CERTIFICATES FOR UNDIRECTED GRAPHS

by

Laura J. Sedgwick

A thesis submitted in conformity with the requirements
for the degree of Masters of Science
Graduate Department of Computer Science
University of Toronto

Certificates for Undirected Graphs

Laura J. Sedgwick

M.Sc Thesis, 1998

Department of Computer Science

University of Toronto

Abstract

An isomorphism certificate (I.C.) for a graph is a set of labeled edges which in combination with an isomorphic copy of a graph allows its errorless reconstruction. Work by Rubinstein, an economist, on binary relations first motivated work on the minimum size of I.C.s for tournaments. This thesis marks the first work on I.C.s for undirected graphs. It provides a structure upon which all further work on I.C.s can be based. As well it introduces the concept of using edges which are in the complement of the graph (non-edges) to provide extra information. The bulk of the work done is on valid and optimal I.C.s for several graphs. We look at wheels, complete binary trees, complete k-ary trees, and grids. For the grid we were unable to prove our I.C. is optimal but an asymptotically correct upper bound was found on the size of the minimum I.C.

Acknowledgments

I wish to thank Professor Michael Molloy for suggesting the material for this thesis, assisting in developing the ideas and theories within, and spending many hours editing it. In addition I'd like to thank Professor Derek Corneil for taking the time to read and comment on the thesis. I'd like to thank the Natural Sciences and Engineering Research Council for the financial support which made this work possible. And finally, I'd like to thank the Department of Computer Science at the University of Toronto, and my fellow graduate students for providing such a pleasant environment to work in.

Contents

List of Figures

Chapter 1

Introduction

In the last decade an economist named Ariel Rubinstein studying binary relations became interested in the most economical way to define these relations. This problem was drawn to the attention of Noga Alon and work on certificates for tournaments began. It is this work that motivated the questions we have sought to answer in this thesis.

A certificate for a graph attempts to capture the minimum amount of information about the graph from which the graph may be reconstructed. The work done on certificates has so far been limited to a class of directed graphs called tournaments. We believe this thesis is the first work extending the concept to arbitrary undirected graphs. It provides an introduction to techniques we have found useful in determining valid and optimal certificates and presents certificates for several classes of graphs. However, as the reader will see, the room for more work is vast.

This chapter will provide the reader with the majority of definitions which they will need to follow the material in this thesis. For any other definitions please consult any basic graph theory textbook. In this chapter we also provide some basic results which are used throughout the entire thesis. The next two chapters provide background material for the reader. Chapter 2 provides an overview of the previous work done on certificates and indicates how some of the work is related to what we have done. Chapter 3 provides more general background on an interesting class of problems which includes the question studied here. And the final three chapters provide the bulk of the original work for the thesis, focusing on certificates for several specific classes of graphs.

The following definitions are important to the understanding of the material in this thesis.

A graph G is identical to a graph G', written $G = G'$ if and only if both graphs have the same vertex and edge sets. A graph G is *isomorphic* to a graph G', written $G \cong G'$, if and only if there is a bijection, $\Theta : V(G) \to V(G')$, for which (v, w) is an edge of G if and only if $(\Theta(v), \Theta(w))$ is an edge of G'. Θ is an isomorphism of G onto G'. An *automorphism*, Θ, is an isomorphism of G onto itself. Note that for any bijection Θ, $G \cong \Theta(G)$ but $G = \Theta(G)$ if and only if Θ is an automorphism.

For a given graph G we say that G *contains* D, $G \succeq D$, provided $D = (D', D'')$ where D' is a spanning subgraph of G and D'' is a spanning subgraph of $\overline{G}$. We refer to the edges of D'' as non-edges of G.

We shall say that D is a *isomorphism certificate, (I.C.),* for G provided that:

1. $G \succeq D$.
2. If $G' \cong G$ and $G' \succeq D$ then $G' = G$.

The focus of this thesis is on determining optimal valid isomorphism certificates for several classes of graphs. This is a type of reconstruction problem. In determining an isomorphism certificate for a graph we are actually determining what information about that graph is crucial to the structure of the graph. Once we have determined the important information we can give that information to someone else and they are able to reconstruct the graph from that information. In chapter 3 we study problems where you are given the information about the graph and the task is to decide whether the information given is sufficient to allow you to reconstruct the original graph. These problems focus on reconstructing the graph up to isomorphism whereas our problem requires an identical reconstruction and as such can be considered the next step in the process.

A *positive isomorphism certificate, (P.I.C.),* of a graph G is an I.C. $D = (D', D'')$ such that $E(D'') = \emptyset$. Similarly, a *negative isomorphism certificate, (N.I.C.),* of a graph G is an I.C. where $E(D') = \emptyset$ and a *mixed isomorphism certificate (M.I.C.)* is an I.C. where $E(D'), E(D'') \neq \emptyset$.

Throughout this thesis we shall relax the distinction between the pair $D = (D', D'')$ and the elements of the pair, D' and D''. Thus we shall often refer to the edge (u, v) as being in D where more properly $(u, v) \in D'$ in the case D is a P.I.C. or a N.I.C. Similarly, we refer to the non-edge (u, v) an being an element of D as well as D''. We refer to a vertex as being isolated in D when in fact we mean isolated in both D' and D''. We may also talk about an automorphism from D onto D which is not precisely correct. For a graph G we are in fact referring to a bijection $\Theta : V(G) \to V(G)$ such that $C = (\Theta(D'), \Theta(D'')) = D$.

The *size* of an I.C. $D = \{D', D''\}$ is $|E(D')| + |E(D'')|$. An I.C. D for a graph G is a *minimum I.C.* if G has no I.C. of smaller size. The corresponding definitions follow for P.I.C.s, N.I.C.s, and M.I.C.s.

The *I.C. number* for a graph is the size of a minimum I.C. We define *P.I.C. number, N.I.C number* and *M.I.C. number* in the obvious manner.

For a graph G, to prove that a pair D, $G \succeq D$, is not a valid I.C. for G it suffices to provide a graph $G' \neq G$ such that $G' \cong G$ and $G' \succeq D$. To prove that a D is a valid I.C. it suffices to show that $\forall G'$, if $G' \cong G$ and $G' \succeq D$ then $G' = G$.

The following lemmas are useful in proofs throughout this thesis. We begin with one very general lemma.

Lemma 1.1 *Let graph $G \succeq D$. Let Θ be a bijection from $V(G)$ onto $V(G)$. If $\Theta(D)$ is an automorphism of D and $\Theta(G)$ is not an automorphism of G then D is not a valid I.C. for G.*

Proof $\Theta(G) \cong G$. $\Theta(G) \succeq \Theta(D)$ and $D = \Theta(D)$ so $\Theta(G) \succeq D$. As $\Theta(G) \neq G$ we know D is not a valid I.C. for G. □

Lemmas 1.2 through 1.7 are subcases of this first lemma.

Two vertices, v_1 and v_2, of a graph G are *interchangeable* if there is an automorphism of G, Θ, such that:

1. $\Theta(v_1) = v_2$.

2. $\Theta(v_2) = v_1$.

3. $\Theta(v) = v, \forall v \in V(G), v \neq v_1, v_2$.

We should note that two vertices v_1, v_2 are interchangeable in G if and only if their neighbourhoods in $G \backslash \{v_1, v_2\}$ are identical.

Lemma 1.2 *Let graph $G \succeq D$. If we have two vertices v_1, and v_2 which are interchangeable in D and not in G then D is not a valid I.C. for G.*

Proof Consider the bijection Θ defined as follows:

1. $\Theta(v_1) = v_2$.

2. $\Theta(v_2) = v_1$.

3. $\Theta(v) = v, \forall v \in V(G), v \neq v_1, v_2$

and then apply lemma 1.1. □

The next three lemmas are frequently used subcases of this lemma which we state explicitly.

Lemma 1.3 (Isolated Vertices) *Let graph $G \succeq D$. If D has two isolated vertices, v_1 and v_2 and v_1 and v_2 are not interchangeable in G then D is not a valid I.C. for G.*

We say that a graph D, contains an *isolated edge* provided D has a component which consists of a single edge.

Lemma 1.4 (Isolated Edge) *Let graph $G \succeq D$. If D has an isolated edge between vertices v_1 and v_2 and v_1 and v_2 are not interchangeable in G then D is not a valid I.C. for G.*

Lemma 1.5 (Dangling Edges) *Let graph $G \succeq D$. If D has two vertices v_1 and v_2 which are not interchangeable in G, and which both have degree one in D; and which are both adjacent to v_3 in D, then D is not a valid I.C. for G.*

Next we present two more useful special cases of lemma 1.1. We say that a graph G contains an *isolated path* provided it has a component which is a path of length three or longer.

Lemma 1.6 (Isolated Path) *Let graph $G \succeq D$. If D contains an isolated path, $v_1, v_2, \ldots, v_n, n \geq 3$ and the bijection Θ defined as follows is not an automorphism of G then D is not a valid isomorphism certificate for G.*

1. $\Theta(v_j) = v_{n-j+1}, 1 \leq j \leq n$.

2. $\Theta(v) = v, v \in V(G), v \neq v_k, 1 \leq k \leq n$.

Lemma 1.7 (Isomorphic Components) *Let graph $G \succeq D$. If D has two isomorphic components C_1 and C_2 with isomorphism $\Theta_C : V(C_1) \rightarrow V(C_2)$ and if the bijection Θ defined as follows is not an automorphism of G then D is not a valid I.C. for G.*

1. $\Theta(v) = \Theta_C(v)$ *if* $v \in V(C_1)$.

2. $\Theta(v) = \Theta_C^{-1}(v)$ *if* $v \in V(C_2)$.

3. $\Theta(v) = v$ *for all other* $v \in G$.

The next lemma is an application of this lemma which will be used in the chapter on grids.

Lemma 1.8 *Assume every automorphism on G except the identity maps at most p vertices in G onto themselves. Then any valid P.I.C. for G can not have two isomorphic components whose combined size is less than $|G| - p$.*

Proof Assume D is a valid P.I.C. for G such that it has two isomorphic components whose combined size is less than $|G| - p$. As every automorphism on G maps at least $|G| - p$ vertices onto different vertices we know that the mapping described in the isomorphic components lemma can not be an automorphism and thus D can not be a valid P.I.C. $\square$

The next two lemmas are each useful in at least one of the chapters of this thesis.

Lemma 1.9 (Certificate Hierarchy) *Let D=(D', D'') be an I.C. for G. For any $C = (C', C'')$ if $D' \subseteq C'$, $D'' \subseteq C''$, and $G \succeq C$ then C is an I.C. for G.*

Proof Consider any graph $G' \cong G$ such that $G' \succeq C$. Then $G' \succeq D$ and since D is an I.C. for G, $G' = G$. Thus C is an I.C. for G. $\square$

In the following lemma we write $D - \{v\}$ when in fact we mean $(D' - \{v\}, D'' - \{v\})$.

Lemma 1.10 *If $G \succeq D$ and vertex v is universal in G then $D - \{v\}$ is an I.C. for $G - \{v\}$ if and only if D is an I.C. for G.*

Proof First assume D is an I.C. for G. Consider $G' \cong G - \{v\}$ such that $G' \succeq D - \{v\}$. Assume $G' \neq G - \{v\}$. Let G'' be the graph created from G' by adding a new vertex v which is adjacent to every vertex in G'. Then clearly $G'' \cong G$, $G'' \succeq D$ but $G'' \neq G$ as $G' \neq G - \{v\}$. However this is a contradiction to the fact D is an I.C. for G and thus $G' = G - \{v\}$. So $D - \{v\}$ must be a valid I.C. for $G - \{v\}$.

Now consider if $D - \{v\}$ is an I.C. for $G - \{v\}$. Consider $G' \cong G$, $G' \succeq D$. Then $G' - \{v\} \succeq D - \{v\}$ and $G' - \{v\} \cong G - \{v\}$. Thus we know $G' - \{v\} = G - \{v\}$ and so $G' = G$. Thus D is a valid I.C. for G. $\square$

The final result in this chapter is a major result bounding below the P.I.C. number of any graph whose only automorphisms involve almost all of the graph. This result is key to the theorems presented in the chapter on grids.

First we present a very loose upper bound on the number of non-isomorphic connected graphs with j edges. This bound can be greatly improved upon but it is sufficient for our purposes and allows for a straightforward proof.

Lemma 1.11 *For any $j > 0$ there are at most 2^{j^2} non-isomorphic connected graphs with j edges.*

Proof Any connected graph with j edges may have at most $j + 1$ vertices. Now we know a graph on i vertices has $\frac{i(i-1)}{2}$ possible edges and thus there are $2^{\frac{i(i-1)}{2}}$ possible graphs on i vertices. Summing this value for all the numbers below $j + 1$ we get that there is a maximum of $\sum_{i=1}^{j+1} 2^{\frac{i(i-1)}{2}}$ possible graphs on j edges. Simplifying, $\sum_{i=1}^{j+1} 2^{\frac{i(i-1)}{2}} < \sum_{i=1}^{\frac{j(j+1)}{2}} 2^i < 2^{\frac{j(j+1)}{2}+1} = 2^{\frac{j^2+j+2}{2}} \leq 2^{j^2}$ if $j \geq 2$. And if $j = 1$ $2^{j^2} = 2 > 1$. $\square$

Theorem 1.12 *Let $\mathcal{G}$ be a class of graphs such that for any graph $G \in \mathcal{G}$ on n vertices the only non-identity automorphisms on G leave at most $o(n)$ vertices fixed. Then the P.I.C. number for any graph $G \in \mathcal{G}$ on n vertices is at least $(1 - o(1))n$.*

Proof Fix any $\epsilon > 0$. We will show that for n sufficiently large, all P.I.C.s for graphs on n vertices have at least ϵn components.

We know by lemma 1.11 that the number of non-isomorphic components of size j is at most 2^{j^2}. Now choose n such that $n > \frac{2}{\epsilon} \sum_{j<\frac{4}{\epsilon}} 2^{j^2}$ and for any graph $G \in \mathcal{G}$ on n vertices all automorphisms on G leave fixed at most $n - \frac{8}{\epsilon}$ vertices.

Now pick any $G \in \mathcal{G}$ which has n vertices. Let c_j be the number of components of size j in a P.I.C. for G of size n or smaller (if no such P.I.C. exists we are done), and let C be the total number of components in the P.I.C. Knowing that we have no automorphisms which leave fixed more then $n - \frac{8}{\epsilon}$ vertices and applying the lemma 1.8 we know $\sum_{j<\frac{4}{\epsilon}} c_j \leq \sum_{j<\frac{4}{\epsilon}} 2^{j^2}$. However by our choice of n this also yields $\sum_{j<\frac{4}{\epsilon}} c_j \leq \frac{\epsilon}{2}n$.

Our goal is to show $C < \epsilon n$. Assume $C = \sum_{j \geq 1} c_j > \epsilon n$, then $\sum_{j \geq \frac{4}{\epsilon}} c_j > \frac{\epsilon}{2}n$.

$$\Rightarrow \sum_{j\geq 1} jc_j \geq \sum_{j \geq \frac{4}{\epsilon}} \frac{4}{\epsilon} c_j \geq \frac{4}{\epsilon} \sum_{j \geq \frac{4}{\epsilon}} c_j \geq \frac{4}{\epsilon}\frac{\epsilon}{2} n = 2n$$

However $\sum_{j \geq 1} jc_j \leq n$.

This contradiction yields that the number of components is less then ϵn. However the P.I.C. number is at least $n - C > n - \epsilon n$ and since this is true for every $\epsilon > 0$ and n sufficiently large in terms of ϵ, the P.I.C. is at least $(1 - o(1))n$. $\square$

Chapter 2

Tournaments

2.1 Introduction

As previously mentioned, all work on certificates prior to this thesis has focused on a class of directed graphs called tournaments. This chapter provides the reader with an overview of the work done in this area. It begins with a discussion of the famous conjecture made by Rubinstein, an economist, which was the original stimulus for all this work. We then look at bounds which have been found for the I.C. number for arbitrary tournaments. Section 2.3 introduces a different type of certificate and more bounds on the certificate number of tournaments. Finally we conclude with several specific results which include finding the I.C. number for two subclasses of tournaments.

A *tournament* is a labeled directed complete graph, i.e. the graph resulting from orienting the edges of a complete graph. The reader should note that this is not the same as a complete directed graph which has two edges between every pair of vertices, one in each direction. One can imagine a tournament on n vertices as a competition between n teams where every pair of teams plays once. The direction of the arc between teams A and B indicates which team won the game. There has been a huge volume of work on tournaments; an introduction to which may be found in [13].

An isomorphism certificate for a tournament consists of a spanning directed sub-

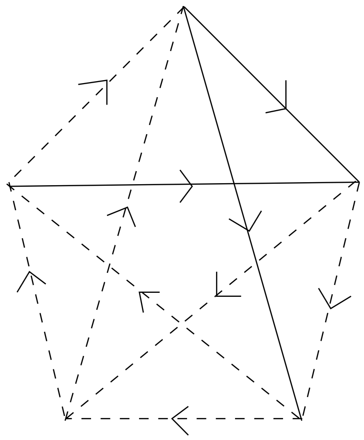

Figure 2.1: H - an I.C. for R_5

graph of the tournament. We measure the size of an isomorphism certificate by the number of arcs in this spanning subgraph. As there is an edge present between every pair of vertices, it is no longer relevant to consider non-edges. However, if we number the vertices in a tournament and consider an arc from a lower numbered vertex to a higher numbered vertex as an edge and an arc from a higher numbered vertex to a lower numbered vertex as a non-edge there are many similarities between tournaments and the graphs we study in this thesis.

In the last five years there has been an increasing amount of interest in certificates for tournaments. This work was first motivated by questions in economics raised by Ariel Rubinstein.

Conjecture 2.1 ([16]) *For $n > 5$ the size of the minimum isomorphism certificate of any tournament on n vertices is at least $n - 1$.*

There are two well-known tournaments which have certificates of size $n-2$. These are the regular tournaments on three and five vertices which we will describe momentarily. It is widely believed that these are the only tournaments with certificates of size smaller than $n - 1$ and the conjecture is true.

The regular tournament on $n = 2m + 1$ vertices, $v_1, v_2, \ldots, v_n$ with arc set $A = \{(v_i, v_j), 1 \leq i \leq n, j = k \bmod n, i + 1 \leq k \leq i + m\}$ is denoted by R_n. That is, each vertex beats the m vertices which succeed it, if we view the vertices cyclicly.

R_3 is a cycle on 3 vertices. Any one of the arcs in this tournament is a valid I.C. for this tournament. Similarly, with a little thought the reader will see that the subgraph H in figure 2.1 is an I.C. for R_5 with only three arcs. The edges of H are solid lines and H has been embedded in a dashed line version of R_5 to assist the reader. Thus the I.C. number for each of these two tournaments is $n-2$.

Though fairly extensive work has been done in this area all attempts to prove the previous conjecture true have been unsuccessful. However, this result was proved to be almost correct by Alon (see [16]) in the following sense:

Theorem 2.2 *The I.C. number of any tournament on n vertices is at least (1 - o(1))n.*

The proof of this theorem follows along the same lines as the proof of theorem 1.12. There are two main differences. The first one being the obvious difference that here we are talking about directed graphs while theorem 1.12 applies to undirected graphs. The second difference is that here we show that we can not have two isomorphic components in any I.C.

Recall that it was crucial to the proof of theorem 1.12 that no P.I.C. had two isomorphic components on a small number of vertices. This was the reason it was necessary to require that all automorphisms (other than the identity) leave fixed only a few vertices. However for tournaments we do not need such a condition since it is straightforward to show that no I.C. has two isomorphic components of any size. Consider an I.C. which has two isomorphic components and then consider the isomorphism which switches the vertices in these two components. It can not be an automorphism as the direction of the arc between two vertices which are swapped has been reversed, and so we could not have had a valid I.C. to begin with. The rest of the proof of this theorem mimics that of theorem 1.12.

Furthermore, Fishburn, Kim and Tetali [4] have examined the I.C. number for all tournaments on less than 8 vertices. The conjecture is true for every such tournament.

The notion of isomorphism certificates has been extended in a very natural way by Fishburn, Kim and Tetali [5]. Recall that a certificate C for a graph G is a subgraph of G which in combination with some extra information I about the graph G allows errorless reconstruction of G. For an isomorphism certificate that information is an unlabeled copy of the graph. In a score certificate that information is the score sequence for the tournament G. (The *score sequence* $S(G) = (s_1, s_2, \dots s_n)$ for a tournament G on n vertices is a list of the out-degree of each of the n vertices. Recalling the view of a tournament as a round-robin competition the reader will see the score for team is the number of teams it beats.) D is a *score certificate (S.C.)* for a tournament $G \succeq D$ provided for every tournament G', if $S(G') = S(G)$ and $G' \succeq D$ then $G' = G$. We define the *score certificate number* for a tournament G to be the size of the minimum score certificate for G.

Though Rubinstein's conjecture has not been proved, several bounds on the I.C. number and the S.C. number for general tournaments have been found. These are presented in the next two sections. The final section introduces the reader to two interesting results on I.C.s for subclasses of tournaments.

2.2 I.C. Number for Tournaments

The proof of Rubinstein's conjecture would provide a lower bound on the value of the I.C. number for an arbitrary tournament. Fishburn, Tetali and Kim succeeded in finding an upper bound on this value for random tournaments [1]. Recently, Alon and Ruszinkò [1] have gone even further and proved that the same bound holds for all tournaments.

Theorem 2.3 *The I.C. number for any tournament on n vertices is at most $n \log_2 n$.*

Proof All logarithms are in base 2. Let T be an arbitrary fixed tournament. By symmetry, it is enough to show that there is some $T' \cong T$ such that the I.C. number of T' is at most $n \log_2 n$.

Recall that $A(T)$ is the arc set of T. We set $H_0 = \emptyset$ and we will recursively define H_i momentarily. For each i, we set $T_i = \{T' | T' \cong T, A(T') \supseteq H_i\}$. Our goal is to find H_i such that $|T_i| = 1$. Note that $|T_0| = n!/|Aut(T)|$ where $Aut(T)$ is the automorphism group of T.

If $|T_i| > 1$ choose any pair of vertices x, y such that $(x, y), (y, x) \notin H_{i-1}$ and such that there are tournaments T^1 and T^2 in T_{i-1} with $(x, y) \in A(T^1)$ and $(y, x) \in A(T^2)$. For every tournament in T_{i-1} either arc (x, y) or (y, x) is in that tournament. If the number of tournaments with arc (x, y) is at most $|T_{i-1}|/2$ then $H_i = H_{i-1} \cup (x, y)$ else $H_i = H_{i-1} \cup (y, x)$. Thus $|T_i| \leq |T_{i-1}|/2$. As $|T_0| \leq n!$ we know $\exists i, i \leq \log(n!)$ such that $|T_i| = 1$. As H_i is a valid I.C. for the unique $T' \in T_i$ and as $T' \cong T$, we know that the I.C. number of T is at most $|H_i| \leq \log(n!) \leq n \log(n)$. $\square$

2.3 Score Certificates

In [5] Fishburn, Kim and Tetali provide the following lower bound on the S.C. number of a tournament:

Theorem 2.4 *Except for the regular tournaments on 3 and 5 vertices the S.C. number of any tournament on n vertices is at least $n - 1$.*

The proof of this theorem is not straightforward. They begin by showing that any S.C. must have at most two components which are trees. They do this by focusing on the bipartite nature of these components and showing that if there are two bipartite components then one of them must be an isolated vertex. Thus if there are three components which are trees by pairing the first two and then the last two it follows that there must be two isolated vertices which is not possible in a valid S.C. This immediately yields the fact that any S.C. has size at least $n - 2$.

In the next section of their paper the authors consider what any minimum S.C. must look like. They show it must consist of two tree components which together are a S.C. for either R_3 or R_5 and a collection of directed cycles of length greater than three. Using a combination of lemmas they show that no S.C. of this form can possibly exists and thus that their theorem holds.

We shall provide the proof of one of the simpler lemmas.

Lemma 2.5 *If D is a valid S.C. for a tournament T then there is no cycle in $\overline{D}$.*

Proof Assume $\overline{D}$ has a cycle C. Let T' be the tournament which results from T by reversing the direction of every edge in C. Then $T' \cong T, T' \succeq D$, and $T' \neq T$. Thus D is not a valid S.C. □

Recently Alon and Ruszinkò [1] were successful in bounding the S.C. number away from the total number of arcs in the tournament.

Theorem 2.6 *There exists an $\epsilon > 0$ so that for any tournament on n vertices, the S.C. number for that tournament is at most $(1/2 - \epsilon)n^2$.*

In their paper they choose $\epsilon = 1/160$ and completed the proof for $n \geq 80$. This could be improved upon. They begin by defining a set of arcs as score reversible if it is possible to reverse the direction of every arc in the set and have the score sequence of the resulting tournament equal the score sequence of the original tournament. All valid S.C.s intersect every score reversible set and conversely any set which intersects every score reversible set is a valid S.C. Thus, to complete their proof it suffices to find a set of ϵn^2 arcs which have no subset which is score reversible. The converse to this set is a valid S.C.

They next introduce the notion of a back edge. If the vertices in the tournament are labeled $v_1, v_2, \ldots, v_n$ a back edge is any arc (v_i, v_j) where $i > j$. They show any set containing only back edges is not score reversible. Thus there can not be more then ϵn^2 back edges or the proof is finished. Then they partition the set of non-back edges. They find a small set of these arcs which in combination with the set of back edges form the desired set and this completes their proof.

2.4 I.C.s for Tournaments

We shall present three other results in this chapter. The first two are both bounds on the I.C. number for special subclasses of tournaments. The final is a disproof of a conjecture by Fishburn, Kim and Tetali [4].

For the following theorem we require the definition of a another special class of tournaments. Define the *transitive tournament* T on n vertices, $v_1, \dots v_n$, to be the tournament where $(v_i, v_j) \in A(T)$ if and only if $i > j$.

Theorem 2.7 ([4]) *The I.C. number for $R_n, n \geq 7, n$ odd, is at most $\frac{3n-9}{2}$.*

Proof Let $n = 2m + 1$ and assume we have vertices $v_1, v_2, \dots v_n$. Fix an arbitrary vertex j and define C_j to have arc set $A(C_j) = A \cup B \cup C$ where:

$A = \{(v_i, v_{i+1}), 1 \leq i \leq 2m, i \neq j-1, m, m+j\}$,

$B = \{(v_{m+1}, v_i), m+3 \leq i \leq 2m+1\}$, and

$C = (v_j, v_{m+j})$.

Then $|C_j| = 2m - 3 + m - 1 + 1 = 3m - 3$. Given that $n = 2m + 1$ we know $\frac{n-1}{2} = m$ and thus $|C_j| = \frac{3n-9}{2}$. So it will suffice to prove C_j is a valid I.C. for R_n.

Consider a tournament $R' \cong R_n$ such that R' has C_j as a subgraph. Then R' has arcs $(v_{m+1}, v_i), m+2 \leq i \leq 2m+1$. As $R' \cong R_n$ we know that we must have arcs into v_{m+1} from the first m vertices. We also know that the vertices which beat v_{m+1} and which are beaten by v_{m+1} must each be a transitive tournament on m vertices. Thus if we had a path of length $m-1$ amongst each of these sets of vertices we would be done.

We are missing one arc in the path $v_1, v_2, \dots, v_m$, the arc (v_{j-1}, v_j). However we can use the arc (v_j, v_{m+j}) to assist us in recapturing the desired information. We have the path $v_j, v_{j+1}, \dots, v_{m+j}$ in R'. Because v_j beats v_{m+j} we know that v_j must beat every other vertex in this path as $R' \cong R_n$. Thus all other vertices in R', particularly v_{j-1} must beat v_j. Thus we have established the first desired path.

Similarly we know that all the vertices in the path $v_j, v_{j+1}, \dots, v_{m+j-1}$ must beat vertex v_{m+j} and thus v_{m+j} must beat v_{m+j+1}. This yields a path on m vertices from v_{m+2} to v_{2m+1} and completes the proof. □

The second theorem in this chapter focuses on the I.C. number for almost transitive tournaments.

Define the *almost transitive tournament* on n vertices, $v_1, v_2, \ldots, v_n$, to be the tournament where v_i beats v_j if $i < j$ except v_n beat v_1. Denote this tournament by ATT_n.

Theorem 2.8 ([4]) *For $n \geq 5$ the I.C. number for ATT_n is n.*

As this proof requires the use of lemmas from Fishburn, Kim and Tetali's paper [4] which we are not providing we have not provided a complete proof. However we shall sketch the main body of the proof.

The authors first show that we must have all the arcs in the path $v_2, v_3, \ldots, v_{n-2}$ in any valid I.C. by an application of one of their lemmas. The reader can easily check that the addition of any two arcs to this set is not a valid I.C. Thus we must only show the subgraph of ATT_n, C with these arcs and $(v_1, v_3), (v_3, v_n), (v_n, v_1)$ is a valid I.C. Consider any graph A' on vertex set $v_{1'}, v_{2'}, \ldots, v_{n'}$, such that $A' \cong ATT_n$ and - relabeling the vertices of C appropriately - $A' \supseteq C$. As $A' \cong ATT_n$ the only cycles of length 3 in A' include the vertices v_1 and v_n. Now $v_{3'}$ has in-degree and out-degree of at least two in C so it can not be v_1 or v_n. And considering the in-degree and out-degree of $v_{1'}$ and $v_{n'}$ we know they must be v_1 and v_n respectively. This information in combination with the path in C identifies all the vertices in A' uniquely and we see $A' = ATT_n$.

Fishburn, Kim and Tetali first raised questions about the class of tournaments which achieve the potential lower bound of $n-1$. We know that the transitive tournaments fall into this class and they conjectured there were few others. They called a tournament T *pseudo-linear* if it can be written as a linear order $A_1 \to A_2 \to \ldots \to A_k$ with each A_i either a single vertex or a directed three cycle. Then they conjectured:

Conjecture 2.9 *For $n \geq 6$, a tournament T on n vertices is pseudo-linear if and only if it has I.C. number $n-1$.*

This conjecture was proven false by Thomasse [18] who provided the counterexample given in figure 2.2. This tournament has R_3 and R_5 as subgraphs with all arcs

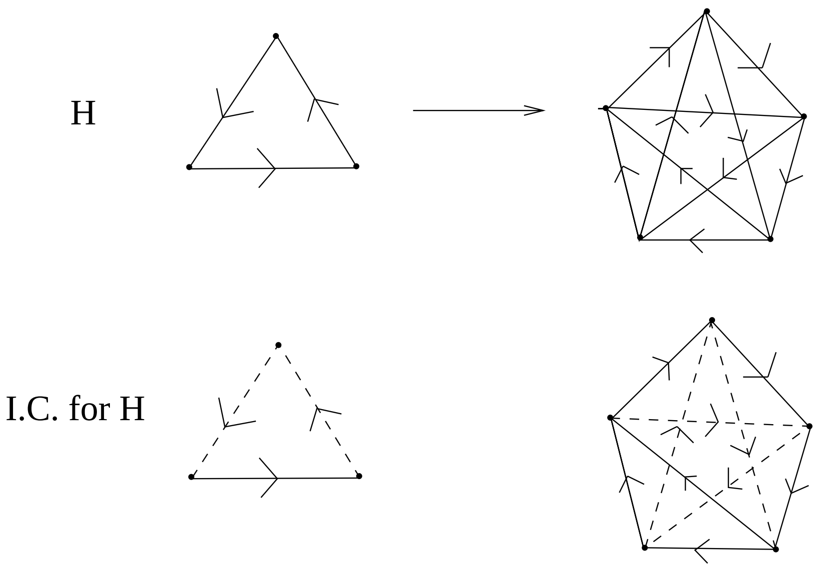

Figure 2.2: Thomasse's counterexample

between vertices in the R_3 subgraph being directed towards the vertices in the R_5 subgraph. In figure 2.2 we also give a valid I.C. for H which has size $n-1$ and as H is clearly not pseudo-linear, it is a counter-example to the conjecture.

Chapter 3

Reconstruction Problems

3.1 Introduction

When we construct a certificate for a graph we are attempting to identify the key information about that graph, i.e. the information that will allow us to reconstruct the graph. In other reconstruction problems we attempt to reconstruct the graph only up to isomorphism. So, the certificate problem takes the reconstruction problem one step farther. In this chapter we will introduce some of the previously studied reconstruction problems.

The first reconstruction problem which we will present was solved by Whitney in 1932 [11]. He showed that given the line graph of a graph one can reconstruct the graph up to isomorphism provided the line-graph has minimum degree of at least four. This is one of the few solved reconstruction problems of which we know.

The most famous reconstruction problem - sometimes even referred to as "The Reconstruction Problem" - is the vertex reconstruction problem. In this problem the objective is to reconstruct a graph from its induced subgraphs. Though it is widely believed this is possible for any graph on at least three vertices this remains an open question for all but a few classes of graphs and we overview the wide range of work on this problem in section 3.3.

A closely related problem is the edge reconstruction problem; instead of being given the induced subgraphs of the graph you are given the subgraphs created by deleting an edge. Work on this problem has been more successful, and for good reason as it has been shown that the vertex reconstruction conjecture implies the edge reconstruction conjecture. Section 3.4 is devoted to this problem.

Finally, we look briefly at the analogous conjecture for directed graphs and present a family of counter-examples to this conjecture found by Stockmeyer [17].

Before we begin to study these problems we would like to draw the reader's attention to the relationship between reconstruction problems and the graph isomorphism problem. Though determining whether two graphs are isomorphic is known to be in NP is not known whether this problem is NP-complete. Determining that this problem is not NP-complete is equivalent to showing that P $\neq$ NP and thus it is commonly believed that this problem falls a class of its own. The following theorem illustrates the importance of reconstruction problems to the isomorphism problem.

We say that a function f is *constant under isomorphism* if for every pair of graphs G, G' with $G \cong G'$, $f(G) = f(G')$. We say that a graph G is *reconstructible up to isomorphism from a function* f if for every graph H such that $f(H) = f(G), H \cong G$.

Theorem 3.1 *If a graph is reconstructible up to isomorphism from a function f which is computable in polytime, then isomorphism is computable in polytime.*

Proof Let G_1 and G_2 be two simple graphs. Compute function f for both of these graphs. If it is identical for both graphs then G_1 and G_2 must be isomorphic as they are reconstructible up to isomorphism from function f. And similarly if the value of function f is different for the graphs then they can not be isomorphic. Thus we can compute whether any two graphs are isomorphic in polytime. $\square$

3.2 Line-graph Reconstruction Problem

In this section we present Whitney's theorem as promised in the introduction.

Theorem 3.2 *Consider any two simple graphs G_1 and G_2 with minimum degrees are at least four. If their line-graphs are isomorphic then $G_1 \cong G_2$.*

Proof We show that given a simple graph G with minimum degree at least four and line-graph L we can always get a reconstruction of G which is isomorphic to G.

Construct the graph G' as follows. For every maximal clique in L with four or more points associate a point x_i of G'. Then join two points x_i and x_j in G' if their corresponding cliques in L share a point. We claim that G' is isomorphic to G.

First consider a point x_i in G. It has degree at least four and thus is the endpoint of at least four edges e_1, e_2, e_3, e_4. Each of these edges corresponds to a point in L and the corresponding four points in L form a clique by the definition of a line-graph.

So, next consider any clique in L of size at least four. Label the vertices in the clique $e_1, e_2, \ldots e_k, k \geq 4$ and let vertex x be the vertex in G which e_1 and e_2 have in common. Now if $e_3, \ldots, e_k$ also all share vertex x as an endpoint we are done as no other edge in G can have x as an endpoint or it would be part of this clique. Thus assume WLOG that e_3 does not have vertex x as an endpoint in G. However it is in a clique with e_1 and e_2 and thus must share an endpoint in G with each of these edges. Thus they must form a triangle in G. However we have e_4 which is adjacent to e_1, e_2 and e_3 in L and thus shares an endpoint with each of them in G. This requires e_4 to have three endpoints which can not be and thus we know all the vertices in our clique share an endpoint in G. □

The reader can easily see that both the star with three edges and the triangle have identical line-graphs, and yet are not isomorphic. It is also clear from the proof why the lower bound on the minimum degree is necessary.

3.3 Vertex Reconstruction Problem

It seems a very natural question to ask whether a graph on n vertices can be uniquely reconstructed -up to isomorphism - if you are given an unlabeled copy of each of the graph's induced subgraphs. It has been noticed however that providing only the induced subgraphs on $n-1$ vertices is an equivalent problem as all the other induced subgraphs can be recreated from this information.

We denote by G_v the graph $G - v$ with $v \in V(G)$ and refer to G_v as a *vertex-deleted subgraph* of G. A *reconstruction* of a graph G is a graph H such that $V(H) = V(G)$ and $H_v = G_v, \forall v \in V(G)$. A graph G is *reconstructible* if every reconstruction $H \cong G$.

Harary [3] introduced an interesting way of viewing the reconstruction problem. He envisioned a deck of cards with each card containing a picture of one of the vertex-deleted subgraphs. Then it became a game to attempt to reconstruct the one (or more) graph(s) that would yield this set of subgraphs. Kelly and Ulam [2] first conjectured that each deck of more than two cards which yields at least one graph would yield exactly one graph.

Conjecture 3.3 *All finite simple graphs with at least three vertices are reconstructible.*

We know that there is a graph on two vertices which is not reconstructible, namely K_2. The graph on two vertices with no edges is a reconstruction of K_2. Thus the conjecture has been formulated for graphs with three or more vertices. All graphs on less than nine vertices have been checked ([9], [7], [2], [12], [15]) and the conjecture still holds.

Further evidence to support this conjecture comes from Muller [2] who shows that the proportion of graphs on n vertices which are not reconstructible tends to 0 as n tends to ∞. However in the last half century no one has been successful in proving the conjecture.

There are two common tactics taken when approaching the problem. The most intuitive approach has the graph theorist trying to show specific classes of graphs are reconstructible in the hope that eventually all graphs will be shown to be reconstructible. Another approach focuses on functions which act on classes of graphs.

We define a function, f, on a class of graphs $\mathcal{G}$ as *reconstructible* if $\forall G \in \mathcal{G}$, $f(G) = f(G')$, for all G' which are reconstructions of G. The following lemma by Kelly [9] shows how knowing a function is reconstructible can be useful.

Lemma 3.4 *For any two graphs F and G such that $|V(F)| < |V(G)|$, the number $s(F, G)$ of subgraphs of G isomorphic to F is reconstructible.*

Proof If we consider a specific subgraph of G isomorphic to F we know that it occurs in all but $|V(F)|$ of the $|V(G)|$ vertex-deleted subgraphs of G. This is true for each of the occurrences of F in G. Thus $s(F,G) = \dfrac{\sum_{v\in V(G)} s(F,G_v)}{V(G)-V(F)}$, and so the number $s(F,G)$ can be calculated using information present in the set of vertex-deleted subgraphs of G and must be the same for all reconstructions of G. □

We can use this lemma to show that the number of edges is a reconstructible function; take $F = K_2$. We can also use it to show the degree sequence of a graph is reconstructible. If we take $F = K_2$ again and calculate $s(F,G) - s(F,G_v)$ for each vertex v in the graph we find the degree sequence. This fact assists us in determining that all regular graphs are reconstructible. If G is a regular graph then we can reconstruct its degree sequence and determine that it is in fact regular with degree δ. Then take anyone of its vertex-deleted subgraphs, add an extra vertex and connect that vertex to the δ vertices which have degree $\delta - 1$ and a graph which must be isomorphic to G has been constructed.

Following the same pattern as the previous proof, proofs of reconstructibility are often broken down into two steps. We call a class of graphs $\mathcal{G}$ *recognizable* provided every reconstruction of a graph $G \in \mathcal{G}$ is isomorphic to a graph in $\mathcal{G}$. Furthermore we call $\mathcal{G}$ *weakly reconstructible* if all reconstructions of a graph $G \in \mathcal{G}$ which are in $\mathcal{G}$ are isomorphic to G. We can immediately see that a class of graphs is reconstructible if and only if it is recognizable and weakly reconstructible. Thus, we often first prove that a class of graphs recognizable and then prove that it is weakly reconstructible. We shall use this technique in the following proofs.

The next theorem is useful in proving a number of classes are reconstructible. If $\mathcal{F}$ is a family of graphs closed under isomorphism then we call a subgraph $F \in \mathcal{F}$ of G, a *$\mathcal{F}$-subgraph* of G. A *maximal $\mathcal{F}$-subgraph* of G is one which is not contained in any other $\mathcal{F}$-subgraph of G. An (F,G)-*chain* is a sequence of subgraphs of G, $X_0, X_1, \ldots, X_n$ such that $X_0 = F$, $X_i \subset X_{i+1}, 0 \leq i \leq n-1$, and $X_n \subset G$. Two such chains are isomorphic if they have the same number of terms and corresponding subgraphs are isomorphic. Finally we refer to the *rank* of F in G as the number of terms in the longest (F, G)-chain. We are now ready to state the counting theorem due to Greenwell and Hemminger [6].

Theorem 3.5 *Let $\mathcal{G}$ be a recognizable class of graphs, and let $\mathcal{F}$ be any class of graphs such that, for every G in $\mathcal{G}$, each $\mathcal{F}$-subgraph of G is (i) vertex-proper; (ii) contained in a unique maximal $\mathcal{F}$-subgraph of G. Then, for every F in $\mathcal{F}$ and every G in $\mathcal{G}$, the number $m(F, G)$ of maximal $\mathcal{F}$-subgraphs of G isomorphic to F is reconstructible.*

Proof (As proved in [2].) We know by the first assumption that $|V(F)| < |V(G)|$. We will show by induction on the rank of F that the following sum holds.

$m(F, G) = \sum_{n=0}^{rankF} \sum (-1)^n s(F, X_1)s(X_1, X_2) \ldots s(X_{n-1}, X_n)s(X_n, G)$ where the inner sum is computed for every non-isomorphic (F, G)-chain of length n.

If rank $F = 0$ then every subgraph of F in G is a maximal $\mathcal{F}$-subgraph and $m(F, G) = s(F, G)$ which can be reconstructed.

Now assume the previous sum holds for all graphs $F' \in \mathcal{F}$ such that the rank of F' is less than r. Let rank $F = r$. Using the second assumption in the theorem we know that

$s(F, G) = \sum_X s(F, X)m(X, G)$

for all non-isomorphic subgraphs X of G. This yields that

$m(F, G) = s(F, G) - \sum_{X \neq F} s(F, X)m(X, G)$

As $s(F, X) = 0$ for any graph X with rank r or greater and thus applying the induction hypothesis to each term yields the desired sum.

Finally we know the left-hand side of the sum is weakly reconstructible and thus so is the right-hand side. □

It is easily shown that disconnected graphs and trees are recognizable classes of graphs. With applications of theorem 3.5 it can be shown - with varying degrees of difficulty - that these classes of graphs are also reconstructible.

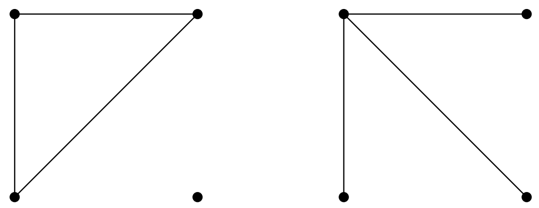

Figure 3.1: A pair of non-edge reconstructible graphs.

3.4 Edge Reconstruction Problem

For any edge $e \in E(G)$ we denote by G_e the subgraph obtained from the graph G by removing the edge e. We say a graph H is an *edge reconstruction* of G provided that there is a bijection $\Theta : E(H) \to E(G)$ such that $H_e \cong G_{\Theta(e)} \forall e \in E(H)$. A graph is *edge reconstructible* provided it is isomorphic to all its edge reconstructions.

We can see from the pair of graphs provided in figure 3.1 that the analogous edge reconstruction conjecture will not hold for all graphs on four vertices. However it is believed to be true for n at least 5.

Conjecture 3.6 *All simple finite graphs on five or more vertices are edge reconstructible.*

This conjecture is not as strong as the vertex reconstruction conjecture. It was proven by Hemminger [8] that a graph is edge reconstructible if and only if its line graph is reconstructible, and so the vertex reconstruction conjecture implies the edge reconstruction conjecture.

The analogous versions of Kelly's lemma and the counting theorem hold for edge reconstructibility. We have gotten much farther solving the edge reconstruction problem. In fact it was proven by Muller in 1977 [14] that all graphs with a sufficient number of edges are edge reconstructible. His proof follows from a theorem of Nash-Williams' [2] based on work of Lovász [10].

3.5 Reconstruction of Directed Graphs

With our interest in tournaments it is worthwhile considering the reconstruction problem for directed graphs. In 1976 Stockmeyer [17] was successful in finding an infinite family of tournaments which are not reconstructible.

First we define *odd(x)* to be the remainder when x is divided by the highest possible power of two. Then for $n \geq 1$ we define the tournament A_n on $p = 2^n$ points by $v_i \to v_j$ if and only if $odd(i - j) \equiv 1 \bmod 4$, for every $1 \leq i, j \leq p, i \neq j$.

Stockmeyer then shows that the following properties hold:

- A_n is self-complementary.
- The first 2^{n-1} points of A_n each have score 2^{n-1} and the remaining 2^{n-1} points have score $2^{n-1} - 1$.
- A_n has only the identity automorphism.

Next he shows that the tournaments resulting from the deletion of either vertex v_k or v_{p+1-k} are isomorphic. Finally he builds two new tournaments from A_n. B_n is the tournament resulting from taking a copy of A_n and adding one vertex which beats all the even numbered vertices in A_n and is beaten by all the odd numbered vertices. Similarly C_n is built from a copy of A_n by adding a vertex which beats the odd-numbered vertices and is beaten by the even numbered vertices. It can be shown using the previously stated properties that B_n and C_n are not isomorphic yet have identical decks. Thus we know that for any value n_0 a counter-example to the desired conjecture of size $2^n > n_0$ can always be found .

Chapter 4

Preliminary Results

4.1 Introduction

This chapter provides the reader with an introduction into the issues that we considered when the work on this thesis first began. When we first investigated this topic we began with simple graphs and considered such issues as what their best P.I.C. was, could this P.I.C. be improved upon using a N.I.C. or a M.I.C., could we find a simple graph which needed all its edges in its P.I.C., etc. Though the bulk of the work in this thesis was done on complete trees and grids and is presented in the two following chapters we are opening with the more basic results.

This chapter begins by introducing the reader the wheel. The results proved on this simple graph will assist the reader in becoming familiar with the techniques used in this thesis. At the end of the chapter we find a P.I.C. for $K_{n,n}$ which is valid and optimal. And sandwiched between these two specific results is a short section discussing graphs which attain the maximal minimum I.C. number for a given n.

4.2 Wheels

Let us begin by defining the wheel. The *wheel* on $n+1$ vertices, W_n, $n \geq 3$, is a cycle on n vertices with a universal vertex. We shall refer to this universal vertex as the *center* of the wheel. The wheel on 7 vertices is depicted in figure 4.1.

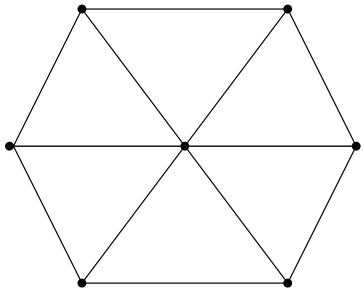

Figure 4.1: W_6

As can be seen immediately from the definition of the wheel, cycles are of utmost importance in understanding the wheel. Thus we shall begin with two theorems on the P.I.C. number of the cycle.

Theorem 4.1 *The P.I.C. number for the cycle, C_n, on n vertices, $n \geq 3$, is $n-2$.*

Proof Consider a graph D which has two components. One is a path on $n-1$ vertices and one is an isolated vertex. Then $C_n \succeq D$. Consider a graph $G' \cong C_n$ such that $G' \succeq D$. Then every vertex in G' has degree two. However in D one vertex has degree zero, two vertices have degree one, and all the other vertices have degree two. Thus we may immediately uniquely identify every edge in G' and $G' = C_n$. Thus D is a valid P.I.C. for C_n and D has size $n-2$.

Now assume we have a P.I.C., D, of size smaller then $n-2$. Then D has at least three components, so we either have two or more isolated vertices and violate the isolated vertex lemma, or we have an isolated edge and violate the isolated edge lemma, or we have an isolated path and violate the isolated path lemma. Regardless, D can not be a valid P.I.C. Thus we know C_n has P.I.C. number exactly $n-2$. □

In fact we can take our discussion of P.I.C.s for cycles even farther.

Theorem 4.2 *The only valid P.I.C.s for C_n are supergraphs of a path on $n-1$ vertices and an isolated vertex.*

Proof Let D be a valid P.I.C. for C_n. Assume D does not contain the path on $n-2$ vertices as a subgraph. Now D must have at least two components. If we

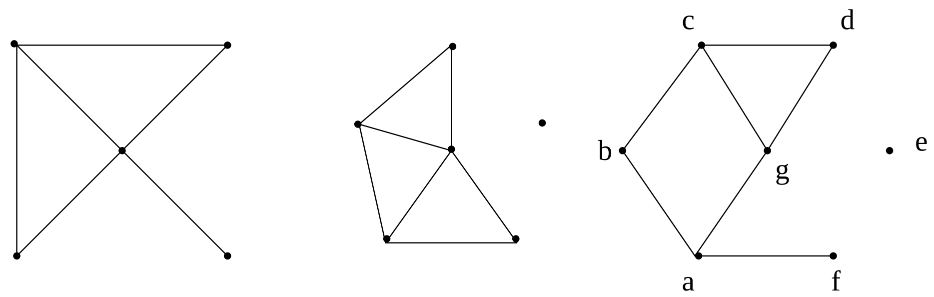

Figure 4.2: U_5, U_6, U_7

have two isolated vertices in D, then D violates the isolated vertex lemma. If D contains an isolated edge, then D violates the isolated edge lemma. And if D has neither of these properties then it must contain an isolated path which violates the isolated path lemma. Thus D is not a valid P.I.C. $\square$

We are now ready to begin our discussion on wheels.

Theorem 4.3 *The P.I.C. number for the wheel on n vertices, $n \geq 4$, is:*

$n = 4$	*0*
$5 \leq n \leq 6$	$n + 1$
$n = 7$	n
$n \geq 8$	$n - 1$

Claim 4.4 *The P.I.C. number for W_3 is zero.*

Proof $W_3 \cong K_4$. As all graphs isomorphic to K_4 are identical, that is all isomorphisms of K_4 are automorphisms, we know that an empty graph will be a P.I.C. for W_3. And as the empty graph can not be improved upon we know this it is the best possible P.I.C. and our P.I.C. number is zero. $\square$

Claim 4.5 *The P.I.C. number for $W_{n-1}, 5 \leq n \leq 6$, is at most $n+1$.*

Proof We shall prove this by showing that the graphs U_5 and U_6 depicted in figure 4.2 are valid P.I.C.s for the corresponding wheels. For $n = 5$ or 6 let U_n be

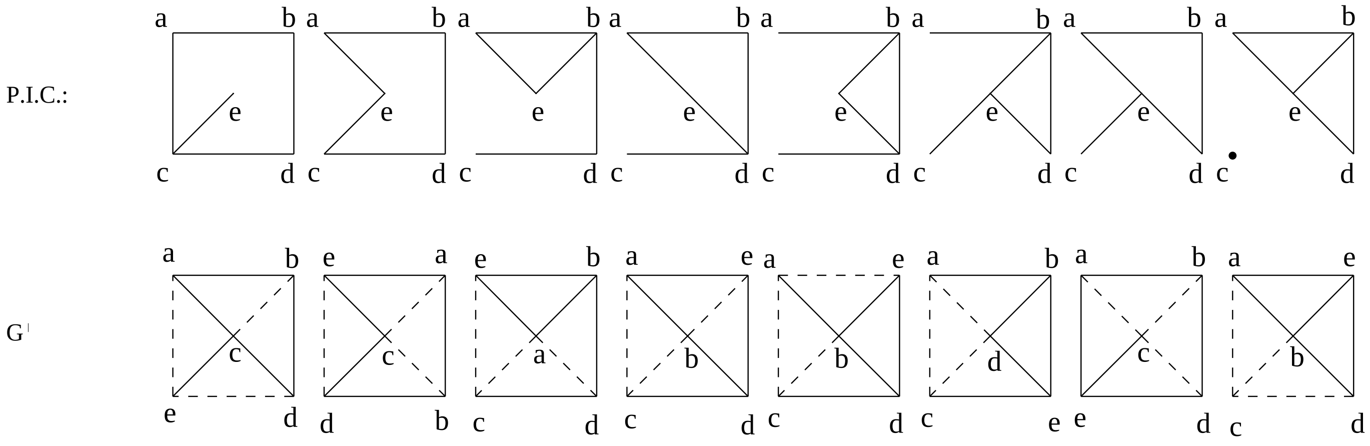

Figure 4.3: Potential P.I.C.s for W_4

the graph on n vertices, $v_1, v_2, \ldots, v_n$, such that $v_1, v_2, \ldots, v_{n-2}$ form a path, and v_n is adjacent to v_1, v_2, v_3, v_4.

Assume we have a graph $G' \cong W_{n-1}$ which contains U_n . As G' is a wheel it must have a center. As v_n has degree four in U_n it must also have degree at least four in G'. As the only vertex in G' with degree four or higher is the center we know v_n is the center of G'. As every other vertex in the graph is adjacent to v_n we may now apply lemma 1.9. Since a path on $n-2$ vertices is a valid P.I.C. for a cycle on $n-1$ vertices, we have that U_n is a valid P.I.C. for W_{n-1}. $\square$

Claim 4.6 *The P.I.C. number for W_{n-1}, $5 \leq n \leq 6$, is $n+1$.*

Proof By applying lemma 1.9 from our introductory chapter we know that any valid P.I.C. for W_{n-1} must contain a valid P.I.C. for the cycle on $n-1$ vertices. From theorem 4.2 we know that includes a path on $n-3$ vertices. In figure 4.3 (4.4) we have shown all subgraphs of W_{n-1} which have 5 (6) edges and contain a path on 2 (3) vertices. For each of these potential P.I.C.s we also show a graph $G' \cong W_{n-1}$ which contains the potential P.I.C. and is not identical to W_{n-1}. Solid lines in G' indicate edges also present in the P.I.C. and are drawn in such a way as to assist the reader. In some cases we have provided a P.I.C. twice. This is done to make it easier for the reader to see we have covered every possible case. If we have a valid P.I.C. for W_{n-1} of size smaller then $n+1$ it must be contained in one of these graphs. As none

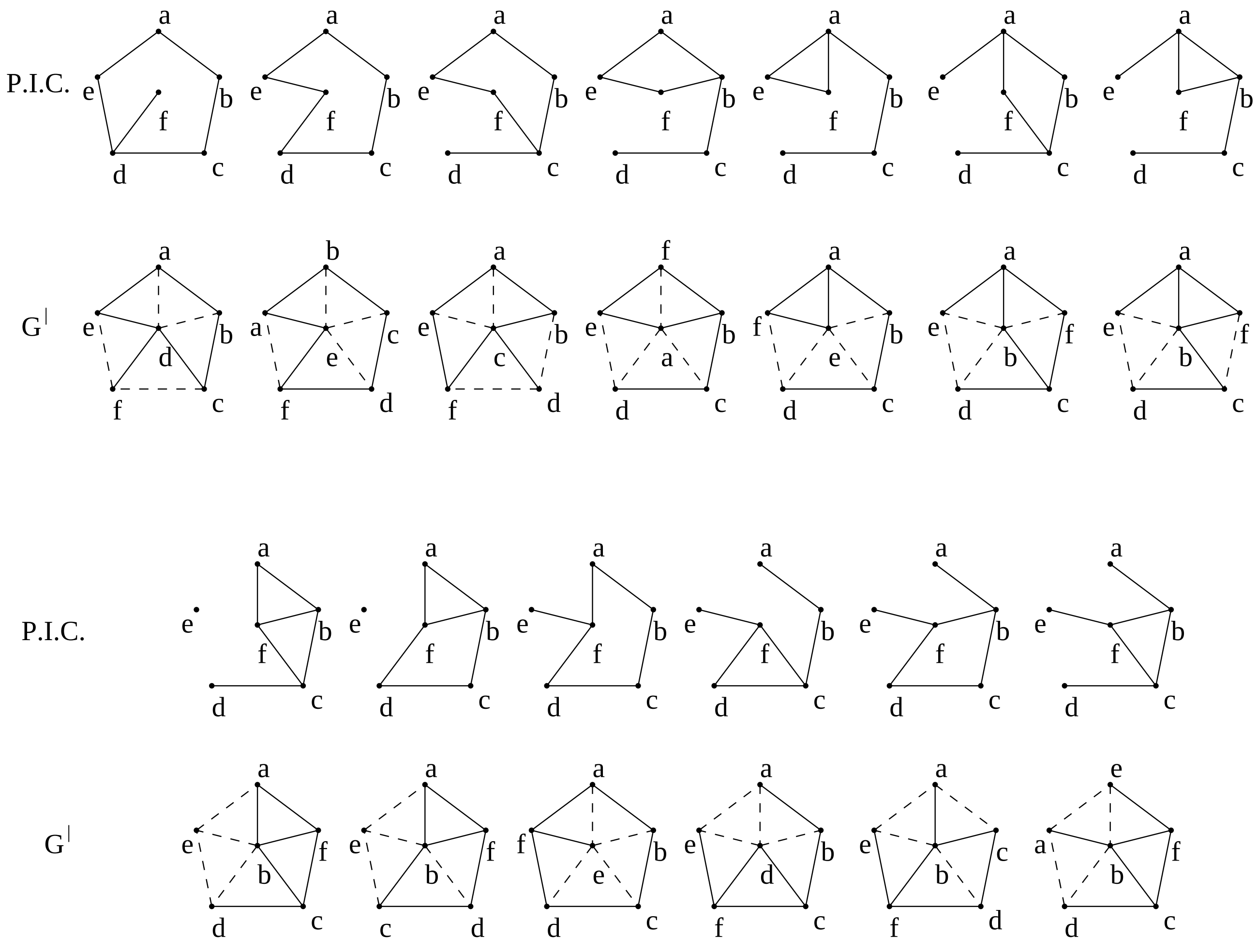

Figure 4.4: Potential P.I.C.s for W_5

of them are valid P.I.C.s however we know that no P.I.C. of size smaller then $n+1$ exists by the certificate hierarchy lemma. $\square$

Claim 4.7 *The P.I.C. number for W_6 is at most 7.*

Proof We shall prove this by showing that the graph U_7 as depicted in figure 4.2 is a valid P.I.C. for W_6. Consider any graph $G' \cong W_6$ such that $G' \succeq D$. Then G' has a center. We know that any vertex of degree three in D must either be the center or be adjacent to the center. The vertices a, c, d and g all have degree three in D and there is no vertex adjacent to each of them. Thus one of these vertices is the center vertex. The only vertex among them which is adjacent to the other three is g and thus it is the center of G'. As we have a path on five of the other vertices which is a valid P.I.C. for the cycle on six vertices by applying lemma 1.10 and theorem 4.2 we know D is a valid P.I.C. for W_6. $\square$

Claim 4.8 *The P.I.C. number for the W_6 is 7.*

Applying lemma 1.9 from chapter 1 and theorem 4.2 again allows us to limit the number of possible P.I.C.s of size smaller then seven. Please refer to figure 4.5 and apply the argument made in claim 4.6. $\square$

Claim 4.9 *The P.I.C. number for W_{n-1}, $n \geq 8$ is at most $n-1$.*

Proof We shall prove this by showing that the graph depicted in figure 4.6 is a valid P.I.C. for the W_{n-1}, $n \geq 8$. Let U_n be the graph on n vertices, $v_1, v_2, \ldots, v_n$ such that $v_1, v_2, \ldots, v_{n-2}$ form a path, and v_n is adjacent to v_2, and v_{n-3}.

Assume we have a graph $G' \cong W_{n-1}$ which contains U_n . We know G' has a center and any vertex in G' which has degree three or more in U_n must either be the center of G' or be adjacent to the center of G'. Both v_2 and v_{n-3} have degree three in U_n and the only common vertex in their neighbourhoods is v_n. Thus v_n must be the center

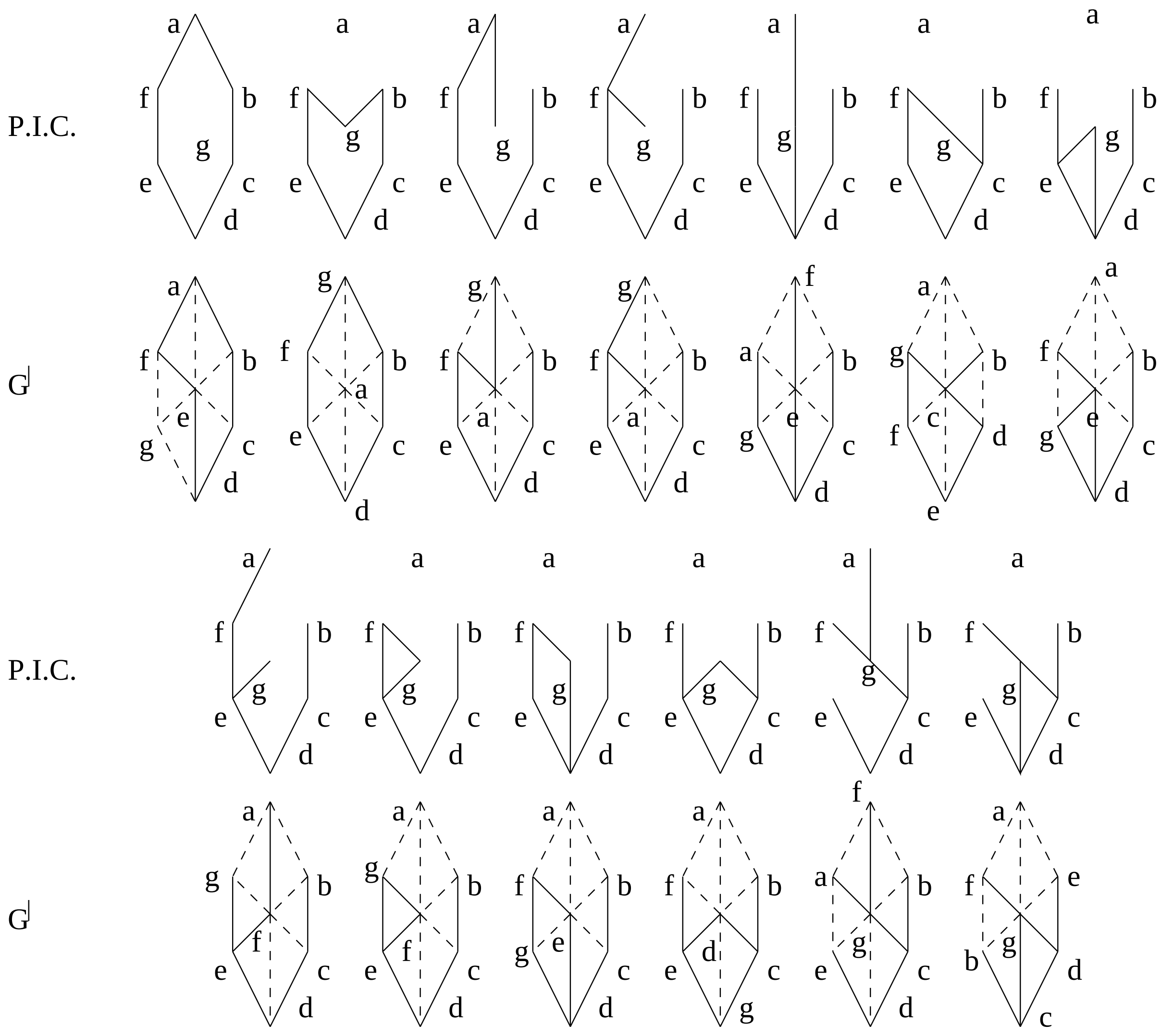

Figure 4.5: Potential P.I.C.s for W_6

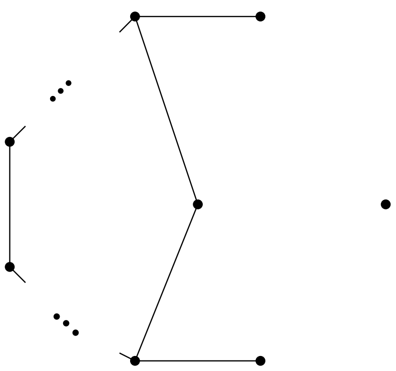

Figure 4.6: U_n

of G'. And by the same reasoning as employed in claim 4.6 we may determine the exact position of the other vertices. Thus U_n is a valid P.I.C. for W_{n-1}. $\square$

Claim 4.10 *The P.I.C. number for W_{n-1}, $n \geq 8$ is $n-1$.*

Proof By an application of lemma 1.9 from chapter 1 and theorem 4.2 we know we need a path on $n-2$ vertices in our P.I.C. It is clear this is not a valid P.I.C. by itself as we have two isolated vertices which are not interchangeable. Thus we must have at least $n-2$ edges in any valid P.I.C. Assume our P.I.C. is a path on $n-1$ vertices. Then any of the vertices may be the center vertex and this is not valid. The only other P.I.C. which has $n-2$ edges and contains a path on $n-2$ vertices is the graph with a path on $n-2$ vertices and an edge from the center to one of the other vertices. If this edge is to the vertex not in the path we have an isolated edge. If this edge is to one of the endpoints of the path we have a path on $n-1$ vertices and the isolated vertex could be the center of the wheel. And finally, if the edge is from the center to one of the internal vertices on the path that vertex could be the center of the graph. Thus we can not have a P.I.C. of size $n-2$. $\square$

Wheels also allow us our first glimpse of the added power that non-edges provide to our certificates. In the following theorem we show that the certificate numbers for wheels may in fact be improved upon if we allow non-edges.

For W_{n-1}, $n \geq 6$ proving the exact value of the I.C. number seems very challenging and has not be done.

Theorem 4.11 *For $n \geq 4$, the I.C. number for W_{n-1} is at most:*

$n=4$	*0*
$n=5$	*2*
$n=6$	*4*
$n \geq 7$	$n-1$

Proof For $n=4$ we achieved this I.C. number in the previous theorem. As it can not be improved upon we know it is the I.C. number. I.C.s for W_{n-1}, $n=5$,

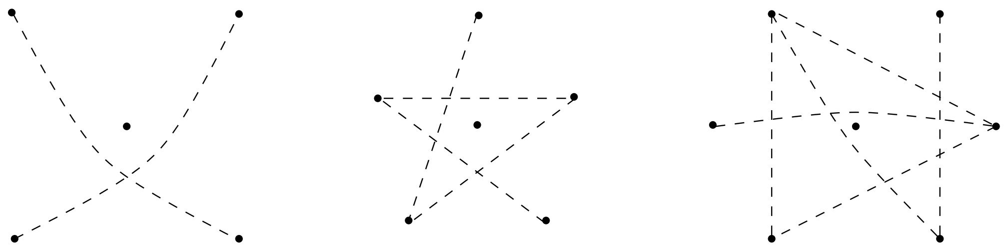

Figure 4.7: Y_5, Y_6, Y_7

6 and 7, are depicted in figure 4.7. The dashed lines denote non-edges. The reader should notice that they are all N.I.C.s.

Claim 4.12 *The I.C. number for W_4 is 2.*

Proof First note that $\overline{W_4}$ has only two edges. As this is a valid N.I.C. for W_4 we know the I.C. number is at most 2. However the isolated edge lemma forbids a single edge from being a valid I.C. and similar reasoning forbids a single non-edge from being a valid I.C. and thus we know the I.C. number is exactly 2. □

Claim 4.13 *The I.C. number for W_5 is at most 4.*

Proof We shall show that Y_6 as shown in figure 4.7 is a valid I.C. for W_5. Let Y_6 be the graph on 6 vertices, $v_1, v_2, \ldots v_6$ which has a non-edge path v_1, v_3, v_5, v_2, v_4. This path has length four.

Assume we have a graph G' isomorphic to W_5 which contains Y_6. Then G' has a center and it must be v_6 as all the other vertices are the endpoint of a non-edge in Y_6. As v_2 has non-edges to v_4 and v_5 we know it is adjacent to v_1 and v_3. Similarly v_3 is also adjacent to v_4 and v_4 is also adjacent to v_5. As all vertices now have degree three or more except v_1, v_2 and v_5 is is easily seen v_1 is adjacent to v_2 and v_5. As $G' = W_5$ we know Y_6 is a valid I.C. for W_5 and thus W_5 has an I.C. number of at most 4. □

Claim 4.14 *The I.C. number for W_6 is at least 6.*

Proof We shall show that Y_7 as shown in figure 4.7 is a valid I.C. for W_6. Let Y_7 be the graph on 7 vertices, $v_1, v_2, \ldots, v_7$ which has non-edges (v_1, v_3), (v_1, v_4), (v_1, v_5), (v_2, v_5), (v_3, v_5), (v_4, v_6). This graph has size 6.

Assume we have a graph G' isomorphic to W_6 which contains Y_7. Then G' has a center and it must be v_7 as all the other vertices are the endpoint of a non-edge in Y_7. As v_1 has non-edges to v_3, v_4, and v_5 we know it must be adjacent to v_1 and v_6. Similarly v_5 must be adjacent to v_4 and v_6. This means that v_3 must be adjacent to v_2 and v_4 and we know $G' = W_6$. Thus Y_7 is a valid I.C. for W_6. $\Box$

Claim 4.15 *The I.C. number for W_{n-1} is at least $n-1$ for $n \geq 8$.*

This follows immediately from the claim 4.10 of theorem 4.3. $\Box$

4.3 Maximal Minimum I.C.s

Having realized that non-edges give us extra power we began to look at worst case scenarios. We began again with P.I.C.s. For a given n we are interested in the size of the largest minimum P.I.C. We discovered that this number may be as high as $\binom{n}{2} - 1$. This is worst possible since if a graph has a P.I.C. of $\binom{n}{2}$ then it must have $\binom{n}{2}$ edges, i.e. it must be K_n which we can easily see has a P.I.C. number of zero.

Theorem 4.16 *For $n \geq 4$, there exists a graph whose P.I.C. number is $\frac{(n)(n-1)}{2} - 1$.*

Proof Denote the graph on n vertices which has every possible edge but one as H_n. As any graph is a P.I.C. for itself we know that the P.I.C. number for H_n is at most $\frac{(n)(n-1)}{2} - 1$. We shall show that this can not be improved on.

Consider any P.I.C., C, for H_n of smaller size. It is missing at least one edge from H_n. Thus it is contained in a graph C' which is missing exactly one edge from H_n. As C is a P.I.C. for H_n so is C' by the certificate hierarchy lemma. Now suppose H_n

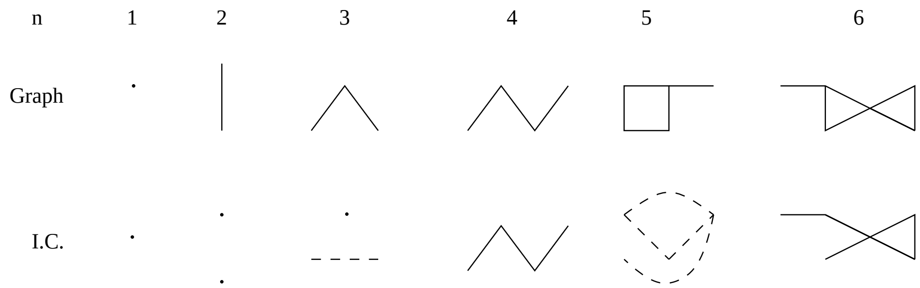

Figure 4.8: Graphs with maximal I.C. numbers

is missing the edge between vertices v_i and v_j. Suppose C' is also missing the edge (v_k, v_l) where at least one of k and l is distinct from i and j. Then the graph created from adding edge (v_i, v_j) to C' is isomorphic to G' but not identical to it and C' is not a valid P.I.C. for H_n. Thus C is not a valid P.I.C. for H_n either. $\square$

However we can see easily that the I.C. number for this same graph is one, as we must only provide one non-edge in our I.C. This leads us to the question of what is the worst possible I.C. number for a given n. For any graph G either G or $\overline{G}$ has an edge set of size less than or equal to $\lfloor \frac{(n)(n-1)}{4} \rfloor$. Thus we know this is an upper bound on the I.C. number for any graph with n vertices. We were not able to improve on this bound for general n. However we did examine the case for small n and noticed even in these cases this bound was not always tight.

Theorem 4.17 *The following are the maximum I.C. numbers for a graph on n vertices:*

n	1	2	3	4	5	6
Max I.C. #	0	0	1	3	4	6

Figure 4.8 gives a graph for each n which achieves this maximum.

The cases when $n = 1$ or 2 are trivial. For $n = 3$ or 4 we have achieved the maximum mentioned in the above paragraph so we know this value can not be improved upon. For $n = 5$ ($n = 6$) the only possible graphs which have a larger minimum I.C. number have 5 (resp. 7) edges. If they have more edges than 5 (resp. 7) the complement of the graph is a N.I.C. of size at most 4 (resp. 6) and if they have fewer edges than 5 (resp. 7) the graph itself is a P.I.C. of size at most 4 (resp. 6). This leaves a small number of graphs to be checked by hand and this is left to the reader.

4.4 P.I.C. for $K_{n,n}$

And finally, we discovered the P.I.C. number for $K_{n,n}$.

Theorem 4.18 *The P.I.C. number for $K_{n,n}$ is n.*

The proof follows from the following two claims.

Claim 4.19 *The P.I.C. number for $K_{n,n}$ is at most n.*

Proof If we label the sides of $K_{n,n}$, $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$ let D be the spanning subgraph with edges $(a_i, b_1), 1 \leq i \leq n$. Then $K_{n,n} \succeq D$.

Let $G' \cong K_{n,n}$ and $G' \succeq D$. Then a_i and a_j are both adjacent to b_1 for any $i, j, 1 \leq i, j \leq n$ in D and thus in G'. As G' is isomorphic to $K_{n,n}$ we know a_i has degree n in G' and it can not be adjacent to a_j for $1 \leq i, j \leq n$. Thus it is adjacent to $b_k, 1 \leq k \leq n$. Thus we have identified every edge in G' and it is identical to $K_{n,n}$. Thus D is a valid P.I.C. for $K_{n,n}$. □

Claim 4.20 *The P.I.C. number for $K_{n,n}$ is at least n.*

Assume we have a P.I.C. , D, for $K_{n,n}$ of size smaller then n. Assume a similar labeling of the vertices as was given in the previous proof. Now any edge in D has a_i, for some $i, 1 \leq i \leq n$ as an endpoint. As we have at most $n - 1$ edges in D there is at least one vertex in D, say $a_i, 1 \leq i \leq n$, which is isolated. Similarly there exists at least one vertex $b_j, 1 \leq j \leq n$, which is isolated. As these two vertices are not interchangeable we know by the isolated vertex lemma that D is not a valid P.I.C. for $K_{n,n}$. □

Chapter 5

Trees

5.1 Introduction

This chapter focuses on optimal P.I.C.s for complete binary trees and complete k-ary trees. Proofs that the P.I.C.s we found are valid and optimal provide the bulk of this chapter with some discussion of N.I.C.s and M.I.C.s for complete trees.

All trees in this chapter are rooted and we say that the root is at level zero of the tree. All vertices adjacent to the root are said to be at level one and for an arbitrary $j > 1$ the vertices at level j are those vertices adjacent to a vertex at level $j-1$ which has not already been assigned to level $j-2$. We say that a binary tree which has no vertices at level higher than i is a binary tree with i levels and which has depth i. Throughout this chapter we shall denote the complete binary tree with i levels by T_i. Similarly we shall denote the complete k-ary tree on i levels by T_i^k. A portrayal of T_i is given in figure 5.1 and the reader can view T_5^3 in figure 5.2. We say that an edge in T_i or T_i^k is at level i provided that it has an endpoint at level i and an endpoint at level $i-1$. For an edge at level i we refer to its endpoint which is at level $i-1$ as the parent vertex of the edge and the endpoint which is at level i as the child vertex of the edge. For a vertex at level i we refer to the vertices at level $i+1$ which are adjacent to that vertex as its children.

We shall begin by briefly examining the P.I.C.s for complete binary trees and complete k-ary trees when i is small. Then follow two sections which prove the P.I.C.

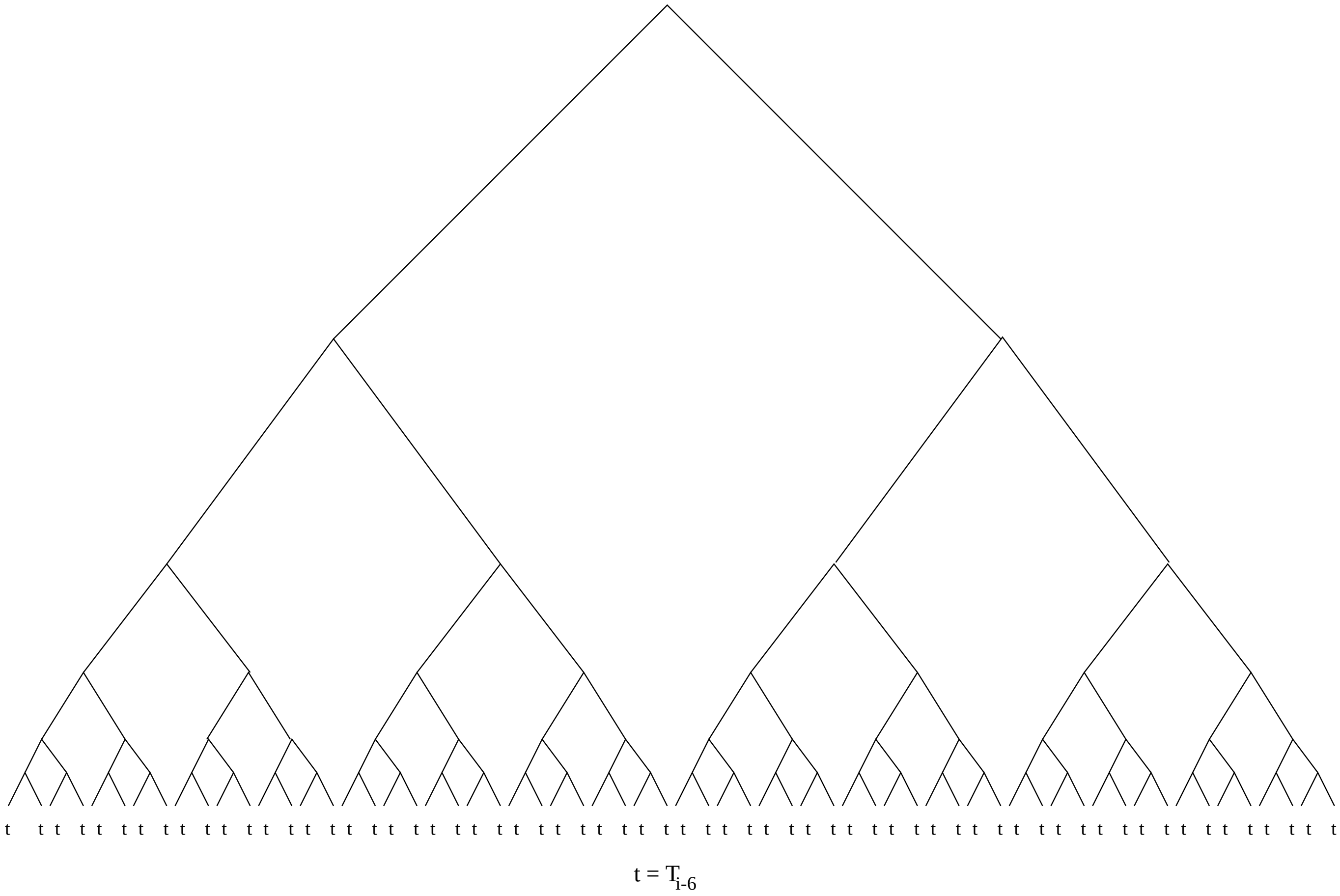

Figure 5.1: T_i

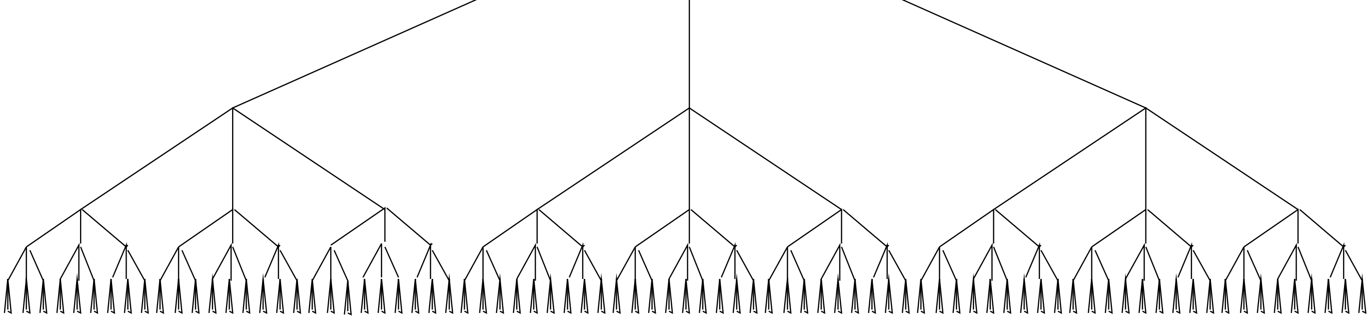

Figure 5.2: T_5^3

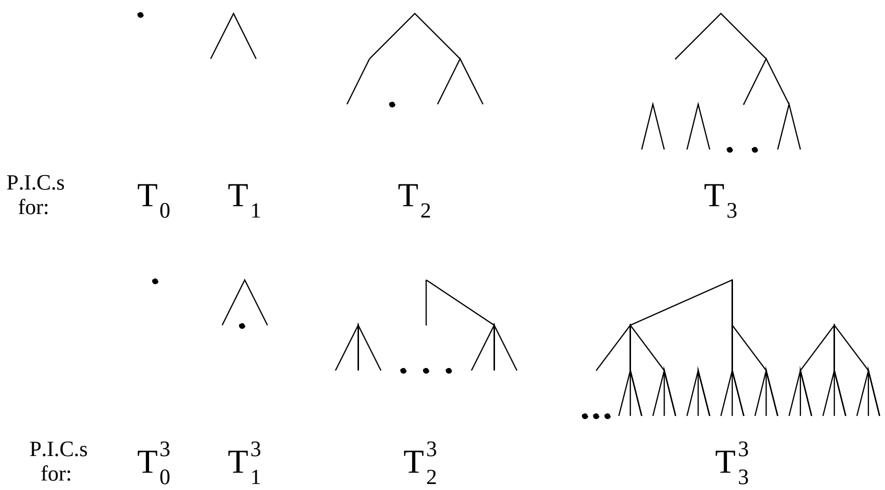

Figure 5.3: P.I.C.s for k-ary trees, $k = 2, 3$.

number for arbitrary T_i and T_i^K. The next short section briefly mentions the complete 3-regular binary tree and finally we complete the chapter with a discussion on the use of non-edges in I.C.s for complete binary and k-ary trees.

5.2 Small binary and k-ary trees.

The large theorems given in this chapter are applicable for any $i \geq 4$. In figure 5.3, we provide P.I.C.s for $T_0, T_1, T_2, T_3, T_0^3, T_1^3, T_2^3$ and T_3^3. The P.I.C.s for arbitrary k are similar. Though the following brief discussion of each of these P.I.C.s provides a justification for why we believe it is valid and optimal we have not provided a proof of this fact.

Here is a list of the minimum P.I.C. numbers for small i. The values of i are across the top of the table and the values for k along the side.

	0	1	2	3
2	0	2	4	10
3	0	2	8	10
$k \geq 4$	0	2	$\sum_{j=1}^{i} k^j - 2k + 2$	$\sum_{j=1}^{i} k^j - 3k + 1$

T_0^k has a P.I.C. of zero as these graphs are always a single vertex and thus their own P.I.C.s.

As T_1 is a path on three vertices we need all of its edges in our P.I.C. in order not to violate the isolated edge lemma.

T_1^3 is a star with three rays. Two of the edges in the graph suffice to determine the center of the star so we know it has a P.I.C. of size two or smaller. However if the P.I.C. only has one edge this violates the isolated edge lemma so we believe the P.I.C. number given is optimal.

For $T_1^k, k \geq 4$ we know that the only vertex with degree greater then one is the root. Thus two edges are sufficient to identify the root which is all that we must do. And the isolated edge lemma forbids a single edge from being a valid P.I.C. so we know this is optimal.

T_2 is the first slightly interesting P.I.C. we encounter. Removing either of its first level edges leaves a component with three dangling edges. Thus both of these edges must be present in our P.I.C. We can not remove two edges at level two with different parent vertices or we violate the isolated vertex lemma. However we may remove two edges at level two which have the same parent vertex so we have a minimum P.I.C. of size four.

T_3 has a P.I.C. of size ten. This P.I.C. is the essentially same as will appear in the proof of theorem 1.1 so we shall not discuss it further. Though the details of the proof require that $i \geq 4$ the reasoning is similar.

T_2^3 again requires some discussion. If we remove two of the edges from the first level the reader will see after a little thought that then only one edge may be removed from the second level. Thus the P.I.C. provided with one edge removed from the first level and three from the second is best possible.

T_3^3 has three edges removed per level except at the first level. Consider a P.I.C. which removes only edges from one level but removes more than three. Then the subtrees created violate the symmetric component lemma. Thus by the certificate hierarchy lemma we may not have a P.I.C. which is a subgraph of a graph of this type. Thus we know removing three edges per level is the best we can do. Consider the first level. Removing all three edges isolates the root vertex which will interfere with our desire to isolate three of the leaves. Removing two of the edges at the first level means we can not remove three edges at the second level without isolating either an edge or a non-leaf vertex. Thus we believe - but have not proven - the P.I.C. shown is the best possible for T_3^3.

$T_2^k, k \geq 4$ may have at most k edges removed from the second level. We may remove at most $k-2$ edges from the first level or we isolate a vertex or an edge. Thus the certificate shown is best possible.

$T_3^k, k \geq 4$ may have at most k edges removed at the second or third level. We may remove at most $k-1$ edges from the first level or else we isolate the root vertex. Thus the P.I.C. number given is believed to be best possible.

5.3 P.I.C. number for T_i.

The following two theorems establish a minimum P.I.C. for complete binary trees. We prove our chosen P.I.C. is valid and optimal.

Theorem 5.1 *For $i \geq 4$ the P.I.C. number for T_i is at most $2^{i+1} - 2i$.*

Proof We shall prove this statement by proving that D, as depicted in figure 5.4, is a valid P.I.C. for T_i. The labels for vertices at low levels of D have been provided in the figure. Other vertices in the graph are labeled in the same manner. D is of the desired size and it is easily seen that $T_i \succeq D$. Thus it remains only for us to show that any graph T' isomorphic to T_i where $T_i \succeq D$ is identical to T_i.

Claim 5.2 *If $T' \succeq D$ and is isomorphic to T_i then $v_{2,0}$ is at level two of T'.*

Proof We know $T' \succeq D$, and T' is a rooted complete binary tree with i levels. As well vertex $v_{2,0}$ is adjacent to $v_{3,0}$, and $v_{3,0}$ is the root of a complete binary tree on $i-3$ levels, so vertex $v_{2,0}$ is at level two or less in T'. Consider the possibility that $v_{2,0}$ is the root of T'. Then it must have degree two or less. However $v_{2,0}$ has degree three in D so it must have degree at least three in T' and therefore can not be the root of T'. Consider next the possibility that $v_{2,0}$ is a vertex at the first level of

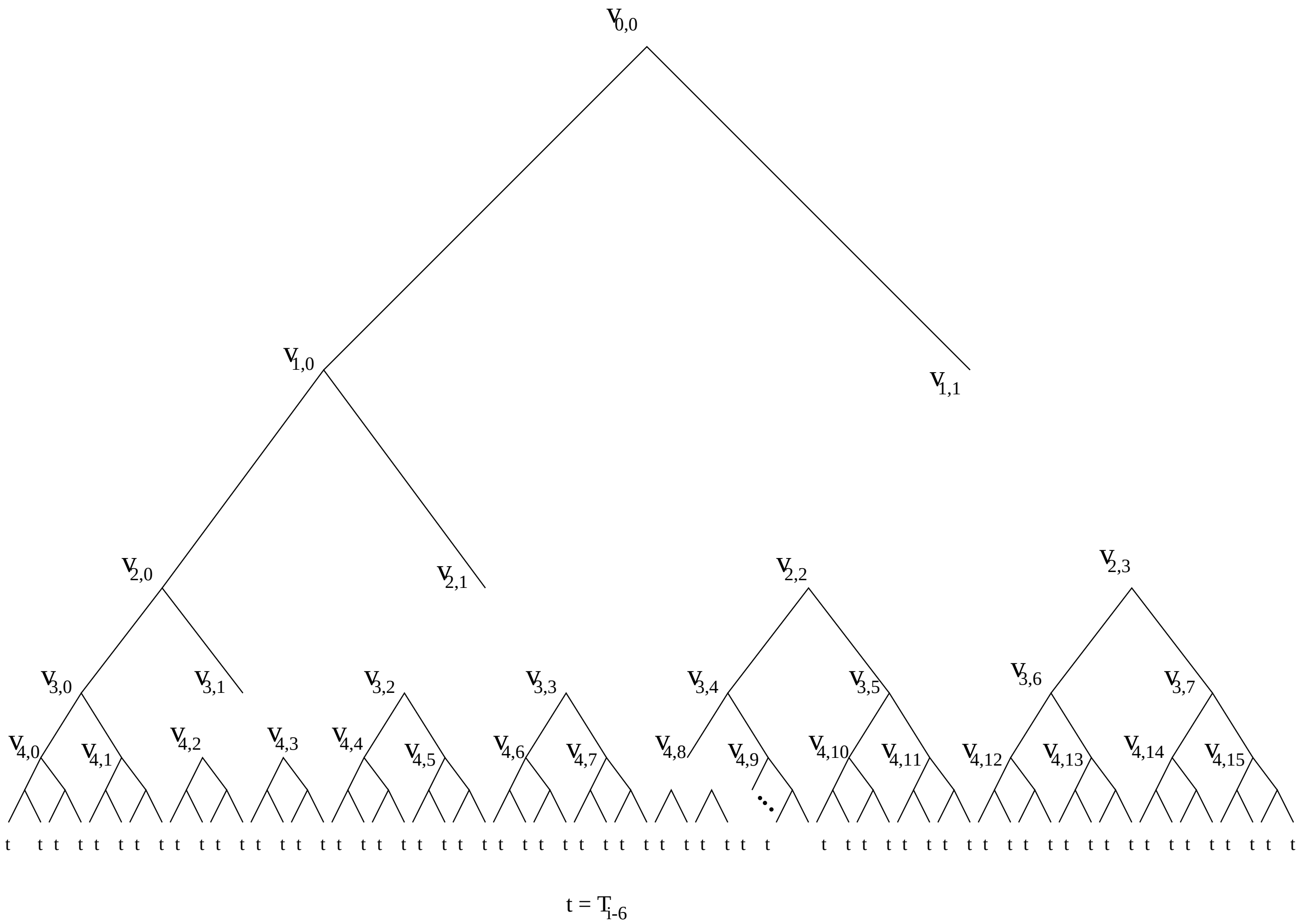

Figure 5.4: D - P.I.C. for T_i

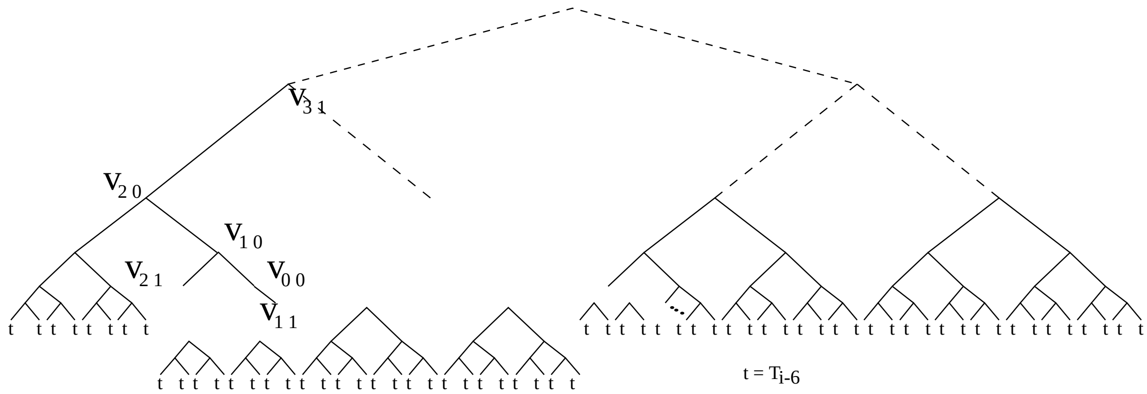

Figure 5.5:

T'. Then $v_{i,0}, v_{i,1}, \ldots, v_{i,2^{i-3}}$ must all be at level $i-1$ in T'. Thus we must have 2^{i-2} vertices which have degree zero in D and are adjacent to these vertices in T'. As we only have two such vertices we know that $v_{2,0}$ must be at level two of T'. □

Claim 5.3 *If $T' \succeq D$ and is isomorphic to T_i then $v_{3,1}$ is at level three of T' and $v_{1,0}$ is at level one of T'.*

Proof In the previous claim we proved $v_{2,0}$ must be at the second level of T' and as $v_{3,1}$ is adjacent to this vertex in D it must be at either the first or third level of T'. Consider the possibility that it is at the first level. Then vertex $v_{1,0}$ must be at the third level, vertices $v_{0,0}$ and $v_{2,1}$ at the fourth level, and vertex $v_{1,1}$ at the fifth level as shown in figure 5.5. Thus T' must contain three subtrees isomorphic to T_{i-4} and two subtrees isomorphic to T_{i-5} whose roots are adjacent to vertices $v_{0,0}, v_{2,1}$ and $v_{1,1}$. Every vertex in one of these subtrees must be in a component of D and in particular must be in a component of size smaller than or equal to the size of T_{i-4}. This is $2^{i-1}-5$ vertices. However we have only $2^{i-1}-2i+6$ such vertices in T'. As $i \geq 4$ we can not meet this requirement and thus we know that vertex $v_{3,1}$ must be at the third level of T'. Thus vertex $v_{1,0}$ must be at the first level of T'. □

Claim 5.4 *If $T' \succeq D$ and is isomorphic to T_i then $v_{0,0}$ is the root of T' and $v_{2,1}$ is at level two of T'.*

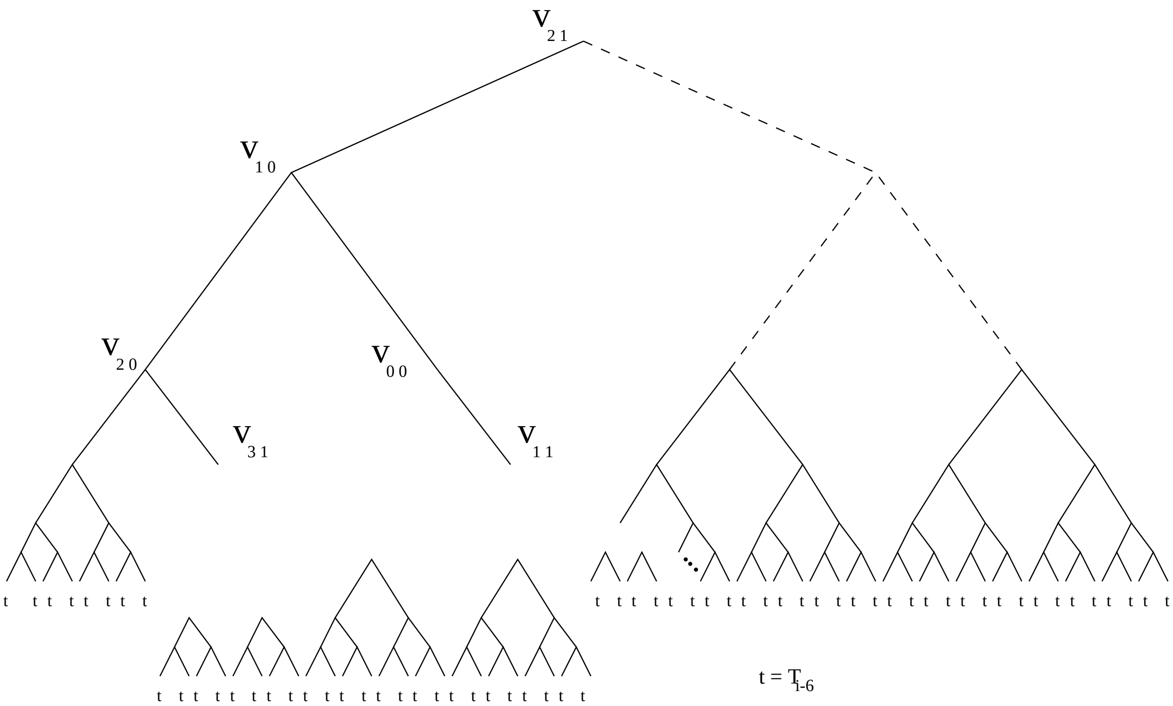

Figure 5.6:

Proof Drawing on the information in the two previous claims we know that vertex $v_{2,1}$ may be at either the zeroth or second level. Assume it is the root of T'. Then vertex $v_{0,0}$ is at the second level and vertex $v_{1,1}$ is at the third level as shown in figure 5.6. Thus we must have four subtrees of T' isomorphic to T_{i-3} to reconnect to vertices $v_{1,1}$ and $v_{3,1}$. This requires that D contain at least $2^i - 4$ vertices which are in components of size smaller or equal to the size of T_{i-3}. D contains only $2^i - 2i$ such vertices however. As $i \geq 4$ we will not have sufficient vertices to be descendants of $v_{3,1}$ and $v_{1,1}$. Thus we know that $v_{2,1}$ is at the second level of T' and hence $v_{0,0}$ is the root of T'. □

Claim 5.5 *If $T' \succeq D$ and is isomorphic to T_i then $v_{1,1}$ is at level one of T' and $v_{2,2}$ and $v_{2,3}$ are at level two of T' and adjacent to $v_{1,1}$.*

Proof As the previous claim shows $v_{0,0}$ is the root of T' $v_{1,1}$ must be at level one of T'. Now we know $v_{2,3}$ must be at the second level (or less) of T' as it is the root of a component of D isomorphic to T_{i-2}. It can not be at the zeroth or first levels as vertices $v_{0,0}, v_{1,0}$ and $v_{1,1}$ already occupy the three spots at these levels in T'. Thus it must be at the second level. Vertex $v_{1,0}$ already has degree three in D and thus can not be adjacent to any vertices other then its neighbours in D. Thus $v_{2,3}$ must be adjacent to $v_{1,1}$.

Next consider the vertex labeled $v_{2,2}$. As it is adjacent to $v_{3,5}$ which is the root of a complete binary tree with $i-3$ levels we know it must also be at level two or less in T'. By the same reasoning as given above it too must be adjacent to $v_{1,1}$. □

Claim 5.6 *If $T' \succeq D$ and is isomorphic to T_i then $v_{3,2}$ and $v_{3,3}$ are at level three of T' and adjacent to $v_{2,1}$; and $v_{4,2}$ and $v_{4,3}$ are at level four of T' and adjacent to $v_{3,1}$.*

Proof Now vertices $v_{3,2}$ and $v_{3,3}$ are the roots of components in D which are isomorphic to T_{i-2}. They must be at level three or less in T'. As the vertices at levels lower than three are already identified we know these vertices must be at level three. Thus they must be adjacent to a vertex, or vertices, at the second level. As all the vertices at the second level or less except $v_{2,1}$ have their neighbours determined we know they must be adjacent to $v_{2,1}$. By similar reasoning $v_{4,2}$ and $v_{4,3}$ must be at level four and adjacent to $v_{3,1}$. □

Now assume $T' \succeq D$ and is isomorphic to T_i. The previous series of claims identifies the label on every vertex in T' at level four or less. As well, except for $v_{4,8}$ we have identified the neighbours of each of the vertices at these levels. Now, we know that at level j, $5 \leq j \leq i-1$, the only vertex known to be at that level which does not have all its neighbours identified is vertex $v_{j,2^{j-1}+2^{j-3}-1}$. We also have two vertices $v_{k,2^{k-1}+2^{k-3}-3}$, and $v_{k,2^{k-1}+2^{k-3}-2}$ which are the roots of components of D isomorphic to T_{i-k} for $6 \leq k \leq i$. Thus we can see that vertices $v_{k,2^{k-1}+2^{k-3}-3}$ and $v_{k,2^{k-1}+2^{k-3}-2}$ must be adjacent to $v_{k-1,2^{k-2}+2^{k-4}-1}$ for $6 \leq k \leq i$. We have now identified the label on every vertex in T' and we can see that this tree is identical to T_i. Thus D is a valid P.I.C. for T_i. □

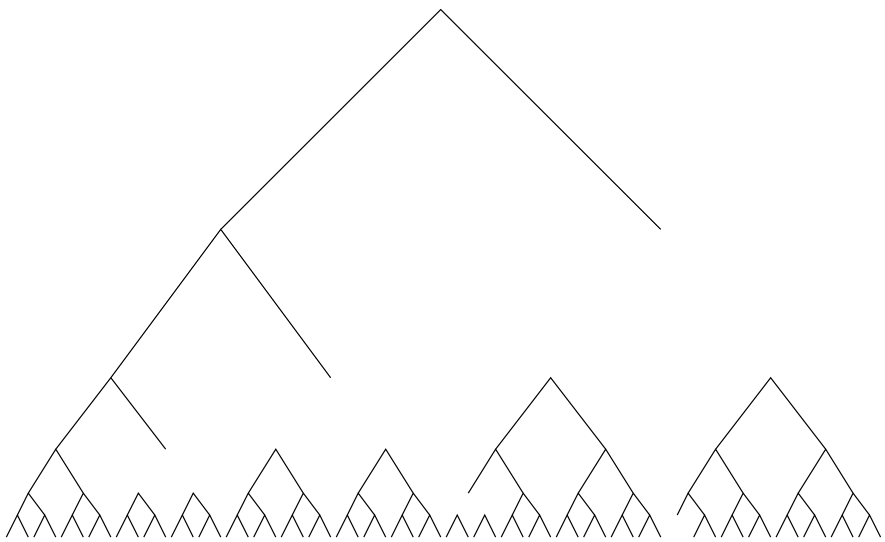

Figure 5.7: Alternate P.I.C. for T_6.

This P.I.C. is not unique as can be seen by figure 5.7, a valid P.I.C. for the complete binary tree with six levels which is not the one discussed in the previous proof.

Theorem 5.7 *The P.I.C. number of T_i is $2^{i+1} - 2i$ for $i \geq 4$.*

Proof We have already shown that this is the largest possible value for this number so all that remains is to show that T_i does not have a P.I.C. of smaller size. Noting that T_i contains $2^{i+1} - 2$ edges we know that this is equivalent to showing that any P.I.C. for T_i contains all but at most $2i - 2$ of the edges from T_i.

Claim 5.8 *Any P.I.C. for T_i is missing at most two edges per level and they share a parent vertex.*

Proof Assume we have a P.I.C. D which is missing three edges at level j $2 \leq j \leq i$. Then the P.I.C. D' which has D as a subgraph and is only missing the edges at level j must also be a P.I.C. by the certificate hierarchy lemma. As D' is missing three edges at level j we have two subtrees isomorphic to T_{i-j} whose roots are adjacent to non-interchangeable vertices. Thus by the isomorphic components lemma D' is not a valid P.I.C. This is a contradiction and thus we know any P.I.C. for T_i is missing at most two edges per level and they share a parent vertex. □

Claim 5.9 *Any P.I.C. for T_i is missing at most $2i-1$ edges.*

Proof We know from the previous claim any P.I.C. is missing at most $2i$ edges. Let us assume we have a P.I.C. D for T_i which is missing $2i$ edges. Then D must contain every edge in T_i except for two edges between each level. Furthermore the missing edges at each level must have a common parent vertex. Thus we know that the root of T_i is an isolated vertex in D as are two of the leaves of T_i. However the root of T_i and its leaves are not interchangeable and thus by the isolated vertex lemma D is not a valid P.I.C. Thus any P.I.C. for T_i is missing at most $2i-1$ edges and in particular must contain either an edge to every leaf, or at least one edge from the root of T_i. □

Claim 5.10 *Any P.I.C. for T_i missing more then $2i-2$ edges has an edge to the root of the tree.*

Proof Consider a P.I.C. which does not have an edge to the root of T_i. Then it must have an edge to every leaf in the tree by our previous discussion. As it may not remove more than two edges from any other level we know at most $2i-2$ edges have been removed from this P.I.C. □

Next let us consider a P.I.C. for T_i which has an edge to the root of the tree. As well, we have removed at most two edges per level. Thus we have removed at most $2i-1$ edges from T_i in creating our P.I.C. Now we have two vertices in T_i at level one, call them $v_{1,0}$ and $v_{1,1}$. One of the vertices can not be adjacent to any vertices at level two as we have removed two edges from level two and they must have a common parent vertex. WLOG call this vertex $v_{1,0}$. Now if $v_{1,0}$ is not adjacent to the root of T_i then it is an isolated vertex. Furthermore it is an isolated vertex which is not identical to any leaf of T_i. However we have two leaves of T_i which are isolated vertices in our P.I.C. and thus by the isolated vertex lemma this is not a valid P.I.C. However if $v_{1,0}$ is adjacent to the root of T_i then the root of T_i is not adjacent to any

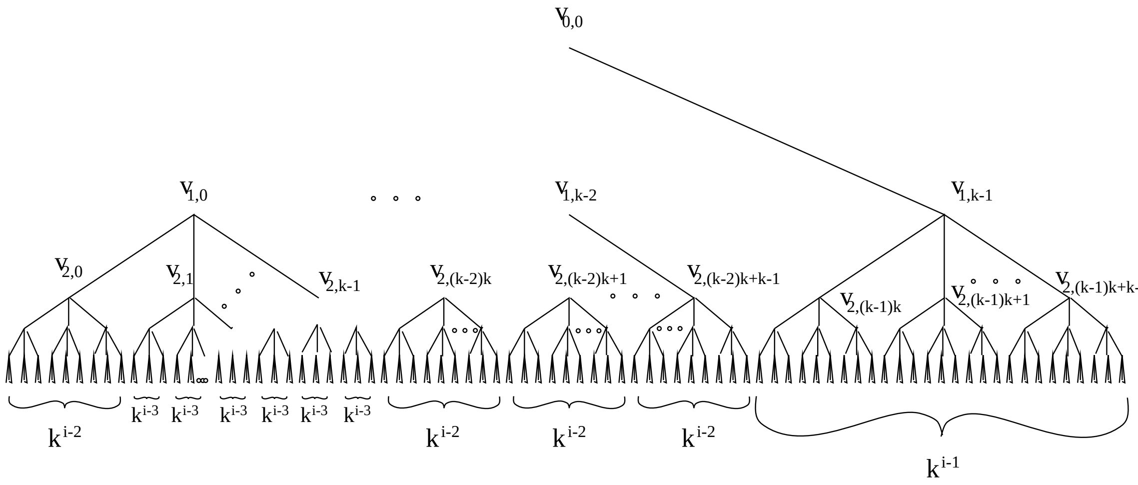

Figure 5.8: C - P.I.C. for T_i^k

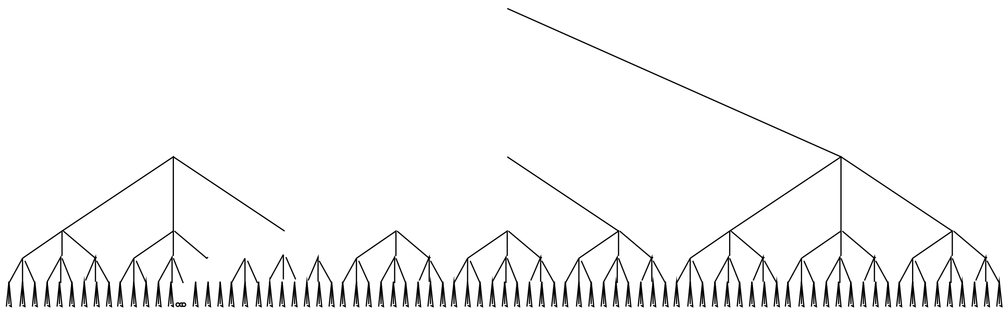

Figure 5.9: P.I.C. for T_5^3

other vertices. Thus we have an isolated edge whose endpoints are not identical with respect to T_i and the isolated edge lemma tells us this P.I.C. is invalid.

Thus any valid P.I.C. must remove at most $2i-2$ edges. Therefore T_i has a minimum P.I.C. of size $2^{i+1}-2i$ and its P.I.C. number is $2^{i+1}-2i$. $\square$

5.4 P.I.C. number for T_i^k.

Theorem 5.11 *For $i \geq 4, k \geq 3$ the P.I.C. number of T_i^k is no larger then $\sum_{j=1}^{i} k^j - ki - 2$.*

Proof We shall prove this theorem by showing that C as shown in figure 5.8 is a valid P.I.C. for T_i^k. Again only some of the labels have been provided. For ease in understanding this P.I.C. please see figure 5.9 as well. This contains a P.I.C. for T_5^3.

It is easily seen that $T_i^k \succeq C$ thus it remains only for us to show that any graph T' isomorphic to T_i^k where $T' \succeq C$ is identical to T_i^k.

Claim 5.12 *If $T' \succeq C$ and is isomorphic to T_i^k then $v_{0,0}$ is the root of T'.*

Proof Vertex $v_{0,0}$ is adjacent to vertex $v_{1,k-1}$ in C and thus in T' as well. $v_{1,k-1}$ is the root of a tree isomorphic to T_{i-1}^k and thus $v_{0,0}$ must be at the zeroth level of T'. $\square$

Claim 5.13 *If $T' \succeq C$ and is isomorphic to T_i^k then $v_{1,j}, 1 \leq j \leq k-1$ is at the first level of T'.*

Proof As $v_{1,k-2}$ is adjacent to $v_{0,0}$ in C the claim follows immediately for this vertex given the previous claim. Now the vertex labeled $v_{1,k-2}$ is adjacent to v_{2,k^2-k-1} in C and thus in T' to the root of a copy of T_{i-2}^k. Thus it must be at the zeroth or first level of T'. As $v_{0,0}$ is the root of T', $v_{1,k-2}$ must be at the first level and adjacent to $v_{0,0}$. As $v_{1,j}, 1 \leq j \leq k-3$ is the root of component of D isomorphic to T_{i-1}^k it too must be at the first level of T' and adjacent to $v_{0,0}$. $\square$

Claim 5.14 *If $T' \succeq D$ and is isomorphic to T_i^k then $v_{1,0}$ is at level one of T' and adjacent to $v_{0,0}$.*

Proof We know that the component of D in which $v_{1,0}$ is contained is a subgraph of the rooted tree T' and thus it has a root. As well this root can have degree no higher then two in D. Thus the root of this component must be the vertex labeled $v_{1,0}$ or one of the vertices labeled $v_{i,j}, 1 \leq j \leq k^{i-2} + 2k^{i-3} - k^{i-4} - 1$. If $v_{i,j}$ is the root of this component then the depth of this component is $2i - 2 > i$ as $i \geq 4$. Thus $v_{1,0}$ is the root of this component and the component has depth $i - 1$. So $v_{1,0}$ is at level one in T' and adjacent to $v_{0,0}$. $\square$

Thus we have identified every vertex at the first level of T'.

Now there is only one vertex at the first level of T' which does not have $k+1$ neighbours in T' identified. In fact it has only two of its neighbours identified and there are $k-2$ copies of T^k_{i-2} in C. Thus the roots of each of these components must be adjacent to this vertex. Similarly for levels two through $i-1$ there is exactly one vertex known to be at this level which does not already have degree $k+1$. As well there are k components of C which are isomorphic to $T^k_{i-j} 3 \leq j \leq i$. Thus the roots of these components must be adjacent to the vertex at level $j+1$ which has degree less than $k+1$. We have uniquely identified every vertex in T' and can see that is identical to T^k_i and thus C is a valid P.I.C. for T^k_i. $\square$

Theorem 5.15 *For $i \geq 4, k \geq 3$ T^k_i has a P.I.C. number $\sum_{j=1}^{i} k^j - ki - 2$.*

Proof We have already shown that T^k_i has a maximum P.I.C. number of this size so all that remains is to show that no P.I.C. with smaller size exists. Noting that T^k_i has $\sum_{j=1}^{i} k^j$ edges we can see this is equivalent to showing that there is no valid P.I.C. for T^k_i which contains all but $ki-2$ edges of T^k_i.

Claim 5.16 *Any P.I.C. for T^k_i may be missing at most k edges per level and the missing edges must all have a common parent vertex.*

As the proof of this claim exactly mirrors the proof of claim 5.8 it has been omitted from this thesis.

Claim 5.17 *Any P.I.C. for T^k_i is missing at most $ki-1$ edges.*

Proof We know that we have at most ki edges missing in any valid P.I.C. by the previous claim. Now assume we have a P.I.C. where we have removed k edges at each level. Then the root of T^k_i is an isolated vertex in this P.I.C. as are k of its leaves. But these vertices are not interchangeable with respect to T^k_i and thus by the isolated vertex lemma this is an invalid P.I.C. Thus we may remove at most $ki-1$ edges from our P.I.C. $\square$

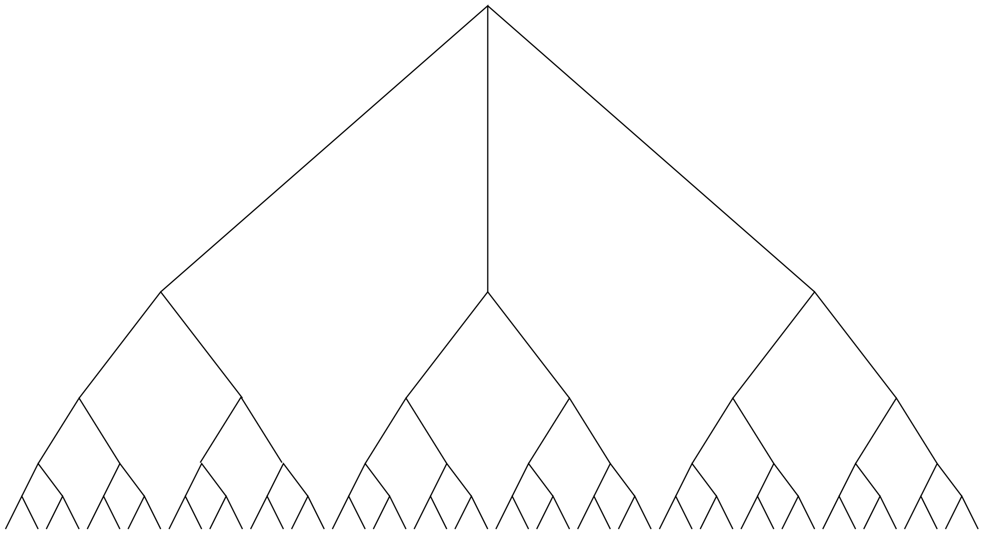

Figure 5.10: 3-regular complete binary tree on 5 levels

If none of the leaves are isolated we can remove at most $ki - k$ edges from T_i^k in creating our P.I.C. As $k > 2$ this P.I.C. is not minimum in size. Thus the root can not be an isolated vertex.

Now we know there are k edges missing at level two and they all have a common parent vertex. This vertex is either an isolated vertex or adjacent only to the root of T_i^k. In the first case another application of the isolated vertex lemma allows us to state that this P.I.C. can not be valid. And in the second case the edge between the root and this vertex is isolated and thus by the isolated edge lemma our P.I.C. is invalid. And so we know at most $ki - 2$ edges have been removed from our P.I.C. and thus any minimum P.I.C. has size $\sum_{j=1}^{i}(k^j) - ki - 2$. □

We should note that theorem 5.15 is also valid for $k = 2$. This was proved as a special case earlier. However the specific P.I.C. given which achieves this bound is not valid when $k = 2$. In fact it is not clearly defined.

5.5 P.I.C. for the complete 3-regular binary tree.

We may also consider the P.I.C. for the complete binary tree which is 3-regular. The only difference in this tree from the standard complete binary tree being that the

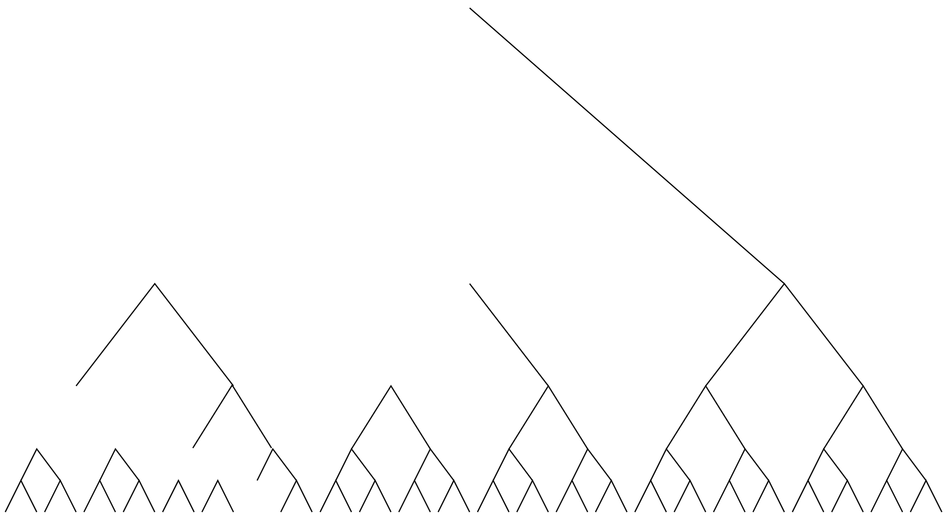

Figure 5.11: P.I.C. for 3-regular complete binary tree on 5 levels

root has degree three. Figure 5.10 depicts this graph for $i = 5$. A minimum P.I.C. for this graph has exactly the same number of edges as a minimum P.I.C. for T_i. As this graph has one more edge we have removed one more edge. As well we have a valid P.I.C. which is very similar to that given in theorem 5.11. This P.I.C. is shown in figure 5.11 for $i = 5$.

5.6 I.C.s and N.I.C.s for T_i and T_i^k.

We conclude with several theorems that show that non-edges do not allow us to improve upon the size of our minimum P.I.C.

Theorem 5.18 *For $i \geq 4, k \geq 3$ a minimum N.I.C. for T_i^k is at least as large as a minimum P.I.C. for T_i^k.*

Proof Let D be a minimum N.I.C. for T_i^k. Consider any two edges at level i in T_i^k which have different parent vertices. Label these edges (v_1, v_2) and (v_3, v_4). Assume neither non-edge (v_1, v_4) nor non-edge (v_2, v_3) is present in D. Then we have a graph T', $T' \cong T_i^k$; $T_i^k \succeq D$; and contains edges (v_1, v_4) and (v_2, v_3). It is not identical to T_i^k and thus D is not a valid N.I.C. As a result we know that for every pair of vertices at the ith level we must have at least one non-edge in our N.I.C. .

This is $(k^i)(k^i-1)/2$ non-edges. It easily seen that this is larger than the number of edges in our minimum P.I.C. for T_i^k. $\square$

In fact we can go even further than this. We can show that even by using an arbitrary combination of edges and non-edges in the I.C. it is impossible to improve on the bounds provided in theorems 5.7 and 5.15. Again this will be done for complete binary trees and k-ary trees in one theorem.

Theorem 5.19 *T_i^k $k \geq 2, i \geq 4$ has a minimum I.C. which is the same size as the minimum P.I.C. for T_i^k.*

Proof We know that a minimum I.C. for T_i^k of the same size as a minimum P.I.C. for T_i^k exists. For this proof we shall show that the number of edges missing from T_i^k + the number of non-edges in our I.C. must be at least $ki-2$. We shall refer to these edges as the missing edges in the same style as we did for theorems 5.7 and 5.15. Let D be a valid minimum I.C. for T_i^k. As shown previously we can not have two edges missing from the same level which have different parents unless we "account" for this with non-edges. That is assume we have two edges missing at level j which have different parent vertices. Label the edges (v_1, v_2) and (v_3, v_4). Now assume neither non-edge (v_1, v_4) or non-edge (v_2, v_3) is present in our I.C. Then we have a graph T' which is isomorphic to T_i^k , contains D and contains edges (v_1, v_4) and (v_2, v_3). Thus it is not identical to T_i^k and D is not a valid I.C. So for every edge missing at level j -except the first k- we must have at least one non-edge present in our I.C. Thus our I.C. must be missing at most ki edges.

Consider a I.C. which has only ki edges missing. Consider the root vertex $v_{0,0}$. It is not only an isolated vertex in our I.C. but there is no non-edge with $v_{0,0}$ as an endpoint. We also have a leaf of the tree, WLOG say $v_{i,0}$, which is an isolated vertex. Though it may be the endpoint of a non-edge it is not the endpoint of a non-edge to a vertex at level one. Thus we may switch the labels on these two vertices and the resulting graph will be isomorphic to T_i^k , contain our I.C. as a subgraph, and not be identical to T_i^k. Thus our I.C. is not valid and we either need to add an edge or a non-edge. Thus we have at most $ki - 1$ edges missing.

Finally consider a vertex at the first level. There exists at least one vertex at level one with no children in the I.C. If it is adjacent to the root of the tree this is an isolated edge and by the isolated edge lemma our I.C. is invalid. If it is not adjacent to the root of the tree it is an isolated vertex and by the same reasoning as used in the previous paragraph our I.C. is invalid. Thus we must remove at most $ki-2$ edges and we know that any valid minimum I.C. is as large as a minimum P.I.C. for T_i^k. $\square$

Chapter 6

Grids

6.1 Introduction

This chapter focuses on determining the value of the P.I.C. number for $m \times m$ grids. We have not been able to exactly identify this number in the general case, but we show that the P.I.C. number for the $m \times m$ grid is asymptotic to the number of vertices. Smaller grids require a case intensive study in order to determine their P.I.C. number and as these cases are of minimal interest we have ignored them.

Let G_m denote the $m \times m$ grid on m^2 vertices. As indicated in figure 6.1 this graph has vertices $g_{i,j}, 1 \leq i, j \leq m$ where

$g_{i,j}$ is adjacent to $g_{i-1,j}, 2 \leq i \leq m, 1 \leq j \leq m$,

$g_{i,j}$ is adjacent to $g_{i+1,j}, 1 \leq i \leq m-1, 1 \leq j \leq m$,

$g_{i,j}$ is adjacent to $g_{i,j-1}, 1 \leq i \leq m, 2 \leq j \leq m$,

$g_{i,j}$ is adjacent to $g_{i,j+1}, 1 \leq i \leq m, 1 \leq j \leq m-1$.

We refer to vertex $g_{i,j}$ as being in the ith row and jth column of G_m.

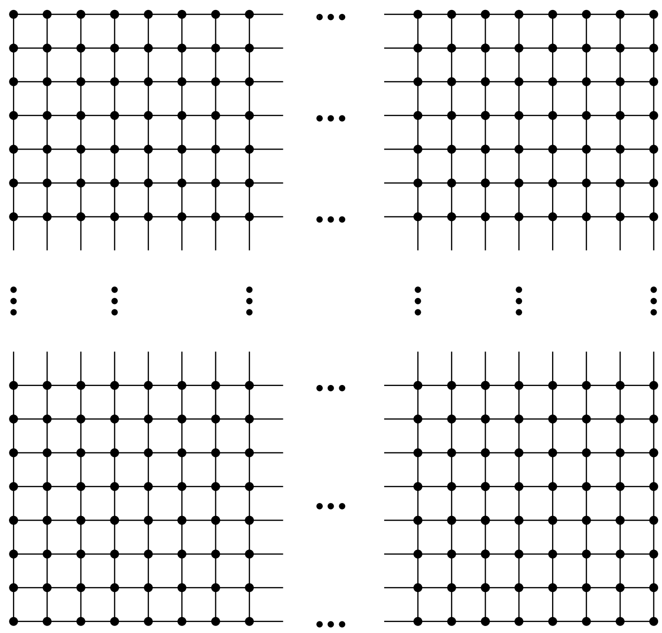

Figure 6.1: G_m

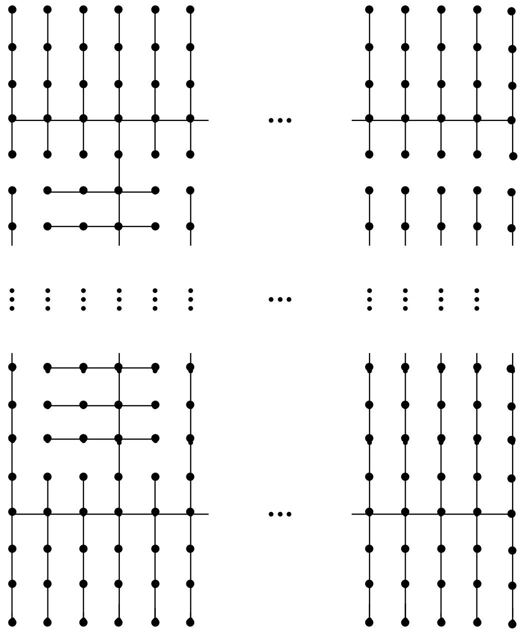

Figure 6.2: D_m

6.2 P.I.C. for G_m

For $m \geq 14$ let D_m denote the spanning subgraph of G_m with edge set $A \cup B \cup C \cup D \cup E \cup F \cup G$ where A, B, C, D are as follows:

$$A =\{(g_{4,j}, g_{4,j+1}), 1 \leq j \leq m-1\}$$

$$B =\{(g_{i,4}, g_{i+1,4}), 1 \leq i \leq m-1\}$$

$$C =\{(g_{i,m-3}, g_{i+1,m-3}), 1 \leq i \leq m-1\}$$

$$D =\{(g_{i,j}, g_{i,j+1}), 1 \leq i \leq m, m-4 \leq j \leq m-1\}$$

$$E =\{(g_{i,j}, g_{i,j+1}), 1 \leq i \leq m, i \neq 2,3,4,5, 1 \leq j \leq m-6\}$$

$$F =\{(g_{i,j}, g_{i,j+1}), 2 \leq i \leq 5, 1 \leq j \leq 4\}$$

$$G =\{(g_{i,j}, g_{i+1,j}), 2 \leq i \leq 4, 6 \leq j \leq m-5\}$$

D_m is depicted in figure 6.2. We will prove that D_m is a valid P.I.C. for G_m, thus showing:

Theorem 6.1 *The P.I.C. number for G_m is no more then $m^2 - 1$ for $m \geq 14$.*

Proof Consider any graph G', $G' \cong G_m$ and $G' \succeq D_m$. We shall show $G' = G_m$.

Throughout this proof we denote by $a_{i,j}$ the vertex at row i, column j of G'. This vertex is equal to a vertex $g_{i',j'}$ for some i', j' and we write $a_{i,j} \equiv g_{i',j'}$.

Our first step if to show that the vertices in the fourth column of D_m must all lie in the same column of G'. (In fact, they may all lie in the same row of G' but in this case we may rotate the entire graph WLOG so this row becomes a column.) We begin by showing all but the first and last vertices in the second column lie in the same column.

Claim 6.2 *$g_{i,4} \equiv a_{i,k}, 2 \leq i \leq m-1$ for some k.*

Proof Consider $g_{i,4}$ for an arbitrary $i, 3 \leq i \leq m-2$. Assume it is at row j, column k in G'. That is $g_{i,4} \equiv a_{j,k}$. Now as $g_{i-1,4}$ is adjacent to $g_{i,4}$ we can say WLOG that $g_{i-1,4} \equiv a_{j-1,k}$. So, consider vertex $g_{i+1,4}$. It too is adjacent to $g_{i,4}$. Thus $g_{i+1,4} \in \{a_{j+1,k}, a_{j,k-1}, a_{j,k+1}\}$. Now if it is equal to $a_{j,k-1}$ or $a_{j,k+1}$ then it must have a common neighbour with $a_{j-1,k} \equiv g_{i-1,4}$. However $g_{i-1,4}$ and $g_{i+1,4}$ both have degree four in D_m and have no neighbours in common in D_m. As every vertex in G' has degree four or less $g_{i-1,4}$ and $g_{i+1,4}$ can have no neighbour in common in G'. Thus $g_{i+1,4}$ must be equal to $a_{j+1,k}$. As $g_{i,4}$ was an arbitrary vertex with $3 \leq i \leq m-2$ we may repeat this process until all but $g_{1,4}$ and $g_{m,4}$ are known to be in column k. Finally consider $g_{2,4}$. As it has degree four in D_m it can not be in the first row of G'. Similarly $g_{m-1,4}$ can not be in the last row of G'. Thus $g_{i,4} \equiv a_{i,k}, 2 \leq i \leq m-1$. □

We can use the same reasoning to show that:

Claim 6.3 $g_{i,m-3} \equiv a_{i,l}, 2 \leq i \leq m-1$ *for some* l.

We next show that the vertices in the fourth row of D_m do indeed form a row in G'. We first focus on the vertices between the columns 4 and $m-3$.

Claim 6.4 $g_{4,j} \equiv a_{4,k+j-4}, 4 \leq j \leq m-3$.

Proof Consider vertex $g_{4,4}$. By claim 6.2 it is equal to $a_{4,k}$ for some k. Next consider $g_{4,5}$. As it is adjacent to $g_{4,4}$ it must be equal to a vertex adjacent to $a_{4,k}$. As it can not be in the same column (from claim 6.2) it must be in the same row. Thus WLOG $g_{4,5} \equiv a_{4,k+1}$. Now consider the next vertex in the row, that is $g_{4,6}$. It has degree four; it is adjacent to $g_{4,5}$; and it has no neighbours in common with $g_{4,4}$; thus we know it must be equal to $a_{4,k+2}$. Next is $g_{4,7}$. We know it adjacent to $g_{4,6}$. If it is equal to $a_{3,k+2}$ or $a_{5,k+2}$ it has a neighbour in common with $g_{3,4}$ or $g_{5,4}$ respectively. However, they both have degree four in D_m and no common neighbours and thus we know $g_{4,7} \equiv a_{4,k+3}$. By the same reasoning, for each $8 \leq j \leq m-5$ we have $g_{4,j} \equiv a_{4,k+j-4}$. Finally, consider $g_{4,m-4}$. It is adjacent to $g_{4,m-5}$. If it is equal to $a_{3,k+m-9}$ or $a_{5,k+m-9}$ then $g_{4,m-3} \in \{a_{3,k+m-10}, a_{3,k+m-8}, a_{5,k+m-10}, a_{5,k+m-8}\}$. We know this because we know $g_{4,m-4}$ and $g_{4,m-3}$ can not be in the same column from claim 6.3. However if $g_{4,m-3}$ is any of these four vertices it has a second neighbour in common with $g_{4,m-5}$ which is not true. Thus $g_{4,m-4} \equiv a_{4,k+m-8}$. And finally $g_{4,m-3} \equiv a_{4,k+m-7}$. $\square$

Our next step is to fix the location of the fourth and fourth last columns. This will be done by finding exact values for the k and l mentioned in claims 6.2 and 6.3 respectively. We begin by showing that k is at least four.

Claim 6.5 *$g_{i,4} \equiv a_{i,k}, 2 \leq i \leq m-1$ for some $k \geq 4$.*

Proof By claim 6.2, $g_{i,4} \equiv a_{i,k}, 2 \leq i \leq m-1$ for some k. For every $m-6 \leq i \leq m-2$, consider $g_{i,3}$ and $g_{i,5}$. Each of these vertices is adjacent to a vertex in column k and is not in column k. Thus we have identified the vertices in the set $\{a_{i,j}, i = k-1, k+1, m-6 \leq k \leq m-2\}$. So consider $g_{i,2}, m-5 \leq i \leq m-3$ and $g_{i,6}, m-5 \leq i \leq m-3$. They must be equal to the six vertices in the set $\{a_{i,j}, i = k-2, k+2, m-5 \leq j \leq m-3\}$. Finally consider $g_{m-4,1}$ and $g_{m-4,7}$. They must be equal to $\{a_{m-4,k-3}, a_{m-4,k+3}\}$. Thus $k-3 \geq 1$ and $k \geq 4$. $\square$

And we get a similar strengthening of claim 6.3:

Claim 6.6 *$g_{i,m-3} \equiv a_{i,l}, 2 \leq i \leq m-1$ for some $l \leq m-3$.*

Proof By claim 6.3 $g_{i,m-3} \equiv a_{i,l}, 2 \leq i \leq m-1$ for some l. First we know the three neighbours of $g_{2,m-3}$ must be equal to the three neighbours of $a_{2,l}$. As well we know $g_{i,m-4}, 3 \leq i \leq 6$ and $g_{i,m-2}, 3 \leq i \leq 6$ must be the eight vertices in the set $\{a_{i,j}, 3 \leq i \leq 6, j = m-4, m-2\}$. Next consider $g_{3,m-5}$ and $g_{5,m-5}$. We know they are adjacent to $g_{4,m-5} \equiv a_{4,l-2}$ and thus they are the two vertices in the set $\{a_{3,l-2}, a_{5,l-2}\}$. Now consider $g_{i,m-1}, 3 \leq i \leq 5$. They are adjacent to vertices in $\{a_{i,j}, 3 \leq i \leq 5, j = l-1, l-2\}$ and are not any of the previous vertices and thus must be $\{a_{i,l+2}, 3 \leq i \leq 5\}$. So finally consider $g_{4,m}$. It is adjacent to $g_{4,m-1} \equiv a_{4,l+2}$ but is not equal to $a_{3,l+2}, a_{5,l+2}$, or $a_{4,l+1}$ and thus must be $a_{4,l+3}$. Thus $l+3 \leq m$ and $l \leq m-3$. $\square$

And finally, using claims 6.4, 6.5, and 6.6 we can determine k and l precisely.

Claim 6.7 $g_{i,4} \equiv a_{i,4}, 2 \leq i \leq m-1$ *and* $g_{i,m-3} \equiv a_{i,m-3}, 2 \leq i \leq m-1$.

Proof From claim 6.4, $g_{4,4}$ and $g_{4,m-3}$ are joined in a row of length $m-6$. Since we know from claims 6.5 and 6.6 $k \geq 4$ and $l \leq m-3$ $k=4$ and $l=m-3$. □

It is finally time to show that the four vertices at the ends of columns 4 and $m-3$ in D_m are in fact in the desired columns of G'.

Claim 6.8 $g_{i,4} \equiv a_{i,4}, 1 \leq i \leq m$ *and* $g_{i,m-3} \equiv a_{i,m-3}, 1 \leq i \leq m$.

Proof Begin by considering $g_{1,4}$. We know it is adjacent to $g_{2,4} \equiv a_{2,4}$. If it is equal to $a_{2,3}$ then either $g_{2,3}$ or $g_{2,5}$ is equal to $a_{1,4}$. As well we know we can identify $g_{4,1} \equiv a_{4,1}$ and $g_{3,2} \equiv a_{3,2}$ by reasoning similar to that used in claim 6.6. Thus we know $g_{1,1}, g_{1,2}, \ldots, g_{1,m-5}$ and $g_{3,1}$ must be the six vertices in the set $\{a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{3,1}\}$. However we have more than six vertices as $m \geq 14$ so this can not be. Using similar counting arguments (where in some cases there are actually fewer then six vertices) we may show $g_{1,4}$ is not equal to $a_{2,5}$, $g_{1,m-3}$ is not equal to $a_{2,m-4}$ or $a_{2,m-2}$, $g_{m,4}$ is not equal to $a_{m-1,3}$ or $a_{m-1,5}$ and $g_{m,m-3}$ is not equal to $a_{m-1,m-4}$ or $a_{m-1,m-2}$ □

With the framework now in place we are ready to complete the proof of the original theorem.

Considering $g_{i,3}$ and $g_{i,5}, m \geq i \geq 1$ in order we know that $\{g_{i,3}, g_{i,5}\} \equiv \{a_{i,3}, a_{i,5}\}$. Again considering the vertices in order, using the already identified vertices and the known adjacencies we can determine that $\{g_{i,2}, g_{i,6}\} \equiv \{a_{i,2}, a_{i,6}\}, m \geq i \geq 1$. The same follows for $g_{i,1}$ and $g_{i,7}, m \geq i \geq 1$. This brings us to $g_{i,8}, m \geq i \geq 1$. We can now determine $g_{i,8} \equiv a_{i,8}, m \geq i \geq 1$ which allows us to uniquely determine the equalities in the previous sets. We know $g_{i,j} \equiv a_{i,j}, 1 \leq i \leq m, 1 \leq j \leq 7$. We may continue this process for all other vertices in the graph and discover that $G' = G_m$ as desired. □

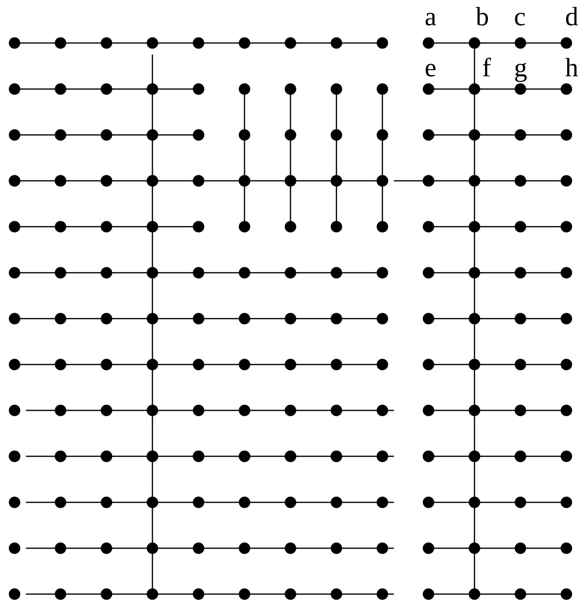

Figure 6.3: Potential P.I.C.

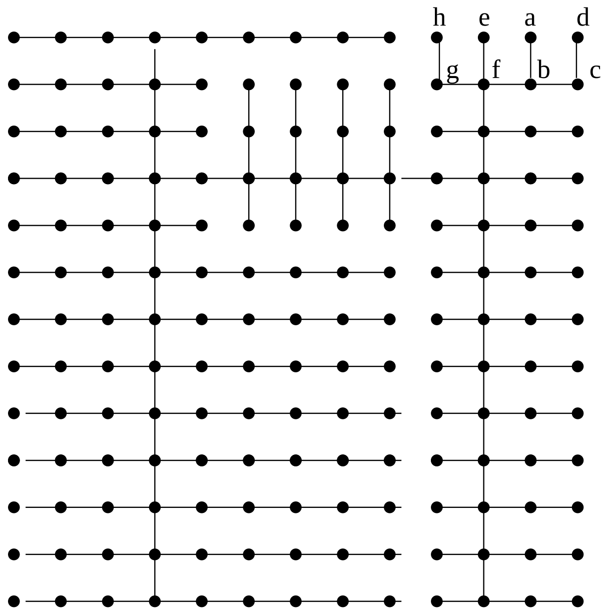

Figure 6.4: Counterexample for potential P.I.C.

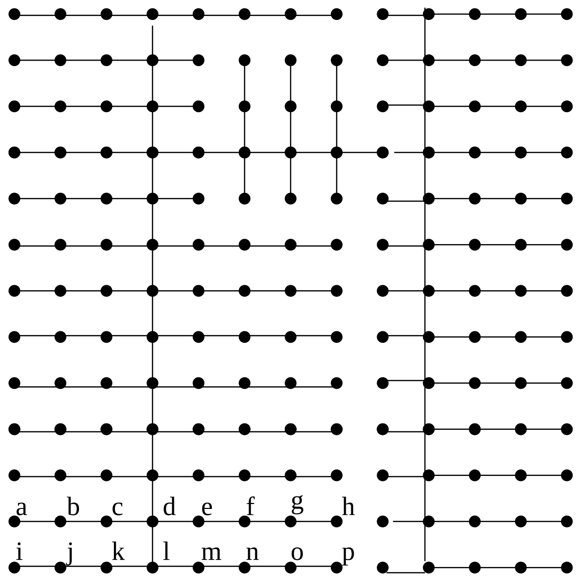

Figure 6.5: Potential P.I.C.

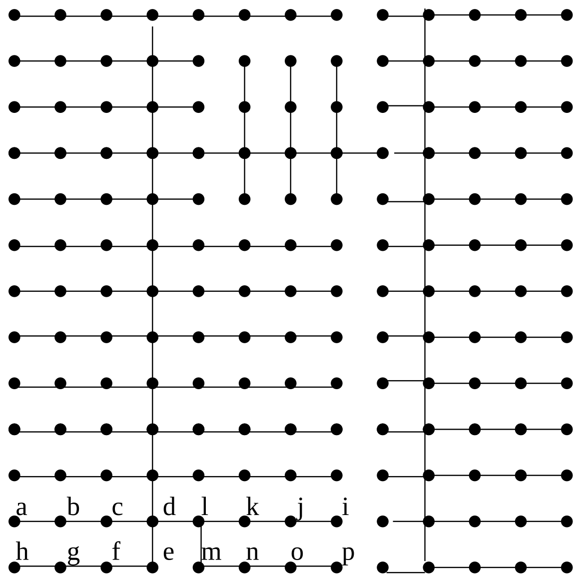

Figure 6.6: Counterexample for potential P.I.C.

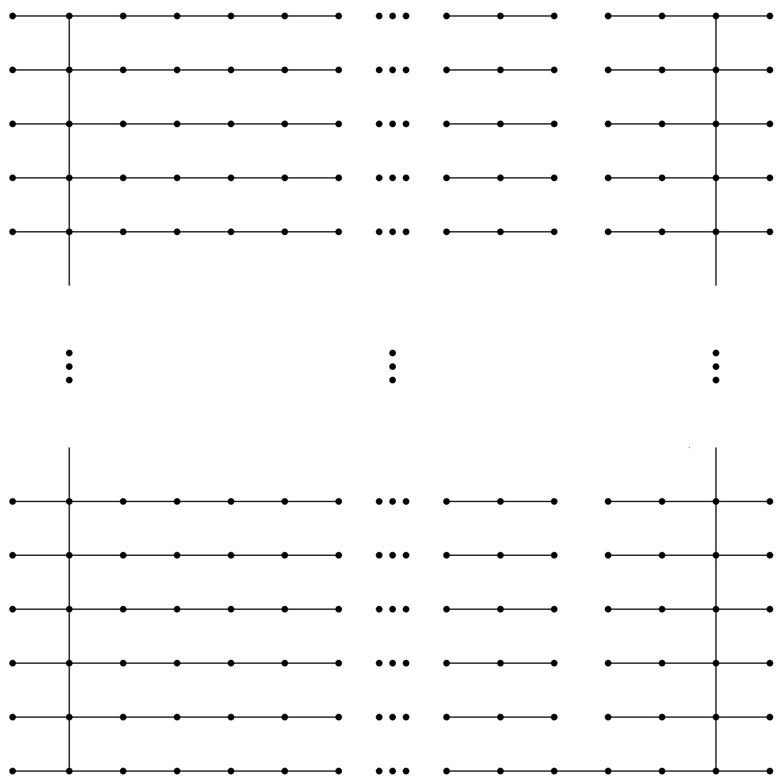

Figure 6.7: Potential P.I.C. for G_m

It is interesting to consider why this theorem is not applicable for some smaller values of m. In figures 6.3, 6.4, 6.5 and 6.6 we have shown two of the problems that occur if the m value is not large enough. In figures 6.3 and 6.5 we have provided the potential P.I.C. and in figures 6.4 and 6.6 a graph G' which is equivalent but not identical to the grid and yet contains the potential P.I.C. The edges of the potential P.I.C. are solid while all other lines are dashed in order to assist the reader.

The reader may also be interested in the potential P.I.C. presented in figure 6.7. This graph was thought to be a valid P.I.C. for G_m for many months. However in figures 6.8 and 6.9 are the counterexamples and the proposed P.I.C.s for $m = 10$ and $m = 12$ respectively. Counterexamples for large m may also be found.

6.3 Lower bounds on the P.I.C. number for G_m.

As we will see, we can apply theorem 1.12 to show that the bound in theorem 6.1 is asymptotically correct, i.e. that the P.I.C. number of G_m is at least $m^2 - o(m^2)$. An

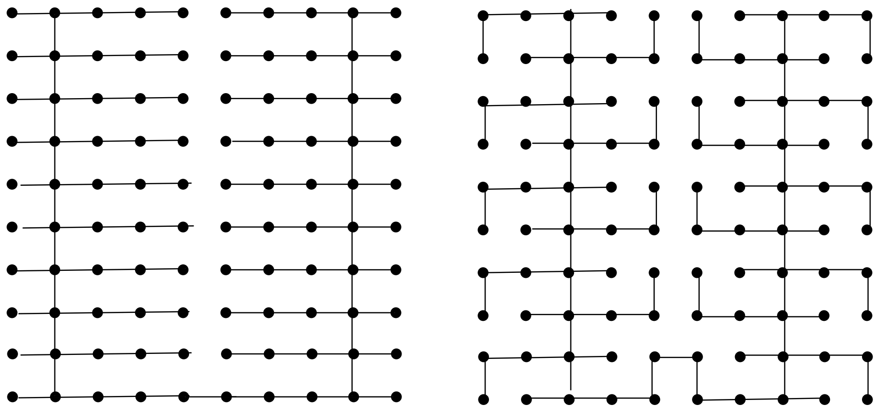

Figure 6.8: Potential P.I.C. and counterexample for $m = 10$.

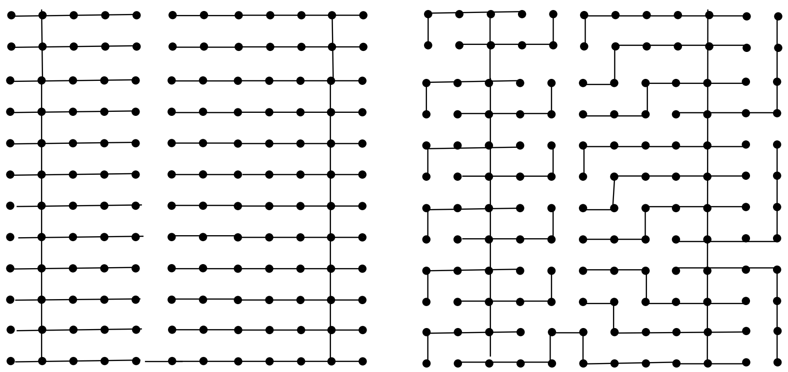

Figure 6.9: Potential P.I.C. and counterexample for $m = 12$.

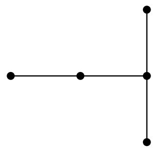

Figure 6.10: B

important idea behind the proof of theorem 1.12 is that if there are a limited number of small components in any P.I.C. then the P.I.C. number must be large. To illustrate this intuition, we will use the simpler fact that no P.I.C. for G_m has more than one component of size less than six to provide a bound on the P.I.C. number of nearly m^2.

Theorem 6.9 *The P.I.C. number for G_m, $m \geq 14$ is at least $\frac{6}{7}m^2 - \frac{6}{7}$.*

First, we prove our claim about the number of small components.

Lemma 6.10 *Any P.I.C. for G_m has no component C with $1 \leq |E(C)| \leq 5$.*

Proof Let C be a P.I.C. for G_m. Consider a component of C. We can not have a component which is a single edge. We know that no two vertices in G_m are interchangeable. Thus the previous statement follows immediately from the isolated edge lemma. Consider a component with two edges. It must be a path on three vertices. Again using the fact that no two vertices are interchangeable we know immediately this component is forbidden by the isolated path lemma.

Next consider a component with three edges. As this is a connected subgraph of G_m this may be either a star, or a path. It can not be a triangle as none are present in G_m. Flipping a path of length less than m^2 results in a bijection which is not an automorphism and thus a path is forbidden by the isolated path lemma. We are left with the possibility this component is a star. Again using the fact no two vertices in G_m are interchangeable, we know a star violates the dangling edge lemma. Thus we can not have a component on three edges either.

Consider a component with four edges. It may be a star, a path, a square, or a graph isomorphic to B in figure 6.10. Both the star and B violate the dangling edge

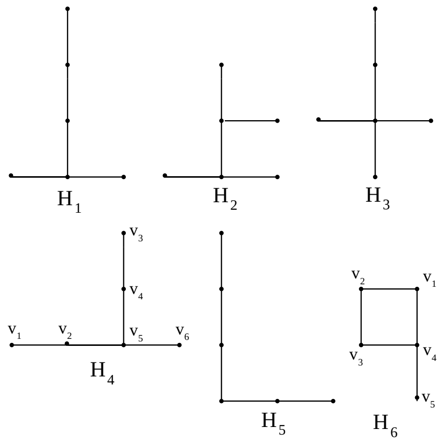

Figure 6.11: Components with 5 edges.

lemma and thus can not be present in our P.I.C. Using the same reasoning as used in the previous paragraph we know the path violates our isolated path lemma. Consider the square. We have a mapping Θ which maps opposite corners of the square onto each other. This is an automorphism on C but not an automorphism on G_m as we have no interchangeable vertices. Thus it violates lemma 1.2. Thus we can have no component with four edges.

Finally consider a component which contains five edges. It must be isomorphic to one of the graphs given in figure 6.11. Graphs H_1, H_2 and H_3 all violate the dangling edge lemma while H_5 violates the isolated path lemma given that flipping just a path is not an automorphism on G_m. Consider H_4. Let Θ be the bijection which maps v_1 onto v_3, v_2 onto v_4, v_3 onto v_1, v_4 onto v_2, and every other vertex onto itself. Then obviously this is an automorphism on the component and not on G_m and thus violates lemma 1.1. Finally, consider H_6. A bijection which maps v_1 onto v_3 and v_3 onto v_1 while mapping every other vertex onto itself is a automorphism on H_6 and not on G_m so again lemma 1.1 is violated and this is not a valid P.I.C. $\square$

And now, as promised, this bound on the number of small components immediately provides a bound on the P.I.C. number of G_m.

Proof of Theorem 6.9

Assume the theorem is false and G_m has a P.I.C., D, of size $l \leq \frac{6}{7}m^2 - \frac{6}{7}$. Since D has m^2 vertices and $l < m^2$, D has $m^2 - l$ components.

As G_m has no interchangeable vertices we know a valid P.I.C. has at most one isolated vertex. This fact in conjunction with the previous theorem yields the fact that all but at most one component of D have at least six edges. Thus D contains at least $(m^2 - l - 1) * 6$ edges. So $6 * (m^2 - l - 1) \leq l$. Manipulating this formula yields $l \geq 6/7m^2 - 6/7$ and as $m \geq 14$ we can easily achieve the lower bound of $m^2/2$ for l. $\square$

To improve this bound to $(1 - o(1))m^2$ the key is to strengthen lemma 6.10 to say that for any fixed k, the number of components of size k in my P.I.C. is bounded above by a function $f(k)$. As implied in the proof of theorem 1.12, to do this it suffices to show the number of vertices which any non-identity automorphism of the grid leaves fixed is at most $o(m^2)$.

Lemma 6.11 *The only automorphisms of the grid are:*

$\Theta_1(g_{i,j}) = g_{i,j}$ *(identity automorphism)*

$\Theta_2(g_{i,j}) = g_{m+1-i,j}$ *(flip across middle row)*

$\Theta_3(g_{i,j}) = g_{i,m+1-j}$ *(flip across middle column)*

$\Theta_4(g_{i,j}) = g_{j,i}$ *(flip across main diagonal)*

$\Theta_5(g_{i,j}) = g_{m+1-j,m+1-i}$ *(flip across other diagonal)*

$\Theta_6(g_{i,j}) = g_{j,m+1-i}$ *(*90° *rotation)*

$\Theta_7(g_{i,j}) = g_{m+1-i,m+1-j}$ *(*180° *rotation)*

$\Theta_8(g_{i,j}) = g_{m+1-j,i}$ *(*270° *rotation)*

Proof Consider the four corners of the grid, $g_{1,1}, g_{m,1}, g_{1,m}$ and $g_{m,m}$. We know that any automorphism, Θ, of the graph must map each of the corners onto another corner as these are the only vertices in the graph of degree two. So, focusing first on $g_{1,1}$ it may be mapped to one of four locations, $g_{1,1}$ $g_{1,m}, g_{m,1}$ and $g_{m,m}$. Now consider $g_{m,m}$. If the location of $g_{1,1}$ is fixed, the location of $g_{m,m}$ must be the only corner which has no path of length m from $\Theta(g_{1,1})$. Then for each of the four possible locations of $g_{1,1}$ there are two locations left for $g_{1,m}$ and this will finalize the location of $g_{m,1}$. So, all in all, we have eight sets of mappings for the four corners. They are:

$$\Theta_1(g_{1,1}) = g_{1,1}; \Theta_1(g_{1,m}) = g_{1,m}; \Theta_1(g_{m,1}) = g_{m,1}; \Theta_1(g_{m,m}) = g_{m,m};$$

$$\Theta_2(g_{1,1}) = g_{m,1}; \Theta_2(g_{1,m}) = g_{m,m}; \Theta_2(g_{m,1}) = g_{1,1}; \Theta_2(g_{m,m}) = g_{1,m};$$

$$\Theta_3(g_{1,1}) = g_{1,m}; \Theta_3(g_{1,m}) = g_{1,1}; \Theta_3(g_{m,1}) = g_{m,m}; \Theta_3(g_{m,m}) = g_{m,1};$$

$$\Theta_4(g_{1,1}) = g_{1,1}; \Theta_4(g_{1,m}) = g_{m,1}; \Theta_4(g_{m,1}) = g_{1,m}; \Theta_4(g_{m,m}) = g_{m,m};$$

$$\Theta_5(g_{1,1}) = g_{m,m}; \Theta_5(g_{1,m}) = g_{1,m}; \Theta_5(g_{m,1}) = g_{m,1}; \Theta_5(g_{m,m}) = g_{1,1};$$

$$\Theta_6(g_{1,1}) = g_{1,m}; \Theta_6(g_{1,m}) = g_{m,m}; \Theta_6(g_{m,1}) = g_{1,1}; \Theta_6(g_{m,m}) = g_{m,1};$$

$$\Theta_7(g_{1,1}) = g_{m,m}; \Theta_7(g_{1,m}) = g_{m,1}; \Theta_7(g_{m,1}) = g_{1,m}; \Theta_7(g_{m,m}) = g_{1,1};$$

$$\Theta_8(g_{1,1}) = g_{m,1}; \Theta_8(g_{1,m}) = g_{1,1}; \Theta_8(g_{m,1}) = g_{m,m}; \Theta_8(g_{m,m}) = g_{1,m};$$

and correspond to the eight automorphisms given in the statement of the lemma.

So we have at most eight different configurations for the corners. We shall next show that if the position of the four corners is known the position of all the vertices is known.

Claim 6.12 *Assume Θ is an automorphism on the grid and the value of Θ is known for the four corners. Then the automorphism is known.*

Proof We will prove this by induction on the size of m.

For $m = 1$ or 2 knowing the four corners is equivalent to knowing the entire graph so the basis step follows.

Now assume the statement is true for all grids of size $m-1$ or smaller and consider the $m \times m$ grid. Assume we know which vertex each of the corners maps to. We refer to two corners as *adjacent* if there is a path of length m between them. This path will be unique. Between any two adjacent corners a and b we label this path

$a, v_2, v_3, \ldots, v_{m-1}, b$. If $\Theta(a)$ and $\Theta(b)$ are fixed there is one unique path of length m between them as well, $\Theta(a), u_2, u_3, \ldots, u_{m-1}, \Theta(b)$ and thus $\Theta(v_i) = u_i, 2 \leq i \leq m-1$.

Next consider the four vertices which are the corners of the square of size $m-2 \times m-2$ which does not include any of the peripheral vertices of the grid. These four vertices all share two neighbours with one of the corners of the grid and are the only such vertices. Thus if the vertices which the corners map to under the automorphism Θ are known then the vertices which each of these four vertices map to must also be known. Thus we can apply the induction hypothesis to the $m-2 \times m-2$ grid with these four vertices as corners and we are know the entire automorphism. $\square$

It is left to the reader to confirm that the eight configurations of the corners correspond to the eight automorphisms given in the statement of the theorem. $\square$

Theorem 6.13 *The P.I.C. number for G_m, $m \geq 14$ is at least $(1-o(1))m^2$.*

Noting that any non-identity automorphism on G_m leaves at most m vertices fixed this follows immediately from theorem 1.12.

Chapter 7

Conclusions

The results presented in this thesis only begin to scratch the surface of the many questions that may be asked - and hopefully someday answered - on certificates. Here we present just a few of these questions.

Though there remain a wide range of questions to be addressed most of the specific graphs considered in this thesis have been fairly thoroughly discussed. We were unable however to determine exactly the P.I.C. for the grid and would be very interested in seeing this done.

To further the collection of work on certificates for undirected graphs a natural approach would be to consider other specific graphs or classes of graphs. Is there anything that can be said about the I.C. number for planar graphs, trees, bipartite graphs etc. We are curious not only about bounds on the value of these I.C. numbers but about algorithms to generate I.C.s for these graphs.

And finally, the reader might also be interested in considering more general questions. Is it possible to bound the I.C. number for arbitrary graphs as a function of some property of the graph? Is it possible to show that most graphs have I.C. numbers bounded above (or below) by a function of the number of vertices in the graph?

We hope that in the next few years these questions shall provide the combinatorics community with many fruitful hours of contemplation and discussion.

Bibliography

[1] N. Alon and M. Ruszinkò. Short certificates for tournaments. *The Electronic Journal of Combinatorics*, 4(R12), 1997.

[2] J.A. Bondy and R.L. Hemminger. Graph reconstruction - a survey. *Journal of Graph Theory*, 1:227–268, 1977.

[3] M. Fiedler, editor. *Theory of Graphs and its Applications*. Academic Press, New York, 1964.

[4] P. Fishburn, J.H. Kim, and P. Tetali. Tournament certificates. Technical report, AT&T Bell Laboratories, 1994.

[5] P. Fishburn, J.H. Kim, and P. Tetali. Score certificates for tournaments. *Journal of Graph Theory*, 24:117–138, 1997.

[6] D.L. Greenwell and R.L. Hemminger. Reconstructing the n-connected components of a graph. *Aequationes Mathematica*, 9:19–22, 1973.

[7] F. Harary and E. Palmer. A note on similar points and similar lines of a graph. *Rev. Roumaine Math. Pures et Appl.*, 10:1489–1492, 1965.

[8] R.L. Hemminger. On reconstructing a graph. *Pocedures of the American Mathematical Society*, 20:185–187, 1969.

[9] P.J. Kelly. A congruence theorem for trees. *Pacific Journal of Mathematics*, 7:961–968, 1957.

[10] L. Lovász. A note on the line reconstruction problem. *Journal of Combinatorial Theory (B)*, 13:309–310, 1972.

[11] L. Lovász. *Combinatorial Problems and Exercise.* North-Holland Publishing Company, 1979.

[12] B.D. McKay. Computer reconstruction of small graphs. *Journal of Graph Theory*, 1:281–283, 1977.

[13] J.W. Moon. *Topics on Tournaments.* Holt, Rinehart and Winston, 1968.

[14] V. Muller. The edge reconstruction hypothesis is true for graphs with more than $n \log_2 n$ edges. *Journal of Combinatorial Theory (B)*, 22:281–283, 1997.

[15] Nijenhuis. Note on the unique determination of graphs by proper subgraphs. *Notices American Mathematical*, 20, 1977.

[16] A. Rubinstein. Why are certain properties of binary relations relatively more common in natural language? *Econometrica*, 64:343–356, 1996.

[17] P.K. Stockmeyer. The falsity of the reconstruction conjecture for tournaments. *Journal of Graph Theory*, 1:19–25, 1977.

[18] S. Thomasse. Counterexample. Private Correspondence.