7.13 Anti-triad 2-splice reduction obstacles

The anti-sets in Figures 7.9, 7.11, 7.12, and 7.13 describe all types of 2-splice reduction obstacles for configurations bounded by a circuit of at most 15 vertices.

<u>Rule #3</u> If more than one edge is missing, then there is a choice of which two vertices have the same colour. A contradiction must be provided for each possibility, as in Example 7.4.

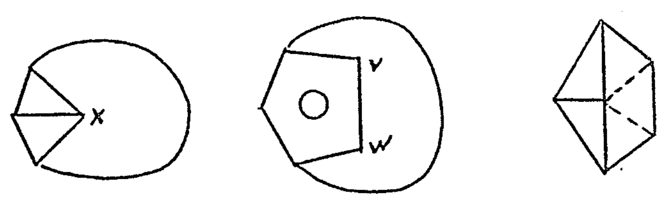


Figure 7.14 The colour on x appears at either v or w

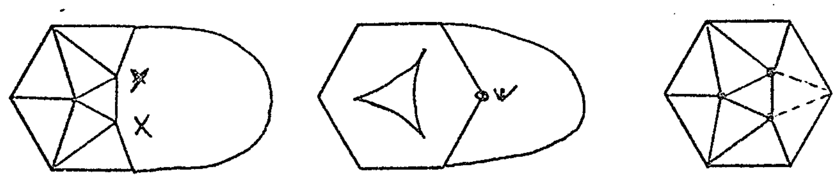


Figure 7.15 The colour on w appears at either x or y
Note: if w=q1 and y≠q1, then x=w=q1.

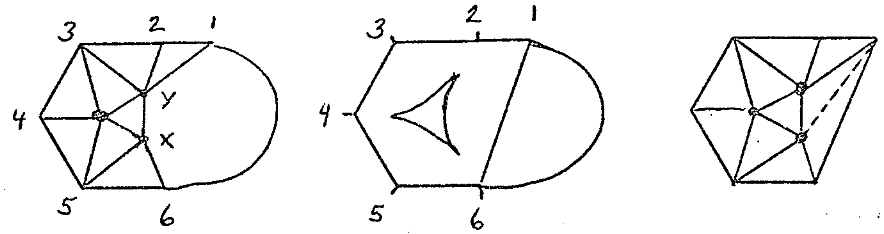
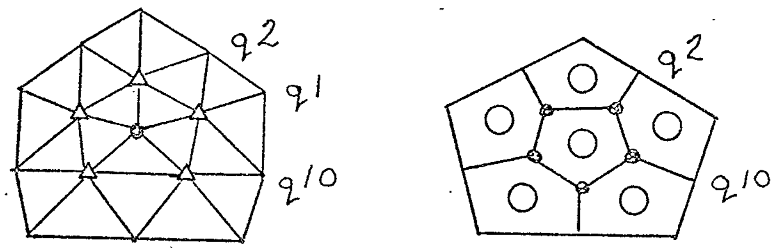


Figure 7.16 The colour on x is the colour on q1

7.5 More Examples

If any of the above subconfigurations is recognised in a cluster, the remaining structure of the splice diagram expressing an anti-set, if one exists, can usually be constructed vertex by vertex until the final contradiction is achieved. Consider Example 7.5, the configuration 7[56565]. From the previous discussion, the colour on x is the same as that on q1. Now vertex y is adjacent to x, q6, q7, and q8; and these four colours appear in the splice diagram at q1, q6, q7 and q8. Therefore join q1 and q8 to a new vertex w, and make this the boundary of an anti-pent. As in Rule 2, w and y must have the same colour. Now z is adjacent to x, y, q8, q9, and q10; and q10 is different in colour from x=q1. These same colours are on q1, w, q8, q9, and q10 in the splice diagram; so a contradiction is obtained if these vertices are the boundary of an anti-pent.

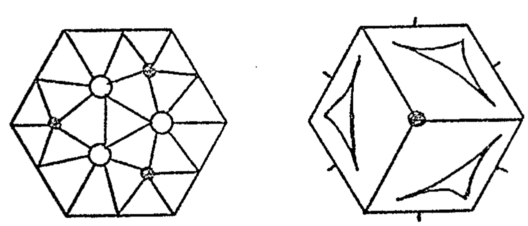
In 5[66666], label the 6-valent vertices x1,x2,x3,x4, and x5; with x1 adjacent to q10,q1, and q2; and xi adjacent to q(2i), i=1,5. In the anti-set, label the inside vertex adjacent to q(2i) by vi, i=1,5.



Example 7.13 5[66666] and an anti-set

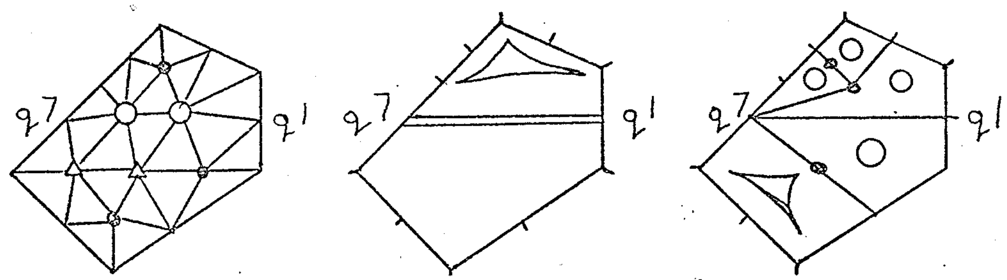
v1 or v5 is coloured as x1. If v1=x1, then v2=x2, v3=x3, v4=x4, and v5=x5. Now x1x2x3x4x5 is a 3-colouring; but v1v2v3v4v5 is full, a contradiction. A similar contradiction follows if v5=x1.

To reduce this configuration, Winn used q2=q4=q6=q8=q10 [28]. Now this colour cannot appear at vi, i=1,5; so any scheme extensible to the reducer cannot be in the anti-set. Another reducer is W10, the 10-wheel. For any scheme extensible to W10 and in the anti-set, the colour on the hub vertex is forced on v1 or v2, v2 or v3, v3 or v4, v4 or v5, and v5 or v1. Since v1v2v3v4v5 is an odd circuit, this is impossible.



Example 7.14 7[57075] and an anti-set

Label the 7-valent vertices x,y, and z. For any scheme in common, the colour on w appears at x or y, y or z, and z or x, an impossibility.



Example 7.15 6[567075] and its exterior

If all the anti-sets have been identified, then a reducer is easily checked. 6[567075]-12 has the above two anti-sets and the exterior set, as computed, has 6959 schemes. The first anti-set has 2741 schemes, all with q1=q7. The second is a 1-splice of the anti-sets for 6[55] and for 5[55*5*], and these configurations have exterior sets of 51 and 44 schemes respectively. Since the 7-circuit has no 2-colouring, the 1-splice of these exterior sets represents 51*44*2=4488 schemes, all with q1≠q7. Since all schemes in the exterior set are accounted for, we can identify three classes of reducers as pictured in Figure 7.17.

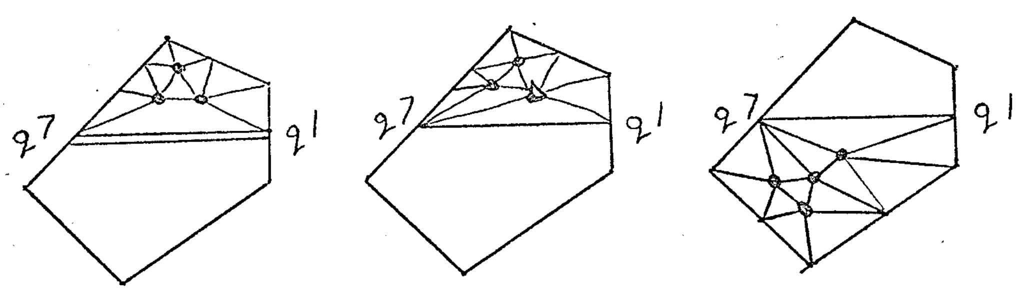
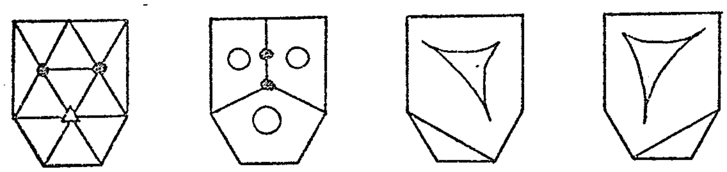


Figure 7.17 Three reducers for 6[567075]

Some anti-sets to configurations cannot be expressed by splices of anti-pents, anti-triads, or anti-diamonds.



Example 7.16 6[55] and three anti-sets

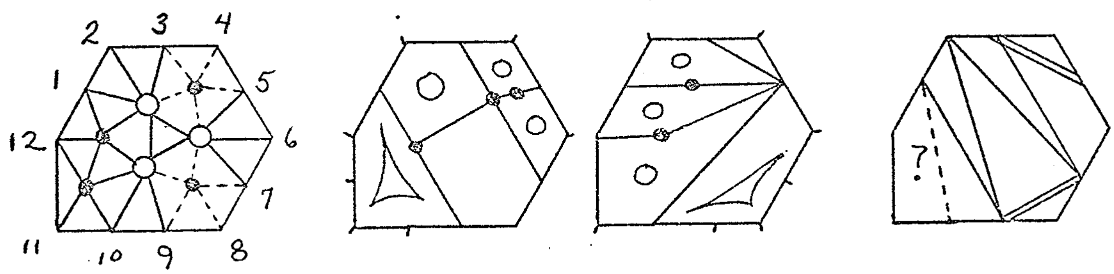
The exterior set for 6[55] has 51 schemes, but only 50 can be expressed by splice diagrams using the pseudo-configurations. In fact, the degenerate anti-sets are subsets of and completely cover the splice of three anti-pents; so they are in a sense superfluous.

Similarly, Ext(*[555]) has 50 schemes of which 48 are represented by a splice of three anti-pents. By using four pseudo-open sets, F. Bernhart was able to describe the missing anti-set as their union [10]. These pseudo-open sets are expressed by splice diagrams, each with a 4-face whose boundary is assigned not an open set of schemes, but a single scheme.

Except when the precise number of schemes in an anti-set is required, as in Example 7.15, we need not worry about these extra schemes. Since 6[55] and *[555] are not reducible, to avoid an open set represented by a splice diagram with a 7-circuit containing three anti-pents, either the configuration must be used, as in Example 7.15, or the anti-set is avoided by some restriction independent of these anti-pents. In Example 7.7, the restriction q8≠q1q2q3 was used. In either case, the missing anti-set will be avoided as a consequence.

Even reducible configurations can have unknown anti-sets. If an anti-set is maintained by some relaxation, then the exterior of the relaxation contains the anti-set. The union of these exterior sets of relaxations will contain all these anti-sets. If this union is not the entire exterior set for the configuration, then there is some anti-set that depends on the entire configuration. It is sometimes difficult to determine an expression for these anti-sets.

Usually the known anti-sets account for all but a few schemes of the exterior set and any configuration that avoids all these anti-sets is in fact a reducer. The smallest exception is 7[5665]-10 which is symmetrically D-irreducible. Since not nearly all the schemes of the exterior set are represented by splice diagrams using the three pseudo-configurations (see Appendix 3), a fourth pseudo-configuration can be introduced to represent this exterior set.



Example 7.17 7[557075], two anti-sets and a reducer

Another exception is 7[557075]-12. The two indicated relaxations contain these anti-sets in their exterior sets, and these exterior sets account for 9420 of the 9523 schemes in the exterior set for the configuration. These anti-sets can be avoided by $q7=q9$, $q7\neq(q1,q2,q3)$, and $q3=q5$ but this is not a reducer. Using it as a seed, the interior of the restricted set has 504 schemes, and it is easily determined that $q1\neq q10$ is the missing restriction for the determination of this reducer.

Chronology 3

The theory of anti-sets is the culmination of the works of several authors, often working independently. Since several small irreducible configurations have anti-sets that can be expressed as a 2-splice of anti-pents, the first case of Rule #2 was probably the first anti-set construction to be discovered. Although worded differently, this construction was known to Y. Shimamoto. He discovered a second construction which was essentially the first case of Rule #3. Assuming the existence of an irreducible graph, and using these constructions, he developed enough theory to indicate that the configuration 8[566665]-14 must have a particular anti-set in its exterior. Shimamoto's computer programme indicated that the configuration was D-reducible. Therefore, he concluded that the original assumption must be incorrect, i.e., irreducible graphs cannot exist. Tutte and Whitney greeted this proof with some misgivings, and then with real skepticism. In [27] they describe what can and cannot be done by the simple examination of Kempe chains. In particular, they concluded that the computer result must be wrong, as a rerun of the corrected program did show.

F. Bernhart introduced the anti-pent and anti-triad to denote sets of schemes, and the idea of splicing two open sets of schemes together. His thesis examined configurations bounded by a circuit of at most 9 vertices, so these pseudo-configurations were sufficient. Rule #1 can be attributed to him. W. Stromquist investigated configurations bounded by a 10-circuit, and discovered that 6[505*606*] and 7[505*565*] have anti-sets that can be expressed as the 2-splice of the anti-triad with the anti-diamond. F. Bernhart generalised these cases to Rule #2 on 2-splices. His proof was not worded in terms of superimposing the two diagrams, nor did he use the idea of missing edges to enumerate the possibilities that must be examined. Since Bernhart did not have a programme to analyse 10-circuit configurations, he discovered the special case of Rule #3, drawn in Figure 7.16, only after I sent him my results listing the sizes of the exterior sets for 7[55606] and 7[56565]. Using the results of my programmes, he was able to develop much of this theory of anti-sets.

Although tightly intertwined with this theory, the results on relaxations and restrictions are due to my investigations. After I had discovered that a reducer for a configuration could be obtained by deleting a vertex, it was a small step to realise that a less radical modification, namely, the deletion of an edge, might work as well. To avoid considering all possible edge deletions, I investigated maximal reducers, and this led to Theorem 7.2. In April, 1975, I was asked by JCT(B) for a second opinion on a paper by Heesch, in which he defined E-reducibility in terms of deleting vertices and/or edges. The term E-reducible is taken from this paper, although my discovery was independent of his. Since this paper gave no explanation of this phenomenon, merely its existence, I withheld approval. The other referee concurred with this decision.

Although the presentation in this thesis may make it seem obvious, it was not apparent that a configuration could be relaxed by splitting a vertex. I discovered this later in 1975, and it rounded out the theory of E-reducers so well that I rewrote a section of [3] to include this new type of reducer. It might be noted that the anti-set for Shimamoto's configuration, 8[566665], is more apparent if the 8-valent vertex is split into two vertices. Although applicable to the configuration that started the theory of anti-sets, this idea of splitting a vertex was the last to be uncovered. Also, this discovery enabled me to construct reduction obstacles for 8[55665] and 8[556655]. In particular, until I found them, no anti-sets were known for 8[556655], and it was believed that the extended form of the Kempe constraints would reduce this configuration.

CHAPTER 8

A Computer Programme to Determine the Reducibility of Configurations

Although it may not be immediately apparent in the description, the implememtation of the Heesch algorithm has the two standard computing bottlenecks, time and space. For every unit increase in the size of the boundary circuit, the space requirements more than treble. The computing required to determine the closure increases by a factor of about five. If all possibly reducible configurations bounded by a circuit of fixed size are to be examined, then, even after excluding those with a 2-splice reduction obstacle, the number of configurations that merit examination approximately trebles. Altogether, the problem escalates by a factor of fifty for every unit increase in the size of the boundary circiut.

This chapter describes the final versions of my programmes, those used to examine configurations bounded by a 12- or a 13-circuit. Although similar to the original versions used to examine configurations bounded by a circuit of six or seven vertices, several major modifications were made to decrease the time and space requirements.

Early in the development of this programme, it became apparent that the main computational effort would be required in the closure routine. Therefore the concept of immersibility was closely examined for an efficient implementation. Further, a data structure was selected on the basis of further efficiencies for this immersion routine.

To each boundary scheme is assigned a single bit of memory which can store a zero/one or false/true value for the scheme. The value of a scheme is the value of its corresponding bit. The set of schemes with value <u>zero</u> is called the <u>current set</u>. To determine the closure of a set of schemes (S), the current set is initialized to (S). As soon as it is determined that a scheme not in the current set is simply immersible in the current set, then the current set is augmented by that scheme, i.e., its corresponding bit is set to zero. When every scheme not in the current set is not simply immersible in the current set, the current set is the closure of the initial set. Because the current set is immediately enlarged, the closure is determined with less effort than that required to determine $f^1(S)$, $f^2(S)$,....

A scheme (∅), with value true, is simply immersible in the current set if it is the left side of a Kempe constraint (∅)-> RHS, where the value of the right hand side, (RHS), is false.

The general 4-break constraints are of the form:

(∅) + (12) <-> (1) + (2).

The immersion can be achieved by the following assignments:

```
TEMP12 <- (1) + (2)
   (∅) <- (∅) * TEMP12
  (12) <- (12) * TEMP12
TEMP12 <- (∅) + (12)
   (1) <- (1) * TEMP12
   (2) <- (2) * TEMP12
```

Excluding fetch and store operations, all possible immersions of these four schemes can be achieved by two OR and four AND operations.

Now consider the 6-break constraint:

(∅) -> (1)(3)(13) +(2)(4)(24) +(1)(4)(14) +(2)(13)(123)
+(3)(24)(234)

For every boundary scheme such that Q(ab/cd) has six components, there is a constraint of this form. For any such scheme (ø) and colour partition (ab/cd), there are fifteen related schemes with similar constraints:

(1) -> (ø)(13)(3) +(12)(14)(124) +(ø)(14)(4) +(12)(3)(23)
+(13)(124)(1234)

(2) -> (12)(23)(123) +(ø)(24)(4) +(12)(24)(124)
+(ø)(123)(13) +(23)(4)(34)

(12) -> (2)(123)(23) +(1)(124)(14) +(2)(124)(24) +(1)(23)(3)
+(123)(14)(134)

(3) -> (13)(ø)(1) +(23)(4)(234) +(13)(34)(134) +(23)(1)(12)
+(ø)(234)(24)

(13) -> (3)(1)(ø) +(123)(134)(1234) +(3)(134)(34)
+(123)(ø)(2) +(1)(1234)(124)

(23) -> ...

(123) -> ...

(4) -> ...

...

...

...

(1234) -> (234)(124)(24) +(134)(123)(13) +(234)(123)(23)
+(134)(24)(4) +(124)(13)(1)

Since most products are different, to evaluate the right sides of all sixteen constraints requires 16*5*2=160 AND operations and 16*4=64 OR operations. A further 16 AND operations are required, one to reset each bit.

A modified form of the constraint is obtained by multiplying (ANDing) each product in the RHS by the scheme on the LHS.

(ø) -> (ø)(1)(3)(13) +(ø)(2)(4)(24) +(ø)(1)(4)(14)
+(ø)(2)(13)(123) +(ø)(3)(24)(234)

...

This does not change the logic of the constraints, but now each product, for example, (ø)(1)(3)(13), appears four times in the sixteen modified constraints. The sets of schemes in these products are called <u>isotopes</u>, and the value of an isotope is the value of the product. Now, the evaluation of all right sides requires only 16*5*3/4=60 AND operations and the same 64 OR operations as before. As a bonus, the old value of the scheme is already ANDed into the sum. So, excluding fetch and store operations, all possible simple immersions can be effected by these operations.

For a set of sixteen 6-break constraints, this immersion routine can be accomplished with a limited number of stores and fetches of intermediate data, in particular, the isotope values. For the 8-break and larger constraints, a convenient pattern for determining, storing, and then ORing these isotopes is needed. This pattern is demonstrated for the 6-break constraints but, it is used only for the larger constraints.

The twenty isotopes can be grouped into five classes of four isotopes each, each class covering the set of sixteen schemes:

I^1: (ø)(1)(3)(13), (2)(12)(23)(123), (4)(14)(34)(134), (24)(124)(234)(1234)

I^2: (ø)(2)(4)(24), (1)(12)(14)(124), (3)(23)(34)(234), (13)(123)(124)(1234)

I^3: (ø)(1)(4)(14), (2)(12)(24)(134), (3)(13)(34)(134), (23)(123)(234)(1234)

I^4: (ø)(2)(13)(123), (1)(12)(3)(23), (4)(24)(134)(1234), (14)(124)(34)(234)

I^5: (ø)(3)(24)(234), (1)(13)(124)(1234), (2)(23)(4)(34), (12)(123)(14)(134)

If the schemes are arranged in the following 4x4 array, these isotope classes cover the array in five patterns.

	∅	3	4	34
∅	(∅)	(3)	(4)	(34)
1	(1)	(13)	(14)	(134)
2	(2)	(23)	(24)	(234)
12	(12)	(123)	(124)	(1234)

M

1	1	3	3
1	1	3	3
2	2	4	4
2	2	4	4

Map of I^1

1	3	1	3
2	4	2	4
1	3	1	3
2	4	2	4

Map of I^2

1	3	1	3
1	3	1	3
2	4	2	4
2	4	2	4

Map of I^3

1	2	3	4
2	1	4	3
1	2	3	4
2	1	4	3

Map of I^4

1	1	3	3
2	2	4	4
3	3	1	1
4	4	2	2

Map of I^5

If each isotope value is replicated in the positions corresponding to the schemes in the isotope, forming five matrices I^1, I^2, I^3, I^4, I^5, then we can write the constraints in matrix form:

$$M \rightarrow I^1 + I^2 + I^3 + I^4 + I^5$$

The matrix M is called a <u>constraint matrix</u>, but the constraints are obtained from M according to the patterns of the isotopes. From the matrix of original scheme values, each isotope value is determined and inserted in the appropriate positions of the isotope class matrix. These matrices are then ORed together, and the result becomes the matrix of new values for the schemes in M. If any bit was originally false, its new value is still false, since every isotope product involving it is false. If a particular bit starts at true, but every product involving it is false, then the scheme corresponding to the true bit is simply immersible in the current set, and the new value for that bit is false. If any isotope value is true, then all the schemes in that isotope have value true, and will remain set to true. These schemes are not simply immersible in the current set by this constraint matrix. For the corresponding colour partition, there exists a Kempe chain disposition such that all possible recolourings are in this isotope, and all these schemes are not yet in the current set. When the current set has grown to the closure of the original set, the open set formed by its complement will be a union of isotopes. Thus, an open set is a union of isotopes.

Separate subroutines are programmed with these isotope patterns. Each accepts a matrix of current values, performs the immersion, and returns a matrix of new values.

The drawback to this arrangement is that not all right sides of each individual scheme constraint needed to be examined in the first place. Only those constraints for which the scheme on the left side has value true are of interest, since only these schemes can be added to the current set. For any constraint in either the simple, modified, or matrix form, the same constraint pattern holds if all the schemes are rotated. Therefore the constraint can be rewritten with each scheme replaced by a vector of n schemes, n the size of the boundary circuit. Since some schemes do not rotate into n distinct schemes, some repetition may occur in this extension, and care must be taken to ensure that these extra bits do not interfere with the computations.

This extension costs nothing if the AND and OR instructions operate on a full word or register, as is true for the IBM 360/370 family of computers. Further, since a register of these computers has 32 bits, if n is at most 16, then the efficiency can be doubled by aligning two such bit vectors on the halfword boundary.

Other costs for using bit vectors are slight. Each rotation class must have a <u>header</u> or zero representative. One possibility is the lexicographic minimum. Instead, I used the following arrangement. If n is even, the scheme abab...ab is (0,0), i.e., the zeroth header, the zeroth (and only) scheme in this class. All other schemes have three consecutive vertices with different colours. The header is the lexicographic minimum of all rotations that start abc. To determine these headers we could examine all ways of colouring the n-3 remaining vertices. In fact, we examine only $(3^{n-3}+(-1)^{n-3})/2$ possibilities, since every scheme has some rotation starting abca or abcb, unless n is a multiple of four. In this case, abcdabcd...abcd is the zeroth and only scheme for the final header. These headers are numbered in ascending lexicographic order.

For the 7-circuit, the 91 schemes are represented by 13 headers. There is one 6-break constraint matrix and five 4-break constraint matrices. The details are listed in Table 8.1. Similar data for the 8-circuit to the 13-circuit are listed in Table 8.2.

	q1	q2	q3	q4	q5	q6	q7
(0,0)	a	b	c	a	b	a	b
(1,0)	a	b	c	a	b	a	c
(2,0)	a	b	c	a	b	a	d
(3,0)	a	b	c	a	b	c	d
(4,0)	a	b	c	a	b	d	b
(5,0)	a	b	c	a	c	a	b
(6,0)	a	b	c	a	c	a	d
(7,0)	a	b	c	a	c	b	d
(8,0)	a	b	c	a	c	d	b
(9,0)	a	b	c	a	d	a	b
(10,0)	a	b	c	a	d	b	d
(11,0)	a	b	c	b	c	a	d
(12,0)	a	b	c	b	c	b	d

	∅	2		
∅	(1,0) |	(2,0)	abcabac |	abcabad
1	(1,2) |	(4,3)	abcbabc |	abcbabd

(2,0) | (6,0)	(3,0) | (3,1)	(4,0) | (10,0)	(5,0) | (9,0)
(7,4) | (8,4)	(10,3) | (7,1)	(9,6) | (8,2)	(5,4) | (6,4)

	∅	3	4	34
∅	(0,0)	(1,6)	(2,0)	(3,0)
1	(0,6)	(5,6)	(4,6)	(8,6)
2	(9,0)	(10,5)	(12,5)	(7,5)
12	(12,3)	(11,2)	(6,5)	(11,5)

Table 8.1 Headers and constraint matrices for Q7

	Boundary Circuit								
	8	9	10	11	12	13			
|Qn|	274	820	2461	7381	22144	66430			
headers	41	94	257	671	1881	5110	A	B	C
4-break	10	14	22	30	43	55	2	2	2
6-break	4	10	22	42	80	132	4	4	5
8-break	1	1	5	15	43	99	8	8	14
10-break			1	1	6	22	16	16	42
12-break					1	1	32	32	132

A number of schemes/isotope

B number of isopopes/isotope matrix

C number of isotope matrices/constraint matrix

Table 8.2 Data for 8-circuit to 13-circuit

From Table 8.2, it is apparent that the amount of work involved in examining a single 12-break constraint matrix is approximately 12 times that required to examine a 10-break constraint matrix, and that is 12 times the work required for an 8-break constraint matrix. For configurations bounded by a 12-circuit or larger, to avoid using excessive amounts of computer time, it is advantageous to examine the small constraints more often than the large ones. To cause some overlap between the schemes examined, so that crude immersion is effected, the 4-, 6-, and 8-break constraint analyses are grouped together. They are examined repeatedly until some pass through all of them fails to immerse at least 5% of the remaining schemes. At this point, the 10-break constraints are examined, and these are followed by the small constraints unless fewer than 5% of the remaining schemes are immersed. In this case, the 12-break constraints are examined. This procedure is repeated until some examination of evey constraint fails to immerse a single scheme, or all the schemes have been immersed. At this point, the current set is the closure of the initial set.

Since a single scheme is not an open set, we can stop if all but one scheme has been immersed in the initial set, and conclude that it is dominant. This avoids the lengthy examination of the 12-break constraint matrix to immerse the 2-colouring, when it is the only scheme not yet in the current set.

Separate families of programmes were written to examine configurations bounded by circuits of each size from 7 to 13. The scheme headers are determined by a first programme and saved. For each header, a second programme determines the three constraint matrices of which it is a member. If a matrix contains a scheme with an earlier header, then that matrix has already been determined. The distinct 4-break and larger constraint matrices are also saved.

Two more programmes determine the reducibility of configurations. They both start with the following steps:

1) Input the scheme headers and the number of distinct schemes for each.
2) Input the constraint matrices.
3) Input a configuration. Since the boundary size is fixed, only the adjacencies of the inside vertices are required. A check is made to ensure there is no obvious error in this description.
4) For each header, there is an n-bit vector that corresponds to the schemes determined by that header. For every scheme that is extensible to the configuration,

set the corresponding bit to zero. Otherwise, set the bit to one.

5) Determine the closure of the current set.

The first programme continues:

6) If the configuration is D-reducible, indicate so and stop.

7) As a matter of course, the interior of the closure is now determined. This step may be delayed until later, and will be unnecessary if a reducer is found before then.

8) Check the configuration for an E-reducer. All relaxations of the following types are considered:

   a) delete a 5-valent vertex;

   b) delete an edge between two inside vertices, each at least 6-valent;

   c) split a vertex 8-valent or greater into two vertices each at least 5-valent. This split is restricted to those forming an empty inside 4-circuit, with one vertex on the boundary circuit.

   If any relaxation is an E-reducer, then this is indicated. After all relaxations have been tested, if any was an E-reducer, then processing is complete.

9) Otherwise, the above relaxations are reconsidered, and their exterior sets determined this time. The union of these exterior sets is also determined. The sizes of these sets are recorded.

This completes the first programme. If the configuration is E-irreducible, then it is investigated for anti-sets. If the configuration is asymmetrically D-irreducible, then one or more reducer candidates are determined. If it is symmetrically D-irreducible, then the object of the anti-set analysis is a reduction obstacle. In any case, if a likely reducer is determined, then this is checked by a second programme. This second programme starts with steps 1) to 5) above, and continues with:

10) Input the candidate reducer, and determine its extensible set. The size of this set is recorded. The closure determined in 5) is restricted by this extensible set, (their bit vectors are ANDed together). The size of the resulting set indicates whether the candidate is or is not a reducer. If it is a reducer, then processing is complete.

11) Otherwise the interior of this restricted set is determined. If the interior is the empty set, then either another candidate must be tried, or the configuration likely has no reducer.

12) Otherwise, this interior set is non-empty, and is likely the extensible set for a reducer. The remaining restrictions on this reducer are determined by the brute force method.

Results are summarised in Appendices 2 to 9. In particular, Appendix 7 lists symmetrically D-irreducible configurations, each of which has neither a 2-splice reduction obstacle nor a direct reducer. This list is complete for boundary circuits of sizes 5 to 11. Every other configuration bounded by a circuit of at most 11 vertices, with no 2-splice reduction obstacle, is reducible. For configurations bounded by a 12- or a 13-circuit, the same claim is limited to confirurations of at most 8 inside vertices.

The codes A1 and A2 are defined in Appendix 4 to specify frequent reducers for E-irreducible configurations. Using these codes, most configurations are in one of the following categories:

D-reducible;

E-reducible (but not D-reducible);

2m[666...6]-4m, directly reducible;

symmetrically D-irreducible, and likely irreducible;

asymmetrically D-irreducible, but with a reducer indicated by A1 or A2.

For the remaining configurations, the anti-set analysis usually leads to a reducer. Only occasionally are there insufficient anti-sets to completely determine the reducer.

These programmes used the computers at the University of Manitoba Computing Centre. The models used were an IBM 370/158, and later, an IBM 370/168. Reduction times can be compared to those given in [5]. For a configuration bounded by a 12-circiut, their programmes determine D-reducibility in about 60 seconds on a model 158. My programmes obtain the same result in approximately 10 seconds. For a configuration bounded by a 13-circuit, they use five minutes to my one. Using the 370/168, these times are all reduced to around one third these values. If a configuration is D-irreducible, then this takes two to three times as long to determine. Determining the interior of the closure further doubles the time used. If the configuration is E-reducible, this can be determined in a negligeable amount of time. Otherwise, the determination of the exteriors of the relaxations,(usually around 8 for a configuration bounded by a 12-circuit, or 11 for one bounded by a 13-circuit), again doubles the time taken. Finally, the candidate reducer can be checked after determining only one closure. If the candidate is not a reducer, then the time is again doubled by checking the n*(n-3) possible diagonals. Fortunately, only a few configurations need this final routine.

CHAPTER 9

9.1 Discharging

Reducible configurations describe which subgraphs an irreducible graph may not have. Excluding these and using a procedure called discharging, several results about irreducible graphs can be proved.

All discharging procedures use Euler's formula relating the numbers of vertices, edges, and faces of a planar connected graph:

$$|V| - |E| + |F| = 2$$

If each vertex and each face is assigned a 'charge' of +1 and each edge is assigned a charge of -1, then we can read the formula as:

'The sum of the charges on the vertices, faces, and edges is +2.'

The discharging rules manipulate these charges conservatively. Assuming that a particular set of configurations does not appear in a planar connected graph, the result of the discharging may be that every vertex, edge, and face is non-positively charged. Thus, planar connected graphs cannot avoid all the configurations in the set. If all the configurations are reducible, then the Four Colour Theorem is proved. If some configurations are not reducible, then every irreducible graph must have at least one of these subconfigurations.

Alternately, if the avoided set contains only reducible configurations, and the charges on the edges and faces are at most zero, and the charge on every vertex is at most 2/k, then there must be at least k vertices in such a planar graph, i.e., $B\# \geq k$.

Using various discharging procedures, the lower bound on the Birkhoff number has been raised from 40 (Ore-Stemple [23]) to 52 (Stromquist [25]). Using reducible configurations determined by my programmes, J. Mayer improved this to $B\# \geq 72$ [21]. Also, it has been shown that an irreducible graph, if one exists, must contain:

1) vertices of degree other than 5 or 7 (Heesch [17]);
2) vertices of degree 6 or 7 (Haken [15]);
3) vertices of degree 6 (Allaire [2]);
4) an adjacent pair of 5-valent vertices (Appel, Haken, and Mayer [6]);

Most recently, ([4] and [5], July 1976), K. Appel, W. Haken, and J. Koch announced that they have found an unavoidable set of approximately 1800 configurations, all of which are reducible. Thus, they have decided the conjecture affirmatively.

9.2 Face Discharging

The Ore-Stemple discharging used about forty reducible configurations to show $B\# \geq 40$. Their actual numerical computations, examining all possible cases, are admittedly cumbersome, and were not included in the publication. The face discharging that follows demonstrates $B\# \geq 20$, and uses only three reducible configurations. It indicates their method, and also the problems with getting better results by this method.

Consider an irreducible graph G. It is triangulated, 5-connected, and every vertex is at least 5-valent. The first step is to transfer all the charges to the faces, leaving the vertices and edges with no charges.

Rule 1 Each vertex distributes its charge (+1) equally to all the triangles of which it is an apex.

Rule 2 Each edge splits its charge (-1) equally between the two triangles of which it is an edge.

In this way the charge assigned to a triangle with vertices of degrees d1, d2, and d3 is:

$$c(d1,d2,d3) = 1/d1 + 1/d2 + 1/d3 - 1/2 - 1/2 - 1/2 + 1$$
$$= 1/d1 + 1/d2 + 1/d3 - 1/2$$

$c(5,5,5)=1/10$	$c(5,6,6)=1/30$
$c(5,5,6)=1/15$	$c(5,6,7)=1/105$
$c(5,5,7)=3/70$	$c(5,6,8)=-1/120$
$c(5,5,8)=1/40$	

Since the maximum charge on a triangle is 1/10, there must be at least 20 triangles. Since every face is a triangle, $3|F| = 2|E|$. From Euler's formula we get $|V| = 2 + |F|/2$. Therefore, there must be at least 12 vertices.

To improve this lower bound, several faces (or parts of faces) are grouped together and their charges averaged over the group. For example, only two triangles have a charge greater than 1/18, the (5,5,5) and (5,5,6) triangles. Since 5[555] is reducible, each outer apex of the edges of a (5,5,5)-triangle must be at least 6-valent. Therefore, these triangles are grouped with the central (5,5,5) triangle. Similarly, two outer apices of a (6,5,5) triangle are at least 6-valent because 6[555] is reducible. These triangles are grouped with the central one.

By evaluating the five cases in Figure 9.1 (these exhaust all possibilities), we conclude that every triangle has a new charge of at most 1/17.5, and B#≥20.

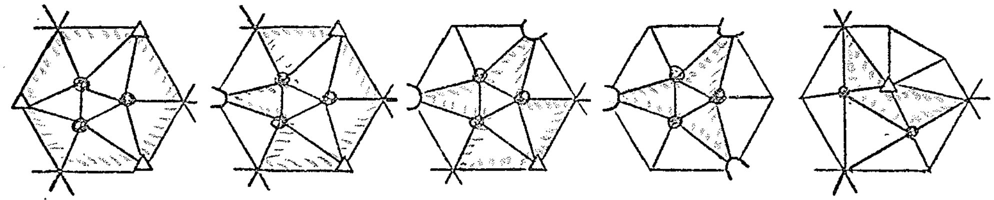


Figure 9.1 The five groups of triangles

What may be overlooked is the fact that every triangle must be assigned to only one group. For example, if the configuration 5[5665] occurred, then the (5,6,6) triangle would be grouped with both (5,5,6) triangles. Rather than investigate this new case, this obstacle is overcome by avoiding the reducible configuration 5[5665].

The grouping used by Ore and Stemple averaged the charge on every triangle to less than 1/37.

Alternately, Heesch's result is proved by discharging any planar graph composed entirely of vertices of degrees 5 and 7, and avoiding a list of 19 reducible configurations. Since 5[555], 7[5555], and 77[55 5 55] are reducible, there are only 7 distinct configurations of connected 5-valent vertices. These form the building blocks for the graph. The total weights of the triangles in these blocks are from the set {-1/14, 0, 1/14, 2/14, 3/14}. To eliminate fractions, all weights are multiplied by 14, thereby obtaining blocks of weights 3,2,1,0, or -1. Heesch showed that to every positive weight building block, enough blocks of weight -1 can be added, so that the total weight is at most zero, and no overlapping occurs in this arrangement. This completely discharges the graph, and we have the required result.

9.3 Vertex Discharging

Similar results can be obtained if the edge and face charges are transferred equally to their incident vertices. After such a transfer, an i-valent vertex has a charge of $1 - i/2 + i/3 = (6-i)/6$ units. Therefore, only vertices of degree 7 or more are charged negatively. These vertices are called <u>major</u>, and may be designated M. A 5- or 6-valent vertex is called <u>minor</u>, and may be indicated by m. To eliminate fractions, multiply all charge values by 6. If there are v^i vertices of degree i, then the charge summation reads:

$$v^5 - v^7 - 2v^8 - \ldots - (6-n)v^n = 12 \qquad (1)$$

where n is the highest vertex degree in the graph.

For Haken's result, assume the existence of an irreducible graph with no vertices of degree 6 or 7. Then 5-valent vertices have a charge of +1, 8-valent vertices are charged -2, 9-valent vertices are charged -3, and so on. The negative charge on a major vertex is distributed according to a weighting scheme to its adjacent 5-valent neighbours. This leaves the major vertices with zero charge. Excluding a short list of reducible configurations, a quick analysis proves that every 5-valent vertex receives enough negative charge to make the new charge on it at most zero. Since the charge summation is at most zero, the contradiction is achieved.

To settle the conjecture, Haken and Appel extended this discharging [4]. The way the extension is determined is interesting. The charge redistribution starts with a few basic rules. All configurations are examined to determine which contain 5-valent vertices that are not discharged by this first set of rules. If the configuration is reducible, or contains a reducible configuration, this reducible configuration is added to the set to be avoided. For each configuration for which no reducible subconfiguration can be found, a new rule is invented. These new rules perturb the previous results, and a new list of exceptions is determined. These are disposed of by more reducible configurations and more rules. For the third iteration, all exceptions are disposed of by reducible configurations, and the discharging is complete.

By the term 'reduction failure', they mean a configuration that may be reducible, but for which their program could not find a reducer. This is not surprising, since their reducer-finding routine is rather plebeian. By using a more sophisticated technique, they may be able to reduce these failures, and thereby shorten their analysis. Further, such a configuration may replace several over-reduced (not minimally reducible) configurations in their list.

Mayer discharges vertices in the opposite direction. At the outset, the only vertices with a positive charge are 5-valent, and that charge is +1. To eliminate future fractions, all charge values are multiplied by 10. Now, an i-valent vertex has an initial charge of 10(6-i) units. The first step distributes 2 units from each 5-valent vertex to each of its neighbours.

If adjacent 5-valent vertices are not allowed, then this completely discharges the 5-valent vertices. Also in this case, no major vertex becomes positively charged, since at most half of its neighbours can be 5-valent. However, a 6-valent vertex may have its charge raised from an initial value of zero to as much as +6. This charge is distributed to its major neighbours in a second step. A surprisingly short analysis demonstrates that no major vertex becomes positively charged by this second step, if fourteen reducible configurations are avoided. Therefore, every irreducible graph must have an adjacent pair of 5-valent vertices.

Using their discharging, Appel and Haken duplicated this result but their proof was much more involved.

To prove B#≥72, Mayer used a similar discharging scheme. It allowed a 6-valent vertex to receive as much as +4 units of charge in one instance, and at most +2 units in other instances. For the special case, some of the charge is redistributed to other 6-valent vertices, until every 6-valent vertex has at most +2 units of charge. Similarly, for the few cases in which a 7-valent vertex receives more than +14 units, and hence has a net charge of more than +4 units, a third set of rules transfers charge from these vertices to other major vertices. Examining all neighbourhoods for a central major vertex and excluding those with reducible configurations, the result of the discharging is that every vertex of degree i has at most 2(i-5) units of charge. Thus

$$120 = \sum_{v \in V} \text{charge on } v \le \sum_{i \ge 5} 2(i-5)v^i$$

$$\text{or } 60 \le \sum_{i \ge 5} (i-5)v^i$$

Adding (1) yields:

$$72 \le \sum_{i \ge 5} (i-5)v^i + (6-i)v^i = \sum_{i \ge 5} v^i = |V| \qquad (2)$$

I employed a similar approach in [2] to prove the non-existence of an irreducible graph with no 6-valent vertex. This paper is reproduced in Appendix 1. My discharging procedure is divided into four stages. The first two stages perfectly discharge every 5-valent vertex, but over-charge some 7- and 8-valent vertices. A third stage discharges most of these, and a fourth stage discharges the final three exceptions. To prove this claim, all neighbourhoods for a central vertex are examined under all possible charge transfer conditions which do not contain a reducible configuration. The determination of a maximum final charge on the central vertex is aided by several theorems which limit the amounts of charge transferred in certain conditions. Every neighbourhood for which this maximum is positive is individually examined. Although allowed separately, combinations of charge transfers are excluded by more reducible configurations. In this way, the final charge is demonstrated to be non-positive for every vertex.

I believe these four stages can be extended to discharge the general case, and thereby provide a second and independent proof of the four-colour conjecture by reducible configurations. The complexity of the two proofs will be comparable, and I expect over a thousand reducible configurations will be necessary.

Conclusions

The proof of the Four Colour Theorem was presented by W. Haken at the Summer Meeting of the American Mathematical Society, held at Toronto in August, 1976. Recalling the several previous false proofs, the general mathematical community was justifiably skeptical. Because this proof depends entirely on computed results, one common complaint is whether these results should be trusted.

While I cannot express any confidence in their discharging scheme, I do believe that the configurations in their list are in fact reducible. They do not develop or use any anti-set theory, but in the light of this theory, there are no glaring errors in their list. In particular, the configurations they list as D-reducible have no obvious anti-sets. More convincing of the correctness of their programmes is the fact that several C-reducible configurations conform to the anti-set theory. In particular, these configurations have exterior sets whose size can be accounted for by anti-sets exactly. Further, for such configurations the reducer they list is disjoint form the corresponding anti-sets.

It remains to provide a good discharging scheme, one whose correctness can be verified with a little patience. I believe that such a scheme can be obtained along the lines used by Mayer and myself ([2], [6], and [21]).

References

[1] M. O. Albertson, On irreducible graphs, Proc. 4th S.E. Conf. Combinatorics, Graph Theory and Computing, 1973, 125-129

[2] F. Allaire, A minimal 5-chromatic planar graph contains a 6-valent vertex, Proc. 7th S.E. Conf. Combinatorics, Graph Theory and Computing, 1976, 61-78

[3] F. Allaire and E. R. Swart, A systematic approach to the determination of reducible configurations, Scientific Report #82, Dept. of Computer Science, U. of Manitoba, 1976, to appear in J. Combinatorial Theory(B)

[4] K. Appel and W. Haken , Every planar map is four colorable, Part I: Discharging, to appear, Illinois J. Math.

[5] K. Appel, W. Haken, and J. Koch, Every planar map is four colorable, Part II: Reductions, to appear, Illinois J. Math.

[6] K. Appel, W. Haken and J. Mayer, Triangulations a v5 separes dans le probleme des quatre couleurs, submitted to J. Combinatorial Theory(B)

[7] A. Bernhart, Six rings in minimal five-color maps, Amer. J. Math., 69(1947), 391-412

[8] F. Bernhart, Some results on open sets of colourings, Research Report CORR 74-7, Faculty of Mathematics, U. of Waterloo, 1974

[9] -"- On the characterization of reductions of small order, Research report CORR 74-17, Faculty of Mathematics, U. of Waterloo, 1974

[10] -"- Summary on splicing diagrams, priv. communication, 1974

[11] G. D. Birkhoff, The reducibility of maps, Amer. J. Math., 35(1913), 114-128

[12] G. D. Birkhoff and D. C. Lewis, Chromatic Polynomials, Trans. Amer. Math. Soc., 64 (1946), 355-451

[13] A. Cayley, On the coloring of maps, Proc. London Math. Soc., 9(1878), 148. See also Proc. Roy. Geographical Soc., 1(1879), 259

[14] P. Franklin, The four color problem, Amer. J. Math., 44(1922), 225-236

[15] W. Haken, An existence theorem for planar maps, J. Combinatorial Theory(B), 14(1973), 180-184

[16] H. Heesch, Untersuchungen zum Vierfarbenproblem, Bibliographiches Institut AG, Manheim, 1969

[17] -"- Chromatic reduction of the triangulation Te, e=e5+e7, J. Combinatorical Theory(B), 13(1972), 46-55

[18] A. B. Kempe, On the geographical problem of the four colors, Amer. J. Math., 2(1879), 193-200

[19] A. G. Kurosh, General Algebra, Chelsea Publishing Co., New York, N.Y., 1963

[20] K. O. May, The origin of the four color conjecture, Isis 56(1965), 346-348

[21] J. Mayer, Probleme des quatre couleurs: un contre-example doit avoir au moins 72 sommets, submitted to J. Combinatorial Theory(B)

[22] O. Ore, The Four Color Problem, Academic Press, New York, 1967

[23] O. Ore and J. Stemple, Numerical calculations on the four color conjecture, J. Combinatorial Theory(B), 8(1970), 65-78

[24] T. L. Saaty, Thirteen colorful variations on Guthrie's four color conjecture, Amer. Math. Monthly, 79(1972), 2-43

[25] W. Stromquist, The four color theorem for small maps, J. Combinatorial Theory(B), 19(1975), 256-268

[26] R. Vaidyanathaswamy, Set Topology (2nd ed), Chelsea Publishing Co., New York, N.Y., 1960

[27] H. Whitney and W. T. Tutte, Kempe chains and the four colour problem, Utilitas Math., 2(1972), 241-281

[28] C. E. Winn, A case of coloration in the four color problem, Amer. J. Math., 59(1937), 515-529

Notes on Appendices

With a few exceptions in Appendix 3, the cluster configuraticns in these appendices are restricted to those without:

1) reducible subconfigurations, and
2) 2-splice reduction obstacles.

In appendices 2 and 5, below each configuration is a number, letter code.

number is the size of the boundary circuit (n)

letter is D for D-reducible
E for E-reducible
A for Assymmetrically D-irreducible
S for Symmetrically D-irreduicble

The codes E1, A1, and A2, used in Appendix 5, are defined in Appendix 4.

If there, the numbers beside a configuration are:

top number - size of extensible set,
middle number - size cf exterior set,
bcttom number - size of Interior of closure of extensible set.

In these numbers, the 2-colouring may be specified as +1.

Degenerate reducers for a configuration are indicated by marking the essential features on the boundary circuit, To this end, the letters a, b, and c are three distinct colours; x, y, and z are three distinct colours independent of a, b, and c. A bar over a letter indicates a colour different from the letter.

Appendix 1

A Minimal 5-Chromatic Planar Graph Contains a 6-Valent Vertex

Frank Allaire
University of Manitoba

This is another of several recent results ([1] to [5]) on minimal 5-chromatic planar graphs obtained through "discharging" (see [2] and [3]). The discharging rules used here are essentially due to Mayer ([4] and [5]) with suitable changes.

To prove this type of result we assume the existence of a 5-chromatic planar graph. It follows that there will be a non-empty family F of such graphs with a minimum number of edges. For any graph G in this family F we have:

1) G has no loops or multiple edges

2) G is a triangulation

3) G has no vertices of valency less than 5

and 4) G contains no reducible configuration.

Table 1 lists the 52 reducible configurations used in this paper. The reducibility of these configurations was determined by the author's computer program. More details on that program are contained in [6].

For any graph $G \in F$, let v_i denote the number of vertices of valency i, V the total number of vertices, E the number of edges, and F the number of faces (all triangles) of G. Then

$$V - E + F = 2$$

$$3E = 2F = 5v_5 + 6v_6 + \ldots + nv_n \ , \text{ and}$$

$V = v_5 + v_6 + v_7 + \ldots + v_n$ where n is the highest valency in G. Combining these we get

$$v_5 - v_7 - 2v_8 - \ldots - (n-6)v_n = 12 \qquad (1)$$

We can consider this equation in another way. Assign a "charge" of (6-i) to each vertex of valency i. Then read (1) as: "The sum of the charges on all the vertices is 12". Through the discharging rules we redistribute the charges until the final charge on every vertex is non-positive. Then we have "The sum of the charges on all the vertices is non-positive." This contradiction invalidates the original assumption of the existence of a minimal 5-chromatic graph.

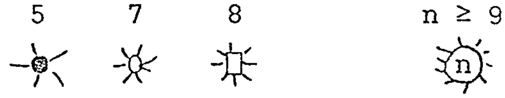
A discharging of a general minimal 5-chromatic planar graph would prove the Four Colour Conjecture. The rules contained in this paper discharge a possible graph in F with $v_6 = 0$.

Preliminaries

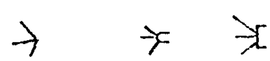
To eliminate fractions, multiply the initial charge values by 10 so that an i-valent vertex starts with 10(6-i) units of charge.

In figures, the following notation is used for vertices:

valency known 5 7 8 $n \geq 9$



Lower limit known ≥ 5 ≥ 7 ≥ 8



In the text, the following conventions are used.

1) A vertex is designated by a small letter a,b,... or its valency.
2) A vertex of valency greater than or equal to 7 is called major and may be designated M.
3) The first neighbourhood (adjacent vertices) of a vertex v is enclosed in square brackets following v. Single vertices more distant from v are usually enclosed by round brackets and inserted between a pair of its consecutive first neighbours. To reduce the number and levels of round

brackets, (5) may be replaced by 0, (7) by 2, (8) by 3. Thus x[a5y(5)52z5b], x = 8, y, z major describes figure 1 as does 5[57z0(b)x(a)0y] x = 8, y ≥ 7, z ≥ 7.

Not all configurations will be diagrammed, so liberal use of a scratch pad is recommended.

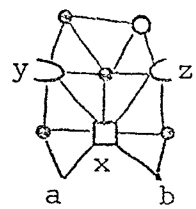


Figure 1.

Discharging Rules

These are divided into four stages. Five-valent vertices are discharged in the first two stages.

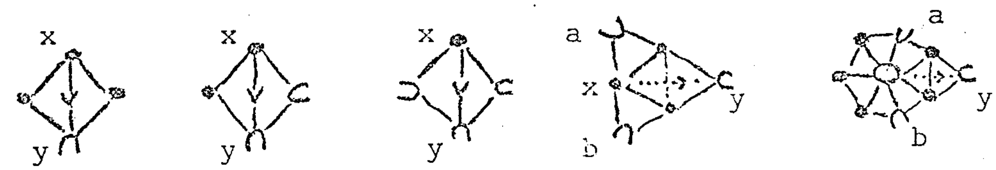
Stage 1 A) For a 5-valent vertex x with consecutive neighbours 5, y, 5 with y major (abbreviated (x = 5)[5 y 5], y ≥ 7), send 4 units of charge from x to y.

B) For (x = 5)[5yM], y ≥ 7, send 3 units of charge from x to y.

C) For (x = 5)[MyM], y ≥ 7, send 2 units of charge from x to y.

Stage 2 A) For (x = 5)[a5(y)5b], send 2 units of charge from x to y.

Figure 2 portrays these rules. In stage 1, y is specified as major. In stage 2A, a, b and y must be major for otherwise the graph G would contain R1 and R1 is a reducible configuration.



1A) 4 1B) 3 1C) 2 2A) 2 2B) 2

Figure 2. Stage 1 and Stage 2 transfer rules

Theorem 1. Every 5-valent vertex has a charge of zero after stages 1 and 2A

Proof. Excluding R1, there are only 5 possible first neighbourhoods for a 5-valent vertex x.

a) (x = 5)[5M55M]

b) (x = 5)[5M5MM]

c) (x = 5)[55MMM]

d) (x = 5)[5MMMM]

e) (x = 5)[MMMMM]

In every case the stage 1 and 2A rules transfer 10 units of charge from x to major vertices. This exactly discharges the initial charge assignment of +10 units of x.

To maintain the result of theorem 1, transfers of charge are never made to a 5-valent vertex. However, as a result of stages 1 and 2A, some major vertices may achieve a positive charge. This in turn is redistributed to other vertices.

Stage 2 B) For (x = 7)[555a5(y)5b] send 2 units of charge from x to y.

Again, a, b and y are major, this time by R2 and R4.

Stage 3 A) For (x = 7)[55y5] send 2 units from x to y. The two exceptions 3B and 3C occur when there is a stage 2 transfer to x through the 5-5 edge.

B) For (x = 7)[5(5)5y5] send 4 units from x to y.

C) For M[50(5)7052(5)y], x = 7 indicated by '2', send 4 units from x to y.

D) For (x=8)[55y55] send 2 units from x to y.

Figure 3. Stage 3 transfer rules.

Lemma 1. Stages 2B and 3 do not transfer charge to a 5-valent vertex.

Proof. R4, R2 and R3.

Lemma 2. In stage 2 there is a net transfer of charge i.e. y cannot simultaneously satisfy the stage 2 conditions and transfer charge to x.

Proof. R1, R4, and R36.

Lemma 3. In stage 3 there is a net transfer of charge, i.e. y cannot simultaneously satisfy the stage 3 conditions and transfer charge to x.

Proof. R8, R9, R14, R26.

These rules are very efficient and effective at discharging most of the vertices of the graph. Theorem 1 shows that a 5-valent vertex is not over-discharged but exactly discharged. For a general graph in F with $v_6 \neq 0$, a 6-valent vertex starts with zero charge. The stage 1 and 2 rules can be modified to ensure that a 5-valent vertex remains exactly discharged while at the same time the 6-valent vertices never receive any charge. Thus the remaining charge transfers are from major vertices to major vertices.

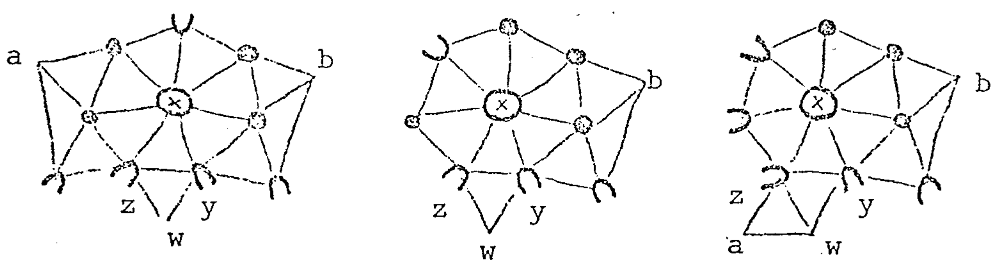
For the stage 3 rules, lemma 1 shows that if the receiving vertex y is 5-valent then the configuration is reducible. The same result holds if y is 6-valent. Not only must y be major, but lemma 3 ensures that there is a net transfer of charge. Thus, the edge from x to y is a natural one-way path for transferring charge between major vertices.

At this point, an exhaustive examination of all neighbourhoods of major vertices is undertaken. For the graphs of interest to this paper, i.e. $v_6 = 0$, only 7-valent vertices with three particular neighbourhoods are left with a positive charge and that charge is two units. Since the original charge assignment totalled 120 there must be at least 60 such vertices, and by equation (1), at least 72 5-valent vertices. Altogether, then, G must have at least 132 vertices. In a similar way, a lower bound on the number of vertices of a general minimal 5-chromatic planar graph can be obtained.

To raise the lower bound from 132 to infinity, i.e. to show that such graphs do not exist, the fourth stage of discharging is performed. For each vertex with a positive charge, an adjacent major vertex is designated to accept the positive charge. If the designated vertex is not negatively charged, then the charge is transferred to an alternate major vertex adjacent to both the designated vertex and the positively charged vertex. The effectiveness of the triad of major vertices can also be heuristically explained. Since each of the three vertices has two consecutive major neighbours, they are not likely to have a positive charge to begin with. On the other hand, if one does have a positive charge and a second is not negatively charged then there are enough 5-valent neighbours in the configuration to force the third major vertex to be 8- or perhaps 9-valent and hence even more likely to be able to absorb the positive charge(s) from its neighbour(s).

Stage 4 A) (x = 7)[y5(b)5M5(a)5z] y, z major, with stage 2 transfers from a to x and from b to x (fig. 4). If vertex y or z is 8-valent or greater, say y, send 2 units of charge from x to y except if (y = 8)[5555M5MM]. In this case z is 8-valent or greater by R47 and does not satisfy the exception by R27, so send 2 units of charge from x to z.

Otherwise both y and z are 7-valent and w is major by R32. Send 2 units of charge from x to y except if (y = 7)[5705$0_x$(5)0z]. In this case send 2 units of charge from x to z since R35 prevents z from simultaneously satisfying the exception.



A) B) C)

Figure 4 Stage 4 rules

The other two configurations can be considered together:

Stage 4 B) (x = 7)[z5M55(b)5y]
4 C) (x = 7)[zMM55(b)5y]

y, z major with a stage 2 transfer from b to x.

If vertex y is 8-valent or greater, send 2 units of charge from x to y. The exception (y = 8)[5555MM5M] does not occur by R34. Otherwise y = 7. If vertex w is major then send 2 units of charge from x to y. Otherwise w = 5 and send 2 units of charge from x to z. In 4B, R32 forces z ≥ 8 and R47 and R49 ensure that (z = 8)[5555MM5M] does not occur. Note that y[wzx5M] i.e. (y = 7)[5MM5M] so if y were to be a stage 4 configuration it would have to be 4B but that is not allowed by R17.

This completes the rules for discharging.

Analysis of Major Vertices

To help prove that every major vertex has a non-positive final charge we investigate the neighbourhoods of those vertices that receive the transfers of charge.

A major vertex receives a stage 1 transfer from every adjacent 5-valent vertex. A stage 2 transfer is sent "across" a pair of 5-valent vertices to a major vertex. Lemma 4 limits this consideration. A stage 3 transfer to a major vertex is from an adjacent major vertex flanked on both sides by adjacent 5-valent vertices. Lemmas 5, 6 and 7 limit the amount of this transfer. Conditions for a vertex to receive a stage 4 transfer are listed in Lemma 8.

Lemma 4. (x = M)[5(a)5(b)5]. At most one stage 2 transfer from a and b is sent to x.

Proof. Otherwise there are three possibilities:

i) a and b satisfy 2A , i.e. a = 5 = b. This is not allowed by R1.

ii) a satisfies 2A , b satisfies 2B. This is not allowed by R2

iii) a and b satisfy 2B. This is not allowed by R9.

Lemma 5. The following vertices x with specified neighbourhoods do not accept any stage 3 transfer from the vertex y.

i) (x = 7)[55y5]

ii) (x = 8)[55y55]

iii) (x = 8)[555y5]

iv) (x = 9)[5555y5]

v) (x = 9)[555y55]

Proofs. i) R8, R9 and R14

ii) R14, R26

iii) R12, R13, R25

iv) R21, R22, R39

v) R23, R24, R40

Lemma 6. (x = 8)[a5y5z5b], both y and z are major. The maximum stage 3 transfer to x from y and z is 4 units each (by definition).

i) If the combined stage 3 transfer from y and z to x is 8 units then both a and b are major and this occurs in only 3 ways.

ii) If the combined transfer is 6 units then at least one of a, b is major. Otherwise the combined transfer is at most 4 units.

Proof. i) We must have y = 7 = z. Excluding R31 and R42 we are left with (x = 8)[a(c)5(5)y5z(5)5(d)b] with stage 2 transfers from c to y and d to z, forcing both a and b to be major. The same three choices for a and b in Lemma 4 are available for c and d here.

ii) We can assume y = 7. Excluding R31, R42 and R43 leaves (x = 8)[a(c)5(5)y5z5b] with a stage 2 transfer from c to y forcing a to be major.

Lemma 7. x[a5y5b]. If y is major and the stage 3 transfer from y to x is 4 units, then either a or b is major. Otherwise the stage 3 transfer is at most 2 units.

Proof. The stage 3 transfer of 4 units occurs only with an adjacent stage 2 transfer and this stage 2 transfer forces one of a, b to be major.

Lemma 8. For a stage 4 transfer from a = 7 to b we must have one of:

$$\left.\begin{array}{rl} \text{i)} & (b = 7)[MMa5M] \\ \text{ii)} & (b = 7)[Ma75M] \end{array}\right\} \quad \text{or} \quad 7[MMM5M]$$

$$\left.\begin{array}{rl} \text{iii)} & (b = 8)[Ma5M] \\ \text{iv)} & (b = 8)[a75M] \end{array}\right\} \quad \text{or} \quad 8[MM5M]$$

$$\begin{array}{rl} \text{v)} & (b = 8)[Ma755M] \end{array} \quad \text{or} \quad 8[MMM55M]$$

$$\left.\begin{array}{rl} \text{vi)} & (b \geq 9)[Ma5M] \\ \text{vii)} & (b \geq 9)[a75] \end{array}\right\} \quad \text{or} \quad (b \geq 9)[MM5]$$

Proof. (see figure 4)

4A	$y \geq 8$ or $z \geq 8$	iii) or vi)
	$y = 7 = z$	i)
4B	$y \geq 8$	iii) or vi)
	$y = 7$, $w = M$	i)
	$y = 7$, $w = 5$	iv) by R48 or vii)
4C	$y \geq 8$	iii) or vi)
	$y = 7$, $w = M$	i)
	$y = 7$, $w = 5$, $z = 7$	ii) by R33
	$y = 7$, $w = 5$, $z = 8$, $a = M$	iv)
	$y = 7$, $w = 5$, $z = 8$, $a = 5$	v) by R47
	$y = 7$, $w = 5$, $z \geq 9$	vii)

The final note in the description of the stage 4 transfers ensures that the occurrence of $(b = 7)[Ma_1a_25M]$, $a_1 = 7 = a_2$ will not result in two stage 4 transfers to b, one from each of a_1 and a_2, i.e. conditions i) and ii) cannot occur simultaneously over the same set of vertices. Similarly iii) and iv) cannot occur together nor can vi) and vii). Thus each of the neighbourhoods 7[MMM5M], 8[MM5M], 8[MMM55M] and $(b \geq 9)$[MM5] signals a possible stage 4 transfer of at most 2 units of charge to the central vertex.

With these aids, all possible first neighbourhoods for a

7-valent vertex are examined in table 2. Only one case (x = 7)[5M5MMMM] can have a positive final charge and that will occur only if there is a stage 3 transfer of 4 units and two stage 4 transfers to x. This forces (x = 7)[57(5)5MMMM] and one of the three possibilities listed in figure 5. The first diagram describes one of the stage 4A exceptions, so the stage 4 transfer does not occur. The second diagram contains R16. The third case is probably reducible, but my program is limited to configurations bounded by a 13-circuit or smaller and the configuration of vertices of known valencies is bounded by a 14-circuit. Instead consider the stage 2 transfer from vertex d. If d = 5 then we have R50. If (d = 7)[555M55M] then the configuration contains R51. Thus the particular combination of a stage 3 transfer of 4 units coupled with two stage 2 transfers cannot occur and 7[5M5MMMM] will have a final charge of at most zero.

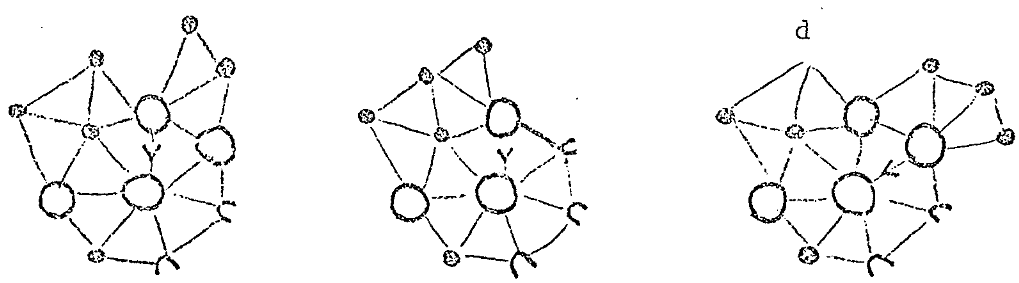


Figure 5 (x = 7)[5705MMMM]

Table 2 with the above discussion and tables 3 to 7 show that the final charge value on every major vertex is non-positive. This and Theorem 1 provide the contradiction that proves the title.

Table 1 Reducible Configurations

R1	5[555]
R2	7[5555]
R3	8[55555]
R4	7[555x505y]

Boundary is a 10-circuit

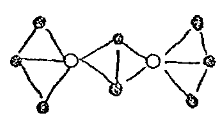
R5	9[555555]
R6	8[555x555y]
R7	8[5555x505y]
R8	7[55705]
R9	5[57075]

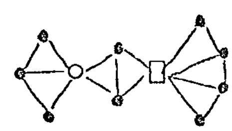
Boundary is an 11-circuit

R10	10[5555555]
R11	9[5555x555y]
R12	8[507555]
R13	8[570555]
R14	8[550755]
R15	7[5705075]
R16	5[5707705]
R17	7[555x7(5)05y]

Boundary is a 12-circuit

R18	11[55555555]
R19	10[55555x555y]
R20	10[5555x5555y]
R21	9[5075555]
R22	9[5705555]
R23	9[5507555]
R24	9[5570555]
R25	8[5080555]
R26	8[5508055]
R27	5[50(5)88(5)05]
R28	7[5075075]
R29	7[5075705]
R30	7[57050805]
R31	8[5705075]
R32	7[507075]
R33	7[50(5)7705]
R34	8[5555x7(5)05y]
R35	7[505x5052(5)7y]

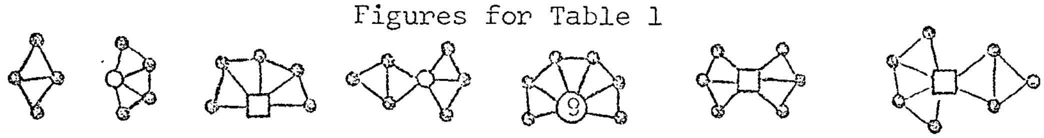
R36 

R52 

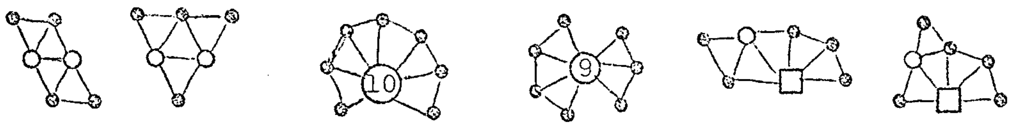
Boundary is a 13-circuit

R37	11[555555x555y]
R38	11[55555x5555y]
R39	9[50805555]
R40	9[55080555]
R41	7[50805705]
R42	8[5705705]
R43	8[57050805]
R44	8[50755705]
R45	8[57055705]
R46	8[57055075]
R47	7[5078(5)05]
R48	7[55570(5)85]
R49	7[5570805]
R50	7[5077(5)2505]
R51	5[5707707]
R52	above

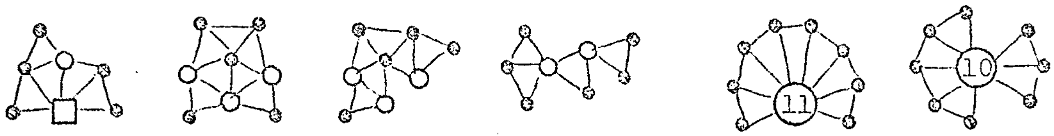
Figures for Table 1



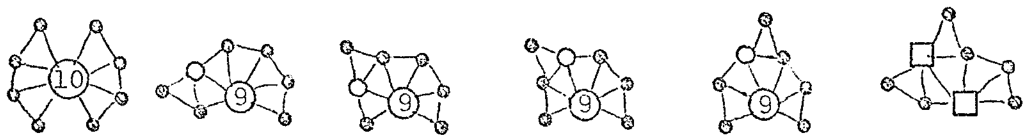
R1 R2 R3 R4 R5 R6 R7



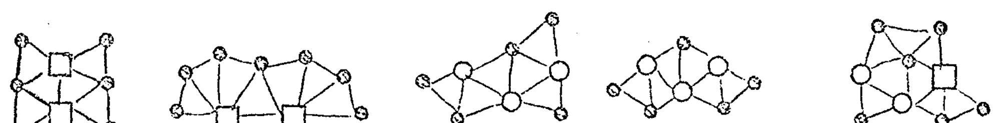
R8 R9 R10 R11 R12 R13



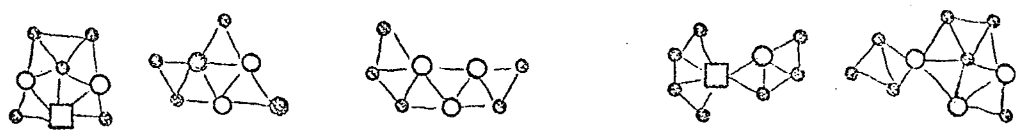
R14 R15 R16 R17 R18 R19



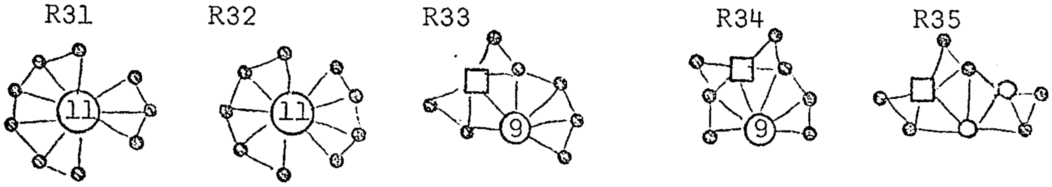
R20 R21 R22 R23 R24 R25



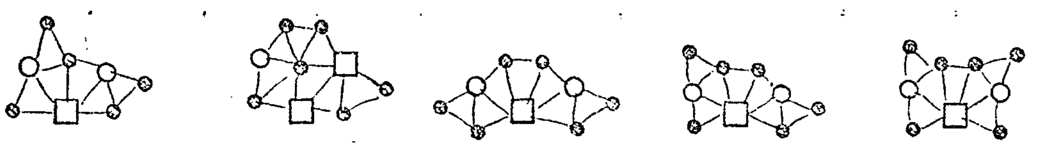
R26 R27 R28 R29 R30



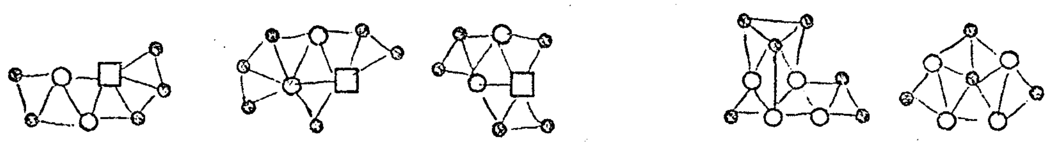
R31 R32 R33 R34 R35



R37 R38 R39 R40 R41



R42 R43 R44 R45 R46



R47 R48 R49 R50 R51

Table 2 7-valent vertices, initial charge = -10

	Stage				Maximum Final	
Exclude R2	1	2	3	4	Charge	Remarks
7[555M55M]	16	-2(+2	-2)-2-2		0	R4,R36
7[555M5MM]	12	(+2	-2)-2		0	
"	12	+2	-2	-2	0	
7[55M55MM]	12	(+2	-2)-2		0	
"		+2(+2	-2)-2	-2	0	
7[55M5M5M]	10	+2	-4-4+4		-2	
"	10		-2-2+4		0	
7[555MMMM]	10				0	
"	10	+2		-2	0	
7[55M5MMM]	8	(+2	-2)-2	+2	-2	
7[55MM5MM]	8	+2			0	
7[5M5M5MM]	6		≤4		0	R15,R28,R29, R30,R41
7[55MMMMM]	6	+2			-2	
7[5M5MMMM]	4		+4	+4	+2	see text
7[5MM5MMM]	4			+4	-2	
7[5MMMMMM]	2			+4	-4	
7[MMMMMMM]					-10	

Table 3 8-valent vertices, initial charge = -20

	Stage				Maximum Final	
Exclude R3, R6	1	2	3	4	Charge	Remarks
8[5555M55M]	20	+2+2	-2-2		0	R7,R52
8[5555M5MM]	16	+2+2			0	stage 4 exception
8[555M55MM]	16	+2+2	-2		-2	
8[555M5M5M]	14	+2	+4		0	
8[55M55M5M]	14	+2+2	-2+4		0	Lemma 6
8[5555MMMM]	14	+2+2			-2	
8[555M5MMM]	12	+2		+2	-4	
8[555MM5MM]	12	+2		+4	-2	
8[55M55MMM]	12	+2+2	-2	+4	-2	

Table 3 8-valent vertices, initial charge = -20

(Continued)

	Stage				Maximum Final	
Exclude R3, R6	1	2	3	4	Charge	Remarks
8[55MM55MM]	12	+2+2			-4	
8[55M5M5MM]	10	+2	+6	+2	0	Lemma 6
8[55M5MM5M]	10	+2	+4	+4	0	R44
"	10		+4+2	+4	0	R44,R45,R46
8[5M5M5M5M]	8		4+4+2+2		0	Lemma 6
"	8		4+2+4+2		0	Lemma 6
8[555MMMMM]	10	+2			-8	
8[55M5MMMM]	8	+2	+4	+4	-2	
8[55MM5MMM]	8	+2		+6	-4	
8[5M5M5MMM]	6		+4+4	+4	-2	
8[5M5MM5MM]	6		+4	+8	-2	
8[55MMMMMM]	6	+2		+4	-8	
8[5M5MMMMM]	4		+4	+4	-8	
8[5MM5MMMM]	4			+8	-8	
8[5MMM5MMM]	4			+8	-8	
8[5MMMMMMM]	2			+4	-14	
8[MMMMMMMM]					-20	

Table 4 9-valent vertices, initial charge = -30

	Stage				Maximum Final
Exclude R5, R11	1	2	3	4	Charge
9[55555M55M]	24	+2+2+2			0
9[55555M5MM]	20	+2+2		+4	-2
9[5555M55MM]	20	+2+2+2		+4	0
9[5555M5M5M]	18	+2+2	+4		-4
9[555M555MM]	20	+2+2		+4	-2
9[555M55M5M]	18	+2+2	+4+4		0
9[55M55M55M]	18	+2+2+2	+2+2+2		0
9[55555MMMM]	18	+2+2		+4	-4

Table 4 9-valent vertices, initial charge = -30

(Continued)

	Stage 1	Stage 2	Stage 3	Stage 4	Maximum Final Charge
9[5555M5MMM]	16	+2+2		+4	-6
9[5555MM5MM]	16	+2+2		+8	-2
9[555M55MMM]	16	+2+2		+4	-6
9[555M5M5MM]	14	+2	+4+4	+4	-2
9[555M5MM5M]	14	+2	+4+4	+4	-2
9[55M55M5MM]	14	+2+2	+4+4	+4	0
9[55MM55M5M]	14	+2+2	+4+4	+4	0
9[55M5M5M5M]	12	+2	+4+4+4+4		0
9[four 5's, five M's]					
5555	14	+2+2		+4	-8
555...5	12	+2	(+4 or	+2)+6	-6
55...5	12	+2+2	(+4 or	+2)+6	-4
55..5..5	10	+2	(4+4 or	2+2)+6	-4
5..5..5..5	8		(4+4+4 or	2+2+2)+4	-6
9[three 5's, six M's]					
555	10	+2		+4	-14
55..5	8	+2	(+4 or	+2)+6	-10
5..5..5..	6		(4+4 or	2+2)+8	-8
9[two 5's, seven M's]					
55	6	+2		+4	-18
5..5	4		(+4 or	+2)+6	-16
9[5MMMMMMMM]	2			+4	-24
9[MMMMMMMMM]					-30

Table 5   10-valent vertices, initial charge = -40

Excluded R10, R19, R20	Stage 1	2	3	4	Maximum Final Charge
10[555555M55M]	28	+8	+2+2		0
10[555555M5MM]	24	+6	+4	+4	-2
10[55555M55MM]	24	+6	+2	+4	-4
10[55555M5M5M]	24	+4	+4+4+4		0
10[5555M555MM]	24	+6	+2	+4	-4
10[5555M55M5M]	22	+6	+2+4+4		-2
10[555M555M5M]	22	+4	+2+4+4		-4
10[555M55M55M]	22	+6	+2+2+2		-6
10[six 5's, four M's]					
555555	22	+6		+4	-8
55555..5	20	+4	(+4 or	+2)+6	-6
5555..55	20	+6	(+4 or	+2)+6	-4
555..555	20	+4	(+4 or	+2)+6	-6
all others	18	+6	(+4+4	+4+4)	0
10[five 5's, five M's]					
55555	18	+4		+4	-14
all others	16	+4	(+4+4	+4+4+4)	0
10[n×5,n≤4,(10-n)×M]	14	+4	(+4+4+4 or	+2+2+2)+6	-4

Table 6   11-valent vertices, initial charge = -50

Exclude R18, R37, R38	Stage 1	2	3	4	Maximum Final Charge
11[5555555M55M]	32	+8	+2+2		-6
11[n×5,n≤8,(11-n)×M]	4n-2	$2\lfloor n/2 \rfloor$		4(11-n)	n-8

Table 7   k-valent vertices, k ≥ 12

Initial charge = 10(6-k)	Stage 1	2	3 or 4	Maximum Final Charge
k[n×5,n≤k,(k-n)×M]	4n	$2\lfloor n/2 \rfloor$	4(k-n)	60-5k

References

[1] O. Ore and J. Stemple, Numerical calculations on the four-color problem, J. Combinatorial Theory 8 (1970), 65-78.

[2] H. Heesch, Chromatic reduction of the triangulations T_e, $e = e_5 + e_7$, J. Combinatorial Theory(B) 13 (1972), 46-55.

[3] W. Haken, An existence theorem for planar maps, J. Combinatorial Theory(B) 14 (1973), 180-184

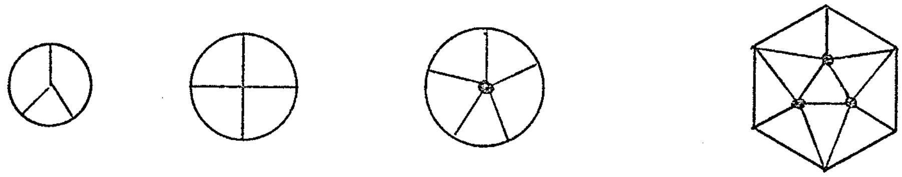
[4] J. Mayer, "Problème des quatre couleurs: un contre-example doit avoir au moins 72 sommets". Private communication.

[5] J. Mayer, "Problème des quatre couleurs: un contre-example doit avoir au moins 96 sommets". Private communication.

[6] F. Allaire and E.R. Swart, A systematic approach to the determination of reducible configurations in the four colour conjecture. J. Combinatorial Theory (B), accepted.

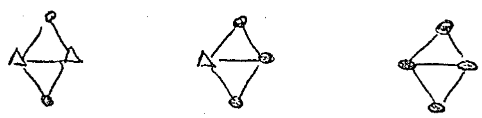
Appendix 2. Small Configurations

At most three vertices



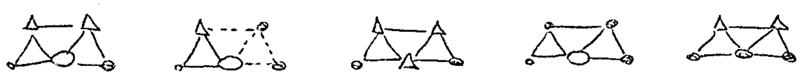
3, D   4, D   5, S   6, S

Four inside vertices (4/n)



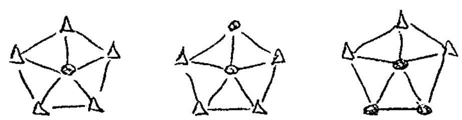
8, A   7, D   6, D

5/n

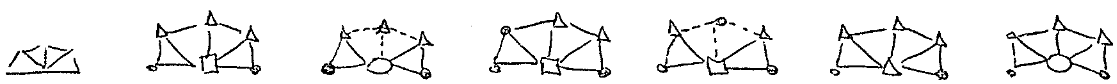


10, S   9, E   9, D   8, D   8, D

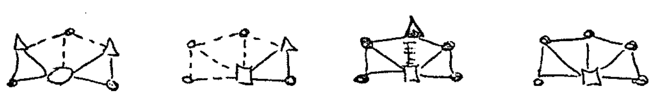
6/n



10, A   9, D   8, D



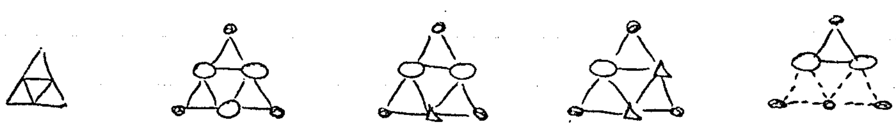
12, S   11, E   11, S   11, E   10, D   10, D



10, E   10, E   10, E   9, D

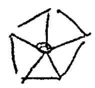
12, S   11, S   11, A   10, E   10, E

10, E   10, D   10, D   9, D



12, S   11, A   10, D   10, E

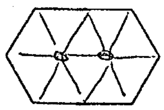
Appendix 3 Antiset Analyses for Small Clusters.



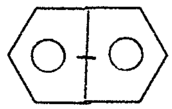
5
5
5

Introduce anti-pent
$|Q5| = 10$

5

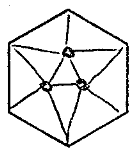


12 + 1
16
13



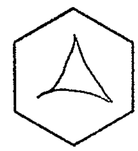
16

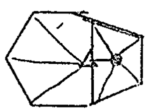
$|Q6| = 31$

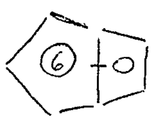


14
13 + 1
14

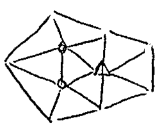
Exterior set cannot be expressed by anti-pents.
Introduce anti-triad



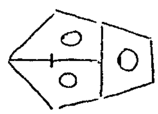




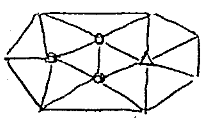
$|Q7| = 91$



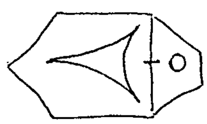
32
51
32



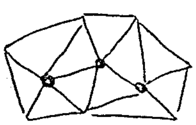
50



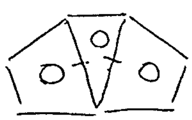
35
39
35



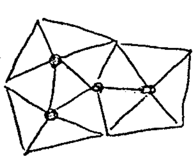
39



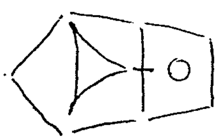
31
50
31



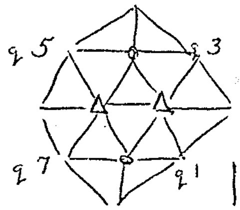
48



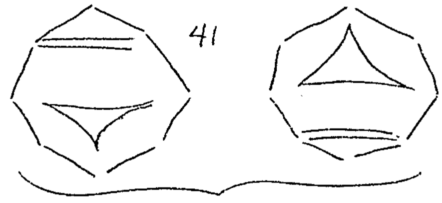
35
44
35



44



81
153+1
87
|Q8| = 274

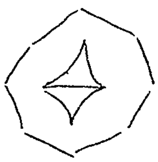


41 41

76

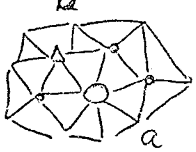
Seed: q1 ≠ q7, q3 ≠ q5 leads to
reducer: q1 = q3 ≠ q5 = q7

Since there must be some other anti-set

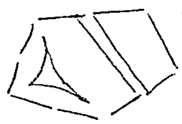


154 and this anti-set must be taken into consideration for a reducer to be valid, introduce the anti-diamond to express Ext(6[565]-8)

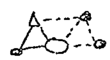
b



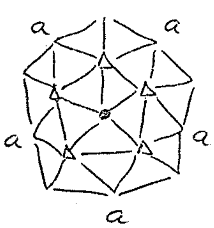
203
81
121



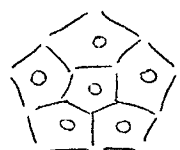
81

reducers: as indicated
or 

|Q9| = 820

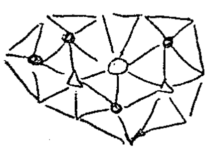


530
1220
676

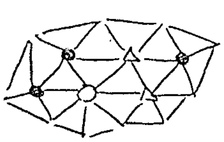


reducers: as indicated or W10

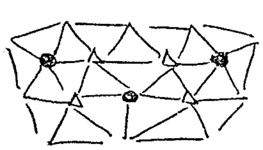
|Q10| = 2461



631

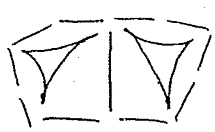


513

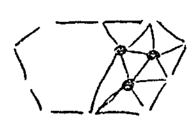
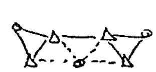


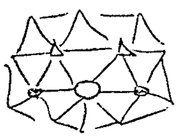
623

365
1916

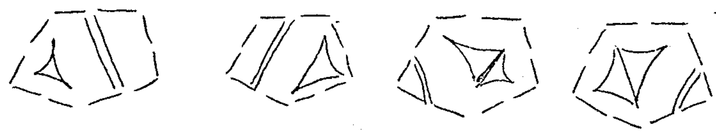


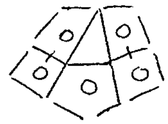
365

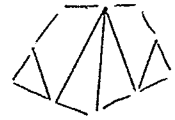
reducers: as indicated
or  or 
or 



433
1885
433



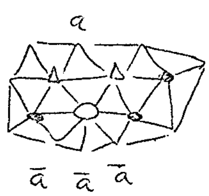


Seed: 

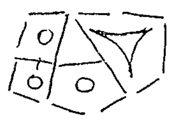
But $7[5665]-10$ is Code S and
$\mathrm{Int}(\mathrm{Seed} \cap \mathrm{Int}(\mathrm{Cl}(7[5665]))) = \phi$

Therefore there is some other anti-set.

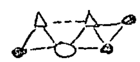
Introduce a pseudo-configuration to represent $\mathrm{Ext}(7[5665])$.

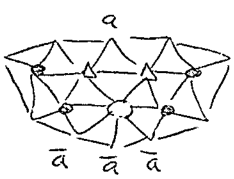


529
747
1173

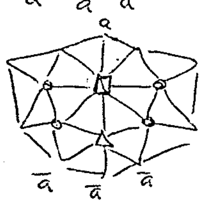


reducers: as indicated

or 

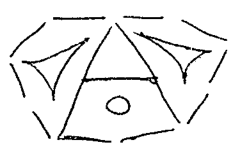


632

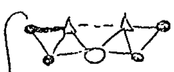


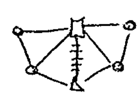
500

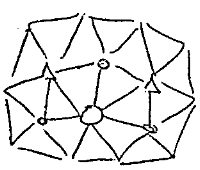
570
1216



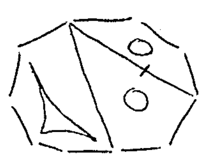
reducers: as indicated or

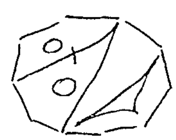




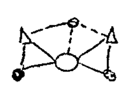


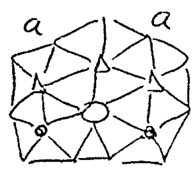
531
672
1205



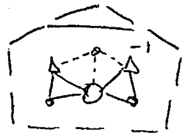


reducer:



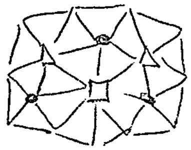


1117
1344
4870
$|Q11| = 7381$



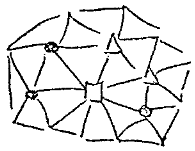
$672 \times 2 = 1344$
$|Q11| = 7381$

reducer:
as indicated
or 

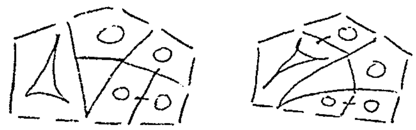


1078
3452
2629

reducer : 

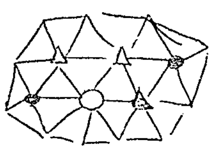


1079
4396
1079

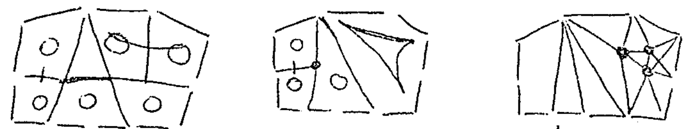


$\underbrace{\qquad\qquad\qquad}_{\text{reduction obstacle.}}$

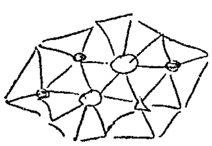
There are several other anti-sets.



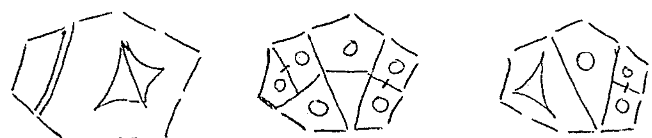
1109
3309
1814



reducer

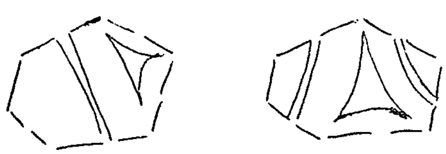


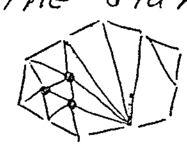
1062
3595
1062



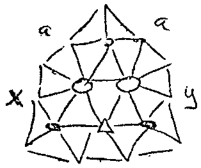
$\underbrace{\qquad\qquad\qquad}_{\text{reduction obstacle}}$

There are several other anti-sets.

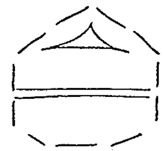


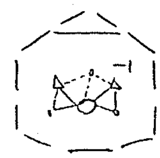
All antisets except the diamond 0-splice are avoided by 

Therefore, the extended constraints may reduce 7[55706]

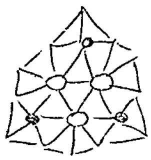


1081
2020
2601

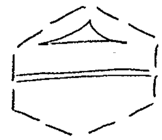


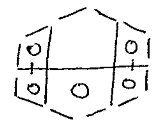


reducer as indicated

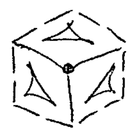


2211
19318
2211





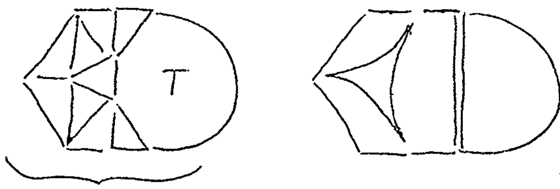
$|Q12| = 22144$



Appendix 4. Frequent Reducers

Let T represent any configuration

1) The configuration below has the indicated anti-set.



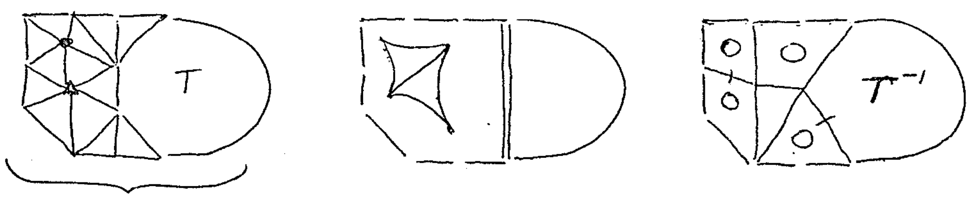
T^*

If this is the only anti-set for T^*, then $Ext(T^*)$ and $Int(Cl(T^*)$ will have the following sizes

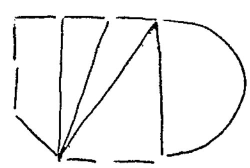
	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
|Ext (T*)|	13+1	0	41	81	284	810	2471	7371	22154
|Int (Cl (T*))|	14	91	224	721	2114	6391	19124	57421	172214
|Qn|	31	91	274	820	2461	7381	22144	66430	199291

Further, T^*/T reduces T^*. This relaxation reducer is indicated in the following appendices by the code E1

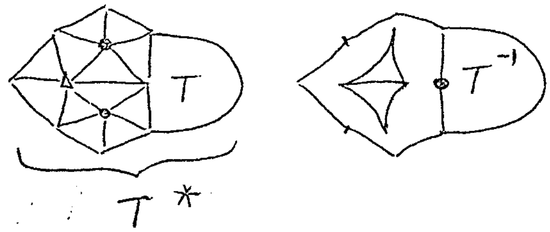
2.) The configuration T^* below has the indicated anti-sets.

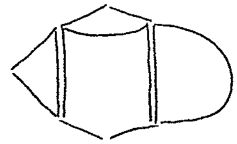


T^*

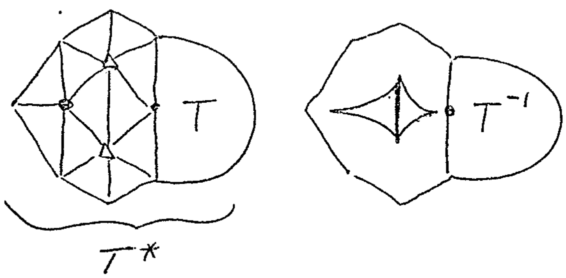
If T^* is reduced by , then the code A1 is used

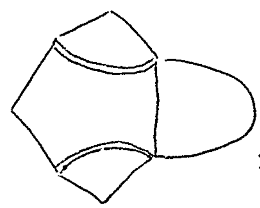
3 a) The configuration below has the indicated anti-set.



If T^* is reduced by , then the code A2 is used.

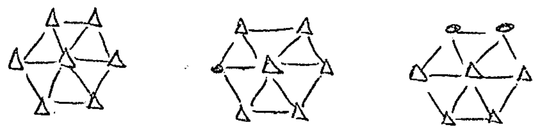
b) The configuration below has the indicated anti-set.



If T^* is reduced by , then the code A2 is used.

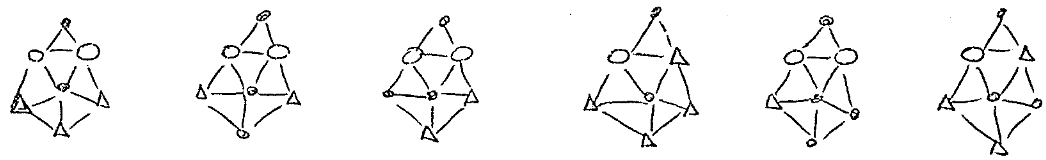
Appendix 5 7/n Clusters

a)

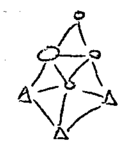


12, S 11, D 10, D

b)

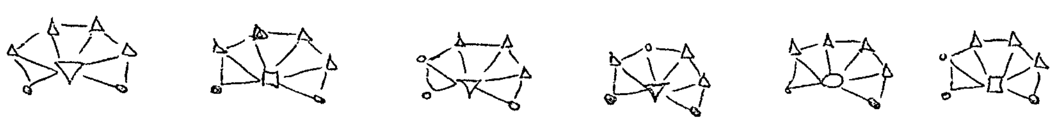


12, S 11, E1 11, E1 11, D 10, E 10, D

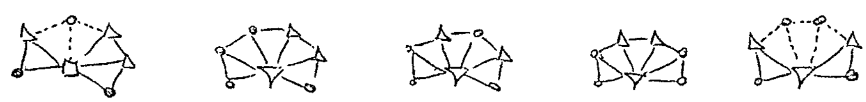


10, D

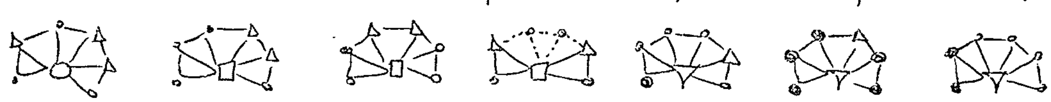
c)



14 13, S 13, S 13, S 12, D 12, S

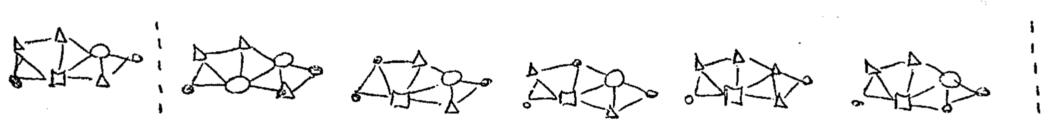


12, E 12, A1 12, S 12, S 12, E

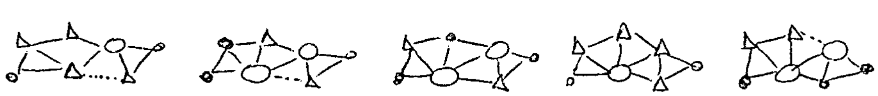


11, D 11, D 11, S 11, E 11, E1 11, D 10, D

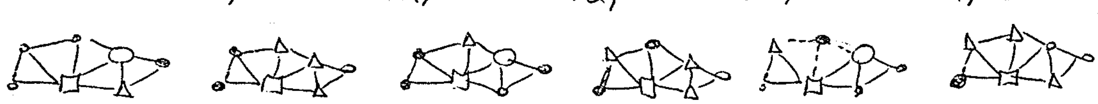
d)



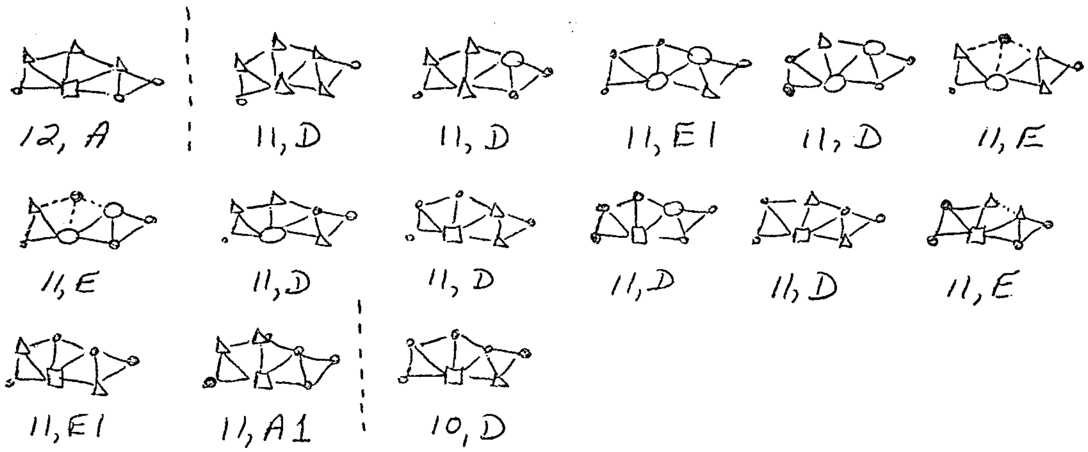
14 13, S 13, S 13, S 13, S 13, S



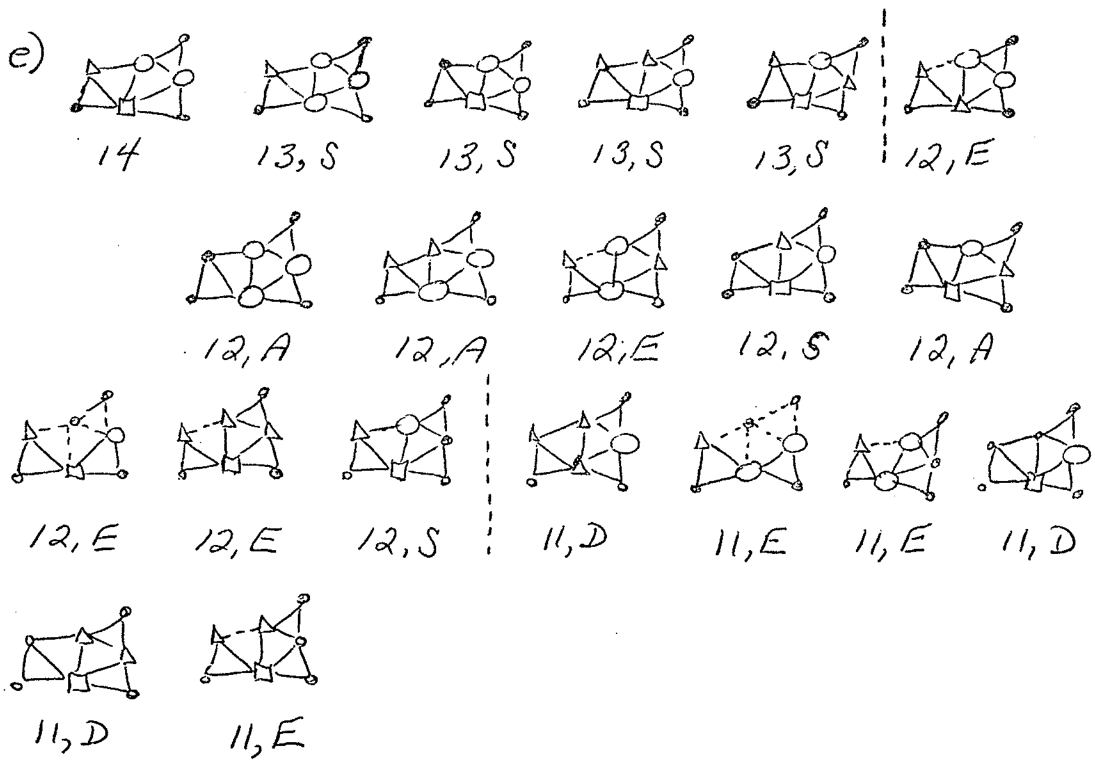
12, E 12, E 12, A 12, A 12, E



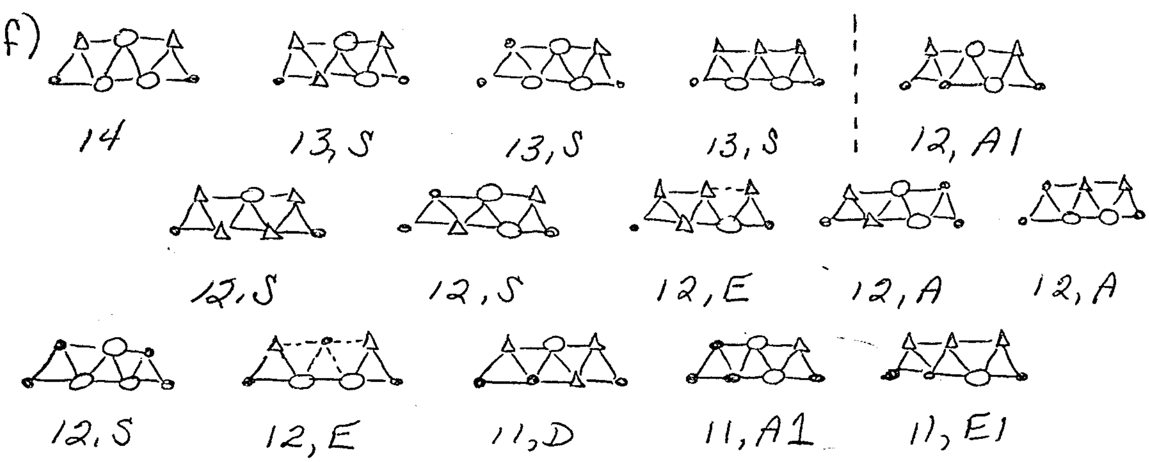
12, A1 12, S 12, S 12, A 12, E 12, A1

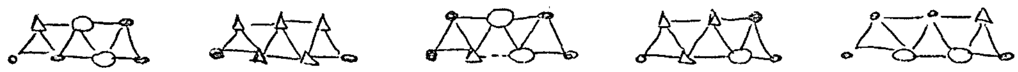

12, A
11, D
11, D
11, E1
11, D
11, E
11, E
11, D
11, D
11, D
11, D
11, E
11, E1
11, A1
10, D

e)

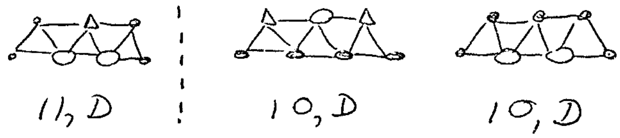

14
13, S
13, S
13, S
13, S
12, E
12, A
12, A
12, E
12, S
12, A
12, E
12, E
12, S
11, D
11, E
11, E
11, D
11, D
11, E

f)


14
13, S
13, S
13, S
12, A1
12, S
12, S
12, E
12, A
12, A
12, S
12, E
11, D
11, A1
11, E1



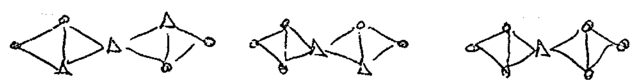
11, D   11, D   11, E   11, D   11, E1



11, D   10, D   10, D

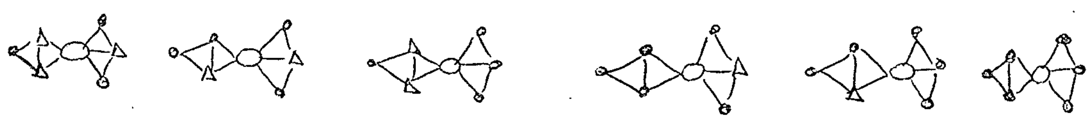
g)

12, S   11, A2   10, A2   10, D



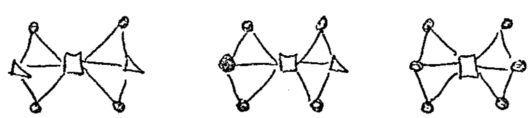
10, D   9, D   8, D

h)



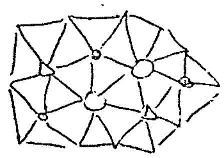
12, S   11, A2   11, A2   10, A2   10, D   9, D

i)

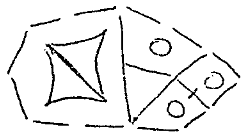


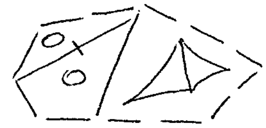
12, S   11, A2   10, D

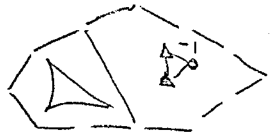
Appendix 6. Anti-sets and Reducers for Code A configurations of Appendix 5.



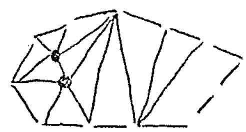
2802
10996
4096



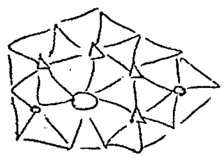




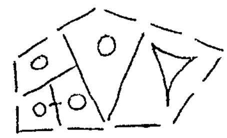
reducer:

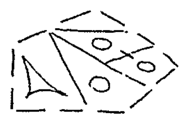


2000

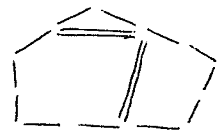


2841
6232
10504

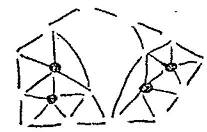


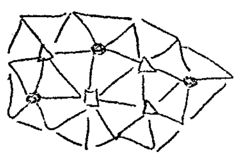


several reducers, for example

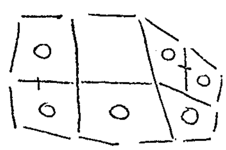


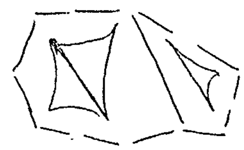
,

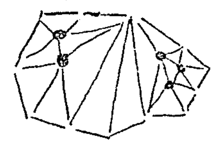


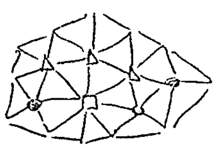


2771
11486
3412

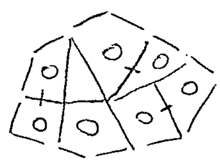


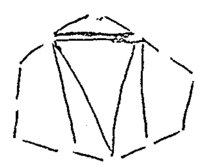


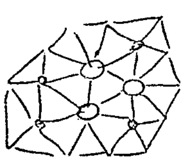




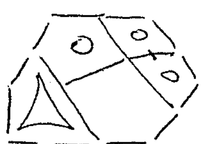
2778
13033
4268

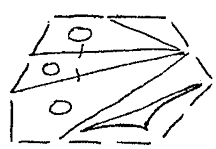


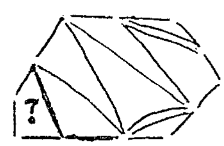


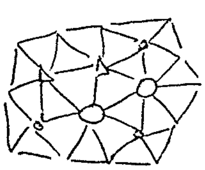


2684
9523
3020

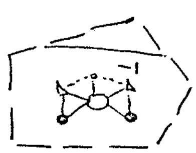




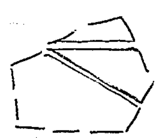


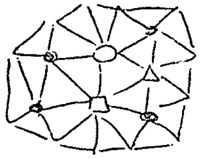


2798
6863
6079

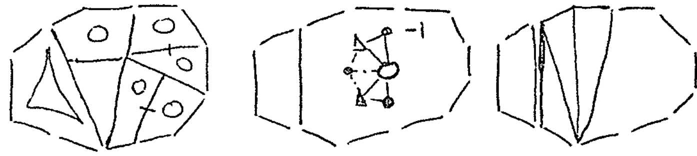


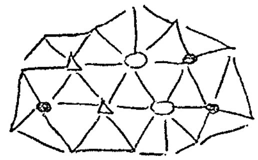
several reducers, for example



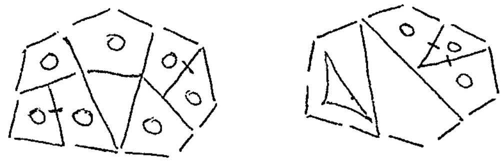


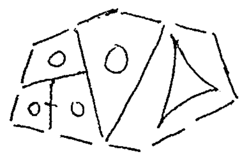
2706
10154
4737



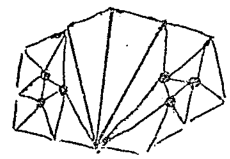


2703
11833
3341

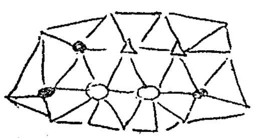




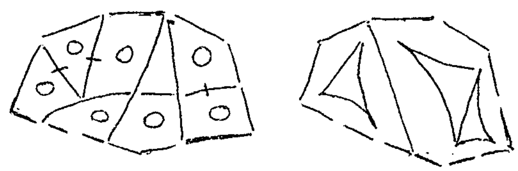
reducer:



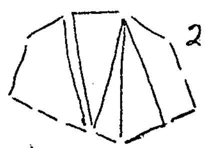
1568



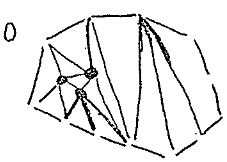
2725
9207
6060



reducers:

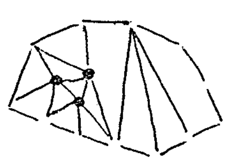


2400

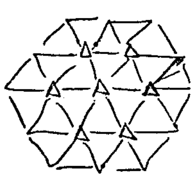


2240

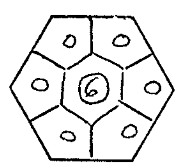
VI



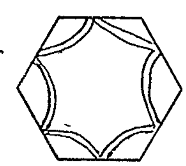
also, one Code S configuration :



2756
14061
2756



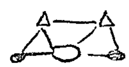
reducer



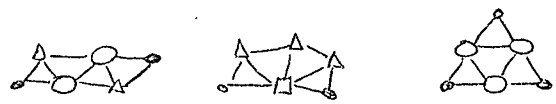
Appendix 7 Symmetrically D-irreducible Clusters (without 2-splice obstacles) (without direct reducers)

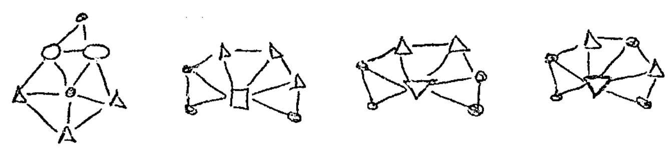
$n = 5$ 

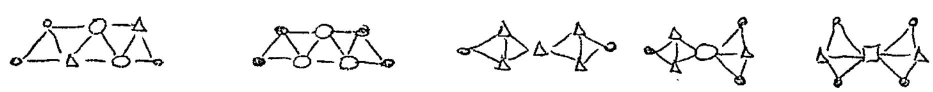
$n = 6$ 

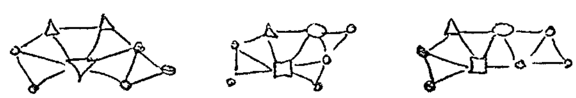
$n = 10$ 

$n = 11$  ← $m = 6$ $m = 7$ → 

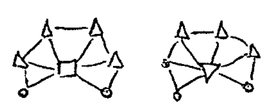
$n = 12$ $m = 6$ 

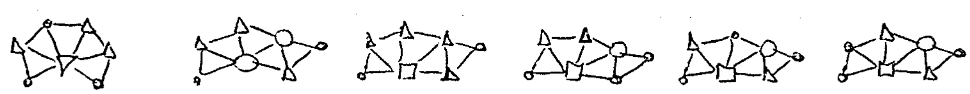
$m = 7$ 

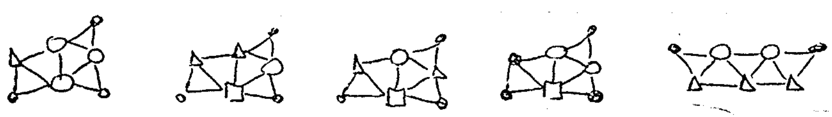


$m = 8$ 

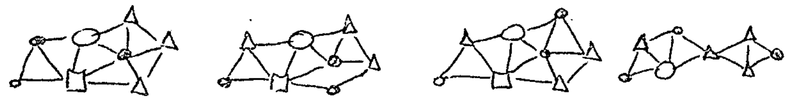
$m \geq 9$?

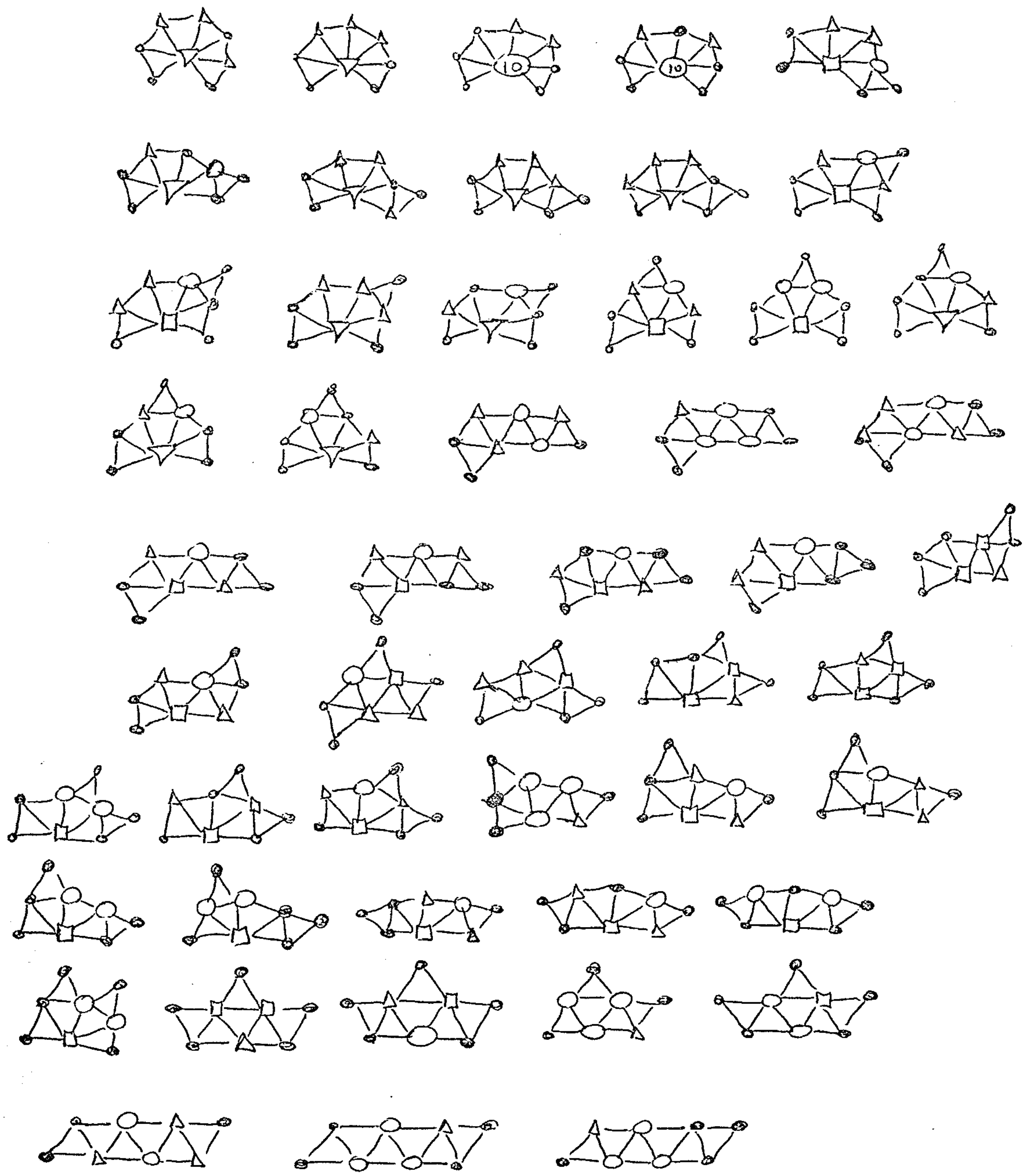
$n = 13$ $m = 7$ (all 15 possibilities) 





$n = 13,\ m = 8$

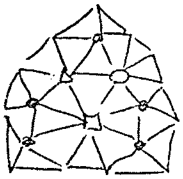




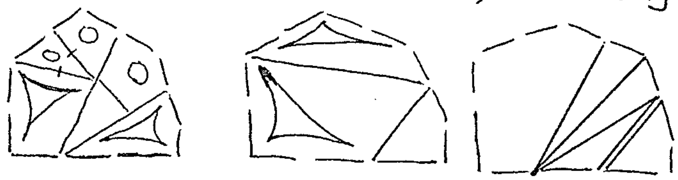
$n = 13,\ m \geq 9$?

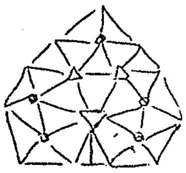
$n \geq 14$?

Appendix 8   Anti-sets and Reducers for 8/12 clusters, Code A only



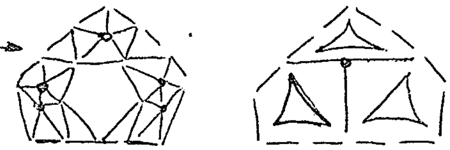
3232
7256
5238

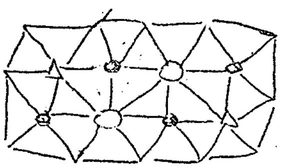




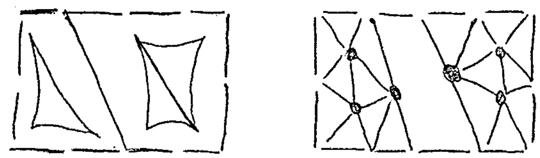
3258
2953
14677

E-reducible: 7094→

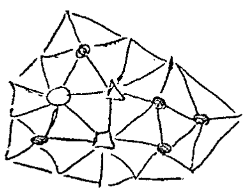




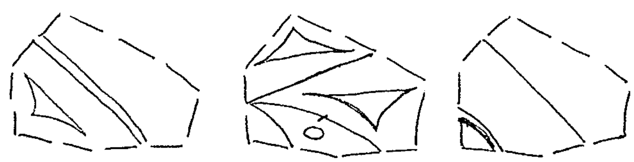
3390
6810
8645



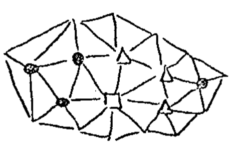
and reflection.   several others.



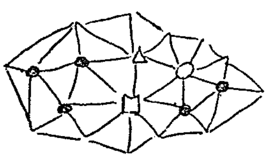
2986
4329
8868



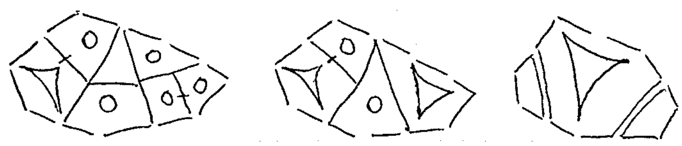
several others.

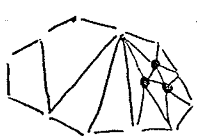


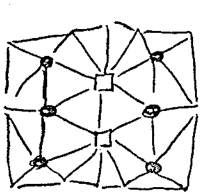
3015
9012
5319



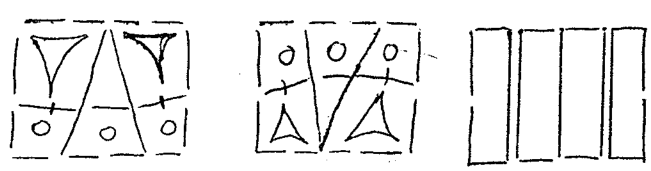
2892
9406
4234

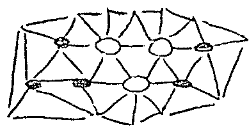




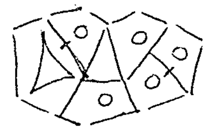


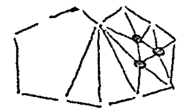
3138
10296
3728

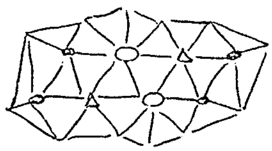




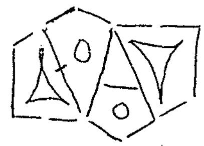
2916
8833
4263

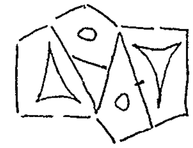


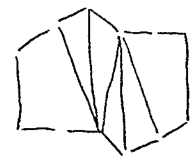




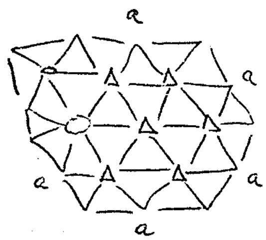
3382
9434
5557



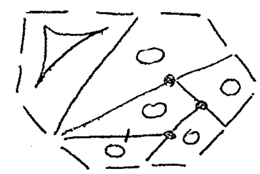


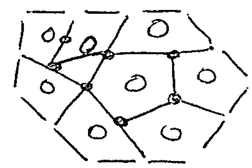


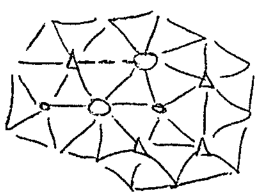
Appendix 9 Anti-sets and Reducers for selected 9/13 clusters



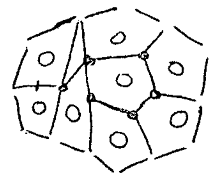
7295
28202
1568

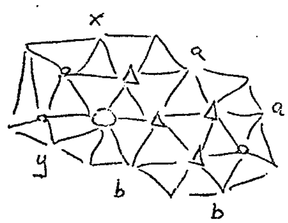




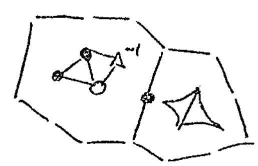


7362
25225
16305

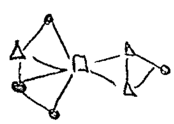
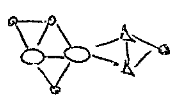
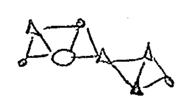




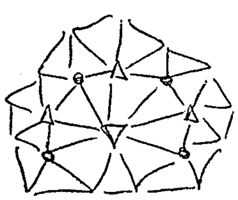
6740
41495
9459



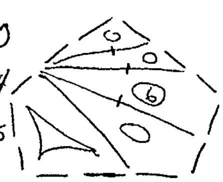
There are similar antisets and reducers for

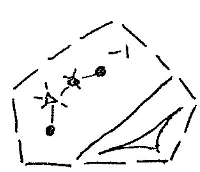
 and . However, 

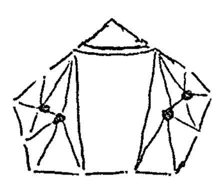
is Symmetrically D-irreducible

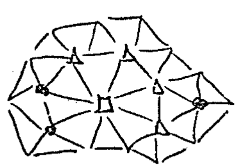


6980
29454
13135

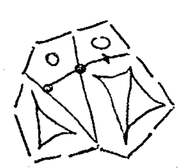


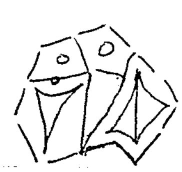


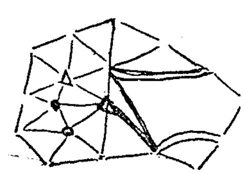


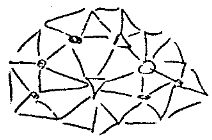


7062
33884
7676

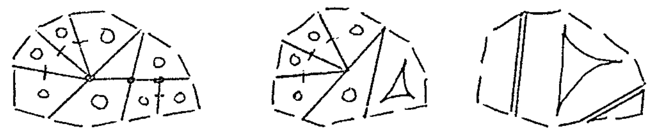




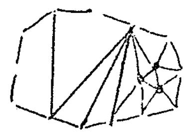




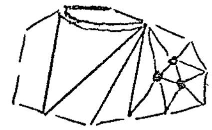
6535
33731
8125

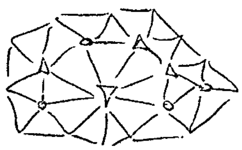


seed:

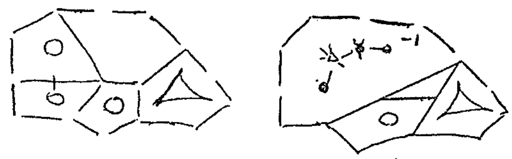


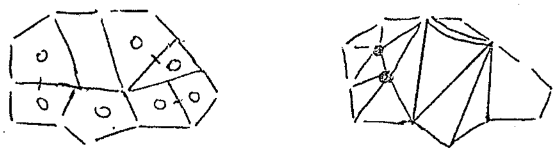
reducer:

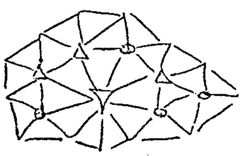




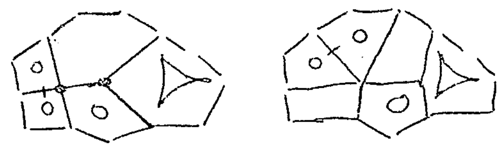
6923
41293
8170

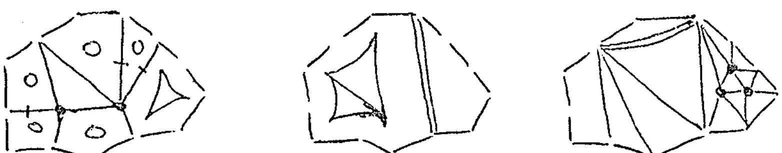


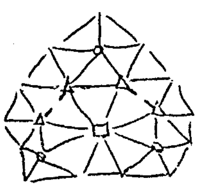




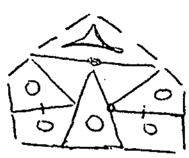
6751
42546
8344

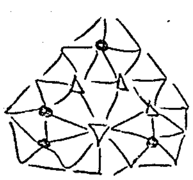




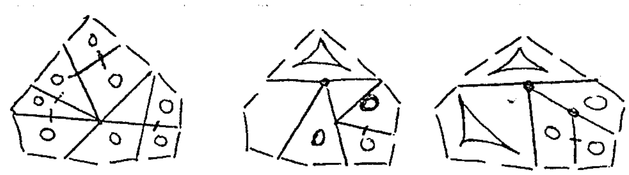


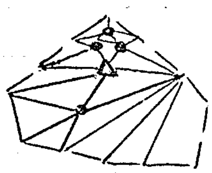
7135
18537
31009

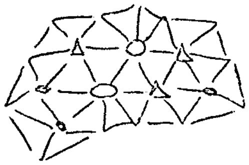




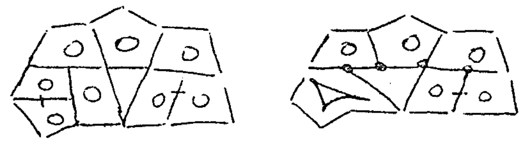
6840
28682
11290

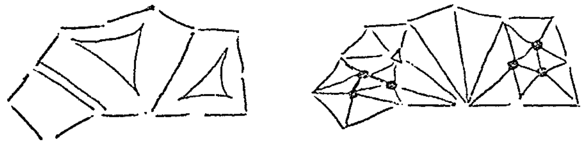


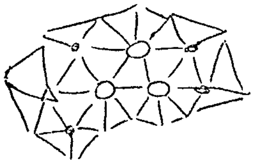




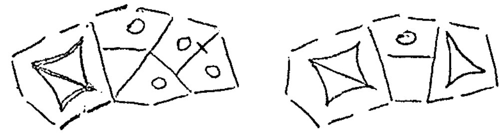
6952
32232
8510

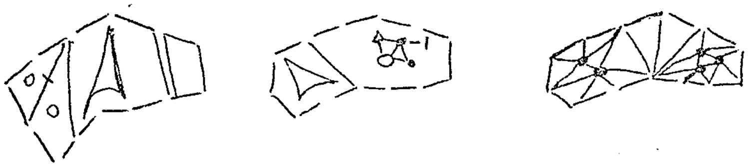


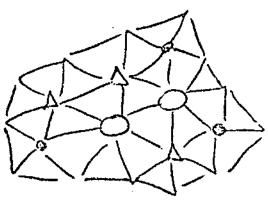




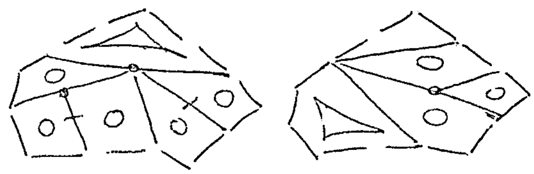
6869
35368
8162

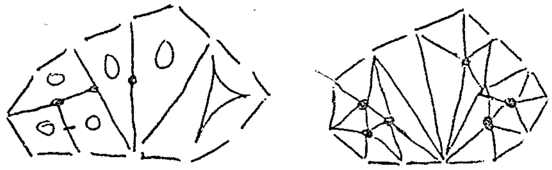


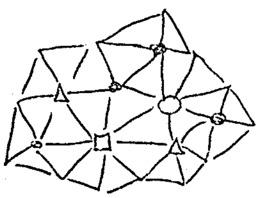




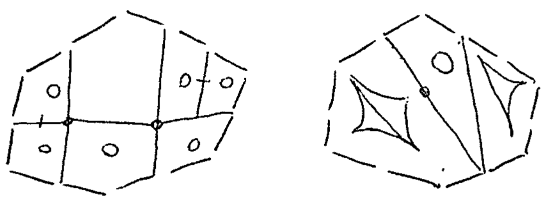
7259
25657
9953

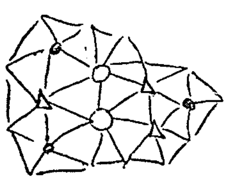




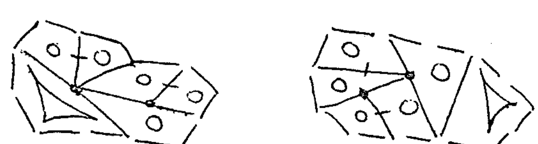


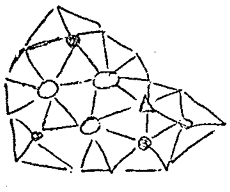
6795
32987
8966



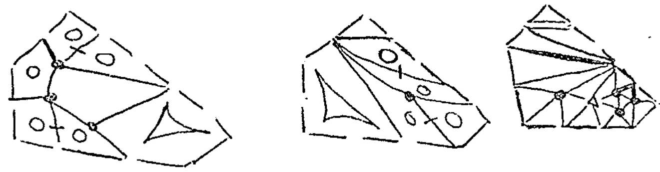


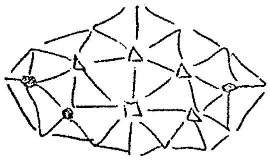
7124
26813
8969



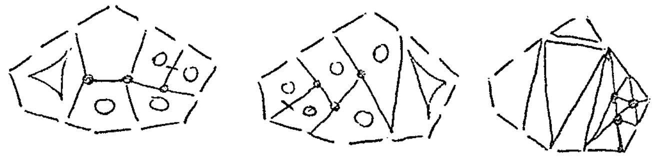


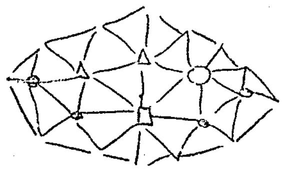
6948
28967
8371



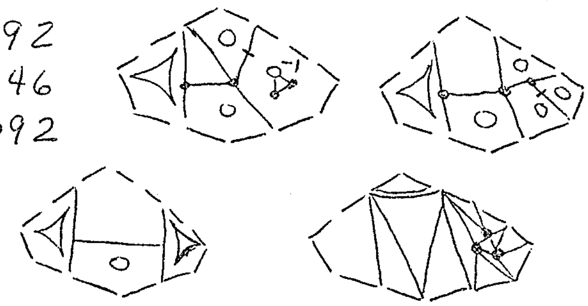


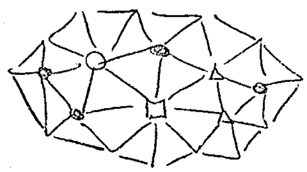
7088
41164
9001



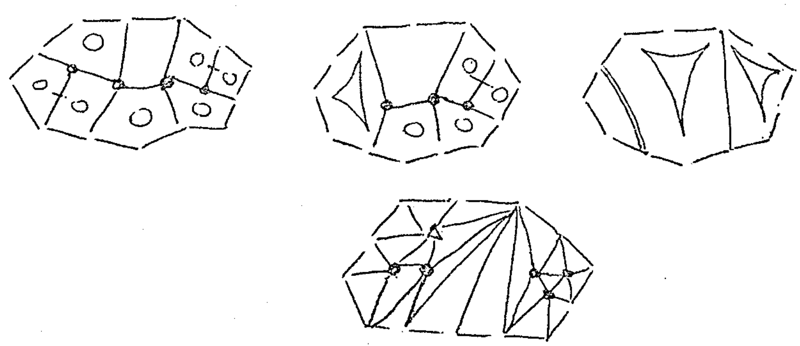


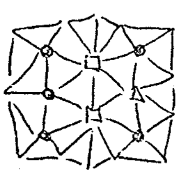
6792
42746
8092



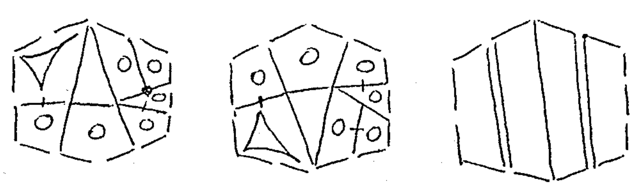


6896
32650
8522





6754
40724
7654



Appendix 10 Extended form of Kempe Constraints

As usual, it is easiest to describe this extended form by using the 6-break constraint as a model. The logic form of this constraint is:

$$(\emptyset) \rightarrow (1)(3)(13) + (2)(4)(24) + (1)(4)(14) + (2)(13)(123) + (3)(24)(234)$$

Using isotopes, this becomes:

$$(\emptyset) \rightarrow (\emptyset)(1)(3)(13) + (\emptyset)(2)(4)(24) + (\emptyset)(1)(4)(14) + (\emptyset)(2)(13)(123) + (\emptyset)(3)(24)(234)$$

From this form, it is a small step to the extended form of the Kempe constraint. Instead of a true/false value, let $(\emptyset)$ denote the number of ways that (U;Qn) can be 4-coloured with scheme $(\emptyset)$ on Qn. For each colour partition, say (ab/cd), every extension of $(\emptyset)$ on Qn to U can be classified according to the way the Kempe chains of U(ab/cd) that meet Qn are arranged. For the 6-break constraint, there are five patterns. In general, there is a pattern for each isotope class. Let $(\emptyset^i)$ denote the number of ways that $(\emptyset)$ on Qn can be extended to U, such that the Kempe chains of U(ab/cd) that meet Qn are arranged in pattern i. This leads to the following constraint:

$$(\emptyset) = \sum_i (\emptyset^i)$$

Let pattern i correspond to an isotope $(\emptyset, x, y, \ldots z)$. For every colouring of (U;Qn) counted by $(\emptyset^i)$, there exists a

Kempe chain interchange with respect to (ab/cd) that transforms the scheme ($\emptyset$) into any other scheme of the isotope. Therefore we have:

$$(\emptyset^{i}) = (x^{i}) = (y^{i}) = \ldots = (z^{i}).$$

Let this value be denoted by $(\emptyset^{i}=x^{i}=y^{i}=\ldots=z^{i})$. The preceeding 6-break constraint becomes:

$$(\emptyset) = (\emptyset^1=1^1=3^1=13^1) + (\emptyset^2=2^2=4^2=24^2) + (\emptyset^3=1^3=4^3=14^3)$$
$$+ (\emptyset^4=2^4=13^4=123^4) + (\emptyset^5=3^5=24^5=234^5)$$

In summary, the extended form has a non-negative integer-valued variable for each scheme, and a similar variable for each isotope. There are up to three constraints for each scheme, and each is of the form:

$$\text{scheme variable} = \sum \text{isotope variables.}$$

While the number of isotopes, and hence the number of variables, is large, for every isotope that contains a forbidden scheme, the associated variable is set to zero, and can be removed from the constraint. Further, whenever a constraint has only one isotope in the summation on the RHS that is not set to zero, the constraint becomes:

$$(\emptyset) = (\emptyset^{i} = x^{i} = y^{i} = \ldots = z^{i}).$$

These two variables can be replaced by one.

Obviously, the logic form of immersion corresponds to the case when every isotope in the RHS has a value that is zero for one reason or another. If, after removing variables corresponding to immersed schemes, and combining variables that must have equal values, we have a constraint of the following form:

$$r = r + s,$$

then we can conclude that the value for variable s must be zero, i.e., the isotope or pattern corresponding to variable s cannot occur. This is one of the ways that the extended constraints can be used to show that certain colourings cannot occur. This conclusion cannot be derived from the logic form of the constraints. After forbidding isotopes in this manner, hopefully one or more schemes can be shown to be forbidden.

If this does not help, then there is one last chance. Consider the system of remaining variables and constraints as a linear programming problem and add the constraint $(\emptyset)\geq 1$. If there is no feasible solution to this problem, then $(\emptyset)=0$, i.e., the scheme $(\emptyset)$ is forbidden. Typically, this method is used in conjunction with a programme that determines closure by the logic form of the constraints.

The above development of the extended constraints is different from that contained in [12], and may be new.