Order Number 9123285

Odd forests, reversing numbers, and discrete representations of interval orders

Isaak, Garth Timothy, Ph.D.

Rutgers The State University of New Jersey - New Brunswick, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Odd Forests, Reversing Numbers, and Discrete Representations of Interval Orders

by Garth T. Isaak

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Operations Research

Written under the direction of

Professor Fred S. Roberts

and approved by

New Brunswick, New Jersey

October 1990

ABSTRACT OF THE DISSERTATION

Odd Forests, Reversing Numbers, and Discrete Representations of Interval Orders

by Garth Isaak, Ph.D.

Dissertation Director: Professor Fred S. Roberts

In this thesis we examine three problems in discrete applied mathematics. Although there is no direct connection between these problems, they are linked by methodology. Each uses techniques common in discrete applied mathematics and in particular techniques of graph theory.

Chapter 2 examines a generalization of maximum matching in a graph called chain packing. In a chain packing, terminals are linked by edge disjoint chains rather than edges. The subgraph induced by the chains must also satisfy given degree constraints. It can easily be shown that a maximum chain packing can be obtained by a forest in which all non-isolated vertices have odd degree. We present a polynomial time algorithm that finds a maximum chain packing by packing odd subtrees. The algorithm builds on an augmenting chain theorem of deWerra and Roberts and is in the spirit of Edmonds' blossom algorithm or matching. More general conditions for blossom detection are used.

Chapter 3 examines sets of arcs in a tournament whose reversal makes a tournament acyclic. The problem of finding a minimum set of arcs in a tournament whose

reversal makes the tournament acyclic is equivalent to the feedback arc set problem, to finding a minimum transversal of the cycles, to finding a maximum acyclic subdigraph, and to finding a ranking which minimizes inconsistencies. Following a question of J.-P. Barthelemy, we examine minimum sets of arcs whose reversal makes a tournament acyclic from a different perspective. Given an acyclic digraph, we define the reversing number to be the number of extra vertices in a smallest tournament in which the given digraph is a minimum reversing set. We examine some basic bounds on the reversing number and exact values on some classes of digraphs.

Finite interval orders are orders which can be represented by 'strictly greater than' on a set of real intervals. Interval orders arise in preference rankings for which indifference is not transitive, orderings of temporal events, and scheduling problems. In Chapter 4, following a question of K.P. Bogart, we examine bounded discrete interval orders (for which the intervals have bounded length and the endpoints must be integral). Using Farkas' Lemma, we reduce the problem to detecting negative cycles or finding shortest paths in an associated digraph. This provides a polynomial algorithm for determining if an order has a bounded discrete representation. Additionally, the digraph model is used as a basis to determine necessary and sufficient conditions for representability in the cases that the length bounds are constant and the lower bound is 0 or 1.

Acknowledgements

There are many people whose support makes the completion of a thesis possible. I will give a somewhat chronological and probably incomplete listing of the groups and individuals who have helped make my education complete:

My parents and my family, for their constant love and support,

Friends for being there,

The many teachers who have tried to pass their knowledge on to me despite my resistance,

Frank Breneman, Richard Rempel, and Arnold M. Wedel, professors at Bethel College who got me started in the Mathematical Sciences,

Professors at Rutgers, who got me interested in discrete mathematics,

Fellow students throughout the years with whom I struggled to understand new ideas and took much needed breaks,

The Center for Operations Research at Rutgers (RUTCOR) for providing support and the atmosphere for working on this thesis,

The secretarial and administrative staff at Rutgers for support and for always knowing the answer,

My thesis advisor Fred S. Roberts, for his invaluable advice and guidance,

The members of my dissertation committee, Endre Boros, Martin Golumbic, and Pierre Hansen, for their many valuable comments in addition to spelling corrections, and their flexibility in scheduling the defense for a Sunday morning,

My new spouse Melissa Hunt, the spice of my life.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Preview of the Thesis

In this thesis we examine three problems in discrete applied mathematics. Although there is no direct connection between these problems, they are linked by methodology. Each uses techniques common in discrete applied mathematics and in particular techniques of graph theory. We briefly preview the three problems in this section, providing a basic idea of what we will do. A more detailed overview of the problems and related work will be given at the beginning of each chapter.

Odd Subtree Packing

Chapter 2 examines a generalization of the problem of finding a maximum matching in a graph called chain packing. In a chain packing, terminals are linked by edge disjoint chains rather than edges. The subgraph induced by the chains must also satisfy given degree constraints. It can easily be shown that a maximum chain packing can be obtained by a forest in which all non-isolated vertices have odd degree. We present a polynomial time algorithm that finds a maximum chain packing by packing odd subtrees. The main result of Chapter 2 is the presentation and proof of correctness and complexity of the algorithm. The algorithm builds on an augmenting chain theorem of deWerra and Roberts [1990] and is in the spirit of Edmonds' [1965] blossom algorithm for matching. More general conditions for blossom detection are used, allowing, for example, blossoms containing cycles of even length.

The Reversing Number of a Digraph

Chapter 3 examines sets of arcs in a tournament whose reversal makes a tournament acyclic. The problem of finding a minimum set of arcs in a tournament whose reversal makes the tournament acyclic is equivalent to the feedback arc set problem, to finding a minimum transversal of the cycles, to finding a maximum acyclic subdigraph, and to finding a ranking which minimizes inconsistencies. (We explain this and other problems in Chapter 3.) Following a question of J.-P. Barthelemy, we examine minimum sets of arcs whose reversal makes a tournament acyclic from a different perspective. Given an acyclic digraph, we define the reversing number to be the number of extra vertices in a smallest tournament in which the given digraph is a minimum reversing set. This can be viewed in the following way: given a set of inconsistencies, what is the smallest tournament in which these arise under the ranking procedure which minimizes inconsistencies?

We present some basic bounds on the reversing number. In particular, we show that the reversing number of a tournament on n vertices is an upper bound on the reversing number of any acyclic digraph on n vertices. The reversing number of a tournament on n vertices is shown to be between $2n - 4\log_2 n$ and $2n-2$. We determine exact values for the reversing number of directed stars, complete bipartite digraphs, sets of disjoint arcs, and alternating paths. We also show that for $n \geq 9$, there exist connected acyclic digraphs with reversing number 0 and examine the largest number of arcs in a graph with reversing number 0.

Bounded Discrete Interval Order Representations

Finite interval orders are partial orders which can be represented by 'strictly to the right of' on a set of real intervals. (A formal definition will be given in Section 1.4.) Interval orders arise in preference rankings for which indifference is not transitive, orderings of temporal events, and scheduling problems. In each of these cases it seems reasonable to place bounds on the length of the intervals and to limit the set of endpoints to a discrete set. Fishburn [1983, 1985a] has examined bounded interval orders. Bogart and Stellpflug [1989, 1990] examine bounded discrete semiorders (semiorders are a special class of interval orders); semiorders that have an interval representation for which each interval has integral endpoints and a given length k. Following a question posed by K.P. Bogart, we examine bounded discrete interval orders, in particular, interval orders that have an interval representation for which the intervals have integral endpoints and length between some given bounds. Making use of Farkas' Lemma and an integer programming formulation, we reduce the problem to detecting negative cycles or finding shortest paths in an associated digraph. This provides a polynomial algorithm for determining if an order has a bounded discrete representation given general bounds on the interval length. Additionally, the digraph model is used as a basis to determine necessary and sufficient conditions for representability in the cases that the length bounds are constant and the lower bound is 0 or 1. In these cases, the family of minimal orders with no representation is also examined.

In the rest of this chapter we review the basic notation used in this thesis, and review necessary terminology and results from graph theory, order relations, and computational complexity.

1.2 Notation

The following table gives some of the basic notation used in this thesis. Notation and definitions for graphs, orders, and computational complexity is described in the following subsections.

1.3 Graphs

In this section we review basic definitions and concepts of graph theory. More information can be found in any standard text on graph theory, such as Bollobás [1979] or Harary [1972]. Due to the lack of agreement on standard notation for some of the

Table 1.1: Notation

Re	reals
$\mathbf{Z}$	integers
$\mathbf{Z}^+$	positive integers
$\mathbf{N}$	non-negative integers
$(\mathbf{Z} \setminus 2\mathbf{Z})^+$	odd positive integers
$\in$	inclusion
$\emptyset$	empty set
$\subseteq$	subset
$\subset$	proper subset
$\cap$	set intersection
$\cup$	set union
$A \setminus B$	setminus , $\{x : x \in A, x \notin B\}$
$\Leftrightarrow$	if and only if
$\|A\|$	cardinality of set A
$A \oplus B$	symmetric difference $(A \cup B) \setminus (A \cap B)$
$\lfloor x \rfloor$	greatest integer $\leq x$
$\lceil x \rceil$	least integer $\geq x$
$\square$	end of proof

terms, we will state carefully the notation used in this thesis.

A *graph* $G = (V, E)$ is a set V of *vertices* together with a set of unordered pairs $\{v, w\} \subseteq V$ called *edges*. A *directed graph* $D = (V, A)$ is a set V of vertices together with a set of ordered pairs $(v, w) \subseteq V \times V$ called *arcs*. We will assume in this thesis that V is finite and that there are no repeated edges or arcs. We will call a directed graph a *digraph*. For an arc (v, w) in a digraph, v is called the *tail* and w is called the *head*. The *reversal* of a set A of arcs is $A^R = \{(w, v) : (v, w) \in A\}$.

A *subgraph* $G' = (V', E')$ of a graph $G = (V, E)$ has $V' \subseteq V$ and $E' \subseteq E$. A subgraph of G *induced* by $V' \subset V$ has edge set $E' = \{\{v, w\} \in E : v, w \in V'\}$. That is, the edge set of an induced sugraph contains all edges with both vertices in V'. *Subdigraph* and *induced subdigraph* are defined similarly. The notation $H|_X$ denotes the subgraph (or subdigraph) induced by the set $X \subset V$.

A *chain* $C = v_1, v_2, \ldots, v_n$ in a graph (digraph) is a sequence of vertices such that $\{v_i, v_{i+1}\}$ is an edge ((v_i, v_{i+1}) is an arc) for $i = 1, \ldots, n-1$. A *closed chain* has

$v_1 = v_n$. A *path* is a chain in which the vertices v_i are distinct. A *cycle* is chain for which $v_1 = v_n$ and all other vertices are distinct. We will call a cycle $v_1, \ldots, v_n$ with $v_1 = v_n$ an $(n-1)$-cycle. Here the number $n-1$ denotes the number of distinct vertices on the cycle. In a path, if $v_1 \neq v_n$, the vertices v_1, v_n are *endpoints* and the vertices v_i, $i = 2, \ldots, n-1$, are *interior*. If $C_1 = v_1, \ldots, v_n$ and $C_2 = w_1, \ldots, w_m$ are two chains such that $\{v_n, w_1\} \in E$ ($(v_n, w_1) \in A$) then C_1, C_2 is the chain $v_1, \ldots, v_n, w_1, \ldots, w_m$. A *Hamiltonian path* contains every vertex $v \in V$.

We will also use another notion of a path in a digraph. An *alternating path* is a digraph whose underlying graph is a path when directions on the arcs are ignored. Following the path, the direction on the arcs will alternate. An alternating path on vertices $\{v_1, \ldots, v_n\}$ has arc set $\{(v_i, v_{i+1}), (v_i, v_{i-1}) : i \text{ is odd, and both vertices are in } V\}$ or the reversal of this arc set.

The *reversal* of a chain (path, cycle) $C = v_1, v_2, \ldots, v_n$ is the chain (path, cycle) $C^R = v_n, v_{n-1}, \ldots, v_2, v_1$ obtained by traversing the vertices in reverse order. In a digraph, the notation for a chain reversal is consistent with the reversal of a set of arcs if the chain is viewed as the set of arcs (v_i, v_{i+1}) on the chain. We will also use this notation in graphs, where the edges have an implied direction along the chain.

If lengths are assigned to the arcs of a digraph, the *length* of a chain (path, cycle) C, denoted *length*(C) is the sum of the lengths of the edges (arcs) in C. If lengths are not assigned, we will assume that all lengths are one, so that the length when no weights are assigned is simply the number of edges (arcs) in C.

The *degree* of a vertex v in a graph G denoted $d_G(v) = |\{\{v, w\} : \{v, w\} \in E\}|$, is the number of edges containing v. Similarly, in a digraph D, the *outdegree* of a vertex v is $d_D^+(v) = |\{(v, w) : (v, w) \in A\}|$, the number of arcs in which v is the tail. The *indegree*, $d_D^-(v) = |\{(w, v) : (w, v) \in A\}|$, is the number of arcs in which v is the head. A *source* in a digraph is a vertex with indegree 0. A *sink* in a digraph is a vertex with outdegree 0. In a graph, a vertex is *isolated* if it has degree 0.

A graph is *connected* if there is a path between every pair of vertices. A digraph will be called *connected* if the underlying graph, when directions on the arcs is ignored, is connected. (This is sometimes called weakly connected.)

A graph $F = (V, E)$ containing no cycles is called a *forest*. A connected forest is called a *tree*. Among the well known properties of trees that we shall use are the following. There is a unique path connecting any two vertices in a tree. Every non-trivial tree contains at least two vertices with degree exactly 1. A vertex of degree 1 in a forest is called a *leaf*.

A tree with a distinguished vertex is called a *rooted tree* and the distinguished vertex is the *root*. A vertex x will be called an *ancestor* of v in a tree T with root v_0 if x lies on the unique path in T from v_0 to v. The vertex x will be called the *parent* of v if x is an ancestor of v and every other ancestor of v is also an ancestor of x. A vertex v is a *child* of x if x is the parent of v. We will say that v is *below* x and x is *above* v if x is an ancestor of v.

A *tournament* is a digraph $T = (V, A)$ such that for each pair $v, w \in V$, exactly one of (v, w),(w, v) is an arc in A.

A digraph containing no cycles is *acyclic*. An *acyclic ordering* σ of the vertices of a digraph is a bijection σ between the vertex set V and $\{1, \ldots, |V|\}$ such that $(v, w) \in A \Rightarrow \sigma(v) < \sigma(w)$. It is well known that a digraph is acyclic if and only if it admits an acyclic ordering (Younger [1963]). We will refer to an acyclic order as the linear order on the vertices specified by σ. Note that for an acyclic tournament, there is exactly one ordering of the vertices which is an acyclic order. Thus, we will talk about *the* acyclic order of a tournament.

1.4 Orders

We review some basic notation and terminology for order relations. For more general information see Roberts [1979]. For more information on interval orders see Fishburn [1985a]. A *binary relation* $\succ$ on a set of elements A is a set of ordered pairs (a, b)

of elements of A, with the condition that (a,b) is in the set denoted by $a \succ b$. Note that the arc set of a digraph may be viewed as a binary relation and vice-versa. A set A together with a binary relation $\succ$ on A is denoted $(A,\succ)$ and called a *partial order* if $\succ$ is transitive and asymmetric on A. A relation is *transitive* if $(a \succ b$ and $b \succ c \Rightarrow a \succ c)$. A relation is *asymmetric* if $(a \succ b \Rightarrow$ not $b \succ aA)$. A *weak order* is a partial order which is *negatively transitive.* A relation is negatively transitive if not $a \succ b$ and not $b \succ c \Rightarrow$ not $a \succ c$. The definitions we use for partial and weak orders are those of strict partial orders and strict weak orders in Roberts [1979]. A *linear order* is a weak order which is complete, i.e., if $a \neq b$ then either $a \succ b$ or $b \succ a$.

An (induced) suborder $(A', \succ')$ of an order $(A,\succ)$ has elements $A' \subseteq A$ and $\succ'$ on A' given by the restriction of $\succ$ to A'. A suborder is *proper* if $A' \neq A$.

A *chain* in a partial order is a sequence $C = a_1 \succ a_2 \succ \cdots \succ a_n$. That is, a chain is a linear suborder. This terminology is consistent with that of a digraph, since a chain in a partial order is a chain in the digraph with arcs corresponding to $\succ$. In fact, this chain is a path; however we will use the terminology chain to be consistent with the partial order literature.

A chain $x_1 \succ x_2 \succ \cdots \succ x_k$ in $(A,\succ)$ will be denoted $x_1 \succ^{k-1} x_k$. Here the superscript for $\succ$ indicates the number of $\succ$ terms appearing in the chain. Similarly, an *incomparability chain* is a sequence $x_1 \sim x_2 \sim \cdots \sim x_k$ and is denoted $x_1 \sim^{k-1} x_k$. We also use this notation for mixed chains. Thus $x \succ^{\eta_1} \sim^{\eta_2} \succ^{\eta_3} y$ would indicate a sequence of relations from x to y with the first η_1 symbols $\succ$, the next η_2 symbols $\sim$, and the last η_3 symbols $\succ$. Elements appearing in the sequence need not be distinct.

In a partial order $(A,\succ)$, the derived relation *incomparability* is given by $(i \sim j \Leftrightarrow$ not $i \succ j$ and not $j \succ i)$. Another derived relation is $(i \succeq j \Leftrightarrow i \succ j$ or $i \sim j)$. An element a_0 in a partial order $(A,\succ)$ is *maximal* with respect to $\succ$ if $a_0 \succeq a$ for all $a \in A$. An element is *minimal* with respect to $\succ$ if $a \succeq a_0$ for all $a \in A$. Note that in a linear order, maximal and minimal elements are unique. A *Hasse diagram* representing a partial order is a drawing in the plane such that if $a \succ b$ and there is no c such that

$a \succ c \succ b$, then a is above b in the Hasse diagram and there is a line from a to b. By transitivity, this diagram specifies the entire relation. There is no line or sequence of lines connecting a and b if $a \sim b$.

An *interval order* $(A, \succ)$ is a partial order such that $\succ$ satisfies ($a \succ x$ and $b \succ y \Rightarrow a \succ y$ or $b \succ x$). It is well known (see, for example Fishburn [1985a]) that $(A, \succ)$ is an interval order if and only if there is a map J from A to a set of closed intervals denoted $J(i) = [l_i, r_i]$ in some linearly ordered set $(Y, >_0)$ such that

$$i \succ j \Leftrightarrow l_i > r_j. \tag{1.1}$$

That is, the interval for i is strictly 'greater than' the interval for j. When A is countable, the linearly ordered set can be taken to be the reals under $>$. We will call such a map a *closed real representation* of the interval order. In the finite case, we can also consider *open real representations* which are maps from A to the set of open intervals denoted $J(i) = (l_i, r_i)$ satisfying $i \succ j \Leftrightarrow l_i \geq r_j$.

In terms of the interval representations, the derived relations $\sim$ and $\succeq$ in an interval order with a closed real representation satisfy

$$i \sim j \Leftrightarrow l_i \leq r_j \text{ and } l_j \leq r_i$$

and

$$i \succeq j \Leftrightarrow \text{ not } j \succ i \Leftrightarrow r_i \geq l_j.$$

A *semiorder* is an interval order which also satisfies ($a \succ b$ and $b \succ c \Rightarrow a \succ d$ or $d \succ c$) for all $a, b, c, d \in A$. A finite semiorder has a real representation (1.1) for which all the intervals have the same length.

The co-comparability graph of a partial order $(A, \succ)$ is the graph with vertex set A and $\{a, b\}$ an edge if and only if $a \sim b$. An *interval graph* is the co-comparability graph of an interval order and an *indifference graph* is the co-comparability graph of a semiorder. Note that there is a unique interval graph which is the co-comparability graph corresponding to an interval order. However, given an interval graph, there may be several different interval orders for which the interval graph is a co-comparability graph of the order.

If $(A, \succ)$ is an interval order and G is its co-comparability graph, then $(A, \succ)$ is called an *agreeing order* for G. An *interval representation* of an interval graph G is a set of intervals $J(a)$ for $a \in V(G)$ such that $\{a, b\}$ is an edge of G if and only if $J(a) \cap J(b) \neq \emptyset$. An interval representation (1.1) of an order agreeing with G is also an interval representation of G. See Hanlon [1982] or Fishburn [1985a] for more information on the relationship between interval orders and interval graphs and Golumbic [1980] for more information on properties of interval graphs.

1.5 Complexity

We briefly review notation and terminology related to computational complexity. For more details see Garey and Johnson [1979] or Aho, Hopkroft, and Ullman [1974]. A non-negative function $f(n)$ is $O(g(n))$ if there exists a constant c such that $f(n) \leq cg(n)$ for all $n \geq 0$. We will use this notation with $f(n)$ equal to the worst case time (over all inputs) required by an algorithm to indicate the complexity of an algorithm.

Informally, a problem of size n is in the class NP if a solution has a certificate which can be verified in time $O(p(n))$ where $p(n)$ is a polynomial in n. A problem is *polynomially solvable* if there is an algorithm finding a solution in time $O(p(n))$ where $p(n)$ is a polynomial in n. A problem P is NP-complete if it is in NP and if the existence of a polynomial algorithm for P would imply the existence of polynomial algorithms for every problem in NP. It is unknown whether there are polynomial algorithms for NP-complete problems.

Chapter 2
Odd Subtree Packing

2.1 Introduction

A *matching* in a graph is a collection of edges, no two of which share a common vertex. That is, a matching is a subgraph consisting of a vertex disjoint collection of edges. Matchings have been extensively studied; see for example Lovász and Plummer [1986] or Lawler [1976] for summaries of work on matchings. Alternatively, a matching can be viewed as a collection of chains containing one edge, such that no two chains share a common edge and such that each vertex in the subgraph induced by the edges has degree at most one. While this second definition may seem redundant or roundabout, it provides a basis for a natural generalization of matching.

A *chain packing* is an edge disjoint collection of chains, each chain joining two distinct endpoints or terminals, with additional constraints at each vertex limiting the incidence of edges from the chains in the collection. More formally, we make the following definition.

Definition 2.1 *Given a graph $G = (V, E)$ and positive integer constraints $b : V \rightarrow \mathbf{Z}^+$, a* **chain packing** *in G is a collection P of edge-disjoint chains such that the endpoints of these chains are all distinct and such that in the subgraph $H = (V, E(P))$ formed by the edge set $E(P)$ of the chains, $d_H(v) \leq b(v)$ for all $v \in V$. The* **size** *of a chain packing is the number of chains $|P|$ in the collection.*

Chain packings for which the chains must have an odd number of edges have been studied in deWerra [1984,1987,1989] and Pulleyblank and deWerra [1989]. The general case of chain packings, which we will examine, has been studied in deWerra and Roberts [1990].

Matching arises as a special case when the degree constraints limit each vertex to a single incident edge. With this constraint, the chains joining the terminals must consist of a single edge, producing a matching.

Remark 2.1 Whenever $b(v) = 1$ for all $v \in V$, a chain packing P is a matching with $|P|$ edges.

A basic problem in the study of matchings is that of determining a maximum cardinality matching. We will be interested in examining maximum size chain packings, a special case of which is maximum cardinality matching. deWerra and Roberts [1990] find an augmenting chain theorem for maximum size chain packing, which generalizes the well known theorem of Berge [1957] for maximum cardinality matching. This generalized augmenting chain theorem will be used to develop a polynomial algorithm for determining a maximum size chain packing that is in the spirit of Edmonds' [1965] blossom algorithm for maximum cardinality matching.

Note that if $H = (V, E(P))$ is the subgraph formed by the edges $E(P)$ of a chain packing P, then H has $2|P|$ vertices of odd degree corresponding to the endpoints of the chains. Conversely, given a subgraph H with $2|P|$ odd degree vertices, we can decompose it into $|P|$ chains plus possibly some cycles using the following decomposition described in deWerra and Roberts [1990]. Find a path in H connecting two odd degree vertices, add this to the collection P and repeat the process for $H \setminus P$, continuing until what is left has only even degree vertices. A graph with only even degree vertices is called *eulerian*. It is well known that eulerian graphs can be decomposed into a collection of (edge disjoint) cycles. The decomposition of H into chains is not unique, but will always contain the same number (one-half the number of odd degree vertices) of chains.

Thus we see that chain packings can be examined from the viewpoint of subgraphs in a graph rather than collections of chains.

Remark 2.2 Given a graph $G = (V, E)$ and constraints $b : V \to \mathbf{Z}^+$, the maximum number of vertices with odd degree in a subgraph H satisfying $d_H(v) \le b(v)$ for all $v \in V$ is equal to twice the maximum size $|P|$ of a chain packing in G.

For simplicity, we will refer to the subgraphs H as chain packings.

Viewing maximum size chain packings from the perspective of subgraphs, we can readily see that cycles in the subgraph are not necessary. Removing a cycle does not change the parity or increase the degree of any vertex. By successively deleting cycles from a chain packing, we obtain a new chain packing of the same size, containing no cycles. Thus it is enough to consider forests when looking for a maximum size chain packing. Furthermore, consider a forest F' containing a non-isolated vertex w with even degree. We can find a path in F' from w to a vertex x with $d_{F'}(x) = 1$ (a leaf in the forest F'). By deleting such a path, we obtain a new forest in which x is isolated and $d_{F'}(w) = d_F(w) - 1$. So w has odd degree. Also, the degree of every other vertex is unchanged or reduced by 2. Removing paths from non-isolated vertices of even degree produces a forest F with the same number of odd degree vertices as F' and $d_F(v) \le d_{F'}(v)$ for all $v \in V$. Thus, a maximum size chain packing can always be realized by a forest in which all the non-isolated vertices have odd degree.

Definition 2.2 *An* **odd forest** *in a graph* $G = (V, E)$ *is a subgraph* $F = (V, E(F))$ *which is a forest such that for all* $v \in V$*, either* $d_F(v) = 0$ *or* $d_F(v)$ *is odd. Given positive odd integer constraints* $b : V \to (\mathbf{Z} \setminus 2\mathbf{Z})^+$*, a* **feasible odd forest** *is an odd forest such that* $d_F(v) \le b(v)$ *for all* $v \in V$*. The* **size** *of an odd forest is the number of vertices with odd degree.*

We will also say that a forest is *feasible* at a particular vertex v if $d_F(v) \le b(v)$. Note that we assume that the constraints are odd integers since all non-isolated vertices have odd degree. From the reductions of chain packings described above, we have the following remark.

Remark 2.3 Given a graph $G = (V, E)$ and constraints $b : V \rightarrow (\mathbf{Z} \setminus 2\mathbf{Z})^+$, the maximum size of a feasible odd forest in G is equal to twice the maximum size $|P|$ of a chain packing in G.

We will show in Section 2.6 that the above equivalence does not hold for a weighted version of chain packing. From the above remark, another (different) view of the generalization of matching which we are considering is that matched edges are replaced by feasible odd subtrees. That is, an odd forest packing is a vertex disjoint packing of feasible odd subtrees. This is similar to the general problem of vertex disjoint packings of subgraphs in a graph. However, because of the feasibility requirements from the degree constraints, odd subtree packing is not equivalent to the general packing problem. Vertex disjoint packings in graphs are studied in Cornuejols, Hartvigsan and Pulleyblank [1982], Hell and Kirkpatrick [1984,1986], and Kirkpatrick and Hell [1983].

The main result of this chapter is the presentation of a polynomial algorithm which yields the following result.

Theorem: *A maximum size feasible odd forest in a graph $G = (V, E)$ can be found in $O(|V|^3)$ time.*

We will call the problem of finding a maximum size feasible odd forest in a graph *odd subtree packing*. In Section 2.2 we review the basic results of deWerra and Roberts [1990] and discuss the idea of blossoms introduced by Edmonds [1965] for his matching algorithm. These will provide the framework for the development of the algorithm for odd subtree packing. In Section 2.3, we will present the algorithm and in Section 2.4 we will prove the correctness of the algorithm. In Section 2.5 we briefly describe a min-max formula providing a certificate that an odd subtree packing does indeed have maximum size.

2.2 Basic Results

We will briefly review some of the ideas from matching which are generalized to the case of chain packing (odd subtree packing). For more details see Edmonds [1965], Lovász and Plummer [1986], Lawler [1976] or just about any textbook on combinatorial optimization. Let M be a matching in a graph G. Let $P = x_1, x_2, \ldots, x_{2n}$ be a path in G such that the edges alternate between edges in M and edges not in M, with $\{x_{2i}, x_{2i+1}\} \in M$ for $i = 1, \ldots, n-1$ and $\{x_{2i-1}, x_{2i}\} \notin M$ for $i = 1, \ldots, n$. Additionally, if no edges from M are adjacent to x_1 and x_{2n}, then the symmetric difference $M \oplus P$ is a matching containing one more edge than M. Such a path P is called an *augmenting path* with respect to M. Clearly, if there is an augmenting path with respect to a matching M, then M is not maximum. Berge [1957] shows that the converse also holds. See also Norman and Rabin [1959] for a related result.

Theorem 2.1 (Berge 1957) *A matching M has maximum cardinality if and only if there is no augmenting path with respect to M.*

Thus an algorithm to find a maximum cardinality matching can be based on a search for an augmenting path from a vertex that is not the end of an edge in the matching. Searching all paths is inefficient and using a search tree presents the potential problem of not detecting an augmenting path. Consider the graph in Figure 2.1. The path $P = x_1, x_2, x_3, x_4, x_5, x_6$ is augmenting. However, if the search tree proceeds along the path $x_1, x_2, x_3, x'_4, x'_5$, the augmenting path P will not be detected. A potential augmenting path from x_1 to one of the vertices on the cycle $x_3, x_4, x_5, x'_5, x'_4$ can be continued with an edge not in M if the cycle is traversed in one direction but not in the other. In general a path with edges alternating in and out of M attached to an odd length cycle with edges alternating except where the cycle meets the path has this same property. Edmonds [1965] called such a structure a *blossom* and developed a labeling scheme to detect and 'shrink' blossoms.

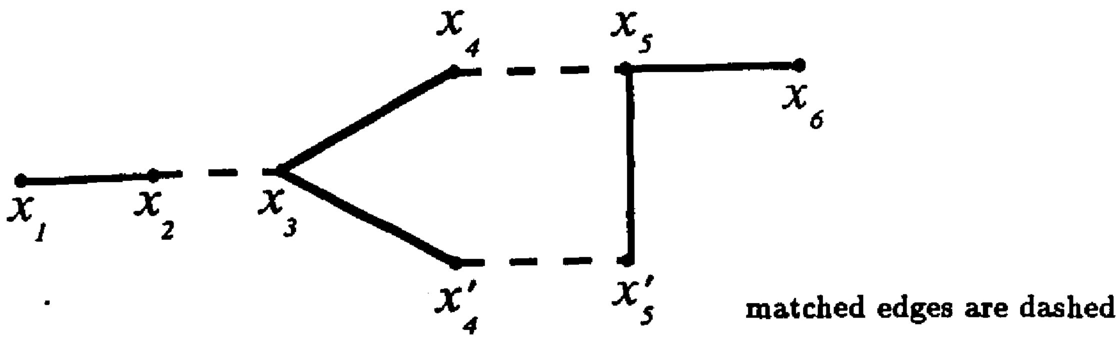


Figure 2.1: A blossom in matching.

An augmentation similar to that in matching occurs in chain packing. In this case it will be necessary to allow chains with repeated vertices in addition to simple paths. If there are no degree constraints, the problem of finding a maximum chain packing is trivial. A *component* of a graph is a maximal connected subgraph. An *odd component* has an odd number of vertices. If there are two vertices having even degree in the current H that are in the same component of G, we find a chain C joining the vertices. Then $H \oplus C$ has two more odd degree vertices. So when there are no degree constraints, a maximum chain packing in G has $|V| - q$ odd vertices where q is the number of odd components of G.

The process of augmentation with C to $H \oplus C$ described above for the case with no degree constraints will fail when constraints are present only if the degree of some vertex is increased over its constraint in $H \oplus C$. Thus, in analogy to matching, a chain C such that that $H \oplus C$ satisfies the degree constraints and such that the endpoints of C have even degree in H will be called *augmenting* with respect to H. Clearly, if there is an augmenting chain with respect to a chain packing H, then H does not have maximum size. deWerra and Roberts show that the converse also holds.

Theorem 2.2 (deWerra and Roberts 1990) *A chain packing H has maximum size if and only if there is no augmenting chain with respect to H.*

Remark 2.4 Theorem 2.2 holds for the the definition of augmenting chain given in the previous paragraph and for a restricted definition of augmenting chains used by deWerra and Roberts which is given in Definition 2.3. The restricted definition limits a vertex to appear at most once, and then only under certain conditions. We will assume in what follows that the definition of augmenting chains is that given in definition 2.3. However, for the purposes of describing the algorithm in the rest of the section and in Section 2.3, it is only necessary to use the property that augmenting chains are such that $H \oplus C$ satisfies the degree constraints and the endpoints of C have even degree in H. For clarity we will delay presentation of the more restrictive definition of augmenting chain until needed for the proof of correctness of the algorithm.

Figure 2.2 gives an example showing that it is necessary to include the possibility of repeated vertices in the augmenting chains.

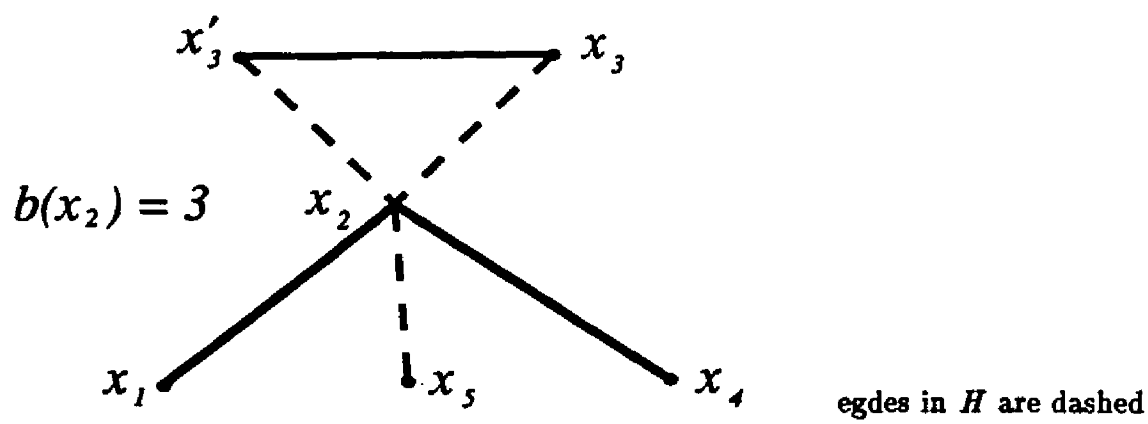


Figure 2.2: A feasible odd forest for which the only augmenting chains $C = x_1, x_2, x_3, x'_3, x_2, x_4$ and $C = x_1, x_2, x'_3, x_3, x_2, x_4$ contain a repeated vertex.

Remark 2.5 By Theorem 2.2 and the equivalence of chain packing and odd subtree packing, a feasible odd forest F has maximum size if and only if there is no augmenting chain with respect to F.

As with matching, the algorithm for finding a maximum size odd forest in a graph G will search for an augmenting chain from an isolated vertex. deWerra and Roberts

[1990] describe such an algorithm in the case that the underlying graph is a tree. The algorithm described in this chapter finds a maximum size feasible odd forest and thus also a maximum size chain packing in general graphs. The algorithm will search for an augmenting chain using an approach which generalizes the notion of blossoms used in matchings. During the search process, certain cycles are detected for which an augmenting chain can be extended from every vertex by following an appropriate direction around the cycle. Such cycles will be called blossoms. Blossoms will be more formally defined by Update 2 in the description of the algorithm.

A labeling scheme for the vertices will be used to implicitly 'shrink' the blossoms when they are detected, allowing traversal of the blossom cycles in either direction. The augmenting chains we consider are more general than those for matching and thus the conditions for formation of blossoms will be extended to guarantee detection of an augmenting chain. See Figure 2.3 for one example of a blossom in the case of chain packings. In this case, x_1, x_2, x_3, x_8 is not an augmenting chain. However, travers-

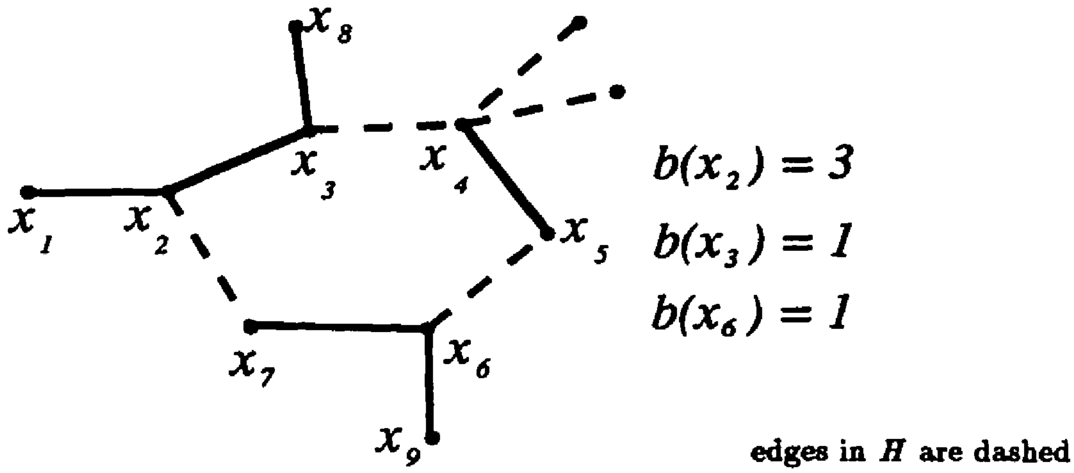


Figure 2.3: An example of a blossom in chain packing.

ing the cycle $x_2, x_3, x_4, x_5, x_6, x_7, x_2$ in the opposite direction yields an augmenting chain $x_1, x_2, x_7, x_6, x_5, x_4, x_3, x_8$. Similarly, x_1, x_2, x_7, x_6, x_9 is not augmenting while $x_1, x_2, x_3, x_4, x_5, x_6, x_9$ is augmenting. Blossoms in chain packings may contain even cycles (while in matching the cycles must have odd length).

Note that for C an augmenting chain with respect to an odd forest F, it is not immediate that $F \oplus C$ is again an odd forest; $F \oplus C$ may contain cycles or non-isolated even degree vertices. The reduction to an odd forest described before Remark 2.3 could be used in the case that $F \oplus C$ is not a feasible odd forest. However, the use of this reduction is not necessary. By specifying a search order, our algorithm will always construct augmenting chains C such that the $F \oplus C$ are feasible odd forests. This is not the case for the chain packings called short chain packings by deWerra and Roberts [1990]. (A *short chain packing* is a chain packing for which there is a decomposition in which every chain has length one or two.) deWerra and Roberts [1990] give an example of a short chain packing for which every augmentation creates a new graph which is not a short chain packing.

Finally, before presenting the algorithm, we briefly note a construction pointed out by an anonymous referee, which reduces the problem of finding a maximum cardinality chain packing to a weighted capacitated b-matching problem.

Given a graph $G = (V, E)$ with 'loops' $x_l(v)$ for each vertex $v \in V$, capacities $b(v)$ for each vertex, and capicities c_e and weights w_e on the edges (including the loops), a weighted capacitated b-matching is an assignment of non-negative integers x_e to the edges (including the loops) such that $x_e \leq c_e$ and for each $v \in V$, $2x_l(v) + \sum x_e = b(v)$ where the sum is over all edges with one end v. The weight of the b-matching is $\sum w_e x_e$ where the sum is over all edges. There are known polynomial procedures for b-matching. See for example Anstee [1987] and the references there.

The construction reducing maximum chain packing to weighted capacitated b-matching a similar to a construction used in Edmonds and Johnson [1973]. Let a graph $G = (V, E)$ and constraints $b(v)$ on the vertices be given. (Assume that the constraints are all odd.) Let each edge in E have weight 0 and capacity 1. Add a new vertex w and new edges $\{w, v\}$ for all $v \in V$. Let $b(w) = |V| + 1$ if $|V|$ is odd and let $b(w) = |V|$ if $|V|$ is even. Let the new edges have capacity 1 and weight 1. Finally, for each vertex (including the new vertex w) add a loop with weight 0 and 'large' capacity (at least $b(v)/2$). Then, it can be seen that a minimum weight capacitated b-matching in the new graph corresponds to a maximum size chain packing in the original graph and vice versa. (The number of new edges which must be used corresponds to the number of vertices which have even degree in the original graph.)

Although this construction implies a polynomial procedure for finding a maximum cardinality chain packing (and thus a maximum cardinality odd subtree packing), our direct algorithm is more efficient and provides insight into the structure of odd subtree packings. In particular, it is used in proving a min-max formula (Theorem 2.13) providing a 'certificate' that an odd subtree packing has maximum size.

2.3 The Algorithm

In this section we first briefly describe the main steps of the algorithm. Then, a detailed statement of the algorithm in a pseudo programming language format will be presented at the end of the section. Some short worked examples are given in Appendix 2.7 in order to illustrate the operation of the algorithm.

If T is a graph and B is a set of edges such that the ends of each edge in B are vertices in $V(T)$, then we will denote by $T \cup B$ the graph with vertex set $V(T)$ and edge set $E(T) \cup B$. If B consists of a single edge $\{x, y\}$, then $T \cup \{x, y\}$ will abbreviate $T \cup B$. It is well known that if T is a tree and $x, y \in V(T)$, then $T \cup \{x, y\}$ contains a unique cycle, consisting of the edge $\{x, y\}$ along with the unique path in T from x to y.

Let a graph $G = (V, E)$ with odd positive integer constraints on $V(G)$ be given. The main subroutine in the algorithm for finding a maximum size feasible odd subforest in G is called SEARCH. SEARCH is a procedure that searches for an augmenting chain with respect to a given feasible odd forest F in G. This is accomplished by *growing* (i.e., by adding new vertices and edges) a rooted search tree T on $V(T) \subseteq V(G)$ with labels on $V(T)$. The initial tree consists of a single vertex v_0, which is isolated in the current forest F. This vertex will become the root as the tree grows.

In the process of growing the tree T, certain cycles, which we will call blossoms, will be detected and used to update the labels. (A description of when blossoms are formed is given by Update 2 below.) The set of edges, which form blossoms will be denoted by B. If an isolated vertex w is added to $V(T)$, we will be able to construct an augmenting chain C between w and v_0 using the edges of $T \cup B$. This chain is constructed by the subroutine CHAIN. We then form a new feasible odd forest $F' = F \oplus C$ with fewer isolated vertices than F. The graph G' used for the next SEARCH procedure is set to the current graph G.

If SEARCH stops before adding an isolated vertex to $V(T)$, then it will be shown that there is no augmenting chain containing v_0 with respect to F. Furthermore, we will show in Lemma 2.11 that if T is the search tree formed at the end of a SEARCH procedure which failed to detect an isolated vertex, then, for C' an augmenting chain with respect to F, there is no augmenting chain with respect to $F \oplus C'$ containing a vertex of $V(T)$. Thus, when SEARCH fails to find an isolated vertex, we can restrict the search for an augmenting chain to the subgraph G' induced by $V(G) \setminus V(T)$. That is, we delete from the remaining SEARCH procedures all vertices of the tree T to get the vertex set $V(G')$. Then, we form a new forest F' which is the subforest of F induced by $V(G')$. Note that F' has fewer isolated vertices than F.

In both the case when an augmenting chain C is detected, with $F' = F \oplus C$, and the case that the search fails to find an augmenting chain, with F' the subforest of F induced by $V(G) \setminus V(T)$, the next iteration of the algorithm consists of searching for an augmenting chain from a vertex which is isolated with respect to F'. The algorithm ends when there are no isolated vertices in F'. The SEARCH subroutine will need to be repeated at most $|V(G)|$ times, since the number of isolated vertices in F' is strictly less than the number of isolated vertices in F. (The number is reduced either by augmentation or by deletion.) A list I is used to keep track of isolated vertices which have not been examined and F^* keeps track of the forest induced in discarded search trees (for reconstruction of F in G at the end of the algorithm).

Vertices in the search tree T will be assigned one of the two labels *free* or *con* (constrained). A vertex labeled con can be updated to free but not vice-versa. These labels are assigned by Updates 1 and 2 described below. Informally, we think of vertices labeled con as those vertices v for which all chains C in $T \cup B$ from the root to v are such that $d_{F \oplus C}(v) = b(v) + 1$. Thus, in attempting to extend C from v to form a new chain C', an edge from F must be used in order that the degree constraint at v is not violated in $F \oplus C'$. The label free applies to vertices v for which a chain C can be found in $T \cup B$ such that $d_{F \oplus C}(v) \leq b(v)$. Chains from the root to free vertices can be extended with F and non-F edges.

In the subroutine SEARCH, constructing T and assigning labels consists of repeatedly applying one of two possible updates. At each step $T \cup B$ is 'improved' either by adding a new edge and vertex to T and labeling the vertex or by adding an edge to B and relabeling some vertices from con to free. We update as follows.

Update 1 — Extending the Search Tree: Examine an edge $\{v, w\} \in E(G)$ with $v \in V(T)$, $w \notin V(T)$ and either $\{v, w\} \in E(F)$ or v labeled free. In this case, add w to $V(T)$ and add $\{v, w\}$ to $E(T)$; label w con if both $\{v, w\} \notin E(F)$ and $d_F(w) = b(w)$. Label it free otherwise.

Update 2 — Forming a Blossom: Examine an edge $\{v, w\} \in E(G)$ with $v, w \in V(T)$ and $\{v, w\} \notin E(T)$ such that either $\{v, w\} \in E(F)$ or both v and w labeled free. In this case update to free all con vertices on the unique cycle formed in $T \cup \{v, w\}$.

We say that an edge is *eligible* for one of the updates if it meets the conditions for being examined in the update. During a call to the SEARCH procedure each edge will be examined at most once for one of the updates. Thus, there can be at most $|E(G)|$

updates during the construction of T. The construction of T ends when an isolated vertex is added to T or when no new updates can be performed.

We will show, by proving the correctness of the algorithm, that the order of selecting edges for updates, from those eligible, will not affect the correctness of the algorithm. However, by specifying that each subtree in the forest F be added as a subtree of T, we can insure that the augmenting chains C which are constructed are such that $F \oplus C$ is again a forest. The conditions on Update 1 allow subtrees of F to appear as subtrees of T, since once an edge from F is added to T, adjacent edges in F become eligible for an update. The proof that this order does indeed create chains which do not form cycles upon augmentation is given in Lemma 2.10.

A list L is used to keep track of edges which are eligible for either Update 1, or Update 2 that have not previously been examined by that update. From the conditions for the updates, the list L will contain edges $\{v, w\}$, which have not been examined for one of the updates, such that $v \in V(T)$, $w \notin V(T)$ and either $\{v, w\} \in E(F)$ or both v and w have label free, and edges $\{v, w\}$ such that $v, w \in V(T)$, $\{v, w\} \notin E(T)$ and either $\{v, w\} \in E(F)$ or both v and w have label free. Edges $\{v, w\} \in E(F)$ will be added to L when one of v or w is first added to $V(T)$. Edges $\{v, w\} \notin E(F)$ are added to L when one of v or w first receives the label free (either initially or during a relabeling) and the other is not in $V(T)$ and the other is not in $V(T)$. An edge $\{x, y\}$ can be removed from L. If $x \in V(T)$ with label free, $y \notin V(T)$, and $\{x, y\} \notin E(F)$, then if an edge $\{z, y\}$ is examined using Update 1, such that y is added to $V(T)$ with the label con, the edge $\{x, y\}$ will be removed from L. If, during a later Update 2, y gets relabeled free, then $\{x, y\}$ will be again added to L. From the conditions for removing an edge from L, note that each edge can be removed from L at most once. If an edge $\{v, w\}$ has been removed from L by examining it for an update, or if the edge is currently in L, we will not allow $\{v, w\}$ to be added to L. We maintain this condition by marking edges currently in L and edges which have been examined for update with *used*. When L becomes empty the search has ended.

In order to keep track of the "blossom detecting" edges used in Update 2 and mark each relabeled vertex with the edge in B causing the relabeling, we maintain two additional labels. If y was relabeled during Update 2 with edge $\{v, w\}$, we use blos1(y) to mark one of v or w which is below y in T. If both are below y (when y is the nearest common ancestor of v and w), then we set blos1(y) to be y if $y = v$ or $y = w$, otherwise, we arbitrarily pick v to be blos1(y). We use blos2(y) to mark the other end of $\{v, w\}$, i.e., the vertex v or w that is not marked as blos1(y). This marking implicitly stores the set B of "blossom forming" edges (edges which have been examinined for Update 2). These marks are also used in the construction of the augmenting chain using the CHAIN procedure.

In order to avoid excess work in determining which vertices to relabel during blossom formation (Update 2), we will maintain an additional mark nca(v) on each vertex v. Initially, nca(v) = v for all vertices. When $\{x, y\}$ is examined for Update 2, we will find the nearest common ancestor z of x and y in the search tree T. Then, for every vertex u on the unique cycle in $T \cup \{x, y\}$, and every v such that nca(v) = u, we set nca(v) to nca(z). It is not difficult to check that the revising of the mark nca(v) maintains the property that nca(v) is on the unique path in T from v to the root of T and that every vertex on this path between v and nca(v) (including v and nca(v)) has been relabeled (if necessary) by an Update 2. Thus, when a blossom is formed when examining $\{x'y'\}$ with nca(x') = nca(y'), every vertex on the cycle $T \cup \{x', y'\}$ has already been relabeled and no relabeling need occur.

When an isolated vertex v_k is added to a search tree T with root v_0, an augmenting chain has been detected. The subroutine CHAIN will be used to construct an augmenting chain C between v_0 and v_k. In the process of constructing C using CHAIN, recursive calls will be made to CHAIN. During each iteration of CHAIN, a vertex y will be under consideration to be added as the next vertex in C. For the first iteration, this vertex will be specified by the initialization. Each iteration will consist of adding y (and possibly some other vertices) to C and determining the new y to be considered during the next iteration. The two ways of extending C are as follows. Recall that during SEARCH, when a vertex y is relabeled to free during Update 2 (blossom formation), $\{\text{blos1}(y), \text{blos2}(y)\}$ denotes the edge examined during that update.

Extension 1 — Moving Up the Search Tree: Add y to C and set the new y to parent(y).

Extension 2 — Traversing a Blossom: Recursively call CHAIN to construct a chain C' from blos1(y) to y. Add the reverse of C' to C and set the new y to be considered to be blos2(y).

When CHAIN is being used to construct a chain C from w to v (where v is not necessarily the root of T, e.g., during a recursive call), the procedure will stop when v is added to C.

During the initial iteration of a call to CHAIN, we will use a boolean variable M, along with the labeling of the first vertex considered in the call, to determine whether to perform Extension 1 or Extension 2 fo rthe first iteration. M will be $\emptyset$ (i.e., not assigned a value) for the first call to CHAIN. Its value (*true* or *false*) for recursive calls to CHAIN will be determined during Extension 2 when the recursive call is made. Note that when Extension 2 is performed during the initial step of a recursive call to CHAIN, the initialization consists of a further recursive call to CHAIN.

During the remaining (non-initial) iterations of a call to CHAIN, the last vertex in the part of C which has already been constructed will be used to determine whether to perform Extension 1 or Extension 2. This vertex (the vertex preceding y in the chain), will be denoted prec(y). Then, if $\{\text{prec}(y), y\} \notin F$ and if y was relabeled free, Extension 2 will be performed. Otherwise, if $\{\text{prec}(y), y\} \in F$ or if y was initially free, Extension 1 will be performed.

We note that recursive calls to CHAIN may not be well defined. If we are constructing a chain from v to u, the chain from v may not include u (for example, if u

is below v in T). However, Lemma 2.6 shows that recursive calls are properly defined and Lemma 2.9 will be used to show that the chains which are produced by CHAIN are indeed augmenting chains.

We now are ready to present a detailed statement of the algorithm in a pseudo programming language format. Variables and lablels which are not initialized are assumed to have an initial value of $\emptyset$ (i.e., are not assigned a value).

Odd Subtree Packing Algorithm

Input: Graph $G = (V, E)$ and odd positive integer degree
  constraints $b(v)$ for $v \in V(G)$.
Initialization: Set $I \leftarrow V(G)$ and $F, F^* \leftarrow \emptyset$.
While $|I| \geq 2$, repeat for any $v_0 \in I$:
  Call SEARCH(v_0,G,F).
  If SEARCH finds an augmenting chain C, then:
    Set $F \leftarrow F \oplus C$,
    remove labels on all vertices and
    set $I \leftarrow I \setminus \{v_0, v_k\}$ (where v_0,v_k are the ends of C).
  Else if SEARCH fails to find an augmenting chain, then:
    For the failed search tree T, add $F \cap T$ to F^*,
    i.e. set $V(F^*)$ to $V(F^*) \cup (V(F) \cap V(T))$ and
    set $E(F^*)$ to $E(F^*) \cup (E(F) \cap E(T))$.
    Set $V(G)$ to $V(G) \setminus V(T)$ and $V(F)$ to $V(F) \setminus V(T)$, and
    delete from $E(G)$ and $E(F)$ edges with at least one end in $V(T)$
    (i.e. set G and F to be the subgraphs induced by $V(G) \setminus V(T)$),
    remove labels on all vertices, and set $I \leftarrow I \setminus v_0$.
Set $F^* \leftarrow F^* \cup F$ and output F^*.

Procedure SEARCH(v_0,G, F)

Input: Graph $G = (V, E)$ and odd positive integer degree

constraints $b(v)$ for $v \in V(G)$,

a feasible odd forest F in G,

and a vertex v_0 with $d_F(v_0) = 0$.

Initialization: Let T be the tree consisting of the single vertex v_0,

and let v_0 have label free.

Let L consist of all edges $\{v_0, w\} \in E(G)$.

Let $\mathrm{nca}(u) = u$ for all $u \in V(G)$.

While L is not empty repeat:

Choose $\{v, w\} \in L$ such that $\{v, w\} \in F$ and

exactly one of v or w is in T;

if there are none choose any $\{v, w\} \in L$.

Update 1: If w is unlabeled, add $\{v, w\}$ to T, (set parent(w) $\leftarrow v$), then:

If $d_F(w) = 0$ (is isolated in F),

label w free and

construct an augmenting chain between v_0 and w

(call procedure CHAIN(w, v_0, T, F, G)), and stop.

Else, if $\{v, w\} \notin E(F)$ and $d_F(w) = b(w)$, then:

Label w con.

Add $\{w, x\}$ to L for all $\{x, w\} \in E(F)$

such that $\{w, x\}$ is not marked *used*,

mark the edges added to L with *used.*

Delete from L edges $\{z, x\}$ such that $\{z, x\} \notin E(F)$

and z is labeled con,

remove the mark *used* from the deleted edges.

Else, if $\{v, w\} \in E(F)$ or $d_F(w) \leq b(w) - 2$, then:

Label w free and add $\{w, x\}$ to L for all $\{x, w\} \in E(G)$

such that $\{w, x\}$ is not marked *used*,

mark the edges added to L with *used*.

(End of Update 1.)

Update 2: Else, if w is labeled (is already in T), and nca(w) = nca(v) then:

do nothing. (Update 2 in this case will cause no relabeling.)

Else, if w is labeled (is already in T), and nca(w) $\neq$ nca(v) then:

Find the unique cycle S in $T \cup \{v, w\}$ and for each vertex

y on S which has label con, relabel y with free.

Set blos1(y) $\leftarrow$ (y if y is v or w, or one of v or w

that is below y in T; if both are below y, then set blos1(y) to be y

if $y = v$ or $y = w$, otherwise, set blos1(y) (arbitrarily) to v).

Set blos2(y) $\leftarrow$ (the vertex v or w not assigned to blos1(y)).

Add $\{x, y\}$ to L for all $\{x, y\} \in E(G)$

such that $\{w, x\}$ is not marked *used*,

mark the edges added to L with *used*.

Additionally, for every z on S, and every u

such that nca(u) = nca(z), set nca(u) to nca(r)

where r is the nearest common ancestor of v and w.

(The edge $\{\text{blos1}(y), \text{blos2}(y)\}$ is implicitly added to B.)

(End of Update 2.)

Stop (when $L = \emptyset$): search from v_0 has failed.

Procedure CHAIN(v_k,v_j,T, M)

Input: Tree $T = (V, E)$, labels free and con on $V(T)$,

labels blos1 and blos2 on a subset of $V(T)$, and edges of T identified as being in the forest F.

A boolean variable M (which may be $\emptyset$, i.e., not assigned a value), and vertices $v_j, v_k \in V(T)$ with v_j above v_k in T.

Initialization:

If $M = \mathit{true}$ and v_k was relabeled free then:

Recursively call CHAIN(blos1(v_k),v_k,T, M), with $M = \mathit{true}$ if $\{\text{blos1}(v_k), \text{blos2}(v_k)\} \notin F$ and *false* otherwise. Initialize C with the reverse of the chain obtained from this call and set $y \leftarrow$blos2(v_k) and prec(y) $\leftarrow$ blos1(v_k).

Else, if $M = \mathit{false}$ or $\emptyset$ or if v_k was initially labeled free, then:

If v_k=v_j set $C = v_k$ and stop: If in a recursive call use C to continue with extension 2, otherwise use C for augmentation.

Else initialize C with v_k and set $y \leftarrow$ parent(v_k); prec(y) $\leftarrow v_k$.

If $y = v_j$ add y to C and stop: If in a recursive call use C to continue with extension 2, otherwise use C for augmentation.

Else (while y is not v_j repeat):

Extension 1: If y was initially labeled free or $\{y, \text{prec}(y)\} \in F$, then:

Add y to C and set prec(y) $\leftarrow y$; $y \leftarrow$parent(y).

Extension 2: Else if y was relabeled free (i.e. blos1(y) $\neq \emptyset$) and $\{y, \text{prec}(y)\} \notin E(F), \mathit{then}:$

Recursively call procedure CHAIN(blos1(y),y,T, M), setting $M \leftarrow \mathit{true}$ if $\{\text{blos1}(y), \text{blos2}(y)\} \notin F$ (else set $M \leftarrow \mathit{false}$). Add to C the reverse of the chain C' obtained from this recursive call and set $y \leftarrow$ blos2(y) and prec(y) $\leftarrow$ blos1(y).

2.4 Correctness of the algorithm

In this section we prove that the algorithm correctly finds a maximum size odd forest in polynomial time. We prove several lemmas to accomplish this. Lemma 2.4 will be used to reach a contradiction if the SEARCH procedure from v_0 stops before detecting an isolated vertex when there is an augmenting chain from v_0. Lemma 2.6 shows that recursive calls to CHAIN are well defined. Lemmas 2.5, 2.7, and 2.8 are technical lemmas used in the proof of Lemma 2.9. Lemma 2.9 shows that chains produced in the CHAIN procedure are indeed augmenting chains. Lemma 2.10 shows that augmentation will not produce cycles. Lemma 2.11 shows that once a search fails from a particular isolated vertex, that vertex and vertices in the tree constructed during the failed search can be deleted without affecting augmentation from other isolated vertices.

We first describe the extra conditions on an augmenting chain that appear in the definition used by deWerra and Roberts [1990]. These allow us to restrict our search to an augmenting chain which is as 'elementary' as possible. Consider any chain C. When v_{i-1},v_i,v_{i+1} appear in sequence in C, then if exactly one of $\{v_{i-1}, v_i\}$, $\{v_i, v_{i+1}\}$ is in F, the degree of v_i is unchanged by augmentation. (A vertex v may appear more than once in the chain; we discuss here only the change caused by each particular appearance v_i.) If both edges incident to v_i are in F, its degree is decreased by two and if neither edge incident to v_i is in F, then its degree is increased by two. Only the last case can lead to a new violation of degree constraints. Call the case when both incident edges to a vertex v_i in a chain are in F a *positive irregular* crossing of v_i by C and the case when both incident edges are not in H a *negative irregular* crossing of v_i by C. A regular crossing occurs when exactly one of the incident edges is in F. It is a *positive regular* crossing of v_i by C when $\{v_{i-1}, v_i\}$ is in F and $\{v_i, v_{i+1}\}$ is not in F and a *negative regular* crossing of v_i by C when $\{v_{i-1}, v_i\}$ is not in F and $\{v_i, v_{i+1}\}$ is in F.

DeWerra and Roberts [1990] gives the following definition for an augmenting chain which restricts negative irregular crossings so that degree constraints will not be violated upon augmentation.

Definition 2.3 *$C = v_0, v_1, \ldots, v_k$ is an* **augmenting chain** *with respect to F if the following hold:*

(i) v_0 and v_k are distinct vertices with even degree in F, both of which are adjacent to exactly one edge of C;

(ii) Each vertex v_j $(0 < j < k)$ has odd degree and one of the following characteristics:

one irregular crossing and no regular crossing by C;

one regular crossing and no irregular crossing by C;

one negative regular crossing by C followed by one positive regular crossing by C and no other crossings;

(iii) If v_j $(0 < j < k)$ has a negative irregular crossing then $d_F(v) \leq b(v) - 2$.

The extra conditions in (ii) are not necessary but keep the chain as elementary as possible. A vertex may appear twice but only when it has a negative regular crossing followed by a positive regular crossing. Note that this condition and in general the definition of augmenting chain holds for the reversal $C^R = v_k, v_{k-1}, \ldots, v_1, v_0$ of C if and only if it holds for C. Condition (ii) simplifies matters by limiting a vertex to at most one negative irregular crossing. Given this, condition (iii) insures that degree conditions will not be violated in $F \oplus C$. Thus if C is an augmenting chain with respect to a packing F, then $F \oplus C$ has two more odd vertices than F and satisfies $d_{F \oplus C}(v) \leq b(v)$ for all vertices v.

We now restate Theorem 2.2 of deWerrra and Roberts [1990] to reflect the restricted definition of augmenting chain and the equivalence of chain packing and odd subtree packing, as noted in Remark 2.5. In the remainder of this chapter, when we refer to an augmenting chain, we will assume that it is an augmenting chain satisfying (i), (ii) and (iii) as in Definition 2.3.

Theorem 2.3 *A feasible odd subforest F has maximum size if and only if there is no augmenting chain as in Definition 2.3 with respect to F.*

Remark 2.6 In the following proofs, we will refer to performing the Updates 1 and 2 and Extensions 1 and 2. This will mean carrying out the steps corresponding to these updates in the formal algorithm description. Also, when referring to Update 2 we will sometimes use the more descriptive phrase *'form a blossom'* and when Extension 2 is performed we will sometimes use *'traverse a blossom'*.

We are now ready to prove the lemmas. If order to avoid an excess of notation when we refer to edges in $E(G)$ for a graph $G = (V, E)$, we will simply use 'edges in G' when there is no chance of confusion. When referring specifically to the vertex set, we will always use $V(G)$.

Lemma 2.4 *Let T be a tree resulting from a call to the procedure SEARCH which ended with L empty. Let $C = u_0, u_1, \ldots, u_k$ be a chain satisfying (ii) and (iii) in the definition of an augmenting chain with respect to F, and also the following conditions:*
(iv) $\{u_{k-1}, u_k\} \in F$
(v) if u_k appears twice with first appearance u_i for $0 < i < k$, then u_i has a negative regular crossing by C
(vi) u_0, u_1, $\ldots$, u_k all appear in T and u_1, $\ldots$, u_k are all below u_0 in T.
Then u_k is labeled free in T.

Proof: We use induction on the number of vertices in the chain C. If $C = u_0, u_1$ then by (iv), $\{u_0, u_1\} \in F$. Since u_0 is in T, the rules for adding edges to L insure that $\{u_0, u_1\}$ was put in L. If parent(u_1)=u_0, then u_1 was added to T with Update 1 while examining u_0. In this case u_1 has the label free since $\{u_0, u_1\} \in F$. If parent(u_1) is not u_0, then examining $\{u_0, u_1\}$ would result in Update 2, forming a blossom, since $\{u_0, u_1\} \in F$. This insures that u_1 has label free.

Now we assume the result is true for chains with k vertices and consider $C = u_0, \ldots u_k$. Suppose first that $\{u_{k-1}, u_k\} \notin T$. The facts that $\{u_{k-1}, u_k\} \in F$ and u_k

is in T insure that $\{u_k, u_{k-1}\}$ was put in L. Then, when $\{u_k, u_{k-1}\}$ was examined, a blossom would have formed giving u_k the label free. (Note that Update 1 could not have been used when examining $\{u_k, u_{k-1}\}$, since then $\{u_k, u_{k-1}\} \in T$.) If $\{u_{k-1}, u_k\} \in T$ then this edge was added by Update 1. If parent(u_k) $= u_{k-1}$ in T, then u_k would have been labeled free when the update was performed. So we are done in this case. Thus we may assume that $\{u_k, u_{k-1}\} \in T$ with parent(u_{k-1})=u_k.

Let f be the largest index from $0, 1, 2, \ldots, k-2$ such that u_f is not below u_k in T. This is well defined since u_0 is above u_k. Thus $u_{f+1}, \ldots, u_{k-1}$ are below u_k. There are two cases.

Case 1: u_k appears twice in C with first appearance $u_i = u_f$. Thus $u_k = u_i$ and $u_{i+1}, u_{i+2}, \ldots, u_{k-1}$ are all below u_k.

Case 2: Not case 1. That is, $u_{f+1}, u_{f+2}, \ldots, u_{k-1}$ are all below u_k, u_f is not below u_k and u_f is not u_k.

We will consider case 2 first.

Case 2:

In case 2, note that $\{u_f, u_{f+1}\} \notin T$ and that since u_k is above u_{f+1} and not above u_f, u_k is on the unique cycle S in $\{u_f, u_{f+1}\} \cup T$. Thus, if $\{u_f, u_{f+1}\} \in F$, then $\{u_f, u_{f+1}\}$ was added to L (whenever the first one of u_f, u_{f+1} was added to $V(T)$). Examining this edge would have formed a blossom, causing the procedure to relabel u_k free. So we may assume $\{u_f, u_{f+1}\} \notin F$. If both u_f and u_{f+1} have label free, then say u_f was labeled free after u_{f+1}. Then, when u_f was labeled free (either initially or via a relabeling), the edge $\{u_f, u_{f+1}\}$ would have been added to L. Then, since this edge was added after both u_f and u_{f+1} were put in $V(T)$, a blossom would have formed when $\{u_f, u_{f+1}\}$ was examined, relabeling u_k to free since it is on the unique cycle in $T \cup \{u_f, u_{f+1}\}$. Similarly if u_{f+1} was labeled free second we get u_k labeled free. Thus the proof of case 2 is complete if we can show that both u_f and u_{f+1} have label free.

If u_f has label con then $d_F(u_f) = b(u_f)$. Then if $\{u_{f-1}, u_f\} \notin F$, since $\{u_f, u_{f+1}\}$ is also not in F, u_f has a negative irregular crossing by C with respect to F and

$d_F(u_f) = b(u_f)$. This contradicts the fact that the augmenting chain C satisfies (iii). Thus $\{u_{f-1}, u_f\} \in F$ and $C' = u_0, u_1, \ldots, u_f$ satisfies condition (iv) of the lemma. Since $u_1, \ldots, u_k$ are below u_0 in T, C' satisfies (vi) and since C' is a subchain of C it satisfies (ii) and (iii). If u_f has a second appearance in C', the conditions on a vertex appearing twice in C insure that its first crossing is negative regular so (v) is also satisfied. Thus C' satisfies all the conditions of the lemma with $f+1 < k+1$ and by induction u_f has label free, a contradiction. So u_f is labeled free.

If u_{f+1} is labeled con then $d_F(u_{f+1}) = b(u_{f+1})$ and in a manner similar to the case when u_f has label con, (iii) for C and $\{u_f, u_{f+1}\} \notin F$ imply that $\{u_{f+1}, u_{f+2}\} \in F$. If u_{f+1} also appears as u_j for $f+1 < j < k-1$, then since the reversal of C satisfies (ii) whenever C does, u_j will have a negative regular crossing by C^R. By the choice of f and the fact that C^R satisfies (ii) and (iii) the chain $u_k, u_{k-1}, u_{k-2}, \ldots, u_{f+2}, u_{f+1}$ satisfies the conditions of the lemma. Moreover, $k-(f+1)+1 < k+1$. By induction u_{f+1} has label free, a contradiction. Thus both u_f and u_{f+1} must have label free and, as described above, u_k must have label free.

Case 1:

We now consider case 1. By the conditions on a vertex appearing twice in a chain, we have $\{u_i, u_{i+1}\} \in F$. If it is not the case that parent$(u_{i+1}) = u_k$ then $\{u_{i+1}, u_i\}$ would have been put in L and the blossom formed when it was examined would give the label free to $u_i = u_k$. Thus we may assume parent$(u_{i+1}) = u_k$. Recall also that parent$(u_{k-1}) = u_k$. Now consider the largest index g from $i+1, i+2, \ldots, k-2$ such that u_g is not below u_{k-1} in T. This is well defined since u_{i+1} and u_{k-1} are children of u_k so that u_{i+1} is not below u_{k-1}. Thus $u_{g+1}, u_{g+2}, \ldots, u_{k-2}$ are all below u_{k-1} in T. Now u_g is not below u_{k-1} but is below u_k, by the definition of case 1. Also, u_{g+1} is u_{k-1} or is below u_{k-1} and parent(u_{k-1})=u_k so u_k is the nearest common ancestor of u_g and u_{g+1}. By similar arguments to those in case 2, either $\{u_g, u_{g+1}\} \in F$ or induction on $u_i, \ldots, u_g$ and $u_k, u_{k-1}, \ldots, u_{g+1}$ shows that both u_g and u_{g+1} get label free. For both these situations a blossom will be formed insuring that u_k gets label free. □

Remark 2.7 For the operation of the SEARCH procedure, we will refer to an *iteration* as the process of carrying out one of the Updates 1 or 2. We will denote the initial tree in SEARCH consisting of one isolated vertex by T_0. The search tree obtained after n iterations will be denoted by T_n. Similarly, B_n will denote the set of edges which were examined resulting in Update 2 (blossom formation) during iterations $1, \ldots, n$. L_n will denote the set of edges eligible for update at the start of the n^{th} iteration. Thus, the edge examined during this iteration is in L_n.

Similarly, for the operation of the CHAIN procedure, we will refer to an *iteration* as the process of carrying out one of the Extensions 1 or 2 from the initial call to CHAIN. Thus one iteration, if it is Extension 2, may include carrying out Extension 1 or 2 a number of times during the recursive call to CHAIN. In a like manner we will refer to iterations within a recursive call to CHAIN.

Definition 2.4 *Let a labeled rooted tree $T_n = (V, E)$ (with some edges in T_n distinguished as edges of a forest F), which is produced during the n^{th} iteration (not necessarily the last) of the SEARCH subroutine, and vertices $v, u \in V(T_n)$, be given. If u_0 is the root of T_n and if u is a vertex on the chain produced by a call to CHAIN(v, u_0, T_n, M), with $M = false$, (i.e., CHAIN(v, u_0, T_n, M) is initialized with Extension 1), then denote by $C(v, u, T_n)$ the part of this chain from v to the first appearance of u. Similarly, if v was initially labeled con, and has been relabeled free, and if u is a vertex on the chain produced by a call to CHAIN(v, u_0, T_n, M) with $M = true$, (i.e., CHAIN(v, u_0, T_n, M) is initialized with Extension 2), then denote by $\tilde{C}(v, u, T_n)$ the part of the chain from v to the first appearance of u.*

Note that for v_0 the root of T_n, and any vertex $v \in V(T_n)$, the chains $C(v, v_0, T_n)$ and $\tilde{C}(v, v_0, T_n)$ always exist since the root will eventually be added a chain produced by CHAIN.

Lemma 2.5 *Let $G = (V, E)$ be a graph with odd positive integer degree constraints on $V(G)$, let F be a feasible odd forest in G and let u be such that $d_F(u) = 0$. Let $T_0, T_1, ..., T_k$ be a sequence of labeled search trees, with T_i produced by the i^{th} iteration of SEARCH. If $C(x, y, T_n)$ exists, then for $j = n, n+1, \ldots, k$, we have $C(x, y, T_j) = C(x, y, T_n)$. Furthermore, if x has been relabeled to free during the i^{th} update for $i \leq n$, and if $\tilde{C}(x, y, T_n)$ exists, then for $j = n, n+1, \ldots, k$, we have $\tilde{C}(x, y, T_j) = \tilde{C}(x, y, T_n)$.*

Proof: We show that $C(x, y, T_j) = C(x, y, T_n)$. The proof for $\tilde{C}(x, y, T_j) = \tilde{C}(x, y, T_n)$ is identical.

By the definition of $C(x, y, T_i)$, the initializations of the procedures producing $C(x, y, T_j)$ and $C(x, y, T_n)$ are the same, namely: consider vertex y and use Extension 1. (In the proof that $\tilde{C}(x, y, T_j) = \tilde{C}(x, y, T_n)$ both start with Extension 2.) If $C(x, y, T_j) \neq C(x, y, T_n)$, then consider the first time when the procedures CHAIN(x, y, T_n, M) and CHAIN(x, y, T_j, M) disagree. (Note that this may be during a recursive call to CHAIN.) In both cases, some vertex w will be under consideration (since all the preceding steps agree), and it must be that Extension 1 is carried out in one procedure and Extension 2 in the other. There are two cases to consider.

Case 1: The disagreeing step is an initialization of a recursive call. Note that the value of M for the recursive call is the same for CHAIN(x, y, T_n, M) and CHAIN(x, y, T_j, M), since all preceding steps agree. (This uses the fact that once blos1 and blos2 are defined, they do not change, and if the recursive call is made in CHAIN(x, y, T_n, M), then the blos1 and blos2 used to set the value of M must already be defined in T_n, and so will be the same in T_j.) Note that M is not null, since we are considering a recursive call to CHAIN. Then for Extension 2 to occur, it must be that w was relabeled free (had the label con and was relabeled) and $M = true$. Then, with $M = true$, for Extension 1 to occur (in the other procedure), w must have been initially labeled free, a contradiction. So the first disagreeing step can not be the initialization of a recursive call.

Case 2: The disagreeing step is not an initialization step.

It must be that prec(w) in the call to CHAIN(x, y, T_n, M) is the same as prec(w) in the call to CHAIN(x, y, T_j, M), since all preceding steps agree. Then, for Extension 1 to occur in one procedure and Extension 2 in the other, it must be that $\{w, \text{prec}(w)\} \notin F$. Also, it must be the case that w is labeled free in T_j and w is labeled labeled con in T_n. (Since labels can only be changed from con to free, it can not be the case that w is labeled free in T_n and labeled con in T_j for $j \geq n$.)

In order for the edge $\{w, \text{prec}(w)\}$ to appear in the chain that is constructed by CHAIN(x, y, T_n, M), it must be either an edge of T_n or an edge in B_n. Then, $\{w, \text{prec}(w)\}$ must have been examined during some iteration of SEARCH, say the i^{th}, with $i < n$. Also, $\{w, \text{prec}(w)\}$ must be in L_i (eligible for update) since it is examined in the i^{th} iteration. If w has the label con in T_n then, if $w \in V(T_{i-1})$, it must have the label con in T_{i-1} as $i - 1 < n$ (and labels can only update from con to free). But, if $\{w, \text{prec}(w)\} \notin F$ and if w is labeled con in T_i, then $\{w, \text{prec}(w)\}$ is not eligible for either Update, a contradiction. So it must be that $w \notin V(T_{i-1})$.

Now, we may assume $w \notin V(T_{i-1})$. In order for $\{w, \text{prec}(w)\}$ to be eligible to be examined for the i^{th} iteration, prec(w) $\in V(T_{i-1})$ with the label free (since also $\{w, \text{prec}(w)\} \notin F$). Then the i^{th} iteration is Update 1. The result is that $\{w, \text{prec}(w)\}$ is added to T_{i-1} to get T_i and prec(w) is the parent of w in T_i. Therefore prec(w) is also the parent of w in T_n. However, it is not difficult to see from the way prec(w) is defined in CHAIN, that either prec(w) is a child of w or $\{w, \text{prec}(w)\}$ is an edge of B_n, contradicting prec(w) being the parent of w in T_i. This completes the proof of case 2 and the proof of the lemma. □

The next Lemma shows that recursive calls to CHAIN are well defined.

Lemma 2.6 *Suppose $T_n = (V, E)$ is a labeled rooted tree (with some edges in T_n distinguished as edges of a forest F), which is produced during the n^{th} iteration (not* *necessarily the last) of the SEARCH subroutine. Let u be the root of T_n. Let $v \in V(T_n)$ be relabeled free during the n^{th} iteration with an Update 2. If, for some z, during some iteration of CHAIN(z, u, T_n, M), Extension 2 is performed while v is being examined, then the recursive call CHAIN(blos1(v), v, T_n, M) is well defined. That is, v is added to CHAIN(blos1(v), u, T_n, M).*

Proof: The n^{th} iteration of SEARCH results in Update 2, so $V(T_{n-1}) = V(T_n)$. Let $y = \text{blos1}(v)$ and $x = \text{blos2}(v)$. For the recursive call, M is true if $\{x, y\} \notin F$ and false otherwise. If M is false, or if y was initially labeled free, then the recursive call is initialized with Extension 1. If M is true and if y was relabeled free, the recursive call is initialized with Extension 2.

Consider the case that M is false, or y was initially labeled free. Since $\text{blos1}(v) = y \in V(T_{n-1})$, $C(\text{blos1}(v), u, T_{n-1})$ exists and by Lemma 2.5, $C(\text{blos1}(v), u, T_n) = C(\text{blos1}(v), u, T_{n-1})$. We must show that v appears on this chain, as the recursive call should produce $C(\text{blos1}(v), v, T_{n-1})$.

On the other hand, if M is true and if y was relabeled free, note first that y must have been relabeled in T_{n-1}. If not, then y was labeled con when $\{y, x\}$ was examined for Update 2. When M is true, $\{x, y\} \notin F$. Then $\{x, y\}$ would not be eligible for Update 2, a contradiction. So y was already relabeled in T_{n-1}, and by Lemma 2.5, $\tilde{C}(\text{blos1}(v), u, T_n) = \tilde{C}(\text{blos1}(v), u, T_{n-1})$. We must show that v appears on this chain, since in this case the recursive call should produce $\tilde{C}(\text{blos1}(v), v, T_n)$.

Let $D = \tilde{C}(\text{blos1}(v), u, T_{n-1})$ when M is true and y was relabeled free and let $D = C(\text{blos1}(v), u, T_{n-1})$ when M is false or y was initially labeled free. Denote $D = v_1, v_2, \ldots, v_k$ with $v_1 = \text{blos1}(v)$ and $v_k = u$. Note that we have shown for both cases that D can be formed in T_{n-1}. Assume that v is not on D, i.e., that $v \neq v_i$ for $i = 1, \ldots, k$. Then, let f be the largest index from $1, \ldots, k-1$ such that v_f is below v in T_{n-1}. Such an f exists, because by definition $v_1 = \text{blos1}(v)$ is below v. Then, $\{v_f, v_{f+1}\} \notin T_{n-1}$. (If it were in T_{n-1}, then either $\text{parent}(v_f) = v_{f+1}$ or parent(v_{f+1}) $= v_f$ and thus if v_f is below v, either v_{f+1} is below v also, or $v_{f+1} = v$. Both are contradictions.)

Any edge in a chain constructed in T_{n-1}, by a call to CHAIN, must either be in T_{n-1} or in B_{n-1}. Since $\{v_f, v_{f+1}\} \notin T_{n-1}$, it must be in B_{n-1}. So, for some $i \leq n-1$, $\{v_f, v_{f+1}\}$ is examined for Update 2 during the i^{th} iteration. T_i is a subtree of T_n, so the unique cycle in $T_i \cup \{v_{f-1}, v_f\}$ is also the unique cycle in $T_n \cup \{v_{f-1}, v_f\}$. Now, since v_f is below v in T_{n-1} and v_{f+1} is not below v in T_{n-1}, v is on the unique cycle in $T_n \cup \{v_{f-1}, v_f\}$ and thus on the unique cycle in $T_i \cup \{v_{f-1}, v_f\}$. But then v is relabeled to free during the i^{th} iteration for $i < n$, a contradiction. So v appears on D. □

Remark 2.8 By Lemmas 2.5 and 2.6, recursive calls to the SEARCH procedure are well defined. In the remaining proofs this fact will implicitly be assumed whenever we assume that a recursive call is well defined. Note that although the proof of Lemma 2.5 refers to recursive calls, it is not essential in that proof that the recursive calls are well defined, since we only need that the two CHAIN procedures disagree at some point, and this would have to occur before the procedures fail. So, the proof of Lemma 2.5 does not rely on the result of Lemma 2.6.

Lemma 2.7 *Suppose $T_n = (V, E)$ is a labeled rooted tree (with some edges in T_n distinguished as edges of a forest F) with root u, that is produced during the n^{th} (not necessarily the last) iteration of the SEARCH subroutine. Also, let $v \in V(T_n)$. Then during the construction of $C(v, u, T_n)$ using the subroutine CHAIN, if $C' = v, v_1, v_2, \ldots, v_{k-1}$ has been constructed after $m-1$ iterations of CHAIN and if v_k is under consideration for the m^{th} iteration, then:*

(a) *If Extension 1 results when considering v_k in the m^{th} iteration,*

$C(v, u, T_n) = C', C(v_k, u, T_n)$.

(b) *If Extension 2 results when considering v_k in the m^{th} iteration,*

$C(v, u, T_n) = C', \tilde{C}(v_k, u, T_n)$.

Proof: When the first iteration of CHAIN(v_k, u_0, T_n, M) and the m^{th} iteration of CHAIN(v, u, T_n, M) agree (with the same vertex under consideration and with the same extension), then all following steps will agree. □

Lemma 2.8 is used in the proof of Lemma 2.9 to show that the chains produced by CHAIN do not include repeated edges.

Lemma 2.8 *Supose $T_n = (V, E)$ is a labeled rooted tree (with some edges in T_n distinguished as edges of a forest F), which is produced during the n^{th} iteration (not necessarily the last) of the SEARCH subroutine. Let u be the root of T_n. Let $v \in V(T_n)$ be relabeled free during the n^{th} iteration of (an Update 2) of SEARCH. Let $\{x, y\}$ be the edge examined for this update and let $y = blos2(v)$ and $x = blos1(v)$. During the construction of $\tilde{C}(v, u, T_n)$ using CHAIN, let D be the chain produced by initialization with Extension 2 and let D' be such that $\tilde{C}(v, u, T_n) = D, D'$. Then, the vertices of D and D', except possibly v, are distinct.*

Proof: We first show that both D and D' can be formed in T_{n-1}. Then, we use this to show that if some vertex other than v appears in both chains, then v would have been relabeled free during some iteration prior to the n^{th}. Note that $V(T_{n-1}) = V(T_n)$ since the n^{th} iteration of SEARCH results in Update 2.

Note that D is either $C(y, v, T_n)$ or $\tilde{C}(y, v, T_n)$ and that D' is either $C(x, u, T_n)$ or $\tilde{C}(x, u, T_n)$. If $\{x, y\} \in F$ (so that M is false), or if y was initially labeled free, the initialization of the recursive call to CHAIN is with Extension 1, and thus D is $C(y, v, T_n)$. Then, since $y \in V(T_{n-1})$, by Lemma 2.5, $C(y, v, T_n) = C(y, v, T_{n-1})$. D is $\tilde{C}(y, v, T_n)$ if $\{x, y\} \notin F$ (so that M is true) and y was relabeled free. (So the recursive call is initialized with a second recursive call.) Since $\{x, y\} \notin F$ and since $\{x, y\}$ was examined for Update 2 during the n^{th} iteration, it must be that y has the label free in T_{n-1} (in order for the edge to be eligible). Thus, y is already relabeled free in T_{n-1} and so by Lemma 2.5, $\tilde{C}(y, v, T_n) = \tilde{C}(y, v, T_{n-1})$. Thus, D could be formed in T_{n-1}.

By similar arguments with $y = \text{prec}(x)$, and $\{x, y\}$ either in F or not in F, it can be seen that $C(x, u, T_n) = C(x, u, T_{n-1})$ and $\tilde{C}(x, u, T_n) = \tilde{C}(x, u, T_{n-1})$. Thus D' can also be formed in T_{n-1}.

Let $D = w_1, w_2, \ldots, w_k$ with $w_1 = y$ and $w_k = v$. Let $D' = w'_1, w'_2, \ldots, w'_{k'}$ with $w'_1 = x$ and $w'_k = u$. Assume that some vertex other than v appears on both chains.

First, note that every vertex in D is below v (or is v). If not, let f be the smallest index from $2, \ldots, k$ that is not below v. (By the definition of $\text{blosl}(v) = y$, y is below v.) Then, v_{f-1} is below v and v_f is not, so v is on the unique cycle in $T_{n-1} \cup \{v_{f-1}, v_f\}$. Since $\{v_{f-1}, v_f\}$ is on D, then $\{v_{f-1}, v_f\}$ is in B_n and for some $i \leq n$. So $\{v_{f-1}, v_f\}$ is examined resulting in Update 2. But, since T_i is a subtree of T_{n-1}, the unique cycle in $T_i \cup \{v_{f-1}, v_f\}$ is the same as the unique cycle in $T_{n-1} \cup \{v_{f-1}, v_f\}$. Then v would have been updated to free when $\{v_{f-1}, v_f\}$ was examined during the i^{th} iteration, a contradiction to the assumption that v is first relabeled free in the n^{th} iteration.

Now, assume that some vertex other than v appears on both D and D'. Let g be the smallest index so that w'_g appears on D. This exists since $w_1 = y$ is not below v. Then w'_{g-1} is not below v and w'_g is below v and by an argument similar to the previous paragraph we get a contradiction to the assumption that v is first relabeled during the n^{th} iteration. □

The next lemma shows that false augmenting chains are not created during the process of blossoming by showing that chains constructed from search trees formed in the algorithm satisfy the conditions for augmenting chains. The proof is by induction on the number of edges examined to form T. The proof of the inductive step is long but consists of straightforward verifications of the induction hypotheses in each of a number of possible (sub)cases.

Lemma 2.9 *Let $G = (V, E)$ be a graph with odd positive integer degree constraints on $V(T)$, let F be a feasible odd forest in G and let u be such that $d_F(u) = 0$. Let*

$T_0, T_1, ..., T_k$ be a sequence of labeled search trees which is produced by the subroutine SEARCH(u, G, F), such that T_0 is the tree consisting of the vertex u and no edges, and such that T_i is the tree produced after the i^{th} iteration. Then, for $n \in \{1, \ldots, k\}$, and for any vertex $v \in T_n$, the subroutine CHAIN(v, u, T_n, M) produces a chain C satisfying (ii) and (iii) in the definition of an augmenting chain.

Proof: We may assume that $u \neq v$ or else the result holds trivially. The proof will be by induction on n, the number of edges from L that have been examined. We will prove a slightly stronger claim which will imply the statement of the lemma.

Use the notation of Definition 2.4. We will prove that:

(a) Both $C(v, u, T_n)$ and $\tilde{C}(v, u, T_n)$ satisfy (ii) and (iii) in the Definition 2.3 of an augmenting chain.
(b) v appears only once in $C(v, u, T_n)$.
(c) If w is the second vertex in $\tilde{C}(v, u, T_n)$, then $\{v, w\} \in F$.
(d) If v appears twice in $\tilde{C}(v, u, T_n)$, the second appearance of v when the chain is traversed from v to u is a positive regular crossing by $\tilde{C}(v, u, T_n)$.

Clearly, (a) implies the statement of the lemma.

Note that the procedure CHAIN stops when an isolated vertex w ($d_F(w) = 0$) is detected (or when L becomes empty). If T_k is the final tree when the SEARCH procedure halts, the only isolated vertex in $V(T_n)$ for $n < k$ is u. The only isolated vertices in $V(T_k)$ are u and w. All other vertices have odd degree in F (since F is an odd forest). When the procedure halts, w will be a leaf in the last tree T_k and thus will not be an internal vertex in the only chain, $C(w, u, T)$ containing w. Thus, the part of (ii) that internal vertices on the chain have odd degree always holds and in the following will will not prove this part of (a).

For $n = 1$, T_1 consists of a single edge $\{u, u'\}$. Note that the first update must be Update 1, so in T_1, u' has not been relabeled by a blossom in Update 2. So (a) for $\tilde{C}(v, u, T)$ and (c) and (d) hold trivially as no $\tilde{C}$ chains exist. Also, $C(u', u, T_1) = u', u$ and it can easily be checked that (a) and (b) hold.

Assume that (a), (b),(c), and (d) hold in T_{n-1}. We will show by induction that (a), (b), (c), and (d) hold in T_n. Let $\{x, y\}$ be the edge that is examined during the n^{th} iteration of SEARCH. There are two cases:

Case 1: Examining $\{x, y\}$ results in Update 1.

Case 2: Examining $\{x, y\}$ results in Update 2, i.e., a blossom is formed by the edge $\{x, y\}$.

Case 1:

In case 1, one of x, y is in $V(T_{n-1})$ and one is not. Assume, without loss of generality that $x \in V(T_{n-1})$ and $y \notin V(T_{n-1})$. By Lemma 2.5, for $v \in V(T) \setminus \{y\} = V(T_{n-1})$, we have $C(v, u, T_n) = C(v, u, T_{n-1})$ and $\tilde{C}(v, u, T_n) = \tilde{C}(v, u, T_{n-1})$. So, by induction (a), (b), (c), and (d) hold when $v \neq y$. Since y is added to $V(T_n)$ during the update producing T_n, it can not be the case that y has been relabeled. So there is no $\tilde{C}(y, u, T_n)$ and (c) and (d) hold vacuously for $v = y$. Thus to complete the proof of Case 1, we need to show (a) and (b) for $C(y, u, T_n)$.

Subcase i: $\{y, x\} \in F$ or x is initially labeled free.

In this subcase, we first show that $C(y, u, T_n) = y, C(x, u, T_{n-1})$. For $C(y, u, T_n)$, the initialization of CHAIN sets the first vertex in the chain to y and considers parent(y)=x next. Since $\{y, x\} \in F$ or x was initially labeled free, Extension 1 will be performed. So, $C(y, u, T_n) = y, C(x, u, T_n) = y, C(x, u, T_{n-1})$. The first equality follows by Lemma 2.7, and the second by Lemma 2.5 (since $x \in V(T_{n-1})$).

Since y does not appear in $C(x, u, T_{n-1})$ (as $y \notin V(T_{n-1})$), y appears only once in $C(y, u, T_n) = y, C(x, u, T_{n-1})$ and (b) holds for $v = y$.

By induction, $C(x, u, T_{n-1})$ satisfies (a). Thus all internal vertices of $C(x, u, T_{n-1})$ satisfy (ii) and (iii) in Definition 2.3. The vertex x is an internal vertex in $C(y, u, T_n)$, so it remains to check that these conditions hold at x. By (b), x appears only once in $C(x,u,T_{n-1})$, so it appears only once in $C(y,u,T_n)$. Thus (ii) holds at x. Condition (iii) can only be violated at x if the crossing by $C(y,u,T_n)$ is negative irregular. If this is the case $\{x, \text{parent}(x)\} \notin F$ and $\{x,y\} \notin F$. Thus, by the conditions of the subcase, x was initially labeled free. The initial label for x was assigned when $\{x, \text{parent}(x)\}$ was examined, resulting in Update 1 (since this edge was added to the search tree). Then, from Update 1, since $\{x, \text{parent}(x)\} \notin F$ and since x was initially labeled free, $d_F(x) \leq b(x) - 2$ and (iii) holds. This completes the proof of Subcase i.

Subcase ii: x was relabeled free and $\{x,y\} \notin F$.

As in the previous subcase, we first show that $C(y,u,T_n) = y, \tilde{C}(x,u,T_{n-1})$. For $C(y,u,T_n)$, the initialization sets the first vertex in the chain to y and then examines $x = \text{parent}(y)$ next. Since x was relabeled free and $\{x,y\} \notin F$, Extension 2 is performed next. So, $C(y,u,T_n) = y, \tilde{C}(x,u,T_n) = y, \tilde{C}(x,u,T_{n-1})$. The first equality follows by Lemma 2.7 and the second by Lemma 2.5 (since x is not relabeled during the n^{th} iteration, it was relabeled already in T_{n-1}). Now $y \notin V(T_{n-1})$, so y appears only once in $y, \tilde{C}(x,u,T_{n-1})$, and condition (b) holds. Also, by induction on $\tilde{C}(x,u,T_{n-1})$ (since (a) holds), (ii) and (iii) are satisfied by all internal vertices of $C(y,u,T_n)$ except possibly x. The conditions (ii) and (iii) at x are satisfied since the extra conditions on x in $\tilde{C}(x,u,T_{n-1})$ hold. Namely, by (c), the first edge of $\tilde{C}(x,u,T_n)$ is in F so the first crossing at x is negative regular and by (d), the second crossing (if any) at x is positive regular. This completes the proof of Subcase ii and thus the proof of Case 1.

Case 2:

In case 2, $x,y \in V(T_{n-1})$ and $V(T_n) = V(T_{n-1})$. By Lemma 2.5, $C(v,u,T_n) = C(v,u,T_{n-1})$ for $v \in V(T_n)$ and by induction the result holds for these chains. Also by Lemma 2.5, if v was relabeled to free by an Update 2 occurring during an iteration prior to the n^{th}, then $\tilde{C}(v,u,T_n) = \tilde{C}(v,u,T_{n-1})$. By induction the result holds for these chains.

Thus, to complete the proof we must show that (a), (c), (d) hold for the new

$\tilde{C}(v,u,T_n)$ which can be formed for vertices v which had their labels updated from con to free by the blossom that is formed when examining $\{x,y\}$, i.e., vertices on the cycle in $T_n \cup \{x,y\}$ which are relabeled by the n^{th} update. We have several subcases to consider.

Subcase i: y is relabeled to free when examining $\{x,y\}$.

The only way for y to be labeled con before $\{x,y\}$ is examined if if $\{x,y\} \in F$, since otherwise, $\{x,y\}$ would not be eligible for Update 2 and thus would not be in L. Note that $\tilde{C}(y,u,T_n)$ is initialized with Extension 2 by recursively calling CHAIN(y,y,T_n), (since blos1(y) = y). It is easy to check that the call to CHAIN(y,y,T_n) trivially produces the chain consisting of the single vertex y. From Extension 2, the next vertex to be considered will be blos2(y)=x with prec(x)=y. Since $\{x, \text{prec}(x)\} = \{x,y\} \in F$, during the next iteration of CHAIN, Extension 1 is used. So, by Lemma 2.7 and by Lemma 2.5, we have $\tilde{C}(y,u,T_n) = y, C(x,u,T_n) = y, C(x,u,T_{n-1})$.

We first show that (a) holds for $\tilde{C}(y,u,T_n)$. By induction, since (a) holds for $C(x,u,T_{n-1})$, all internal vertices in $\tilde{C}(y,u,T_n)$ except possibly x satisfy (*ii*) and (*iii*) in Definition 2.3. By induction, (b) holds for $C(x,u,T_{n-1})$ and thus x appears only once in $C(x,u,T_{n-1})$ and (*ii*) holds for x. Since $\{x,y\} \in F$ the crossing at x is not negative irregular so (*iii*) holds, completing the proof for (a).

Note that (c) holds for $\tilde{C}(y,u,T_n)$ since x is the second vertex in this chain and $\{x,y\} \in F$.

Finally, we show that (d) holds for $\tilde{C}(y,u,T_n)$. Thus, we assume that y appears twice in $\tilde{C}(y,u,T_n)$. Let $C(x,u,T_{n-1}) = v_1, v_2, \ldots, v_m, u$ with $x = v_1$. Then, $\tilde{C}(y,u,T_n) = y, v_1, v_2, \ldots, v_m, u$.

Consider first the case that y is not above x in T_{n-1}. Let the second appearance of y be as v_k. Let f be the smallest index from $2, \ldots, k$ such that either v_f is below y in T_{n-1} or $v_f = y$. Then, the edge $\{v_{f-1}, v_f\}$is not in T_{n-1}, since if it were, v_{f-1} would either be y or be below y. Thus, $\{v_{f-1}, v_f\} \in B_{n-1}$. In order for $\{v_{f-1}, v_f\}$ to be in B_{n-1}, it must have been examined for an Update 2 at some iteration $i < n$. Then all con vertices on the cycle in $T_i \cup \{v_{f-1}, v_f\}$ are relabeled free during the i^{th} update, and thus have the label free in T_{n-1}. Note that since T_i is a subtree of T_n, the unique cycle in $T_i \cup \{v_{f-1}, v_f\}$ is also the unique cycle in $T_n \cup \{v_{f-1}, v_f\}$. Since y is con in T_{n-1} (as it is relabeled by the n^{th} update), y can not be on the unique cycle in $T_n \cup \{v_{f-1}, v_f\}$. Then, since y is above v_f and it is not on the cycle in $T_n \cup \{v_{f-1}, v_f\}$, it must be that y is above v_{f-1} contradicting the choice of f. So we have shown that y must be above x in T_n.

Now, we consider the case that y is above x in T_n. Recall that we denote $\tilde{C}(y, u, T_n) = y, v_1, v_2, \ldots, v_m, u$ with $x = v_1$ and the second appearance of y as v_k. Consider the edge $\{v_{k-1}, v_k\}$. It must either be an edge of T_{n-1} or and edge in B_{n-1} since it is an edge in $C(x, u, T_{n-1})$. If $\{v_{k-1}, v_k\} \in B_{n-1}$, then Update 2 was performed on $\{v_{k-1}, v_k\}$ during iteration i for some $i \leq n-1$. This update would relabel $v_k = y$ to free, contradicting the assumption that y is first relabeled in the n^{th} iteration. So $\{v_{k-1}, v_k\} \in T_{n-1}$.

Note that $\{v_{k-1}, v_k\}$ is not added to $C(x, u, T_{n-1})$ as part of a chain from a recursive call to CHAIN. If it were, then for some z it is on the chain $D(\text{blos1}(z), z, T_{n-1})$, (where D can be either C or $\tilde{C}$), formed by Extension 2 when z is examined in CHAIN(x, u, T_{n-1}, M). Further, assume that z is picked so that it is not the case that $\{v_{k-1}, v_k\}$ is added to $D(\text{blos1}(z'), z', T_{n-1})$ as part of a chain during a second recursive call when z' is examined while constructing $D(\text{blos1}(z), z, T_{n-1})$. (Such a z exists since there are a finite number of recursive calls.) Then, v_k is on the unique cycle in $T_{n-1} \cup \{\text{blos1}(z), \text{blos2}(z)\}$. Also, $\{\text{blos1}(z), \text{blos2}(z)\}$ was examined for Update 2 during the i^{th} iteration of SEARCH for some $i \leq n-1$ (since it is in B_{n-1}). But then examining $\{\text{blos1}(z), \text{blos2}(z)\}$ would cause $v_k = y$ to be relabeled free during the i^{th} iteration of SEARCH, contradicting the assumption that it is first relabeled in the n^{th} iteration.

Thus, $\{v_{k-1}, v_k\}$ is added to $C(x, u, T_{n-1})$ by Extension 1, and furthermore, $v_k =$ parent(v_{k-1}). This last point follows since from Extension 1, either $v_k =$ parent(v_{k-1})

or v_{k-1} = parent(v_k), the second occurring only when this edge is added as part of the reversal of some chain produced by a recursive call to CHAIN. Note that prec(v_k) is v_{k-1} and the next vertex to examine (which is v_{k+1}) is set to parent(v_k) by this extension.

Now, we have v_k = parent(v_{k-1}). Since $\{v_k, v_{k-1}\} \in T_{n-1}$ it must be that this edge was examined with Update 1 at some iteration $j \leq n$ of SEARCH. Thus, $v_k \in V(T_{j-1})$ and $v_{k-1} \notin V(T_{j-1})$. Also, v_k has the label con in T_{j-1} since it is first relabeled free during the n^{th} iteration. Thus, in order for $\{v_k, v_{k-1}\}$ to be eligible for Update 1 in the j^{th} iteration, it must be that $\{v_k, v_{k-1}\} \in F$.

Finally, since parent (v_k) $= v_{k+1} \in T_{n-1}$, it must be that $v_{k+1} \in V(T_{j'-1})$ and $v_k \notin V(T_{j'-1})$ for some $j' < n-1$ and $T_{j'}$ is formed from $T_{j'-1}$ by examining $\{v_{k+1}, v_k\}$ with Update 1. Then, since v_k was initially labeled con (during this update), it must be that $\{v_{k+1}, v_k\} \notin F$. So we have $\{v_{k-1}, v_k\} \in F$ and $\{v_{k+1}, v_k\} \notin F$. Thus the second appearance of y as v_k when $\tilde{C}(y, u, T_n)$ is traversed from v to u has a positive regular crossing. This completes the proof that (d) holds for $\tilde{C}(y, u, T_n)$.

Subcase ii: x is relabeled free when examining $\{x, y\}$.
The proof is symmetric to the proof of subcase i.

Subcase iii: $v \neq x, y$ is relabeled free when examining $\{x, y\}$.
Without loss of generality, we may assume that blos1(v) is y. The chain $\tilde{C}(v, u, T_n)$ is initialized with Extension 2, a recursive call to CHAIN(y, v, T_n, M). For the recursive call, M is true if $\{x, y\} \notin F$ and false otherwise. If M is false, then the recursive call produces $C(y, v, T_n)$ and if M is true and y was relabeled, the recursive call produces $\tilde{C}(y, v, T_n)$.

Then, $\tilde{C}(v, u, T_n)$ begins with $[C(y, v, T_{n-1})]^R$ or $[\tilde{C}(y, v, T_{n-1})]^R$ followed by x. The chain building process continues from x to u using either $C(x, u, T_n)$ or $\tilde{C}(x, u, T_n)$ depending on whether or not $\{x, y\} \in F$ and if not, depending on whether or not x was relabeled by an earlier blossom. Thus, $\tilde{C}(v, u, T_n) = [D(y, v, T_{n-1})]^R, D'(x, u, T_{n-1})$, where D and D' can be either C or $\tilde{C}$.

We first show that (a) holds for $\bar{C}(v,u,T_n)$. By Lemma 2.8, the vertices of $D(y,v,T_{n-1})$ and $D'(x,u,T_{n-1})$ except possibly v are distinct. Thus, by induction (*ii*) and (*iii*) hold at all internal vertices except possibly at x and y.

We will show that (*ii*) and (*iii*) hold at y. A similar proof shows (*ii*) and (*iii*) hold at x. If $D(y,v,T_{n-1})$ is $\bar{C}(y,v,T_{n-1})$ then M for the recursive call to CHAIN, when y is examined, is true and $\{x,y\} \notin F$. The crossing of y by $\bar{C}(v,u,T_n) = [\bar{C}(y,v,T_{n-1})]^R, D(x,u,T_{n-1})$ with leaving edge $\{x,y\}$ has entering edge the last edge of $[\bar{C}(y,v,T_{n-1})]^R$. This first edge of $\bar{C}(y,v,T_{n-1})$ is in F since (c) holds by induction. So the crossing is positive regular. If y has two crossings, the first crossing (which is the second appearance of y on $\bar{C}(y,v,T_{n-1})$) is negative regular since inductively by (d) the reverse of this crossing is positive regular. In either case (*ii*) and (*iii*) hold for y.

If $D(y,v,T_{n-1})$ is $C(y,v,T_{n-1})$ then either M is true for the recursive call ($\{x,y\} \in F$) or y was initially labeled free. By (b) and induction, y appears only once in $C(y,v,T_{n-1})$. So (*ii*) holds at y. If $\{x,y\} \in F$ then the crossing at y is either positive irregular or negative regular, satisfying (*iii*). If y was initially labeled free then either $d_F(y) \leq b(y) - 2$ or $\{y, \text{parent}(y)\} \in F$. In the first case (*iii*) holds trivially and in the second case, the crossing at y is either positive irregular or positive regular again satisfying (*iii*). So (a) holds.

We next check that (c) holds in $\bar{C}(v,u,T_n)$. Note that the first edge $\{v,v'\}$ in $\bar{C}(v,u,T_n)$ is the last edge in $D(y,v,T_{n-1})$ (i.e., the first edge of $D(y,v,T_{n-1})^R$). The edge $\{v,v'\}$ must be an edge from T_{n-1}, since otherwise, it is in B_{n-1} because it is used in a chain formed by CHAIN. Then, if $\{v,v'\} \in B_{n-1}$, Update 2 occured when $\{v,v'\}$ was examined during the i^{th} iteration for some $i \leq n-1$. But this would have relabeled v to free during the i^{th} iteration, contradicting the assumption that v is first relabeled during the n^{th} iteration.

So we have $\{v,v'\}$ appearing as an edge of T_{n-1}. Then if parent$(v) = v'$, the edge is added to $D(y,v,T_{n-1})$ during a recursive call to CHAIN (within the recursive call to form $D(y, v, T_{n-1})$). Again, this can not occur because this would imply that v was relabeled free during an earlier iteration. Thus, parent(v') = v. So, this edge was added during Update 1 during some iteration, say the j^{th} for $j \leq n-1$, with $v \in V(T_{j-1})$ and $v \notin V(T_{j-1})$. For Update 1 to occur in this case, since v has the label con in T_{j-1}, it must be that $\{v, v'\} \in F$. So (c) holds.

Finally, we check that (d) holds for $\tilde{C}(v, u, T_n)$, so we assume that v appears twice in $\tilde{C}(v, u, T_n)$. Since CHAIN(y, v, T_{n-1}) stops when v is first considered, v appears only once in $D(y, v, T_{n-1})$. Thus, for v to appear twice in $\tilde{C}(v, u, T_n)$, it must be that v appears as an internal vertex of $D(x, u, T_{n-1})$. Say that v', v, v'' appear, in that order when traversing $D(x, u, T_{n-1})$ from x to u. In a manner analogous to the proof of (c) in the two previous paragraphs, it can be shown that $\{v, v'\} \notin B_{n-1}$ (since if it was, v would have been relabeled by an earlier iteration), and thus $\{v, v'\} \in F$ in order for Update 1 (adding the edge to the search tree), to occur when v is labeled con.

Then, during the CHAIN procedure constructing $D(y, v, T_{n-1})$, when v is considered, we have prec(v) = v' and so Extension 1 occurs. This adds v to $D(y, v, T_{n-1})$ and considers parent(v) next. So, v'' = parent(v). Then, for some $k \leq n-1$, we have $v'' \in V(T_{k-1})$ and $v \notin V(T_{k-1})$, and $\{v'', v\}$ is examined for Update 1 during the k^{th} iteration. Since v is initially labeled con (during this iteration) it must be that $\{v'', v\} \notin F$.

So, we have $\{v', v\} \in F$ and $\{v, v''\} \notin F$ for the second appearance of v in traversing $\tilde{C}(v, u, T_n)$ from v to u. Thus (d) holds and the proof is complete. □

We next show that by forming the search tree T so that subtrees in F are subtrees in T, the augmenting chains which are detected will not form cycles upon augmentation.

Lemma 2.10 *Let $G = (V, E)$ be a graph with odd positive integer constraints on the vertices, let F be a feasible odd forest in G, and let v_0 be a vertex with $d_F(v_0) = 0$. Let SEARCH(v_0, F, G) halt in the n^{th} iteration when v_k with $d_F(v_k) = 0$ is added to $V(T_n)$*

as a result of Update 1. Let C be the chain produced by CHAIN(v_k, v_0, T_n, M). Then $F \oplus C$ is a feasible odd forest.

Proof: Note that M is null for the initial call to CHAIN, so $C = C(v_k, v_0, T_n)$, i.e., Extension 1 is performed for the initialization. From part (b) of the proof of Lemma 2.9, v_k will appear only once in C. The call to CHAIN ends when v_0 is first encountered, so v_0 appears only once in C. Also, by assumption, $d_F(v_0) = 0$ and $d_F(v_k) = 0$. So (i) in Definition 2.3 of an augmenting chain holds for C. By Lemma 2.9, (ii) and (iii) also hold. So C is an augmenting chain and thus $F \oplus C$ satisfies the degree constraints. Also, the procedure halts when an isolated vertex is first detected; then the only vertices in T_n and thus in C that have degree 0 (are isolated) in F are v_0 and v_k. These vertices have odd degree after augmentation. Interior vertices on the chain have odd degree before and after augmentation. The degrees of vertices not on C are unchanged. Thus, the vertices of $F \oplus C$ have odd degree or are isolated. It remains to check that $F \oplus C$ is indeed a forest, i.e., that it contains no cycles.

Assume that there is a cycle $S \subseteq (F \oplus C)$. There are two cases to consider.

Case 1: S contains an edge which is in $F \setminus C$.

Case 2: $S \subseteq C \setminus F$.

Case 1:

In case 1, let $P = y_1, y_2, \ldots, y_j$ be any maximal path of edges from F along the cycle S, and let y_{j+1} be the next vertex following S in the direction from y_1 to y_j. Note that y_{j+1} may be y_1. Since P is maximal, $\{y_j, y_{j+1}\} \notin F$. Thus, $\{y_j, y_{j+1}\}$ appears as an edge in C.

All of the edges of P are in F, so P is contained in a subtree of F. Let $T(v)$ be this subtree. By the order for examining edges, $T(v)$ will be a subtree of T_n. By relabeling, if necessary, we may assume that y_1 is not below y_j in T_n. Some vertex y_a of P is the first vertex added to the search tree, i.e., for some i, $y_a \in V(T_i)$ and for $a' \neq a$, $y_{a'} \notin V(T_i)$. Then, in T_n, for $j \geq b > a$, parent(y_b) $= y_{b-1}$ and for $1 \leq c < a$, parent(y_c) = y_{c+1}. Thus in T_n, we have parent(y_j) = y_{j-1} (since y_1 is not below y_j and so $a \neq j$). Then, for some g, $y_{j-1} \in V(T_{g-1})$ and $y_j \notin V(T_{g-1})$ and the g^{th} iteration results in Update 1 when $\{y_{j-1}, y_j\}$ is examined. Thus, since this edge is in F, y_j is initially labeled free.

We will first assume that edge $\{y_j, y_{j+1}\} \in T_n$ and reach a contradiction. If $\{y_j, y_{j+1}\} \in T_n$, then parent(y_{j+1}) = y_j, since otherwise, y_j would be the first vertex of $T(v)$ contrary to our assumption that y_1 is not below y_j.

During the call to CHAIN, the edge $\{y_{j+1}, y_j\}$ is added to C by putting y_{j+1} in C and considering y_j next (possibly during a recursive call to CHAIN). Since y_j was initially labeled free the edge $\{y_j, \text{parent}(y_j)\} = \{y_j, y_{j-1}\}$ will be added to C by examining y_{j-1} next. This contradicts $\{y_j, y_{j-1}\} \in P \subseteq (F \setminus C)$.

Now we assume that $\{y_j, y_{j+1}\} \notin T_n$ and reach a contradiction. In this case, $\{y_j, y_{j+1}\} \in B_n$. Since it is in C, we have $\{y_j, y_{j+1}\} = \{\text{blos1}(z), \text{blos2}(z)\}$ for some vertex z initially labeled con. In this case, if u is the root of T, then C contains $D(\text{blos2}(z), u, T)$ and $[D'(\text{blos1}(z), z, T)]^R$ where D and D' can be either C or $\bar{C}$. If y_j is blos2(z), then since y_j was initially labeled free we have $D(\text{blos2}(z), u, T) = C(\text{blos2}(z), u, T)$ and the initialization of this chain shows that its first edge is $\{y_j, y_{j-1}\}$ since parent(y_j) = y_{j-1}. Similarly, if y_j is blos1(z), it can be shown that $\{y_j, y_{j-1}\}$ is put in C. Both contradict $\{y_j, y_{j-1}\} \in P \subseteq (F \setminus C)$.

We have reached a contradiction for both $\{y_j, y_{j+1}\} \in T_n$ and $\{y_j, y_{j+1}\} \notin T_n$. So, case 1 can not occur.

Case 2:

For case 2 every edge of S is part of C but not in F. Let y_1 be the first vertex of S which is added to C when C is constructed using CHAIN and let y_k and y_2 be its neighbors in S. Both $e = \{y_1, y_k\}$ and $f = \{y_1, y_2\}$ are part of C and must appear on C after the entering edge for the first appearance of y_1 by the choice of y_1. Then, since y_1 appears at most twice in C, by the conditions for an augmenting chain in Definition 2.3, one of these two edges is the leaving edge for the first appearance of y_j or both are the adjacent edges at the second appearance of y_j. For the first of these possibilities there must be a second crossing in order for both e and f to appear in C. Thus by the augmenting chain conditions, the first crossing must be negative regular with leaving edge in F. This contradicts e and f not in F. For the second possibility, the augmenting chain conditions imply that the second crossing must be positive regular, with entering edge in F. This again is a contradiction since one of e or f must be the entering edge in this case and neither one is in F. This completes the proof that case 2 cannot occur and the proof of the lemma. □

The final lemma shows that if the search procedure from a vertex fails with respect to a forest F, then no vertex of the failed search tree T will appear as a vertex of an augmenting chain at a later stage of the algorithm.

Lemma 2.11 *If procedure SEARCH(v_0,G,F) fails to find an augmentation and T_m is the search tree at the end of the failed search, then, for any forest F' formed during a later stage of the algorithm and for any augmenting chain C with respect to F', we have $V(T_m) \cap V(C) = \emptyset$. That is, no vertex of T_m will be appear on such an augmenting chain.*

Proof: As long as no vertex of T_m appears on augmenting chains at later stages of the algorithm the status of edges in T_m and edges incident to T_m with respect to the F's obtained at each stage will be unchanged. That is, the subforests of F and F' induced by $V(T_m)$ are identical, and furthermore, F and F' agree on the set of edges with one end in $V(T_m)$ and one end not in T_m.

Assume by way of contradiction that for some forest F' produced during a later stage of the algorithm, there is an augmenting chain C which contains some vertex of T_m. Let F' be the first packing formed for which there is such an augmenting chain, and let $C = w_0, w_1, \ldots, w_n$.

Pick a vertex $w_j \in T_m \cap C$. By Lemma 2.9 there is a chain $D = v_0, v_1, \ldots, v_k$

(where $v_k = w_j$ and v_0 is the root of T_m), using vertices $V(T_m)$ that satisfies (ii) and (iii) in the Definition 2.3 of augmenting chains relative to the packing F and therefore, relative to F' (since F and F' agree on $V(T_m)$). Choosing D to be the chain from v_0 to a vertex of C with the fewest vertices, we may assume that either $D = v_0$ (a trivial chain) or $v_i \notin C$ for $i = 0, 1, \ldots, k-1$. In the first case, when $D = v_0$, then by definition, C is augmenting with respect to F'. In the second case, consider

$$P^+ = v_0, v_1, \ldots, v_{k-1}, w_j, w_{j+1}, \ldots, w_n$$

$$P^- = v_0, v_1, \ldots, v_{k-1}, w_j, w_{j-1}, \ldots, w_0.$$

We will first show that one of P^+ or P^- is augmenting with respect to F'.

Clearly P^+ and P^- satisfy (ii) and (iii) relative to F', except possibly at $w_j = v_k$, since the other vertices partition into two distinct subchains. We have already noted that $v_0, \ldots, v_{k-1}$ satisfies (ii) and (iii) relative to F'. Also, $w_j, w_{j-1}, \ldots, w_0$ and $w_j, w_{j+1}, \ldots, w_n$ satisfy (ii) and (iii) as they are subchains of an augmenting chain.

Since w_j appears at most twice on C, it appears at most once on either P^+ or P^-. If $\{v_{k-1}, v_k\} \in F'$ or $d_F(w_j) \leq b(w_j) - 2$ then the one of P^+ or P^- on which w_j appears once satisfies $(ii), (iii)$ at w_j. Otherwise, if $d_F(w_j) = b(w_j)$ (and $\{v_{k-1}, v_k\} \notin F'$) then at least one of $\{w_j, w_{j-1}\}, \{w_j, w_{j+1}\}$ is in F'. If $\{w_j, w_{j-1}\} \in F'$, then P^- satisfies (ii) and (iii) at w_j. Note here that w_j may appear again on P^-. However the two appearances will be consistent with (ii), because in going from w_0 to w_j, the first crossing of w_j is negative regular. The case $\{w_j, w_{j+1}\} \in F'$ is similar. Also, v_0, w_n and w_0 are isolated in F'. Thus, one of P^+ or P^- is augmenting with respect to F'.

Thus we have an augmenting chain C' with respect to F' from v_0 to either w_n or w_0. In the case that $D = v_0$ (is trivial), $C' = C$ is such a chain. Otherwise C' is one of P^+ or P^-. Assume that $C' = P^+$, the cases $C' = P^-$ and $C' = C$ are similar. Let l be the smallest index such that w_{l+1} is not in $V(T_m)$. Then since F' and F agree on edges in T_m and edges incident to T_m and since $\{w_l, w_{l+1}\}$ was not added to T_m, (the edge was not eligible for an update), w_l is labeled con (in T_m) and

$\{w_l, w_{l+1}\} \notin F$. Thus $\{w_l, w_{l-1}\} \in F$, since P^+ is augmenting. We shall show that the chain $v_0, \ldots, v_{k-1}, w_j, w_{j+1}, \ldots, w_l$ satisfies the conditions (ii) and (iii) and also (iv), (v), and (vi) of Lemma 2.4. We have just showed that condition (iv) holds.

If w_l has a second appearance among $v_0, \ldots, v_{k-1}, w_j, \ldots, w_{l-1}$ then condition (ii) for augmenting chains and the fact that the w_j and v_i are distinct insure that condition (v) of Lemma 2.4 is satisfied by $v_0, \ldots, v_{k-1}, w_j, \ldots, w_l$. By the choice of l, $w_i \in V(T_m)$ for $i = j, \ldots, l-1$. By definition, $v_i \in V(T_m)$ for $i = 0, 1, \ldots, k-1$. Since D is not trivial, $v_0 \neq w_i$ for $i = j, \ldots, l-1$, and by the minimality of D, $v_0 \neq v_i$ for $i = 1, \ldots, k-1$. Thus, since v_0 is the root of T_m, w_i for $i = j, \ldots, l-1$ and v_i for $i = 1, \ldots, k-1$ are all below v_0 in T_m. Thus, condition (vi) of Lemma 2.4 holds. Conditions (ii) and (iii) also hold since P^+ is an augmenting chain. Thus the conditions of Lemma 2.4 are satisfied by $v_0, \ldots, v_{k-1}, w_j, \ldots, w_l$ and the conclusion of the lemma yields a contradiction to the assumption that w_l has label free. This shows that there can be no augmenting chain meeting T. □

We now use the lemmas to show the correctness of the algorithm.

Theorem 2.12 *The odd subtree packing algorithm finds a maximum size feasible odd forest in a graph $G = (V, E)$ in $O(|V|^3)$ time.*

Proof: The algorithm starts with a feasible odd forest $F = \emptyset$ and, by Lemma 2.10, the F's obtained at each stage of the algorithm are indeed feasible odd forests. Let F^* denote the feasible odd forest obtained in G when the algorithm ends. Let G' the graph induced by the vertices remaining in $V(G)$ when the algorithm ends. That is, $v \in V(G')$ if and only if v has not appeared in a failed SEARCH tree during any stage of the algorithm.

It is not difficult to check that the list I contains exactly those vertices $v \in V(G')$ that are isolated with respect to the forest F^* and have not been used as the root in a call to SEARCH at some stage of the algorithm. Then, since $|I| < 1$ when the algorithm stops, there is at most one vertex $v \in V(G')$ with $d_{F^*}(v) = 0$, i.e, such that v is isolated in F^*. Thus, there is no augmenting chain in G' with respect to the subforest of F^* induced by $V(G')$ (since there is only one isolated vertex in $V(G')$). Then, by Lemma 2.11, there can be no augmenting chain with respect to F^* in G since the vertices of such a chain would be contained in $V(G')$. So, by Theorem 2.3, the feasible odd forest F^* has maximum size in G.

For the complexity, note that there can be at most $|V|$ calls to the SEARCH subroutine, since each time SEARCH is called, the number of vertices in I is reduced by either augmentation or deletion (when the search fails), and since no vertices are added to I. Thus, since the CHAIN subroutine is called at most once for each SEARCH, the number of calls to CHAIN is at most $|V|$. The complexity of CHAIN is bounded by the number of edges in the chain which is constructed, so one call to CHAIN is $O(|V|)$ and the total time for calls to CHAIN is $O(|V|^2)$.

Now consider the complexity of SEARCH. Maintaining the staus of the edges in L takes $O(|E|)$ time over the course of a call to SEARCH since each edge is considered at most twice for each of its endpoints; once when the vertex is first labeled and once when it is relabeled free.

Each time Update 1 is performed, a new vertex w is added to the search tree. Thus Update 1 can be performed at most $|V|$ times and the time spent adding new vertices to the tree is $O(|V|)$.

For Update 2, let $\{x, y\}$ be the edge examined for this update. Recall that when $\text{nca}(x) = \text{nca}(y)$ no relabeling is necessary for Update 2. Thus constant time is required for Update 2 if $\text{nca}(x) = \text{nca}(y)$. Each edge is examined at most once during a call to SEARCH. So the total time spent on Update 2 when $\text{nca}(x) = \text{nca}(y)$ is $O(|E|)$.

Finally, we consider Update 2 when $\text{nca}(x) \neq \text{nca}(y)$. Call this case a *non-trivial* blossom formation. Let $N(z)$ denote the set of vertices x such that $\text{nca}(x) = z$. Each time a non-trivial blossom is formed, a new $N(z)$ is formed which is the union of the $N(v)$'s for v on the cycle in $T \cup \{x, y\}$. Thus, at most $|V| - 1$ non-trivial blossoms can be formed before every vertex v has the same nca(v). That is, at most $|V| - 1$ non-trivial blossoms are formed.

Relabeling a vertex w requires constant time. Each vertex is relabeled at most once, so the total time spent relabeling over all non-trivial blossoms is $O(|V|)$. The additional time required for a non-trivial blossom when $\{x, y\}$ is examined is $O(|V|)$; to construct the cycle in $T \cup \{x, y\}$ and revise the nca labels. Since at most $|V|$ non-trivial blossoms are formed, the total time over all non-trivial blossoms spent revising nca labels and finding cycles is $O(|V|^2)$. This dominates the complexity of SEARCH. So one iteration of SEARCH requires $O(|V|^2)$ time, and the total time spent in SEARCH is $O(|V|^3)$ since there are at most $|V|$ calls to SEARCH (as $|I|$ is reduced by each iteration).

It is not difficult to see that for each call to SEARCH, revising the list I, discarding vertices from the graph under consideration in the next stage and storing the forest induced by discarded vertices, and forming a new forest with an augmenting chain can be done in $O(|E|)$ time. Thus, the total time spent during the algorithm on revising I and maintaining F is $O(|V||E|)$ since there are at most $|V|$ calls to SEARCH.

Finally, from the preceeding discussion of the complexities of the various parts of the algorithm, we see that the overall time complexity of the algorithm is $O(|V|^3)$. □

Remark 2.9 We note that the time complexity in the previous theorem is dominated by the time spent updating the nca labels during the formation of non-trivial blossoms. The sets with the same nca label can be maintained by a fast disjoint set union procedure exactly as described for the case of matching in Tarjan [1983] and Gabow and Tarjan [1985]. We will not go into the details of these data structures. However, we note that it is not difficult to see that the disjoint set union described in Tarjan [1983] and Gabow and Tarjan [1985] can be used in the case of odd subtree packing. Maintaining blossoms and sub-blossoms is exactly the same in both cases. If these fast disjoint set union procedures are used, the complexity bound for odd subtree packing can be improved to $O(|V||E|)$.

2.5 A Min-Max Formula

In this section we describe a min-max formula for odd subtree packing. The result of Theorem 2.13 was obtained as the final version of this thesis was being prepared. Thus, we include only a statement and a brief sketch of the proof. The proof makes use of the failed search trees from the algorithm.

In matching, the Berge-Tutte formula states that the maximum size of a matching is equal to the minimum over all subsets S of the vertex set V of $(|V|+|S|-odd(S))/2$ where $odd(S)$ is the number of odd components in the graph induced by $V \setminus S$. (See for example Lovász and Plummer [1986] for details on the Berge-Tutte formula.)

We can state a similar formula for odd subtree packing. Note that this formula provides a 'certificate' that a given odd subtree packing is indeed of maximum size. Let $odd(S)$ denote the number of odd components in the graph induced by $V \setminus S$. Also, let $b(S) = \sum_{v \in S} b(v)$.

Theorem 2.13 *Let a graph $G = (V, E)$ and non-negative integer constraints b on V be given. The maximum size of an odd subtree packing in G is equal to the minimum over all subsets $S \subseteq V$ of*

$$|V| - (odd(S) - b(S)).$$

We will sketch the idea of the proof.

We first note that for any $S \subseteq V$, $|V| - (odd(S) - b(S))$ provides an upper bound on the size of an odd subtree packing. Let F be a maximum size feasible odd forest. In general, any edge with exactly one end in an odd component of the graph induced by $V \setminus S$ must have the other end in S (by the definition of a component). By parity considerations, and since there are an odd number of vertices in an odd component, either some vertex of each odd component is isolated or there is an edge of F with exacly one end in the component. The other end of such an edge must be in S. At most $b(S)$ edges of F can have one end in S, so at least $(odd(S) - b(S))$ odd components must contain an isolated vertex. This shows that the maximum is less than or equal to the minimum.

To show equality, we construct a set S' for which equality holds. For each vertex v which is isolated with respect to F, construct a tree with v as the root using the SEARCH procedure of the algorithm. The trees for the isolated vertices will be disjoint, (or else an augmenting chain with respect to F would be detected). Furthermore, it can be shown using parity arguments and the fact that no edges are eligible for Update 2 in the trees, that the set of vertices which are labeled con in these trees will be the set S'.

2.6 Further Research

The development of the algorithm for odd subtree packing leaves open some natural questions. Most of these concern generalizing important concepts in matching to the cases of chain packing or odd subtree packing. Details on the matching results stated below can be found in Lovász and Plummer [1986].

1. It would be interesting to see if there is any special structure when the set of degree constraints is limited. For example if $b(v) = 1$ or 3 for all v.

2. The collection of sets of vertices that are subsets of the set of odd degree vertices in some chain packing form a matroid. This is shown in deWerra and Roberts [1990], following from their augmenting chain theorem. In matching there is a similar matroid. Additionally, the collection of sets of vertices that are subsets of unmatched vertices in some matching form a matroid. It is unknown whether a similar property holds for the isolated vertices in chain packing.

3. It would be interesting to explore the possibility of a general structure for chain packings along the line of the structure theory for matching developed by Gallai

and Edmonds. (See for example Lovász and Plummer [1986] for details on this structure theory.)

4. It would be interesting to find a short proof of Theorem 2.13 that does not rely on the algorithm.

5. There is no known simple description of the polytope formed by the convex hull of incidence vectors of odd subtree packings.

6. In generalizing weighted matching, we must consider weights on vertices rather than edges (as in matching). If weights are assigned to vertices, the problem of maximizing the sum of the weights of odd vertices can be considered. We note that in the weighted version, the problems of chain packing and odd subtree packing are not equivalent. For example, in Figure 2.2, delete the edge $\{x_5, x_2\}$, let x_1, x_3, x_3' and x_4 have weight 100 and let x_2 have weight 1. Then a maximum odd forest must contain x_2 and so its weight is 301. However, the chain packing containing the edge $\{x_3', x_4\}$ and the chain of two edges $\{x_1, x_2\}, \{x_2, x_4\}$ is feasible and has weight 400. Are there polynomial algorithms for weighted chain packing or weighted odd subtree packing?

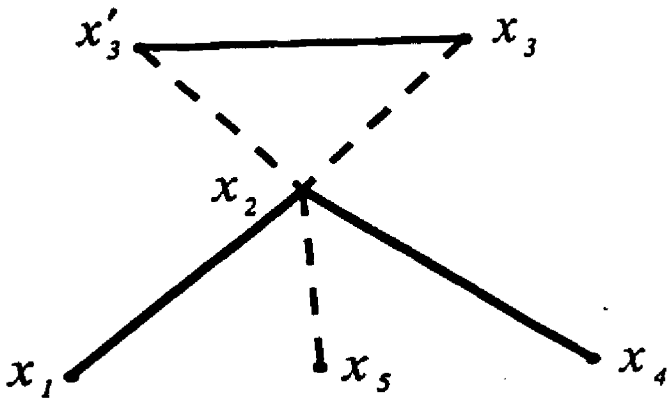
2.7 Appendix: Examples of the Chain Packing Algorithm

In this appendix we give for illustration several partially worked examples of the chain packing algorithm.

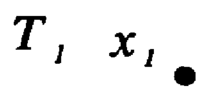
Example 1:

An iteration of the SEARCH procedure on the graph G of Figure 2.2 with respect to the forest F in that figure. Let every degree constraint b be 3.

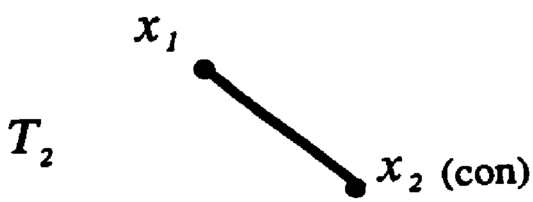
We exhibit the graph below. Dashed edges are from F.



We show the construction of the search tree in stages, along with the list L of edges eligible for update, and labels on the vertices.

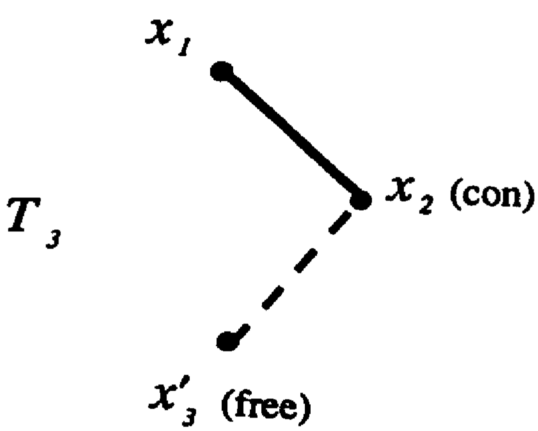


$L = \{\{x_1, x_2\}\}$

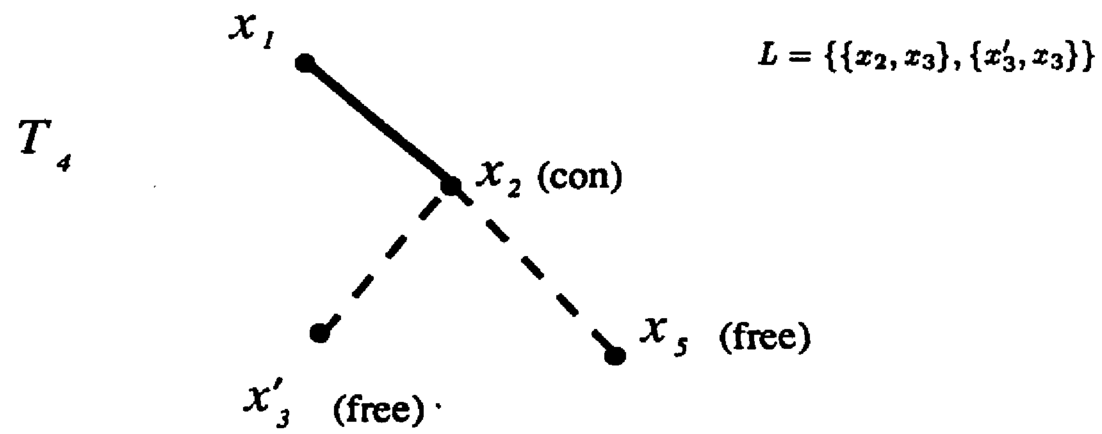


$L = \{\{x_2, x_5\}, \{x_2, x_3\}, \{x_2, x'_3\}\}$

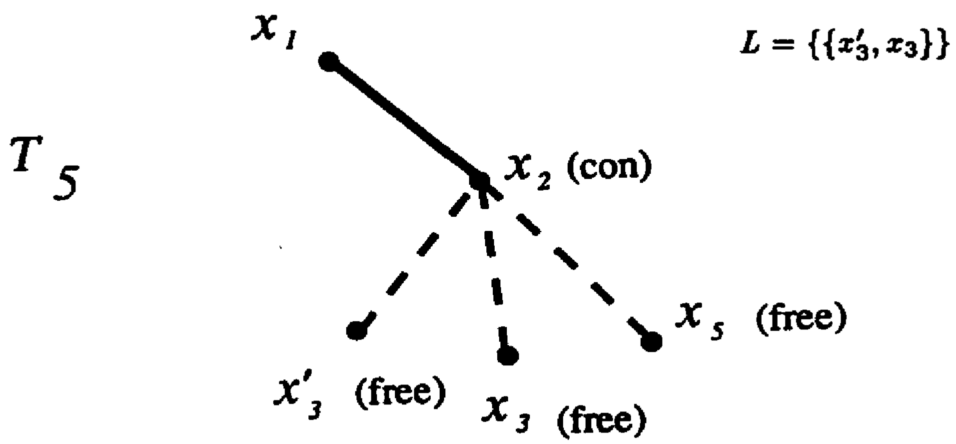
$\{x_1, x_2\}$ is examined for Update 1; x_2 gets label con

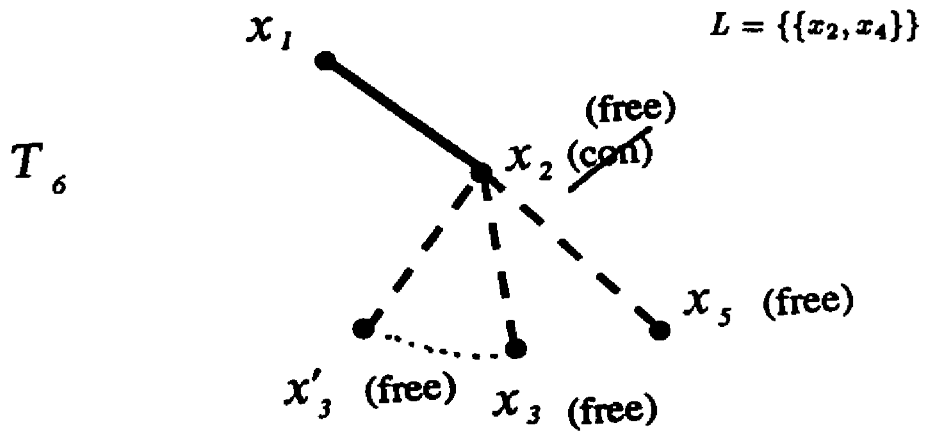


$L = \{\{x_2, x_5\}, \{x_2, x_3\}, \{x'_3, x_3\}\}$

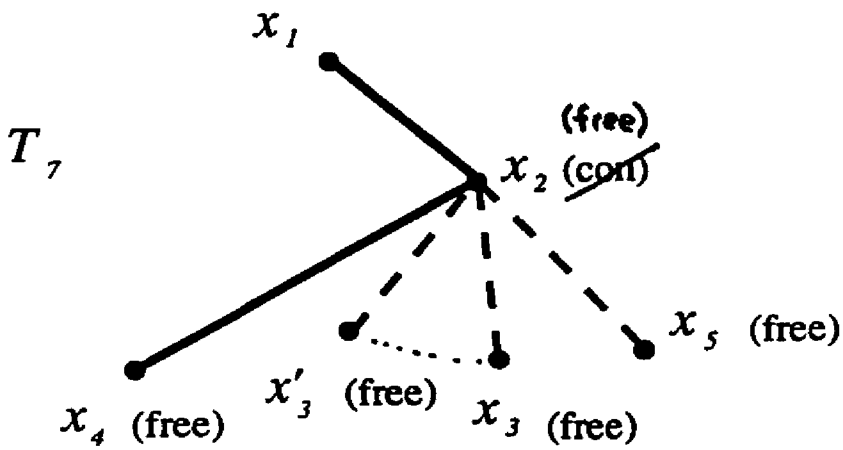


Note that even though $\{x'_3, x_4\}$ is in L, it would not be examined for update at this point since the subtree of F on vertices x_2, x_4, x'_3, x_5 is added to the search tree first under the conditions for selecting arcs from L.





Update 2 is performed on edge $\{x'_3, x_3\}$, causing x_2 to be relabeled free and $\{x_2, x_4\}$ to be added to L. Also, blosl(x_2) is arbitrarily set to x_3; adding edge $\{d, x_3\}$ to B.

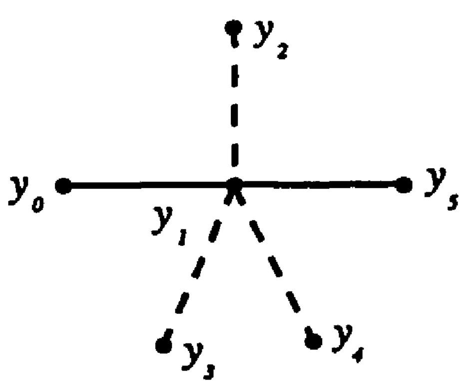


Vertex x_4 with $d_F(x_4) = 0$ is added to T and an augmenting chain is detected. The algorithm then calls CHAIN to construct the augmenting chain.

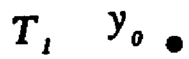
CHAIN is initialized with $C = x_4$ and x_2 under consideration. Since $\{x_2, x_4\} \notin F$ and since x_2 was relabeled free, Extension 2 is performed next, i.e., a recursive call to CHAIN is made to construct a chain from $\mathrm{blos1}(x_2) = x_3'$ to x_2. This recursive call produces a chain x_3', x_2. The reverse x_2, x_3' of x_3', x_2 is added to C to form $C = x_4, x_2, x_3'$ and x_3 is the next vertex to consider. Extension 1 occurs adding x_3 to C (so $C = x_4, x_2, x_3', x_3$) and x_2 is considered next. Again, Extension 1 occurs, adding x_2 to C (so $C = x_4, x_2, x_3', x_3, x_2$) and x_1 is considered. Finally, since x_1 is the root, x_1 is added to C and CHAIN outputs $C = x_4, x_2, x_3', x_3, x_2, x_1$, an augmenting chain.

Example 2:

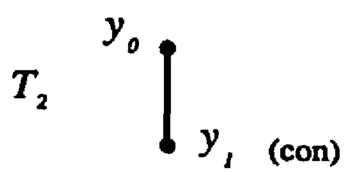
A simple example of a failed search with G given below. Forest edges are dashed. Let $b(y_1) = 3$ and all other degree constraints be 1.



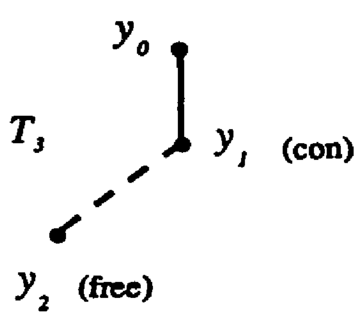
Consider a search from y_0. We show the trees at each stage and the corresponding L. The search fails when L becomes empty.



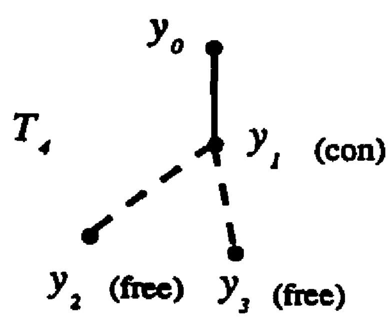
$L = \{\{y_0, y_1\}\}$



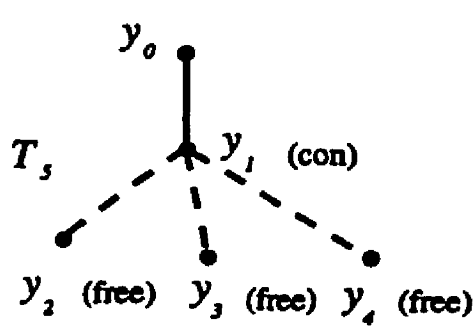
$L = \{\{y_1, y_2\}, \{y_1, y_3\}\{y_1, y_4\}\}$



$L = \{\{y_1, y_2\}, \{y_1, y_3\}\}$



$L = \{\{y_1, y_4\}\}$



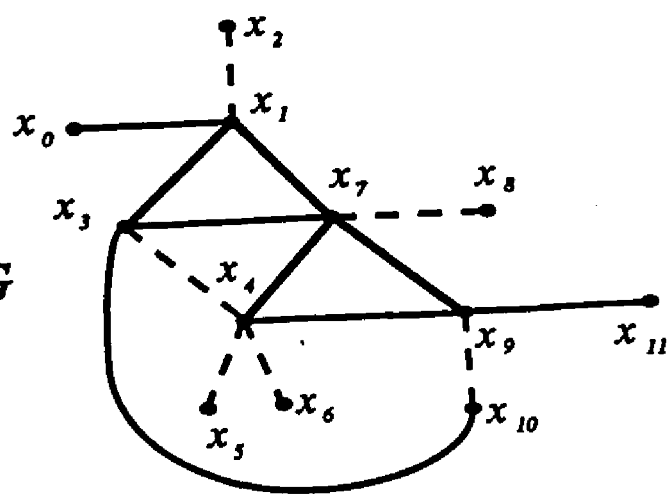
$L = \emptyset$

The procedure halts when $L = \emptyset$. The vertices of T_4 are discarded and G' consists of the isolated vertex y_5.

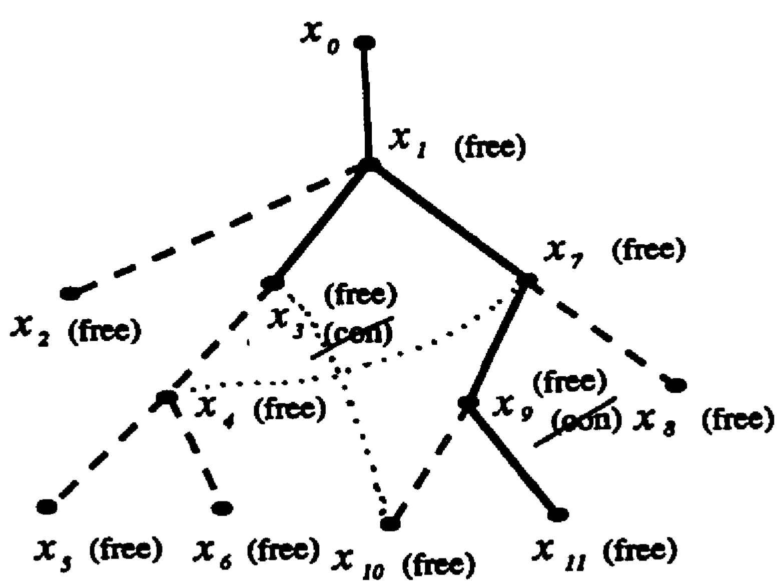
Example 3:

For the last example, we show the graph G (along with dashed forest edges) and the final SEARCH tree and simply list the updates. Blossom edges for the search tree are $\{x_3, x_{10}\}$ and $\{x_4, x_7\}$. These are indicated by dotted lines. Let $b(x_1) = b(x_4) = b(x_7) = 3$ and all other degree constraints be 1.

G



T



Initially T_0 consists of the isolated vertex x_0 and $L = \{\{x_0, x_1\}\}$. At each stage the edge which is examined is removed from L.

Iteration	Edge Examined	Update	Changes in L	Comments
1	$\{x_0, x_1\}$	1	add $\{x_1, x_2\}$, $\{x_1, x_3\}$, $\{x_1, x_7\}$	x_1 gets label free
2	$\{x_1, x_2\}$	1	—	edge $\{x_1, x_2\}$ is examined before other edges in L as it is an edge of the subtree of F containing x_1
3	$\{x_1, x_3\}$	1	add $\{x_3, x_4\}$	x_3 gets label con so $\{x_3, x_7\}$ and $\{x_3, x_{10}\}$ are not added to L
4	$\{x_4, x_5\}$	1	—	either $\{x_4, x_5\}$ or $\{x_4, x_6\}$ must be examined at this step
5	$\{x_4, x_6\}$	1	—	—
6	$\{x_1, x_7\}$	1	add $\{x_7, x_4\}$, $\{x_7, x_9\}$, $\{x_7, x_8\}$	x_7 gets label free, do not add $\{x_7, x_3\} \notin F$ to L since x_3 has label con
7	$\{x_7, x_8\}$	1	—	—
8	$\{x_7, x_9\}$	1	add $\{x_9, x_{10}\}$	x_9 gets label con; at this point $\{x_7, x_4\}$ could also be examined
9	$\{x_9, x_{10}\}$	1	—	—
10	$\{x_4, x_7\}$	2	add $\{x_3, x_7\}$, $\{x_3, x_{10}\}$	form a blossom; relabel x_3 to free; set nca(x_3) and nca(x_7) to x_1

11	$\{x_3, x_7\}$	2	—	nothing happens since $\mathrm{nca}(x_3) = \mathrm{nca}(x_7) = x_1$, i.e., Update 2 causes no new relabeling
12	$\{x_3, x_{10}\}$	2	add $\{x_9, x_{11}\}$, $\{x_9, x_4\}$	form blossom, relabel x_9 to free
13	$\{x_9, x_{11}\}$	1	—	x_{11} is islolated, an augmenting chain has been detected

From the SEARCH tree constructed above, the procedure CHAIN produces the augmenting chain $C = x_{11}, x_9, x_{10}, x_3, x_4, x_7, x_1, x_0$ with two recursive calls, which produce x_{10}, x_9 and x_4, x_3 whose reversals are added to C.

Chapter 3

The Reversing Number of a Digraph

3.1 Introduction

Recall that a tournament is a directed graph such that for each pair x, y of vertices exactly one one of the arcs (x, y) or (y, x) is present. Slater [1961] and Younger [1963] introduced the study of minimum sized sets of arcs which when reversed make a directed tournament acyclic. Call such a set a *minimum reversing set.* We[1] investigate a related question: Given an acyclic digraph D, what is the size of a smallest tournament T which has the arc set of D as a minimum reversing set. The *reversing number* of D is the number of 'extra vertices' in T. More formally, we make the following definitions.

Definition 3.1 *A* **reversing set** *of a tournament T is a set of arcs F, such that $(T \backslash F) \cup F^R$ is acyclic. A* **minimum reversing set** *in T is a reversing set of minimum size.*

The notation $(T \backslash F) \cup F^R$ will be used often in this chapter and indicates the tournament obtained by reversing the direction of the arcs in the set F.

Since an acyclic tournament has a unique acyclic order, we will talk about the acyclic order obtained after reversing the arcs of a reversing set. Given a general

[1] The material in this chapter is joint work contained in Barthelemy et al. [1990]. For completeness we include all material in this manuscript. Results which are not substantially the author's will be indicated with Barthelemy et al. in this chapter. Theorem 3.6 is the author's proof of a result of J.-P. Barthelemy. Theorem 3.7 was proved independently by F.S. Roberts and J.-P. Barthelemy and by the author. Source and sink extensions and Lemma 3.16 are generalizations of an idea due to O. Hudrey and J.-P. Barthelemy. The calculations for Table 3.1 are joint with B. Tesman. Theorem 3.23 is joint with B. Tesman. The cases $n \leq 7$ of Theorem 3.24 are due to O. Hudrey and the rest of the theorem is due independently to O. Hudrey and to B. Tesman and the author.

ordering σ of the vertices of a tournament, we define the set of *backwards arcs* relative to σ to be arcs (v, w) in the tournament such that $\sigma(w) < \sigma(v)$. With this notation, a reversing set F is the set of backwards arcs relative to the acyclic order of the tournament obtained by reversing the arcs in F.

Definition 3.2 *Given an acyclic digraph D, the* **reversing number** *$r(D)$ of D is $\min(|V(T)| - |V(D)|)$, where the minimum is taken over all tournaments T such that D is a minimum reversing set of T.*

The notion of reversing number is due to J.-P. Barthelemy (personal communication), who asked for which digraphs this concept is well defined. We show in Theorem 3.7 that the reversing number is well defined if and only if D is an acyclic digraph, justifying the definition. This chapter will establish basic bounds on the reversing number and examine exact values of the reversing number on some classes of digraphs.

In our study of reversing numbers we will make use of results on minimum reversing sets. Reversing sets have been studied by a number of authors in different contexts using different terminologies. In the electrical engineering literature *feedback arc sets*, sets of arcs whose removal makes a digraph acyclic, have been studied. Given a digraph D, it is easy to see that a minimum set of arcs whose removal makes D acyclic is also a minimum set of arcs whose reversal makes D acyclic and vice-versa, so the minimum feedback arc problem and the minimum reversing set problem are equivalent.

To see that the minimum feedback arc problem and the minimum reversing set problem are equivalent, note that it is obvious that any set of arcs whose reversal creates an acyclic digraph also creates an acyclic digraph by its removal (since the remaining arcs form an acyclic digraph). Conversely, let F be a minimal subset of the arc set A of a tournament whose removal makes the tournament acyclic. By minimality, if $(x, y) \in F$, then (x, y) is contained in a cycle $C = y, v_1, \ldots, v_k, x, y$ in $(A \setminus F) \cup \{(x, y)\}$. If there is a cycle C' in the tournament $(A \setminus F) \cup F^R$ obtained by reversing the arcs of F, then replace each arc $(y, x) \in F^R$ which is on C' with the path $y, v_1, \ldots, v_k, x$

from a cycle C containing (x, y) in $(A \setminus F) \cup \{(x, y)\}$. This results in a closed directed chain in $A \setminus F$. Such a chain contains a cycle, contradicting the fact that removal of F creates an acyclic digraph. Thus, the equivalence is established.

Runyon first suggested study of the feedback arc set problem. (His question is cited in the list of problems in Seeshu and Reed [1960] and is called the feedback cut set problem.) Tucker [1960] gave an integer programming formulation and Younger [1963] began the analysis of the structure of the feedback arc sets. Lawler [1964] formulated the problem of finding a minimum feedback arc set as a quadratic assignment problem. Hakimi [1965], Lempel and Cederbaum [1966], Kamae [1967], and Yau [1967] continued analysis of the structure of these sets and suggested algorithms and heuristics for finding minimum feedback arc sets in general. In addition, Karp [1972] showed that finding the size of a minimum reversing set, i.e. a minimum feedback arc set, is NP-complete in general.

In the statistics literature, Slater [1961] first suggested the study of *minimum sets of inconsistencies* of a preference ordering (ranking) with the observed relations from a complete paired comparison experiment. The graph theoretic model of paired comparison experiments has the objects being compared as vertices of a digraph and an arc from x to y if and only if x is preferred to y. A *nearest adjoining order* is a linear order such that the number of preferences inconsistent with that order is minimized. Since preferences in a linear order induce a complete acyclic tournament, minimizing the set of inconsistencies is the same as finding a minimum set of arcs whose reversal makes the preference digraph acyclic and vice-versa. Slater [1961] sought to determine the probability distribution over every tournament (outcomes of all possible paired comparisons) of the size of a minimum set of inconsistencies over all possible orderings. This work was continued by Alway [1962], Thompson and Remage [1964], Remage and Thompson [1966], Bermond [1972], Bermond and Kodratoff [1976], Monjardet [1973], Hubert [1976], and Baker and Hubert [1977], to name a few, with suggestions for algorithms and study of more general questions with different weightings on the amount of inconsistency. Hubert [1976] is a survey uniting the electrical engineering and statistics literature.

A third source of interest in minimum reversing sets arises in the mathematics literature. Erdös and Moon [1965] introduced the question of finding the greatest integer k such that every tournament on n vertices has a set of k consistent arcs (i.e., an acyclic subdigraph with k arcs). The study of this value has been continued by Reid [1969], Reid and Parker [1970], Spencer [1971,1980], and de la Vega [1983]. A number of authors have studied the computational aspects of determining a largest acyclic subdigraph of a digraph. The complement in a digraph of the arc set of a largest acyclic subdigraph is a minimum reversing set of the digraph and vice-versa. The polytope of the largest acyclic subdigraph problem has been studied by Grötschel, Jünger, and Reinelt [1984,1985] and Jünger [1985]. Korte [1979] examines approximation algorithms for this problem.

As we have already remarked, the problems mentioned above are all equivalent. (This has been proved by a number of authors.) Since reversing the arcs in a minimum reversing set makes a digraph acyclic, every cycle in the digraph must contain an arc from the minimum reversing set. That is, the arcs of a minimum reversing set are a *transversal* of the cycles in the digraph. In fact the minimum size of a transversal of cycles in a digraph is equal to the size of a minimum reversing set. (This follows from the fact that removing the arcs of a transversal of cycles creates an acyclic digraph and from the equivalence of minimum feedback arc sets and minimum reversing sets.) This has been shown by Dambit and Gindberg (cited in Bermond [1975]) and Remage and Thompson [1966].

Remark 3.1 In a tournament, the problems of finding a minimum reversing set, a minimum set of inconsistencies, a minimum feedback arc set, a largest acyclic subdigraph, and a minimum transversal of cycles are all equivalent.

See Jünger [1985] for more information on equivalent versions of the problem of finding a minimum reversing set and for applications.

Since a minimum reversing set is also a minimum transversal of the cycles, every arc in a minimum reversing set is contained in a cycle. In fact, we show in Theorem 3.6 that every arc of a minimum reversing set in a tournament must be contained in some cycle on three vertices (a 3-cycle). However, while the largest collection of arc disjoint cycles in a digraph provides a lower bound on the size of a minimum reversing set, this bound is not tight. Kotzig [1975] and Bermond and Kodratoff [1976] have shown that for $n \geq 10$ the bound is not tight even for tournaments, i.e., for $n \geq 10$ there exist tournaments on n vertices such that the size of a minimum reversing set is strictly greater than the largest collection of disjoint cycles in the tournament (see also Bermond and Thomassen [1981]).

In Section 3.2, we review basic results on reversing sets which are useful in the study of reversing numbers. We also show that the reversing number is well defined. In Section 3.3, we develop some basic bounds on the reversing number. In particular, we show that the reversing number of a tournament on n vertices is an upper bound on the reversing number of any acyclic digraph on n vertices. We also show a lower bound of $n-1$ on the reversing number of an acyclic digraph on n vertices if the digraph contains a Hamiltonian path. Graphs with reversing number 0 are studied in Section 3.4. Using a technique to extend a digraph on n vertices to a digraph on $n+1$ vertices without increasing the reversing number, we show that there are connected acyclic digraphs with reversing number 0 for $n \geq 8$. A parameter $d(n,r)$ giving the size of the largest arc set of an acyclic digraph on n vertices with reversing number r is also introduced in Section 3.4. Bounds on $d(n,1)$ and $d(n,0)$ are examined. Section 3.5 shows that the reversing number of an acyclic tournament on n vertices is between $2n - 4\log_2 n$ and $2n-2$. Finally, Section 3.6 establishes exact values of the reversing number for directed stars, disjoint arcs, paths, and complete bipartite digraphs.

3.2 Basic Results on Minimum Reversing Sets

The following lemmas regarding reversing sets will be useful in the study of reversing numbers. The first three are from Younger [1963]. All follow easily from the definitions above.

Lemma 3.1 (Younger 1963) *If F is a minimum reversing set of a tournament T then, for $F' \subseteq F$, F' is a minimum reversing set of $T' = (T \setminus B) \cup B^R$ where $B = F \setminus F'$.*

Lemma 3.2 (Younger 1963) *If a vertex v is a source or sink in a tournament T, then F is a minimum reversing set of T if and only if F is a minimum reversing set of $T \setminus \{v\}$.*

Recall that an acyclic tournament has a unique acyclic order.

Lemma 3.3 (Younger 1963) *If T is a tournament and F is a minimum reversing set such that $\pi(v_1) < \pi(v_2) < \cdots < \pi(v_n)$ is the acyclic ordering after reversal of the arcs in F, then for any segment $v_i, v_{i+1}, \ldots, v_{i+j} = S$, $F|_S$ is a minimum reversing set of $T|_S$.*

Lemma 3.1 says that if F is a minimum reversing set of a tournament T then for any subset F' of F, if we reverse in T the arcs which are in F but not in F' the new tournament T' has F' as a minimum reversing set. If T' had a smaller reversing set B then $(F \setminus F') \cup B$ would be a reversing set of T smaller than F. Lemma 3.2 states that no arc in a minimum reversing set of a tournament T has a tail which is a source in T or a head which is a sink in T. Lemma 3.3 is a direct consequence of Lemmas 3.1 and 3.2.

Lemma 3.4 *If T is a tournament and W is any subset of the vertices of T, then for a minimum reversing set F of T, the number of arcs in F joining vertices of W is greater than or equal to the size of a minimum reversing set of $T|_W$.*

Proof: The arcs in F with both ends in W form a reversing set of T restricted to W. $\square$

Lemma 3.5 *If τ is a collection of arc disjoint cycles in a tournament T, then for each reversing set F in T,*

$$|\tau| \leq |F|.$$

Proof: If $C \cap F = \emptyset$ for a cycle C in T, then C is a cycle in $(T \setminus F) \cup F^R$, contradicting the assumption that F is a reversing set. So each cycle contains at least one arc from F. Since the cycles are arc disjoint the bound follows. $\square$

We have mentioned in the introduction that each arc of a minimum reversing set of a tournament T is in a 3-cycle of T. The proof of this is given in Theorem 3.6. Note that with our notation, x, y, z, x is a 3-cycle since there are three distinct vertices.

Theorem 3.6 (Barthelemy et al. 1990) *Let T be a tournament and let F be a minimum reversing set of T. Then every arc of F belongs to some 3-cycle of T.*

Proof: Consider an arc $(y, z) \in F$. Reversing the arcs of F which do not meet y or z will not affect inclusion of (y, z) in a 3-cycle of T. By Lemma 3.1 reversing these arcs yields a new F and T with $(y, z) \in F$ and F a minimum reversing set of T. Assume that the vertices are labeled so that the acyclic ordering π of $(T \setminus F) \cup F^R$ is $\pi(x_1) < \pi(x_2) < \cdots < \pi(x_n)$. So every arc of F goes from x_j to x_k for some $j > k$. Note that deleting vertices v such that $\pi(v) < \pi(z)$ or $\pi(v) > \pi(y)$ will not form new 3-cycles. Thus, we may assume that $(y, z) = (x_n, x_1)$. It also follows from the reversal of arcs not meeting y or z that every arc of F has the form (x_j, x_1) or (x_n, x_j) since arcs (x_j, x_i) for $1 < i < j < n$ do not meet $y = x_n$ or $z = x_1$.

For $k = 1, n$, let

$$X_k^+ = \{(x_k, x_j) \in T : 1 < j < n\}$$

$$X_k^- = \{(x_j, x_k) \in T : 1 < j < n\}$$

Note that the four sets described above are all disjoint and that $F = X_1^- \cup X_n^+ \cup \{(x_n, x_1)\}$. Also, since all arcs of T which join x_i to x_j, $1 < i < j < n$, go from x_i to x_j, it follows that $\left[T \setminus (X_1^+ \cup X_n^-)\right] \cup (X_1^+ \cup X_n^-)^R$ is acyclic with acyclic ordering π' satisfying $\pi'(x_n) < \pi'(x_2) < \cdots < \pi'(x_{n-1}) < \pi'(x_1)$. Since F is a minimum reversing set, we have

$$|X_1^+| + |X_n^-| = |X_1^+ \cup X_n^-| \geq |F| = |X_1^-| + |X_n^+| + 1$$

Thus, since $|X_1^+| + |X_1^-| + |X_n^+| + |X_n^-| = 2(n-2)$, we have $|X_1^+| + |X_n^-| > (n-2)$. So, by the pigeonhole principle, there exists a j with $1 < j < n$ such that both (x_1, x_j) and (x_j, x_n) are in $X_1^+ \cup X_n^- \subset T$. Then x_1, x_j, x_n, x_1 is a 3-cycle in T containing (x_n, x_1). □

We have noted that minimum reversing sets are necessarily acyclic. The next theorem shows that every acyclic digraph arises as a minimum reversing set of some tournament.

Theorem 3.7 (Barthelemy et al. 1990) *Let D be a digraph. The following two conditions are equivalent:*

(i) D is acyclic.

(ii) D is a minimum reversing set of some tournament.

Proof: If D contains a cycle then so does D^R; thus every reversing set must be acyclic. Conversely, assume that D is acyclic. Assume also that the vertices $V(D) = \{u_1, u_2, \ldots, u_n\}$ are labeled so that there is an acyclic ordering π of D satisfying $\pi(u_1) < \pi(u_2) < \cdots < \pi(u_n)$. We construct a tournament T with minimum reversing set D as follows. Let $V(T) = V(D) \cup \{v_{ij} : (u_i, u_j) \in D\}$. Let T' be an acyclic tournament on $V(T)$ with acyclic ordering π' satisfying $\pi'(u_n) < \pi'(u_{n-1}) < \cdots < \pi'(u_1)$ and $\pi'(u_j) < \pi'(v_{ij}) < \pi'(u_i)$ for all v_{ij}. This can be done since $v_{ij} \in V(T) \Leftrightarrow (u_i, u_j) \in D \Rightarrow i < j$. Thus corresponding to each arc (u_i, u_j) of D there is an extra vertex v_{ij} which falls between the ends of the arc in the ordering π'.

Note that $D^R \subseteq T'$, so we can define $T = (T' \setminus D^R) \cup D$, i.e. , $T' = (T \setminus D) \cup D^R$. Since T' is acyclic, D is a reversing set of T. Also, $\tau = \{u_i, u_j, v_{ij}, u_i : (u_i, u_j) \in D\}$ is a collection of arc disjoint 3-cycles in T with $|\tau| = |D|$. Therefore, by Lemma 3.5, D is a minimum reversing set of T. □

It follows from Theorem 3.7 that the reversing number $r(D)$ is well defined and that

$$r(D) \leq |D|. \tag{3.1}$$

Here and in the remainder of this chapter, we use the notation $|D|$ to indicate the size of the arc set of D when there is no chance of confusion. This notation is consistent with the idea that we are viewing the arc sets of the digraphs as reversing sets.

Given an acyclic digraph D and tournament T, if D is a minimum reversing set of T and no tournament with fewer vertices than T has D as a minimum reversing set then we say that T *realizes* D. If T realizes an acyclic digraph D, then $r(D)$ is the number of vertices in $V(T) \setminus V(D)$.

3.3 Basic Results on Reversing Number

In this section we make use of basic results on minimum reversing sets to establish some elementary facts about the reversing number. We first get a bound on the reversing number of an acyclic digraph in terms of the reversing number of a tournament by using a more general bound on the reversing number of subdigraphs.

Theorem 3.8 *Let $D' \subseteq D$ be acyclic digraphs on n vertices. Then $r(D') \leq r(D)$.*

Proof: By Lemma 3.1, if T is a tournament having D as a minimum reversing set then there is a tournament T' on the same number of vertices having D' as a minimum reversing set. □

Note here that it is important that both D and D' have the same number of vertices; otherwise Theorem 3.8 is not true. For example a single arc has reversing number 1 (Theorem 3.13), but many nontrivial acyclic digraphs have reversing number 0 (Theorem 3.17).

Corollary 3.9 *For an acyclic digraph D on n vertices, $r(D) \leq r(T_n)$ where T_n is the acyclic tournament on n vertices.*

Theorem 3.20 will give some bounds on the reversing number of acyclic tournaments. These together with Corollary 3.9 will give general bounds on the reversing number of any acyclic digraph.

We next take note of several basic results for getting bounds on the reversing number of an acyclic digraph D.

Lemma 3.10 *For an acyclic digraph D, $r(D) = r(D^R)$.*

Proof: For any tournament T, $(T \setminus D) \cup D^R$ is acyclic if and only if $(T^R \setminus D^R) \cup D$ is acyclic. Thus D is a minimum reversing set of T if and only if D^R is a minimum reversing set of T^R. $\square$

Lemma 3.11 *Let D be an acyclic digraph and let T realize D. If $\pi(v_1) < \pi(v_2) < \cdots < \pi(v_n)$ is the acyclic ordering of $(T \setminus D) \cup D^R$, then for any segment $S = v_i, v_{i+1}, \ldots, v_{i+j}$, the number of vertices v in S such that $v \notin V(D)$ is greater than or equal to the reversing number of $D|_S$.*

Proof: By Lemma 3.3, $D|_S$ is a minimum reversing set of $T|_S$. Thus $T|_S$ has at least as many vertices which are not in $V(D)$ as a tournament realizing $D|_S$. $\square$

Let D be an acyclic digraph with vertex set V. For any $v \in V$, suppose $V \setminus \{v\}$ can be partitioned as $V_1' \cup V_2'$ such that in every acyclic ordering of D, the vertices of

V_1' come before v and the vertices of V_2' come after v. Suppose also that there are no arcs from V_1' to V_2'. Then v will be called an *order splitting vertex* of D and V_1' is its *opening set* and V_2' its *closing set.* By the definition of acyclic orderings, there are also no arcs from V_2' to V_1'.

Lemma 3.12 *If v is an order splitting vertex of an acyclic digraph D, and V_1' and V_2' its opening and closing sets respectively, then $r(D) = r(D_1) + r(D_2)$, where D_1 and D_2 are the digraphs induced by $V_1 = V_1' \cup \{v\}$ and $V_2 = V_2' \cup \{v\}$, respectively.*

Proof: Let T realize D and π be an acyclic ordering of $(T \setminus D) \cup D^R$. Let W_1 be those vertices x of $V(T)$ with $\pi(x) \geq \pi(v)$ and W_2 the vertices x of $V(T)$ with $\pi(x) \leq \pi(v)$. Note that v is in both of these sets and that $V_1 \subseteq W_1$ and $V_2 \subseteq W_2$. By Lemma 3.11, $r(D_1) \leq |W_1 \setminus V_1|$ and $r(D_2) \leq |W_2 \setminus V_2|$ and so $r(D_1) + r(D_2) \leq |W_1 \setminus V_1| + |W_2 \setminus V_2| = r(D)$.

To show the reverse inequality, we construct a tournament T' on $r(D_1) + r(D_2) + |V(D)|$ vertices having D as a minimum reversing set. Let T_1 realize D_1 and T_2 realize D_2. For $i = 1, 2$ denote the vertex set of T_i by W_i. We can choose W_1 and W_2 so that $W_1 \setminus \{v\}$ and $W_2 \setminus \{v\}$ are disjoint. Then $(T_1 \setminus D_1) \cup D_1^R$ is an acyclic tournament. Let π' be the acyclic ordering of $(T_1 \setminus D_1) \cup D_1^R$ and let w denote the (unique) source in $(T_1 \setminus D_1) \cup D_1^R$. If $w \in V(T_1) \setminus V_1'$, then by Lemma 3.3, D_1 is a minimum reversing set of $T_1 \setminus \{w\}$, contradicting the assumption that T_1 realizes D_1. If $w \in V_1'$ then the reverse σ of the ordering on V_1 defined by π' is an acyclic ordering of D_1 for which v is not the last vertex. Since there are no arcs between V_1' and V_2' in D, we can combine σ with any acyclic ordering (with respect to $D_2|_{V_2'}$) of V_2' to follow σ. This gives an acyclic ordering of D for which not all the vertices of V_1' appear before v, contradicting the fact that V_1' is the opening set for the order splitting vertex v. Thus the source w in $(T_1 \setminus D_1) \cup D_1^R$ must be v. In a similar manner, it can be shown that $(T_2 \setminus D_2) \cup D_2^R$ is an acyclic tournament with v as a sink.

Let T' be the tournament formed by joining T_1 and T_2 at v with all arcs between T_1 and T_2 going from T_2 to T_1. Note that the arc set of T' can be partitioned into three parts, the arc set of T_1, the arc set of T_2, and the set of arcs between $W_1 \setminus \{v\}$ and $W_2 \setminus \{v\}$, all of which are directed from $W_2 \setminus \{v\}$ to $W_1 \setminus \{v\}$.

Since there are no arcs between V_1' and V_2', the arc set of D is partitioned into the arc set of D_1 and the arc set of D_2. So $D = D_1 \cup D_2$ and $|D| = |D_1| + |D_2|$ since these sets are disjoint. Consider $T = (T' \setminus D) \cup D^R$. Since D_i is a reversing set of T_i for $i = 1, 2$, $T|_{W_1}$ and $T|_{W_2}$ are acyclic. (This uses the fact that the arc sets of $T|_{W_1}$ and $T|_{W_2}$ are disjoint.) Since also all arcs in T between W_2 and W_1 are directed from W_2 to W_1, T is acyclic. Thus D is a reversing set of T'.

Finally, we show that every reversing set of T' has size $|D|$, and thus that D is a minimum reversing set of T'. If F is a minimum reversing set of T', then $\left|F|_{W_1}\right| \geq |D_1|$ by Lemma 3.4 and the fact that D_1 is a minimum reversing set of $T'|_{W_1}$. Similarly, $\left|F|_{W_2}\right| \geq |D_2|$. Since the arc sets $F|_{W_1}$ and $F|_{W_2}$ are disjoint, $|F| \geq \left|F|_{W_1}\right| + \left|F|_{W_2}\right| \geq |D_1| + |D_2| = |D|$. The last equality follows since there are no arcs between V_1' and V_2' in D. Thus, D is a minimum reversing set of T' and $r(D) \leq |D_1| + |D_2|$. □

Recall that the directed path P_n on n vertices is the digraph with vertex set $\{v_1, \ldots, v_n\}$ and arc set $\{(v_i, v_{i+1}) : i = 1, \ldots, n-1\}$.

Theorem 3.13 *Let P_n be the directed path P_n on n vertices. Then, $r(P_n) = n - 1$.*

Proof: A single arc P_2 has $r(P_2) = 1$ since it is not a minimum reversing set of itself (the only tournament on 2 vertices) and it is a minimum reversing set of a 3-cycle. By repeated application of Lemma 3.12 the result follows since every vertex of P_n is order splitting. □

Corollary 3.14 *If D is an acyclic digraph on n vertices containing a Hamiltonian path, then $r(D) \geq n-1$.*

Proof: Apply Theorem 3.8 to the result of Theorem 3.13. □

Note that if a digraph has a unique acyclic ordering, then it contains a Hamiltonian path. Then by the Corollary, a digraph on n vertices with a unique acyclic ordering has reversing number at least $n-1$. However, when there is not a unique acyclic ordering, the reversing number can be small. The next theorem states a necessary condition for the reversing number to be 0.

Theorem 3.15 *If $r(D) = 0$, then D has a least two distinct sources and at least two distinct sinks.*

Proof: Let $V(D) = \{v_1, v_2, \ldots, v_n\}$, let T realize D, and let π be the acyclic ordering of $T' = (T \setminus D) \cup D^R$. Note that $(v_i, v_j) \in D \Rightarrow \pi(v_i) > \pi(v_j)$. Since $r(D) = 0$ and T' is acyclic, we may assume that $\pi(v_i) = i$, $i = 1, 2, \ldots, n$. Thus, v_1 is a sink of D. If $(v_2, v_j) \in D$ then $j = 1$. However, if $(v_2, v_1) \in D$ then by Lemma 3.3 applied to $v_1, v_2 = S$, $D|_S = (v_2, v_1)$ is a minimum reversing set of the (acyclic) tournament on 2 vertices, a contradiction. Thus v_2 must also be a sink of D. By a similar argument there must be at least two distinct sources. □

3.4 Small Reversing Numbers

We will next consider the smallest reversing number among digraphs on n vertices. For $n \geq 2$, let $r_n = \min r(D)$, where the minimum is taken over all acyclic digraphs D on n vertices having no isolated vertices. Also for $n \geq 2$, let $r'_n = \min r(D)$, where the minimum is taken over all connected acyclic digraphs D on n vertices. Clearly we have $r_n \leq r'_n$, for every $n \geq 2$.

In order to calculate these parameters we introduce conditions under which extending certain digraphs will produce new digraphs without increasing the reversing number. Let D be an acyclic digraph, let T realize D and let τ be a collection of $|D|$ arc disjoint cycles in T. (Note that it is not necessary that a T realizing D contain such a collection τ.) Also let $S = \{(x_i, y_i) \in T : i = 1, 2, \ldots, k\}$ be a collection of arcs from T none of which is an arc of one of the cycles in τ and assume that S is vertex disjoint, i.e., the x_i and y_i are all distinct. Let z be any element not in $V(T)$. We define two new digraphs: D', the *sink extension of D with respect to S,* and D'', the *source extension of D with respect to S,* as follows.

$$V(D') = V(D'') = V(D) \cup \{z\}$$

$$A(D') = A(D) \cup \{(z, x_i) : i = 1, 2, \ldots, k\}$$

$$A(D'') = A(D) \cup \{(y_i, z) : i = 1, 2, \ldots, k\}$$

We also define T', the D' extension of T with respect to S, and T'', the D'' extension of T with respect to S, as follows. Let $M = \{x_1, \ldots, x_k\} \cup \{y_1, \ldots, y_k\}$ be the set of vertices which are endpoints of the arcs in S. Let T' and T'' have vertex sets $V(T') = V(T'') = V(T) \cup \{z\}$ and arc sets

$$A(T') = A(T) \cup \{(z, x_i), (y_i, z) : i = 1, 2, \ldots, k\} \cup \{(v, z) : v \in V(T) \setminus M\}$$

$$A(T'') = A(T) \cup \{(z, x_i), (y_i, z) : i = 1, 2, \ldots, k\} \cup \{(z, v) : v \in V(T) \setminus M\}.$$

Finally, we define the extensions τ' and τ'' of τ with respect to S by,

$$\tau' = \tau'' = \tau \cup \{x_i, y_i, z, x_i : i = 1, 2, \ldots, k\}.$$

Lemma 3.16 *Let D be an acyclic digraph with reversing number $r(D)$, and let T realize D. Assume also that there is a collection τ of $|D|$ arc disjoint cycles in T and a set S of vertex disjoint arcs in T, none of which is an arc of a cycle from τ. Let D' be the sink extension of D with respect to S, T' be the D' extension of T with respect to S,* *and τ' the extension of τ with respect to S. Also, let D'' be the source extension of D with respect to S, T'' be the D'' extension of T with respect to S, and τ'' the extension of τ with respect to S. Then the following hold:*

(i$'$) *τ' is a collection of $|D| + |S|$ arc disjoint cycles in T',*

(ii$'$) *D' is a minimum reversing set of T',*

(iii$'$) $r(D') \leq r(D)$.

and

(i$''$) *τ'' is a collection of $|D| + |S|$ arc disjoint cycles in T'',*

(ii$''$) *D'' is a minimum reversing set of T'',*

(iii$''$) $r(D'') \leq r(D)$.

Proof: Let $S = \{(x_i, y_i) \in T : i = 1, \ldots, k\}$. The cycles added to τ to obtain $\tau' = \tau''$ are arc disjoint from τ by the choice of S and since $z \notin V(T)$. Thus $|\tau'| = |\tau''| = |\tau| + |S|$. Also by the definitions of T', T'', τ', and τ'', each of the cycles in τ' is in T' and each of the cycles of τ'' is in T''. Thus (i$'$) and (i$''$) hold.

Note that $(T' \setminus D') \cup (D')^R$ is acyclic since $(T \setminus D) \cup D^R$ is acyclic, and that z is a sink in $(T' \setminus D') \cup (D')^R$. Analogously, $(T'' \setminus D'') \cup (D'')^R$ is acyclic with source z. Thus D' is a reversing set of T' and D'' is a reversing set of T''. By Lemma 3.5 applied to τ' and τ'', minimum reversing sets of T' and T'' have size at least $|\tau'|$ and $|\tau''|$, i.e., each has size at least $|\tau| + |S|$. Then, since D' is a reversing set of size $|D'| = |D| + |S| = |\tau| + |S| = |\tau'|$, D' is a minimum reversing set of T' and (ii$'$) holds. Similarly, D'' is a minimum reversing set of T'' and (ii$''$) holds.

Note that $|V(T')| = |V(T'')| = |V(T)| + 1$. Since D' is a minimum reversing set of T', $r(D') \leq |V(T')| - |V(D')| = |V(T)| + 1 - (|V(D)| + 1) = r(D)$, and similarly $r(D'') \leq r(D')$. So (iii$'$) and (iii$''$) hold. □

We make use of this lemma to construct connected digraphs with reversing number 0 for $n \geq 7$.

Determining r_n and r'_n for $n < 7$ requires some case analysis. In order to do this we review a result of Bermond and Kodratoff [1976]. We look at the following upper bounds on the size of a minimum reversing set of a tournament on n vertices. Let m_n denote the maximum size of a minimum reversing set, where the maximum is taken over all tournaments on n vertices. Bermond and Kodratoff [1976] show that $m_2 = 0$, $m_3 = m_4 = 1$, $m_5 = 3$, $m_6 = 4$ and $m_7 = 7$.

Theorem 3.17 (Barthelemy et al. 1990) $r_2 = r'_2 = 1; r_3 = r'_3 = 2; r_4 = r'_4 = r_5 = r'_5 = 1; r_6 = 0, r'_6 = 1$ *and for* $n \geq 7, r_n = r'_n = 0$.

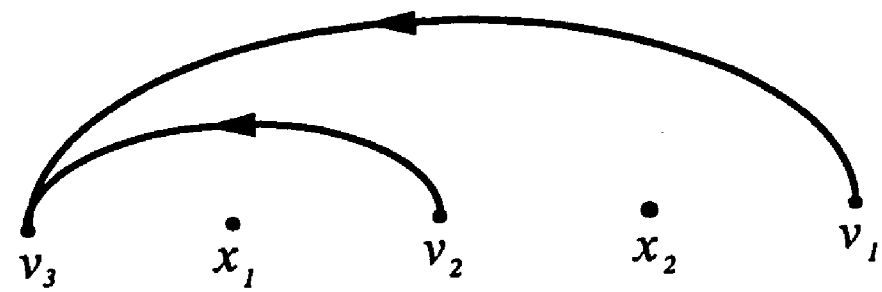
Proof: We first consider cases when n is small.

Case $n = 2$: The only acyclic digraph on 2 vertices with no isolated vertices is an arc which is not a minimum reversing set of itself and is a minimum reversing set of a 3-cycle. Thus $r_2 = r'_2 = 1$.

Case $n = 3$: Every digraph on 3 vertices with no isolated vertices has at least two arcs and is connected. So $r_3 = r'_3$. Since $m_3 = m_4 = 1$ there is no tournament on 3 or 4 vertices having a connected digraph on three vertices as a minimum reversing set. Figure 3.1 shows a tournament on five vertices, with a connected digraph on three vertices as a minimum reversing set, so $r_3 = r'_3 = 2$.

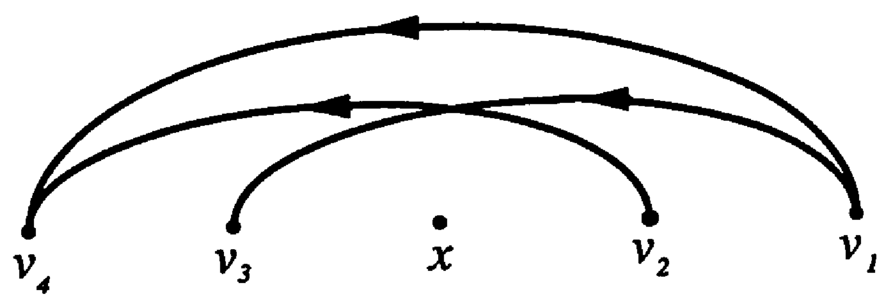
Case $n = 4$: An acyclic digraph on 4 vertices with no isolated vertex has at least 2 arcs. Since $m_4 = 1$, $r_4 \geq 1$ and $r'_4 \geq 1$. Figure 3.2 shows a connected digraph on 4 vertices and a tournament realizing it, so $r_4 = r'_4 = 1$.

Case $n = 5$: Any acyclic digraph on 5 vertices with no isolated vertex has at least 3 arcs. Recall that the outdegree $d^+_T(x)$ of vertex x in T is the number of arcs $(i, j) \in T$. Consider any tournament T on 5 vertices. If some vertex x in T has outdegree 4 then



All arcs which are not shown are directed from left to right in the figure.

Figure 3.1: A tournament realizing a connected digraph on three vertices, containing disjoint cycles (v_1, v_3, x_2, v_1) and (v_2, v_3, x_1, v_2).



All arcs which are not shown are directed from left to right in the figure.

Figure 3.2: A tournament on five vertices realizing a connected digraph on four vertices, containing disjoint cycles (v_2, v_4, v_3, v_2), (v_1, v_4, x, v_1), and (v_1, v_3, x, v_2, v_1).

x is a source and by Lemma 3.2, a minimum reversing set of T is a minimum reversing set of $T \setminus \{x\}$. Since $m_4 = 1$, the maximum size of a minimum reversing set of such a tournament is 1 and thus T cannot realize a digraph on 5 vertices containing no isolated vertex.

Consider tournaments T on 5 vertices having no vertex with outdegree 4 and some vertex x with $d_T^+(x) = 3$. Then reverse the arc for which x is the head to obtain a new tournament T' which has a vertex of degree 4 and as above a minimum reversing set of size at most 1. Thus T has a reversing set of size at most 2. Then a minimum reversing set of T has size at most 2 and T cannot realize a digraph on 5 vertices.

Finally, if T is a tournament on 5 vertices such that $d_T^+(x) \leq 2$ for all vertices x in T, then T is a regular tournament with all 5 vertices having degree 2. All such tour-

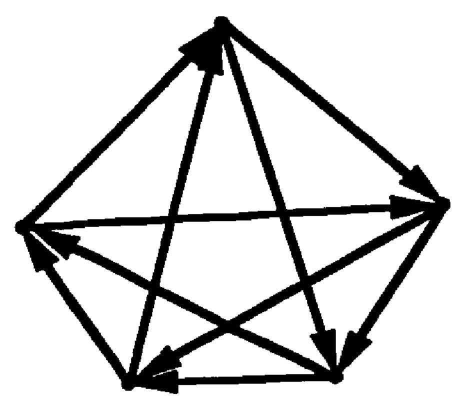
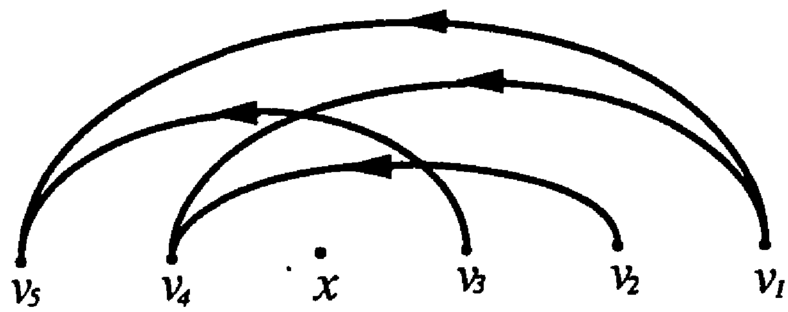


Figure 3.3: A regular tournament on five vertices.

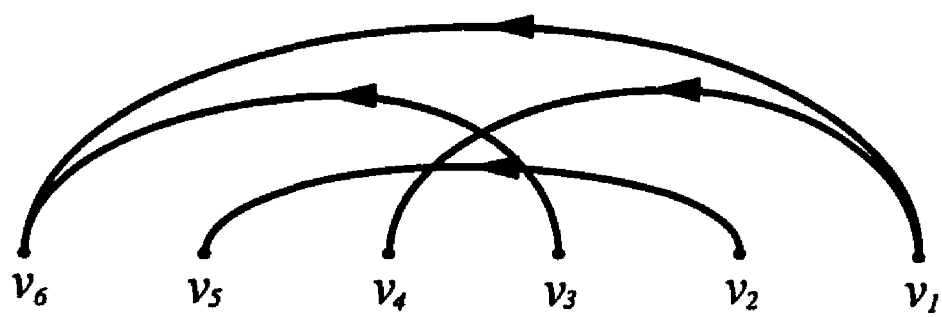
naments are isomorphic to the tournament shown in Figure 3.3. It is straightforward to show that all of its minimum reversing sets have three arcs and hence contain an isolated vertex. Thus $r_5 \geq 1$ and $r'_5 \geq 1$. Figure 3.4 gives an example to show that $r_5 = r'_5 = 1$.

Case $n = 6$: Figure 3.5 shows that $r_6 = 0$. Any connected digraph on 6 vertices has at least 5 arcs and since $m_6 = 4$ no tournament on 6 vertices realizes a connected digraph on 6 vertices. Thus $r'_6 \geq 1$ and Figure 3.6 gives an example to show that $r'_6 = 1$.



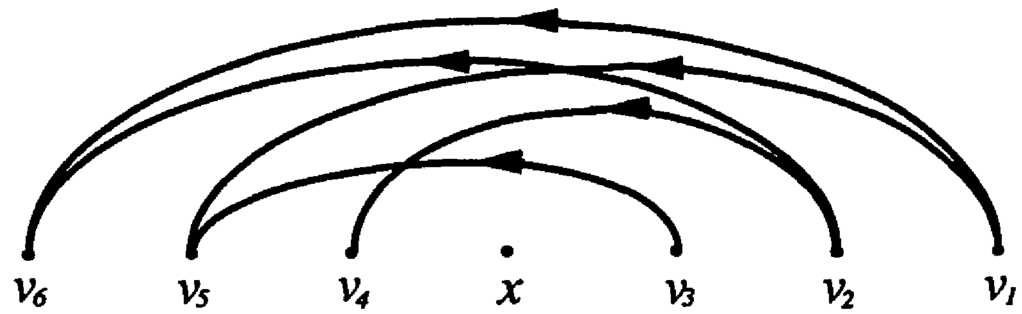
All arcs which are not shown are directed from left to right in the figure.

Figure 3.4: A tournament on six vertices realizing a connected digraph on five vertices, containing arc disjoint cycles (v_2, v_4, v_3, v_2), (v_1, v_4, x, v_1), (v_1, v_5, v_2, v_1), and (v_3, v_5, x, v_3).



All arcs which are not shown are directed from left to right in the figure.

Figure 3.5: A tournament on six vertices realizing a digraph on six vertices, containing arc disjoint cycles (v_3, v_6, v_5, v_3), (v_1, v_6, v_2, v_1), (v_2, v_5, v_4, v_2),and (v_1, v_4, v_3, v_1).



All arcs which are not shown are directed from left to right in the figure.

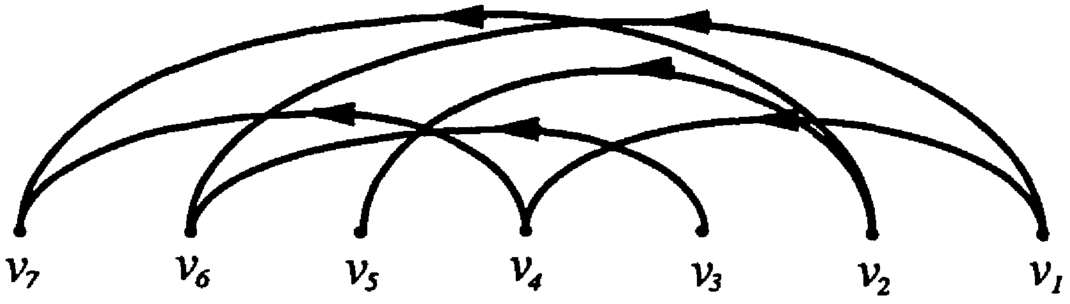
Figure 3.6: A tournament on seven vertices realizing a connected digraph on six vertices, containing arc disjoint cycles (v_3, v_5, v_4, v_3), (v_2, v_4, x, v_2), (v_2, v_6, v_3, v_2), (v_1, v_5, v_4, v_1), and (v_1, v_6, v_4, v_1).

Case $n = 7$: We exhibit in Figure 3.7 a connected acyclic digraph D_7 on 7 vertices, along with a T having D_7 as a reversing set, and a collection τ of $6 = |D_7|$ arc disjoint cycles in T. Thus $r(D_7) = 0$. This shows that $r_7 = r'_7 = 0$.

Case $n \geq 8$: We will show that alternating paths on n vertices, $n \geq 8$, have reversing number 0. Recall that each alternating path can be labeled so that it is the following digraph A_n or its reversal: $V(A_n) = \{v_1, \ldots, v_n\}$ and the arc set $A(A_n) = \{(v_i, v_{i-1}), (v_i, v_{i+1}) : i \text{ is odd, and both vertices are in } V\}$. Recall also that Lemma 3.10 says that $r(D) = r(D^R)$ for all D. Thus, in order to prove the result for all alternating paths it is enough to consider A_n.

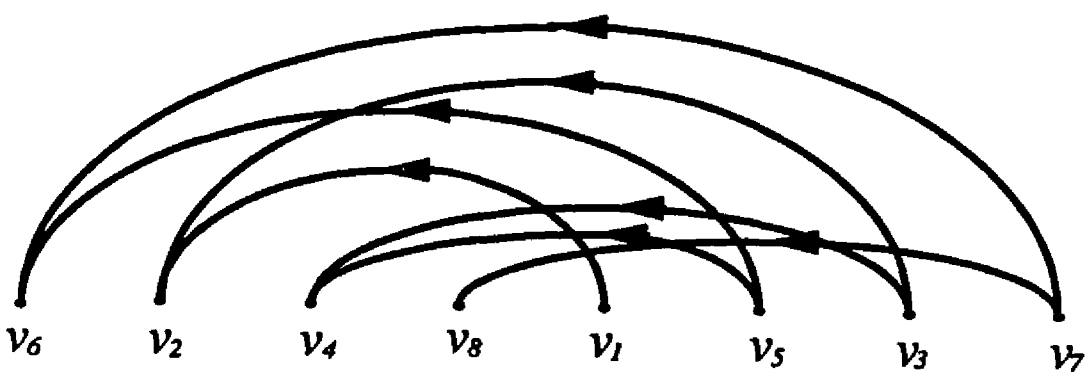
By our convention of denoting the size of the arc set by $|A_n|$, we have $|A_n| = n-1$. Figure 3.8 exhibits a tournament $T(A_8)$ with A_8 as a reversing set. This tournament contains a set τ_8 of seven arc disjoint cycles and so, by Lemma 3.5, A_8 is a minimum reversing set of the tournament and $r(A_8) = 0$.

Note that $(v_8, v_6) \in T(A_8)$ and this arc is not an arc of any cycle in τ_8. Denoting



All arcs which are not shown are directed from left to right in the figure.

Figure 3.7: A tournament on seven vertices realizing a connected digraph on seven vertices, containing arc disjoint cycles (v_1, v_6, v_5, v_1), (v_2, v_7, v_6, v_2), (v_4, v_7, v_5, v_4), (v_2, v_5, v_3, v_2), (v_3, v_6, v_4, v_3), (v_1, v_4, v_2, v_1).



All arcs which are not shown are directed from left to right in the figure.

Figure 3.8: A tournament realizing A_8 and the set $\tau_8 = (v_1, v_2, v_8, v_1)$, (v_3, v_2, v_5, v_3), (v_3, v_4, v_1, v_3), (v_5, v_4, v_8, v_5), (v_5, v_6, v_1, v_5), (v_7, v_6, v_2, v_7), (v_7, v_8, v_3, v_7).

the new vertex z in the sink extension by v_9, the sink extension of A_8 with respect to $S = (v_8, v_6)$ has vertex set $V(A_8) \cup \{v_9\}$ and arc set $A(A_8) \cup \{(v_9, v_8)\}$. Thus, the sink extension is A_9. By Lemma 3.16, $r(A_9) \leq r(A_8) = 0$. So $r(A_9) = 0$.

For $n \geq 9$ we prove by induction that there exist tournaments $T(A_n)$ and collections τ_n of arc disjoint cycles in $T(A_n)$ satisfying:

(a) $V(T(A_n)) = V(A_n)$,

(b) $T(A_n)$ has A_n as a minimum reversing set,

(c) $|\tau_n| = n - 1$,

(d) If n is odd, there is exactly one arc (v_n, v_{n-1}) in $T(A_n)$ with v_n as its tail, and for n even, there is exactly one arc (v_{n-1}, v_n) in $T(A_n)$ with v_n as its head.

(e) There is exactly one cycle in τ_n containing the vertex v_n. This is v_{n-1}, x, v_n, v_{n-1} if n is odd, and x, v_{n-1}, v_n, x if n is even for some $x \neq v_n, v_{n-1}$.

By (a), $|V(T(A_n))| = |V(A_n)|$. By (b), $r(A_n) \leq |V(T(A_n))| - |V(A_n)| = 0$. So, proving that (a) and (b) hold for all $n \geq 9$ will complete the proof.

Let $T(A_9)$ be the $D' = A_9$ extension of $T(A_8)$ and τ_9 the extension of τ_8, both with respect to (v_8, v_6). By the definition of the D' extension $T(A_9)$ and since $V(T(A_8)) = V(A_8)$, we have $V(T(A_9)) = V(A_9)$. So (a) holds. By Lemma 3.16, $T(A_9)$ has A_9 as a minimum reversing set. So (b) holds. Also, since $|\tau_8| = 7$ and by the definitions of the D' extension $T(A_9)$ and the (v_8, v_6) extension τ_9 of τ_8, it is easy to check that (c), (d), and (e) hold for $n = 9$.

Assume by way of induction that the result holds for n. Consider $n + 1$ even (and thus n odd), $n + 1 \geq 10$. By (e), and since $|V(T(A_n))| \geq 3$, there exists a vertex $y \neq x$ which is not on the unique cycle $v_{n-1}, x, v_n, v_{n-1} \in \tau_n$ containing v_n. By (d), $(y, v_n) \in T(A_n)$ since $y \neq v_{n-1}$ and (v_n, v_{n-1}) is the only arc in the tournament with v_n as its tail. By (b), $r(A_n) \leq |V(T(A_n))| - |V(A_n)| = 0$. Since the reversing number is non-negative, $r(A_n) = 0$ and thus $T(A_n)$ realizes A_n. By (c), $|\tau_n| = n - 1 = |A_n|$. Thus, $T(A_n)$ and τ_n satisfy the conditions necessary to take the source extension of

A_n with respect to (y, v_n). This source extension D'' of A_n with respect to is A_{n+1}. This follows since if we denote the new vertex in the extension by v_{n+1}, the new arc is (v_n, v_{n+1}) and since n is odd.

The $D'' = A_{n+1}$ extension $T(A_{n+1})$ of $T(A_n)$ has A_{n+1} as a minimum reversing set by Lemma 3.16. So (b) holds. By induction $V(T(A_n)) = V(A_n)$. Then by the definition of the $D'' = A_{n+1}$ extension, $V(T(A_{n+1})) = V(A_{n+1})$ and (a) holds.

Additionally, from the construction of the tournament $T(A_{n+1})$, $T(A_{n+1})$ contains exactly one arc (v_n, v_{n+1}) with v_{n+1} as its head. So (d) holds. Finally, the extension of τ_n with respect to (y, v_n) is $\tau_{n+1} = \tau_n \cup \{y, v_n, v_{n+1}, y\}$ and the new cycle is arc disjoint from the cycles of τ_n. So, $|\tau_{n+1}| = |\tau_n| + 1 = n$. The last equality follows by induction. So (c) holds. By construction, τ_{n+1} has exactly one cycle y, v_n, v_{n+1}, y containing the new vertex v_{n+1}. Thus (e) holds.

In a similar manner, for $n + 1$ odd, by (d) and (e) for n, there is a vertex y in $V(T(A_n))$ such that (v_n, y) is an arc in $T(A_n)$ and such that (v_n, y) is not contained in any cycle of τ_n. By (c) and the fact that (a) and (b) imply that $T(A_n)$ realizes A_n, the sink extension of A_n with respect to (v_n, y) is defined. Then this sink extension of A_n with respect to (v_n, y) is A_{n+1} and in a manner similar to the case when $n + 1$ is even, it can be checked that the $D' = A_{n+1}$ extension $T(A_{n+1})$ of $T(A_n)$ and the extension τ_{n+1} of τ_n, both with respect to (v_n, y), satisfy (a) through (e). $\square$

An interesting question is to determine the largest number of arcs a digraph on n vertices with reversing number 0 can have. A similar question can be asked for reversing number r. To study this we introduce the parameter $d(n, r) = \max |A(D)|$ (the maximum number of arcs), where the maximum is taken over all connected acyclic digraphs with $|V(D)| = n$ and $r(D) = r$. If no such D exists for a given n and r, then we say that $d(n, r)$ does not exist.

Since we are considering connected digraphs on n vertices, $n - 1 \leq d(n, r) \leq \binom{n}{2}$. By Equation 3.1, $d(n, r) \geq r$. Since a minimum reversing set of any tournament contains at most half the arcs in the tournament, $d(n,r) \leq \frac{1}{2}\binom{r+n}{2}$. Thus we get

$$\max\{r, n-1\} \leq d(n,r) \leq \min\left\{\frac{1}{2}\binom{r+n}{2}, \binom{n}{2}\right\}. \tag{3.2}$$

Corollary 3.9 and Theorem 3.20 (below) show that $d(n,r)$ is undefined for $r > 2n-2$. By Theorem 3.17, $d(n,0)$ is defined if and only if $n \geq 6$.

Let $f(n)$ be the largest k such that every tournament on n vertices contains an acyclic digraph with k arcs. It appears that upper bounds on $f(n)$ might provide graphs with reversing number 0 and a large number of arcs, since there exists some tournament with n vertices containing no acyclic digraph with $f(n)+1$ arcs, i.e., minimum reversing sets of this tournament have at least $\binom{n}{2} - f(n) - 1$ arcs. The upper bound $f(n) \leq \frac{1}{2}\binom{n}{2} + cn^{3/2}$, c constant, determined by Erdős and Moon [1965] and Spencer [1971], would then give digraphs with reversing number 0 and $\frac{1}{2}\binom{n}{2} - cn^{3/2}$ arcs. However, this reasoning does not necessarily work since we assume that our digraphs have no isolated vertices, while the digraphs obtained as minimum reversing sets of tournaments providing the upper bounds on $f(n)$ may have isolated vertices.

Making use of Lemma 3.16 we can improve values of $d(n,0)$.

Theorem 3.18 *For $n \geq 7$, $d(n,0) \geq \left\lceil \frac{(n-2)^2}{\alpha} \right\rceil$, where $\alpha = 5+\sqrt{21}$.*

Proof: For $n = 7$ the result follows from the example constructed in proving the case $n = 7$ of Theorem 3.17. For $n = 8, 9, 10, 11$, the result holds since, for the alternating paths A_n with $|A_n| = n-1$, we have $r(A_n) = 0$. This was proved in the case $n \geq 8$ of the proof of Theorem 3.17. (Actually, the same proof takes care of the case $n = 12$, but we will need more out of $n = 12$.)

For $n \geq 12$, we prove by induction that there exists a digraph H_n on n vertices, a tournament $T(H_n)$, and a collection τ_n of arc disjoint 3-cycles in $T(H_n)$ satisfying;

(a) $|H_n| = \left\lceil \frac{(n-2)^2}{\alpha} \right\rceil$

(b) $V(T(H_n)) = V(H_n)$,

(c) $T(H_n)$ has H_n as a minimum reversing set,

(d) $|\tau_n| = |H_n|$.

By (b) and (c), $r(H_n) \leq |V(T(H_n))| - |V(H_n)| = 0$. Then (a) gives the desired result.

For $n = 12$, let $H_{12} = A_{12}$, where A_{12} is the alternating path constructed in the inductive proof of the $n \geq 8$ case of Theorem 3.17. Also let $T(H_{12})$ be $T(A_{12})$ and τ_{12} be the τ_{12} from that proof. Note that τ_{12} contains only 3-cycles, for τ_8 of Figure 3.8 has only 3-cycles and τ_{12} is obtained from τ_8 by successively adding 3-cycles. Then (b), (c), and (d) hold from the induction in Theorem 3.17. Also, for $H_{12} = A_{12}$, (a) holds since $\left\lceil \frac{(n-2)^2}{\alpha} \right\rceil = \lceil \frac{100}{\alpha} \rceil = 11 = |A_{12}|$.

Assume by induction that (a),(b),(c) and (d) hold for n (where $n \geq 12$). Suppose we can find a set S of vertex disjoint arcs in $T(H_n)$, none of which is an arc of one of the cycles of τ_n, such that

$$|S| = \left\lceil \frac{(n-1)^2}{\alpha} \right\rceil - \left\lceil \frac{(n-2)^2}{\alpha} \right\rceil . \tag{3.3}$$

By (b) and (c) for n, $r(H_n) \leq |V(T(H_n))| - |V(H_n)| = 0$. Thus, $r(H_n) = 0$ and $T(H_n)$ realizes H_n. This along with (d) for n insure that the sink extension of H_n with respect to S is defined.

Define H_{n+1} to be the the sink extension of H_n with respect to S, $T(H_{n+1})$ to be the $D' = H_{n+1}$ extension of $T(H_n)$ with respect to S, and τ_{n+1} to be the extension of τ_n with respect to S. Note that be Lemma 3.16 and the fact that τ_n is a collection of 3-cycles, τ_{n+1} is a collection of arc disjoint 3-cycles in $T(H_{n+1})$, and H_{n+1} is a minimum reversing set of $T(H_{n+1})$. Thus, (c) holds for $n+1$. Also, (a) holds for $n+1$ because

$$\begin{aligned} |H_{n+1}| &= |H_n| + |S| \\ &= \left\lceil \frac{(n-2)^2}{\alpha} \right\rceil + \left(\left\lceil \frac{(n-1)^2}{\alpha} \right\rceil - \left\lceil \frac{(n-2)^2}{\alpha} \right\rceil \right). \end{aligned}$$

By induction and the construction of τ_{n+1}, we have $|\tau_{n+1}| = |\tau_n| + |S| = |H_n| + |H_{n+1}| - |H_n| = |H_{n+1}|$. So (d) holds for $n+1$. By induction and the construction of H_{n+1}, $V(T(H_{n+1})) = V(H_{n+1})$ and (b) holds for $n+1$.

Thus, the proof is complete if we can find a collection S of vertex disjoint arcs in $T(H_n)$, none of which is an arc in a cycle of τ_n such that (3.3) holds. Let u_n be the right side of (3.3). It suffices to show that we can find u_n vertex disjoint arcs in $T(H_n)$ among those not in τ_n. Note that

$$u_n \leq \frac{(n-1)^2-(n-2)^2}{\alpha}+1=\frac{2n-3+\alpha}{\alpha}.$$

The cycles in τ_n are all 3-cycles and hence contain $3|H_n| \leq 3\left(\frac{(n-2)^2}{\alpha}+1\right)$ arcs from $T(H_n)$. Thus there are at least

$$\begin{aligned} \binom{n}{2}-3\left(\frac{(n-2)^2}{\alpha}+1\right) &= \frac{n^2-n}{2}-\frac{3n^2-12n+12+3\alpha}{\alpha} \\ &= \frac{n^2(\alpha^2-6\alpha)+n(24\alpha-\alpha^2)+(-24\alpha-6\alpha^2)}{2\alpha^2} \end{aligned} \tag{3.4}$$

arcs in $T(H_n)$ which are not in τ_n. In order to show that we can find u_n vertex disjoint arcs in $T(H_n)$ among those not in τ_n, it is enough to show that every undirected graph G on n vertices with the number of edges given in (3.4) contains a matching of size $\frac{2n-3+\alpha}{\alpha} \geq u_n$. For then we pick a matching of size u_n to give as the u_n vertex disjoint arcs in $T(H_n)$ none of which is in τ_n.

In general if the largest matching in a graph G on n vertices contains m edges, then G contains at most $\max\left\{\binom{2m+1}{2}, \binom{m}{2}+m(n-m)\right\}$ edges (see e.g. Bollobás [1979, pg 65]). For $m < \frac{2n-3+\alpha}{\alpha}$, the second of these two values is larger. Also, for a given n and $m < \frac{2n-3+\alpha}{\alpha} < n$, it is easy to check by taking the first derivative that $g(m) = \binom{m}{2}+m(n-m) = mn-\frac{1}{2}(m+m^2)$ is increasing in m. Thus the maximum value of $g(m)$ for $m < \frac{2n-3+\alpha}{\alpha}$ is strictly less than the value when $m = \frac{2n-3+\alpha}{\alpha}$. Hence, by evaluating $g(m)$ at $m = \frac{2n-3+\alpha}{\alpha}$ we get the following strict upper bound on the number of edges that G can contain if it has no matching of size at least $\frac{2n-3+\alpha}{\alpha}$.

$$\begin{aligned} g\left(\frac{2n-3+\alpha}{\alpha}\right) &= \frac{2n^2-3n+\alpha n}{\alpha}-\frac{2n-3+\alpha}{2\alpha}-\frac{4n^2+9+\alpha^2-12n+4n\alpha-6\alpha}{2\alpha^2} \\ &= \frac{n^2(4\alpha-4)+n(2\alpha^2-12\alpha+12)+(9\alpha-2\alpha^2-9)}{2\alpha^2} \end{aligned} \tag{3.5}$$

If G has no matching of size at least $\frac{2n-3+\alpha}{\alpha}$, then (3.4) < (3.5). This implies (since, by the choice of α, $\alpha^2-10\alpha+4=0$),

$$0 > n^2(\alpha^2-10\alpha+4)+n(-3\alpha^2+36\alpha-12)+(-4\alpha^2-33\alpha+9)$$

$$\Leftrightarrow 0 > n(-3\alpha^2 + 36\alpha - 12) + (-4\alpha^2 - 33\alpha + 9)$$

$$\Rightarrow 11.8 > n$$

Thus for $n \geq 12$, H_n must contain a matching of size at least $\frac{2n-3+\alpha}{\alpha}$. □

We are also able to get bounds on $d(n,1)$.

Theorem 3.19 (Barthelemy et al. 1990) *For n even,* $\frac{n^2+n}{4} \geq d(n,1) \geq \frac{n^2+2n}{8}$

Proof: Note that $d(2,1) = 1$ since the only connected acyclic digraph on 2 vertices is a single arc, which has reversing number 1 by Theorem 3.13. Thus the result holds for $n = 2$. For $n > 2$ and $r = 1$, $\frac{1}{2}\binom{n+1}{2} \leq \binom{n}{2}$, so the upper bound follows from equation 3.2. Next, let $n = 2m$. We will construct a digraph D on n vertices with $\frac{n^2+2n}{8}$ arcs having reversing number 1. Let D have vertex set $X \cup Y$ where $X = \{1, 2, \ldots, m\}$ and $Y = \{m+2, m+3, \ldots, 2m+1\}$. Let (i,j) be an arc of D if $i \in Y$, $j \in X$ and $j \leq i-m-1$. Thus there are arcs from $2m+1$ to $1, 2, \ldots, m$; from $2m$ to $1, 2, \ldots, m-1$; and so on. Note that D is acyclic and that D has $\frac{m(m+1)}{2} = \frac{n^2+2n}{8}$ arcs. Let T' be the acyclic tournament on $\{1, 2, \ldots, 2m+1\}$ with acyclic order $\pi(1) < \pi(2) < \cdots < \pi(2m+1)$. Note that $V(T') = V(D) \cup \{m+1\}$ and that $D^R \subset T'$. Thus D is a reversing set of $T = (T' \setminus D^R) \cup D$. If we can show that D is a minimum reversing set of T then $r(D) \leq 1$.

Note that the outdegrees $d_T^+(i) = m$ for all vertices i of T. To see this note that from the construction of T and the definition of D, $(i,j) \in T \Leftrightarrow j = i+1, i+2, \ldots, i+m (\text{mod } 2m+1)$. An acyclic tournament on $2m+1$ vertices has exactly one vertex with outdegree j for $j = 0, 1, 2, \ldots, 2m$. If T and T'' are two tournaments on $\{1, 2, \ldots, 2m+1\}$, then the number of arcs which must be reversed to obtain T'' from T is at least $\frac{1}{2}\sum_{i=1}^{2m+1} |d_T^+(i) - d_{T''}^+(i)|$. This follows by counting the number of arcs which must be reversed to obtain the new outdegree and noting that by this process each arc reversal is counted twice. Using this bound, the fact that $d_T^+(i) = m$ for all

i, and the fact that any acyclic tournament T'' obtained by reversing some arcs from T has outdegrees $0, 1, 2, \ldots, 2m$, we see that any reversing set of T contains at least $\frac{1}{2}\sum_{j=0}^{2m} |j-m| = \sum_{k=1}^{m} k = \frac{m(m+1)}{2}$ arcs. Thus any reversing set of T contains at least as many arcs as D, so D is a minimum reversing set of T. This shows that $r(D) \leq 1$.

Finally we show that $r(D) \neq 0$. Assume that some tournament T on the vertex set of D has D as a minimum reversing set. Since there are no arcs in D between vertices of X, $T|_X$ must be acyclic. Similarly $T|_Y$ must be acyclic. Let Z be the set of arcs which have one end in X and the other in Y and which are not in D. Note that

$$|Z| = m^2 - \frac{m(m+1)}{2} = \frac{m(m-1)}{2} < \frac{m(m+1)}{2} = |D|$$

Also $(T \setminus Z) \cup Z^R$ is acyclic since it is acyclic on X and Y and all arcs between X and Y are directed from Y to X. This contradicts the assumption that D is a minimum reversing set of T, so $r(D) \neq 0$. □

3.5 Tournaments

The reversing number of acyclic tournaments is important since it gives an upper bound on the reversing number of general digraphs as noted in Corollary 3.9.

Theorem 3.20 *For the acyclic tournament T_n on n vertices, $2n - 4\log_2 n \leq r(T_n) \leq 2n - 2$.*

Proof: In this proof, all logarithms will be base 2. Let T_n be an acyclic tournament with vertex set $V(T_n) = \{v_1, v_2, \ldots, v_n\}$ such that the acyclic ordering of T_n is $\pi'(v_n) < \pi'(v_{n-1}) < \cdots < \pi'(v_1)$.

In order to obtain a lower bound on the reversing number of the acyclic tournament T_n on n vertices we consider a smallest tournament $T(T_n)$ having T_n as a minimum reversing set. Since T_n is a minimum reversing set of $T(T_n)$, the acyclic order π of $T(T_n)$ after reversal of the arcs in T_n satisfies $\pi(v_1) < \pi(v_2) < \cdots < \pi(v_n)$. By Lemma 3.3

we may assume that for all vertices u in $T(T_n)$, $\pi(v_1) < \pi(u) < \pi(v_n)$ since otherwise there would be a smaller tournament having T_n as a minimum reversing set. Denote the extra vertices (those not in T_n) of $T(T_n)$ by u_{ij} where $\pi(v_i) < \pi(u_{ij}) < \pi(v_{i+1})$ for $1 \leq i < n$ and for a given i, $\pi(u_{ij}) < \pi(u_{ij'})$ for $1 \leq j < j' \leq x_i$. Thus we have denoted the number of extra vertices between v_i and v_{i+1} in the acyclic order π by x_i. Using this notation, the reversing number of T_n is $\sum_{h=1}^{n-1} x_h$.

Recall that the backwards arcs relative to an ordering σ in $T(T_n)$ are arcs $(y, z) \in T(T_n)$ with $\sigma(z) < \sigma(y)$. For any ordering σ of the vertices of $T(T_n)$ the number of backwards arcs relative to σ is at least as large as the number of arcs in T_n, i.e., at least $\frac{n(n-1)}{2}$. This holds since T_n is a minimum reversing set of $T(T_n)$. By Lemma 3.3 a similar condition holds for certain subtournaments of $T(T_n)$. For any ordering σ of the vertices of $T(T_n)$ restricted to a segment (in the order π) $V'_{jk} = \{v_j, v_{j+1}, \ldots, v_k\} \cup \{u_{rs} : j \leq r < k, 1 \leq s \leq x_r\}$, the number of backwards arcs in the segment relative to σ is at least as large as $\frac{(k-j+1)(k-j)}{2}$, the number of arcs in T_n restricted to the segment.

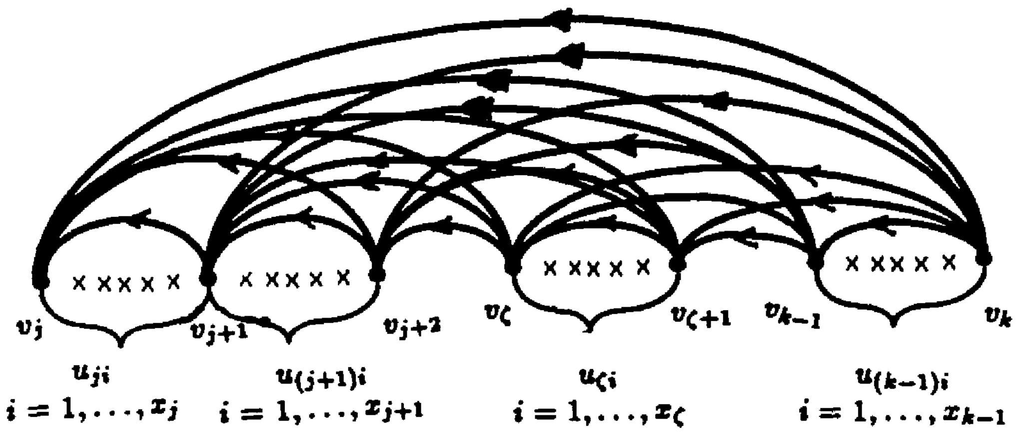
We make use of one 'bad' ordering to get a set of inequalities on the x_i which can then be combined to get a lower bound on the reversing number. This ordering applied to the subtournament of $T(T_n)$ induced by V'_{jk} places all the extra vertices u_{rs} to the 'right' or 'left' (in their natural order consistent with π), and the vertices v_c which appear in T_n in the 'middle' in the acyclic order π' of T_n. That is, for a given $j < k$, for $0 \leq a, a' < k-1$, $1 \leq b \leq x_{j+a}$, $1 \leq b' \leq x_{j+a'}$ and for $c = j, j+1, \ldots, k$, the ordering σ on V'_{jk} is given by

$$\sigma(u_{ab}) < \sigma(u_{a'b'}) \Leftrightarrow a < a' \text{ or } a = a' \text{ and } b < b'$$

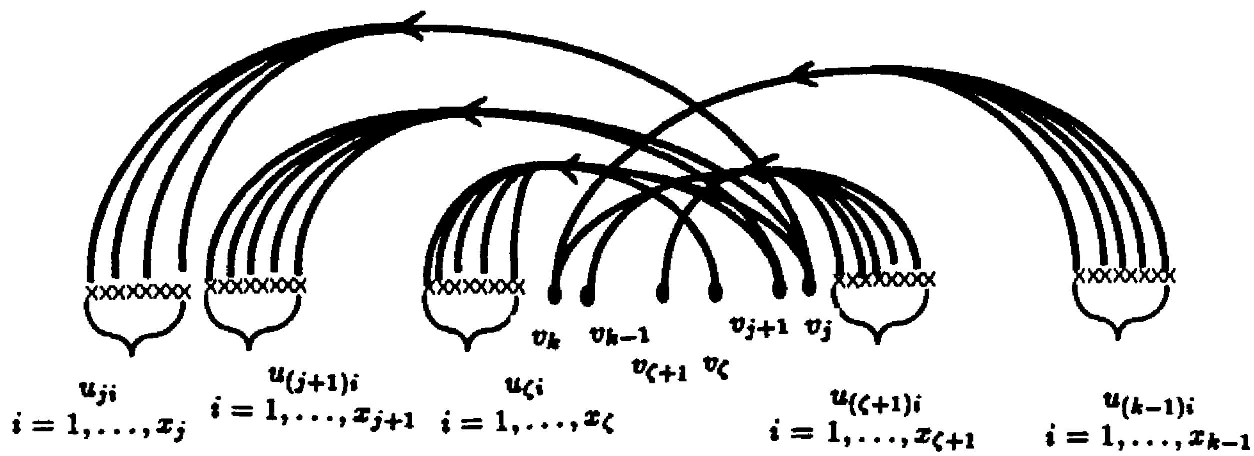
$$\sigma(u_{(j+a)b}) < \sigma(v_c) \Leftrightarrow a \leq \left\lfloor \frac{k-j-1}{2} \right\rfloor$$

$$\sigma(v_c) < \sigma(v_{c'}) \Leftrightarrow c > c'$$

Figure 3.9(a) shows the backwards arcs in the subtournament of T_n on V'_{jk} relative to the ordering π and Figure 3.9(b) shows the backwards arcs in the subtournament of T_n on V'_{jk} relative to the ordering σ. From Figure 3.9(b) (or from the definitions of T_n,



(a) $T(T_n)\Big|_{v_{j'k}}$ under the ordering π.



(b) $T(T_n)\Big|_{v_{j'k}}$ under the ordering σ.

All arcs which are not shown are directed from left to right in the figure.

Figure 3.9: Backwards arcs in the subtournament of T_n on V'_{jk} relative to π and σ

π and σ), it can be checked that the backwards arcs in T_n restricted to V'_{jk} are: for each $0 \leq a \leq \lfloor\frac{k-j-1}{2}\rfloor$, $(v_c, u_{(j+a)b})$ for $j \leq c \leq j+a$ and $1 \leq b \leq x_{j+a}$ and, for each $\lfloor\frac{k-j-1}{2}\rfloor < a < k-j$, $(u_{(j+a)b}, v_c)$ for $j+a+1 \leq c \leq k$ and $1 \leq b \leq x_{j+a}$.

Making use of the fact that for each i, there are x_i vertices u_{ij}, we have the following count on the number z of backwards arcs relative to σ. For given k, j, we have

$$\begin{aligned} z &= \sum_{a=0}^{\lfloor\frac{k-j-1}{2}\rfloor}\sum_{c=j}^{j+a}\sum_{b=1}^{x_{j+a}} 1 + \sum_{a=\lfloor\frac{k-j-1}{2}\rfloor+1}^{k-j-1}\sum_{c=j+a+1}^{k}\sum_{b=1}^{x_{j+a}} 1 \\ &= \sum_{a=0}^{\lfloor\frac{k-j-1}{2}\rfloor}(a+1)x_{j+a} + \sum_{a=\lfloor\frac{k-j-1}{2}\rfloor+1}^{k-j-1}(k-j-a)x_{j+a} \\ &= \sum_{i=1}^{\lfloor\frac{k-j-1}{2}\rfloor+1} i x_{j+i-1} + \sum_{i=1}^{\lceil\frac{k-j-1}{2}\rceil} i x_{k-i} \end{aligned}$$

In the last line, we have made the change of counters $i = a+1$ in the first sum and $i = k-j-a$ in the second sum. When $k-j$ is even, both sums have the same number of terms. Combining these we get

$$z = \sum_{i=1}^{\frac{k-j}{2}} i(x_{j+i-1} + x_{k-i}).$$

When $k-j$ is odd, the first sum has one more term then the second. Writing the last term of the first sum separately and combining the remaining terms from both sums, we get

$$z = \left[\sum_{i=1}^{\frac{k-j-1}{2}} i(x_{j+i-1} + x_{k-i})\right] + \frac{k-j+1}{2}x_{j+\frac{k-j-1}{2}}.$$

Since the number of backwards arcs relative to σ is at least as large as $\frac{(k-j+1)(k-j)}{2}$, we get the following inequalities.

$$\sum_{i=1}^{\frac{k-j}{2}} i(x_{j+i-1} + x_{k-i}) \geq \frac{(k-j+1)(k-j)}{2} \text{ for } k-j \text{ even,} \tag{3.6}$$

$$\left[\sum_{i=1}^{\frac{k-j-1}{2}} i(x_{j+i-1} + x_{k-i})\right] + \frac{k-j+1}{2}x_{j+\frac{k-j-1}{2}} \geq \frac{(k-j+1)(k-j)}{2} \text{ for } k-j \text{ odd,} \tag{3.7}$$

where the first term in the sum is interpreted as 0 if $k - j = 1$.

At this point, we have inequalities (3.6) and (3.7) which provide lower bounds on expressions inolving the number of extra vertices x_i. By taking appropriate positive multiples of these inequalities and then summing we can obtain an inequality which provides a lower bound on $\sum_{h=1}^{n-1} x_h$, which is the reversing number. In order to describe the multipliers for the inequalities, we will recursively construct a collection of inequalities (3.6) and (3.7) for which the number of copies of each particular inequality will provide the multiplier.

For a given $p = p_0$, we consider the collection C_p of inequalities defined as follows. Include an inequality for each $0 \leq h \leq \lfloor \log p \rfloor$. To obtain the h^{th} inequality, define p_h recursively by $p_h = \lfloor \frac{p(h-1)}{2} \rfloor$. Set $j = 1$ and $k = p_h$. Then use inequality (3.6) if $k - j$ is even, and $k \neq j$; the empty inequality $0x_1 \geq 0$ if $k = j$ and the inequality (3.7) if $k - j$ is odd, in each case multiplied by 2^h.

For example, with $p = 4$ the inequalities in C_4 are

$$x_1 + 2x_2 + x_3 \geq \frac{(4)(3)}{2} = 6 \quad (h = 0)$$
$$2\left(x_1 \geq \frac{(2)(1)}{2} = 1\right) \qquad (h = 1)$$

(There is no inequality for $h = 2$, since here $p_2 = 1$, and $j = k = 1$.)

Summing the inequalities in C_p we obtain an inequality of the form

$$\sum_{m=1}^{p-1} c_m x_m \geq f(p).$$

We demonstrate by induction the following bounds on the values of the coefficients c_m and the right hand side $f(p)$.

(a) $c_m \leq p - m$

(b) $f(p) \geq p^2 - 2p \log p$

For $p = 2, 3$ one can easily check that (a) and (b) hold. For $p = 4$, summing the inequalities noted above gives

$$3x_1 + 2x_2 + x_3 \geq 8$$

which satisfies (a) and (b).

Assume that (a) and (b) hold for numbers smaller than p. Given $p \geq 5$ the collection C_p contains one copy of (3.6) or (3.7) for $j = 1$ and $k = p$ and for each inequality appearing in $C_{\lfloor \frac{p}{2} \rfloor}$ the inequality multiplied by two.

Thus, for $m > \lfloor \frac{p}{2} \rfloor$, the coefficient c_m is $p - m$ by construction. For $m \leq \lfloor \frac{p}{2} \rfloor$,

$$c_m \leq \left[2\left(\left\lfloor \frac{p}{2} \right\rfloor - m\right)\right] + m = 2\left\lfloor \frac{p}{2} \right\rfloor - m \leq p - m.$$

Here the term in brackets follows by induction on the inequalities in $C_{\lfloor \frac{p}{2} \rfloor}$ which are multiplied by two, and the final m is the coefficient in the new inequality. (Note that in the new inequality, we have $k \neq j$ since $k = p_h$.) This proves that (a) holds for all p.

Now, we show (b). We also have that $f(p) \geq 2f\left(\lfloor \frac{p}{2} \rfloor\right) + \frac{p(p-1)}{2}$. The first term follows from the inequalities in $C_{\lfloor \frac{p}{2} \rfloor}$ which are multiplied by two, and the final term for the new inequality with $j = 1$ and $k = p$. We now use the inductively assumed bound for $f\left(\lfloor \frac{p}{2} \rfloor\right)$. For p even, $p \geq 6$, we get

$$\begin{aligned} f(p) &\geq 2\left(\left\lfloor \frac{p}{2} \right\rfloor^2 - 2\left\lfloor \frac{p}{2} \right\rfloor \log \left\lfloor \frac{p}{2} \right\rfloor\right) + \frac{(p)(p-1)}{2} \\ &= 2\left(\frac{p^2}{4} - p(\log p - 1)\right) + \frac{p^2}{2} - \frac{p}{2} \\ &= p^2 - 2p\log p + \frac{3}{2}p \\ &\geq p^2 - 2p\log p. \end{aligned} \tag{3.8}$$

For p odd, $p \geq 5$, we get,

$$\begin{aligned} f(p) &\geq 2\left(\left\lfloor \frac{p}{2} \right\rfloor^2 - 2\left\lfloor \frac{p}{2} \right\rfloor \log \left\lfloor \frac{p}{2} \right\rfloor\right) + \frac{(p)(p-1)}{2} \\ &= 2\left(\frac{(p-1)^2}{4} - (p-1)(\log(p-1) - 1)\right) + \frac{p^2}{2} - \frac{p}{2} \\ &= p^2 - 2p\log(p-1) + \frac{p}{2} - \frac{3}{2} + 2\log(p-1) \\ &\geq p^2 - 2p\log p. \end{aligned} \tag{3.9}$$

Thus (b) holds.

Similarly to C_p, we can define for a given n, collections C'_p. These include an inequality for each h, $0 \le h \le \lfloor \log p \rfloor$. To obtain the h^{th} inequality, let $p_h = p$ and recursively define $p_h = \lfloor \frac{p(h-1)}{2} \rfloor$ as before. Set $j = n - p_h + 1$ and $k = n$ and use for the h^{th} inequality (3.6) if $k - j$ is even and $k \neq j$; the empty inequality $0x_i \geq 0$ if $k = j$; and the (3.7) if $k - j$ is odd.

The sets of inequalities C_p and C'_p are symmetric in the sense we now make precise. Consider C_p when $j = 1$ and $k = p_h$ and C'_p when $j = n - p_h + 1$ and $k = n$. Then $k - j$ is $p_h - 1$ in both cases, so we use the same inequality (3.6) or (3.7) in each case. Whenever in (3.6) or (3.7) in C_p there is a term $ix_i = ix_{j+i-1}$, then in (3.6) or (3.7) in C'_p there is a corresponding term $ix_{k-i} = ix_{n-i}$. Whenever in (3.6) or (3.7) in C_p there is a term $ix_{k-i} = ix_{p_h-i}$, then in (3.6) or (3.7) in C'_p there is a corresponding term $ix_{j+i-1} = ix_{n-(p_h-i)}$. Whenever in (3.7) in C_p there is a term $\frac{k-j+1}{2} x_{(j+\frac{k-j+1}{2})} = \frac{p_h-1+1}{2} x_{1+\frac{p_h-1+1}{2}} = \frac{p_h}{2} x_{\frac{p_h}{2}}$, then in (3.7) in C'_p there is a corresponding term $\frac{k-j+1}{2} x_{(j+\frac{k-j+1}{2})} = \frac{n-(n-p_h+1)+1}{2} x_{n-p_h+1+\frac{(n-(n-p_h+1)-1}{2})} = \frac{p_h}{2} x_{\frac{n-p_h}{2}}$. In all cases, whenever there is a term x_m in the set of inequalities C_p, there is a corresponding term x_{n-m} with the same coefficient in the set of inequalities C'_p.

As with C_p, summing the inequalities in C'_p we obtain an inequality of the form

$$\sum_{m=n-p+1}^{n-1} c'_m x_m \geq f'(p)$$

where

(a′) $c'_{n-m} \leq p - m$

(b′) $f'(p) \geq p^2 - 2p \log p$.

By the symmetry to C_p, with x_{n-m} replacing x_m, (a′) and (b′) hold.

Finally to get a bound on $\sum_{i=1}^{n-1} x_i$ we use the following collection of inequalities.

(i) One copy of inequality (3.6) or (3.7) for $j = 1$ and $k = n$

(ii) One copy of the collection $C_{\lfloor \frac{n}{2} \rfloor}$

(iii) One copy of the collection $C'_{\lfloor \frac{n}{2} \rfloor}$.

Summing inequalities from (i),(ii) and (iii) we get an inequality

$$\sum_{m=1}^{n-1} d_m x_m \geq \frac{(n)(n-1)}{2} + f\left(\left\lfloor\frac{n}{2}\right\rfloor\right) + f'\left(\left\lfloor\frac{n}{2}\right\rfloor\right). \tag{3.10}$$

The right hand side of this inequality is the sum of the bounds for (i), (ii), and (iii).

For the coefficients d_m on the left side of the inequality, note that in $C_{\lfloor\frac{n}{2}\rfloor}$ the only non-zero coefficients are for $\{x_1, \ldots, x_{\lfloor\frac{n}{2}\rfloor-1}\}$ and in $C'_{\lfloor\frac{n}{2}\rfloor}$ the only non-zero coefficients are for $\{x_{n-\lfloor\frac{n}{2}\rfloor+1}, \ldots, x_{n-1}\}$. Note that $n-\lfloor\frac{n}{2}\rfloor+1 = \lceil\frac{n}{2}\rceil+1$ so the non-zero coefficients from (ii) and (iii) do not overlap. Consider the coefficient d_m for $m < \lfloor\frac{n}{2}\rfloor$. In this case, $d_m \leq (\lfloor\frac{n}{2}\rfloor - m) + m \leq \frac{n}{2}$. Here the first term is the coefficient from (ii) with the bound (a) and the final m is the coefficient of x_m in (i). For d_m if $m > \lceil\frac{n}{2}\rceil$, we get the same bound from (iii) and (a′) and (i). When n is even $x_{\frac{n}{2}}$ appears only in (i) and has coefficient $\frac{n}{2}$. For n odd, $x_{\lfloor\frac{n}{2}\rfloor}$ and $x_{\lceil\frac{n}{2}\rceil}$ appear only in (i) with coefficient $\lfloor\frac{n}{2}\rfloor$. So the coefficients d_m are all less than or equal to $\frac{n}{2}$.

Also note that substituting the bounds (b) and (b′) for $f(\lfloor\frac{n}{2}\rfloor)$ and $f'(\lfloor\frac{n}{2}\rfloor)$ into the right hand side of (3.10) we get the same right hand side as in (3.8) and (3.9)with n instead of p. Thus, as in (3.8) and (3.9), we get the right hand side of (3.10) greater than or equal to $n^2 - 2n\log n$. Using this bound and the bound $d_m \leq \frac{n}{2}$, we get from (3.10) that

$$\frac{n}{2}\sum_{m=1}^{n-1} x_m \geq \sum_{m=1}^{n-1} d_m x_m \geq \frac{(n)(n-1)}{2} + f\left(\left\lfloor\frac{n}{2}\right\rfloor\right) + f'\left(\left\lfloor\frac{n}{2}\right\rfloor\right) \geq n^2 - 2n\log n.$$

Hence,

$$\sum_{m=1}^{n-1} x_m \geq \frac{2}{n}\left(n^2 - 2n\log n\right) = 2n - 4\log n,$$

giving the desired lower bound on the reversing number of T_n.

For the upper bound we construct a tournament T on $3n-2$ vertices having T_n as a minimum reversing set. Let T_n have acyclic ordering $\pi(v_n) < \pi(v_{n-1}) < \cdots < \pi(v_1)$. Let T' be an acyclic tournament with vertex set $V(T) = V(T_n) \cup \{u_{ij} : 1 \leq i \leq n-1, j = 0, 1\}$ and acyclic ordering π' satisfying $\pi'(v_i) < \pi'(u_{i0}) < \pi'(u_{i1}) < \pi'(v_{i+1})$ for $1 \leq i \leq n-1$. Since $T_n^R \subset T'$ we can define $T = (T' \setminus T_n^R) \cup T_n$. Thus T_n is a reversing set of T. To show that T_n is a minimum reversing set of T we consider the following set of $\frac{n(n-1)}{2}$ triples, which we show to be 3-cycles.

$$\tau = \{v_j, v_i, u_{k_{ij}\delta_{ij}}, v_j : i < j\}$$

$$\text{for } k_{ij} = i + \left\lfloor \frac{j-i-1}{2} \right\rfloor \text{ and } \delta_{ij} = (j-i) \bmod 2.$$

Since $i \leq k_{ij} < j$, we have $\pi'(v_i) < \pi'(u_{k_{ij}\delta_{ij}}) < \pi'(v_j)$. So $(v_i, u_{k_{ij}\delta_{ij}}) \in T'$ and thus it is in T. Similarly $(u_{k_{ij}\delta_{ij}}, v_j) \in T$. Also since $i < j$, $(v_j, v_i) \in T$. Thus the entries of τ are indeed cycles on three distinct vertices, i.e. 3-cycles. Also these 3-cycles are arc disjoint since $k_{ij}\delta_{ij} = k_{i'j'}\delta_{i'j'}$ if and only if $i = i'$ and $j = j'$. Thus by Lemma 3.5, T_n is a minimum reversing set of T. □

We note at this point that we could set up an integer linear program to minimize the sum of the x_i subject to inequality (3.6) or (3.7) for all j and k with $1 \leq j < k \leq n$. The solution of this would provide a bound on the reversing number. It would be interesting to see if the bound derived from this integer program is tight. The multipliers used in the collection of inequalities used in the proof of the lower bound can be viewed as variables in a dual feasible solution to the linear program obtained by relaxing the integer constraints. With a little more work we can improve the upper bound to $2n - 4$, which is useful in getting exact values of $r(T_n)$ for small n. However, this upper bound is not tight in all cases, as can be seen in Table 3.1, which lists exact values of $r(T_n)$ for small n. The values in this table have been calculated by special cases of the techniques in the proof.

Table 3.1: Exact Values of $r(T_n)$ for Small n

n	2	3	4	5	6	7	8	9	10	11	12
$r(T_n)$	1	3	4	6	8	10	11	14	15	17	19

3.6 Reversing Numbers of Acyclic Digraphs in Some Special Classes

In this section we compute the reversing number for acyclic digraphs in various special classes.

3.6.1 Stars

Let a *directed star* S_n be a digraph on n vertices with a distinguished vertex v such that all arcs in S_n contain v as either head or tail. Note that S_n contains $n-1$ arcs and by our convention of denoting by $|S_n|$ the size of the arc set of S_n, we have $|S_n| = n-1$.

Theorem 3.21 *If S_n is a directed star on n vertices then $r(S_n) = n-1$.*

Proof: By Lemmas 3.10 and 3.12, we may assume that S_n is the directed star in which $v = v_0$ is the head of all arcs, i.e., $S_n = \{(v_i, v_0) : i = 1, 2, \ldots, n-1\}$. Let T realize S_n and let π be the acyclic ordering of $(T \setminus S_n) \cup S_n^R$. Since $(v_0, v_i) \in S_n^R$, $\pi(v_0) < \pi(v_i)$, $i = 1, 2, \ldots, n-1$. Without loss of generality, $\pi(v_0) < \pi(v_1) < \pi(v_2) < \cdots < \pi(v_{n-1})$. Also, by Lemma 3.2, we may assume that there are no 'extra' vertices w, i.e., vertices in $V(T) \setminus V(S_n)$, such that $\pi(w) > \pi(v_{n-1})$ or $\pi(w) < \pi(v_0)$. For $i = 1, 2, \ldots, n-1$, let there be k_i extra vertices $\{x_{i1}, x_{i2}, \ldots, x_{ik_i}\}$ between v_{i-1} and v_i in π, i.e., $\pi(v_{i-1}) < \pi(x_{ij}) < \pi(v_i)$, for $j = 1, 2, \ldots, k_i$.

Note that $\sum_{i=1}^{n-1} k_i = r(S_n)$. Let $X = \{(v_0, x_{ij}) : i = 1, 2, \ldots, n-1,\ j = 1, 2, \ldots, k_i\} \subseteq T$. Then $X \subseteq T$ and $(T \setminus X) \cup X^R$ is acyclic, with the acyclic order π' obtained from π by making v_0 a sink instead of a source and maintaining the acyclic order among the other vertices. That is, $\pi'(u) = \pi(u) - 1$ for $u \neq v_0$ and $\pi'(v_0) > \pi'(v_{n-1}) > \pi'(u)$ for all $u \in V(T)$. Since S_n is a minimum reversing set of T,

$$|X| \geq |S_n| = n-1.$$

Note that $|X| = \sum_{i=1}^{n-1} k_i = r(S_n)$. Therefore, $r(S_n) \geq n-1$. Letting $k_i = 1$ for all i gives a tournament of $2n-1$ vertices containing the $n-1$ arc disjoint 3-cycles

x_{i1}, v_i, v_0, x_{i1}, $i = 1, \ldots, n-1$, with S_n as a reversing set and thus a minimum reversing set by Lemma 3.5. □

3.6.2 Disjoint Arcs

As mentioned above, there exist digraphs whose reversing number is 0. An example will be the disjoint union of n arcs, the graph we denote by E_n.

Theorem 3.22 (Barthelemy et al. 1990) *$r(E_1) = r(E_2) = 1$, and $r(E_n) = 0, n \geq 3$.*

Proof: Note that $E_1 = P_2$. Therefore, by Theorem 3.13, $r(E_1) = 1$.

By Theorem 3.17, $r(E_2) > 0$ since E_2 has only 4 vertices. Let T' be given by the digraph in Figure 3.10.

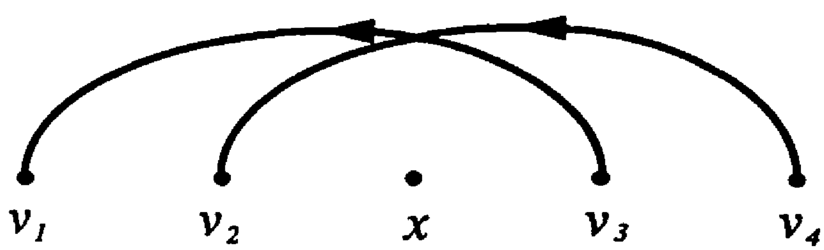


Figure 3.10: T' Realizing E_2.

E_2 is clearly a reversing set of T'. Also, the two arc disjoint 3 cycles v_3, v_1, x, v_3 and v_4, v_2, v_3, v_4 imply that the reversal of one arc of T' will not produce an acyclic tournament. Therefore, T' realizes E_2 and $r(E_2) = 1$.

Let $n \geq 3$ and let the E_n be defined by: $V(E_n) = \{v_1, v_2, \ldots, v_{2n}\}$, $A(E_n) = \{(v_{n+1}, v_1), (v_{n+1}, v_2), \ldots, (v_{2n}, v_n)\}$. Let T' be the acyclic tournament on $V(E_n)$ with acyclic ordering π such that $\pi(v_i) = i$. Note that $E_n^R \subseteq T'$. Let $T = (T' \setminus E_n^R) \cup E_n$.

Hence, E_n is a reversing set of T. Next, we will exhibit n arc disjoint 3-cycles in T. Since there are n arcs in E_n, this will imply by Lemma 3.5 that E_n is a minimum reversing set of T, i.e., T realizes E_n. Therefore, since $|V(T)| = |V(E_n)|, r(E_n) = 0$.

Let

$$\begin{aligned}\tau &= \{(v_1, v_2, v_{n+1}, v_1), (v_2, v_3, v_{n+2}, v_2), \ldots, (v_{n-1}, v_n, v_{2n-1}, v_{n-1})\} \\ &\qquad \cup \{(v_n, v_{2n-2}, v_{2n}, v_n)\}.\end{aligned}$$

It is an easy exercise to see that τ contains n arc disjoint 3-cycles from T, provided that $n \geq 3$. □

3.6.3 Bipartite Digraphs

In this section we compute $r(K_{m,n})$, where

$$V(K_{m,n}) = \{v_1, v_2, \ldots, v_m\} \cup \{u_1, u_2, \ldots, u_n\}$$

and

$$A(K_{m,n}) = \{(v_i, u_j) : i = 1, \ldots, m, j = 1, \ldots, n\}.$$

$K_{m,n}$ will be called a *complete bipartite* digraph.

We will make use of latin rectangles in the next proof. An $m \times n$ *latin rectangle* with entries from a set S of n distinct elements is an array with entries from S such that no element of S appears twice in the same row or in the same column. It is not difficult to show, using for example Hall's marriage theorem, that $m \times n$ latin rectangles exist for $m = 1, \ldots, n$. (See for example Roberts [1984].)

Theorem 3.23 (Barthelemy et al. 1990) $r(K_{m,n}) = \max\{m, n\}$.

Proof: By Lemma 3.10, we may assume that $\max\{m, n\} = m$.

First we show that $r(K_{m,n}) \leq m$. Let T' denote the acyclic tournament on

$V(K_{m,n}) \cup \{x_1, x_2, \ldots, x_m\}$ with acyclic ordering π such that

$$\begin{aligned} \pi(u_i) &= i & i &= 1, 2, \ldots, n \\ \pi(x_j) &= n + j & j &= 1, 2, \ldots, m \\ \pi(v_k) &= n + m + k & k &= 1, 2, \ldots, m \end{aligned}$$

Note that $K_{m,n}^R \subseteq T'$. Let $T = (T' \setminus K_{m,n}^R) \cup K_{m,n}$. Hence, $K_{m,n}$ is a reversing set of T.

Since there are mn arcs in $K_{m,n}$, if we can exhibit mn arc disjoint cycles in T, this will imply by Lemma 3.5 that $K_{m,n}$ is a minimum reversing set of T and hence $r(K_{m,n}) \leq m$. Let L be an $m \times n$ Latin rectangle with entries from $x_1, x_2, \ldots, x_m$. Consider the mn 3-cycles u_j, L_{ij}, v_i, u_j for $1 \leq i \leq m$ and $1 \leq j \leq n$. Since L is a latin rectangle, $i \neq i' \Rightarrow L_{ij} \neq L_{i'j}$ and $j \neq j' \Rightarrow L_{ij} \neq L_{ij'}$. Thus the mn 3-cycles are arc disjoint.

Next, suppose that $r(K_{m,n}) < m$. Therefore, there exists a tournament T with minimum reversing set $K_{m,n}$ such that $|V(T)| < m+n+m$. Without loss of generality, we may assume that the acyclic ordering π' of the vertices of $T' = (T \setminus K_{m,n}) \cup K_{m,n}^R$ satisfies

$$\pi'(u_1) < \pi'(u_2) < \cdots < \pi'(u_n) < \pi'(v_1) < \pi'(v_2) < \cdots < \pi'(v_m).$$

Let $\{x_1, x_2, \ldots, x_k\}$ be the extra vertices in T, i.e., $\{x_1, x_2, \ldots, x_k\} = V(T) \setminus V(K_{m,n})$, and note that $k < m$. Also note that every directed cycle in T must contain an arc of the form (u_i, x_j) where $\pi'(u_i) < \pi'(x_j)$. Let $X = \{(u_i, x_j) : \pi'(u_i) < \pi'(x_j)\} \subseteq T$. Thus X is a transversal of the cycles and by the remarks in the introduction, the minimum size of a transversal is equal to the size of a minimum reversing set. Thus the size of a minimum reversing set of T is at most $|X| \leq kn < mn = |K_{m,n}|$. This contradicts the assumption that $K_{m,n}$ is a minimum reversing set of T. Therefore $r(K_{m,n}) \geq m$. Combining the two inequalities we have $r(K_{m,n}) = m$. $\square$

3.6.4 Alternating Paths

We have shown in the case $n \geq 8$ of the proof of Theorem 3.17 that the reversing number of alternating paths on eight or more vertices is 0. We now determine the reversing number of all alternating paths.

Theorem 3.24 (Barthelemy et al. 1990) *Let A_n be an alternating path on n vertices. Then,*

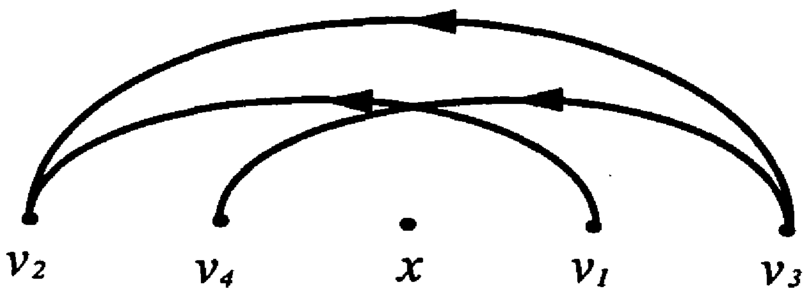
$$r(A_n) = \begin{cases} 1 & \text{if } n = 2,4,5,6,7 \\ 2 & \text{if } n = 3 \\ 0 & \text{if } n \geq 8. \end{cases}$$

Proof: As noted in the proof of Theorem 3.17, Lemma 3.10 says that $r(D) = r(D^R)$ for all D. Thus, we may assume that A_n is labeled with vertex set $\{v_1, \ldots, v_n\}$ and arc set $\{(v_i, v_{i+1}), (v_i, v_{i-1}) : i \text{ is odd, and both vertices are in } V\}$. The cases $n \geq 8$ were shown in the proof of the case $n \geq 8$ of Theorem 3.17. Thus we must consider the cases $n \leq 7$.

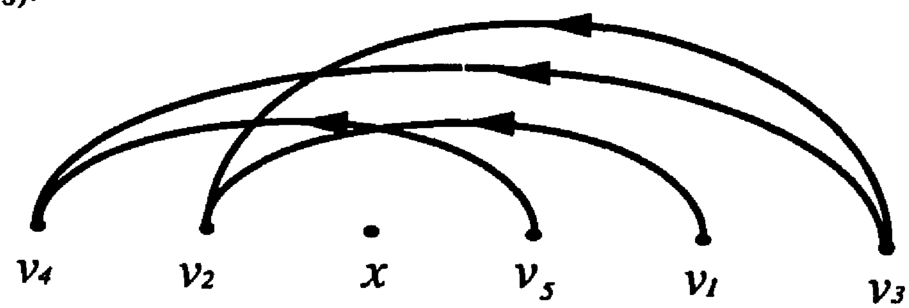
Case $n = 2, 3$: Note that A_2 and A_3 are directed stars on two and three vertices respectively. Thus, by Theorem 3.21, $r(A_2) = 1$ and $r(A_3) = 2$.

Case $n = 4, 5, 6$: By Theorem 3.17, $r(A_4), r(A_5), r(A_6) > 0$. Figure 3.11 shows directed tournaments $T'(A_4), T'(A_5)$, and $T'(A_6)$ on 5, 6, and 7 vertices, respectively, which can easily be shown to have reversing sets A_4, A_5, and A_6, respectively. Also, in Figure 3.11 we list 3, 4, and 5 arc-disjoint cycles from $T'(A_4), T'(A_5)$, and $T'(A_6)$, respectively, to show that $T'(A_4), T'(A_5), T'(A_6)$ realize A_4, A_5, A_6, respectively. Thus $r(A_4) = r(A_5) = r(A_6) = 1$.

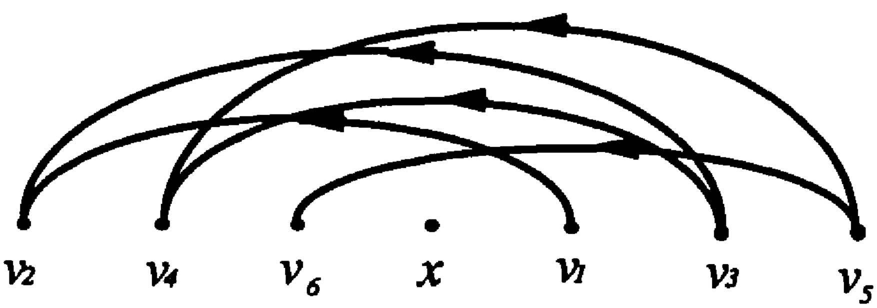
Case 3: $n = 7$: We show that $r(A_7) \leq 1$, by the tournament in Figure 3.12.



(a) $T'(A_4)$ containing arc disjoint cycles (v_1, v_2, v_4, v_1), (v_3, v_2, x, v_1, v_3), and (v_3, v_4, x, v_3).



(b) $T'(A_5)$ containing arc disjoint cycles (v_1, v_2, x, v_1), (v_3, v_2, v_5, v_3), (v_3, v_4, v_1, v_3), and (v_5, v_4, x, v_5).



(c) $T'(A_6)$ containing arc disjoint cycles (v_1, v_2, v_6, v_1), (v_3, v_2, x, v_3), (v_3, v_4, v_1, v_3), (v_5, v_4, x, v_5), and (v_5, v_6, v_3, v_5).

All arcs which are not shown are directed from left to right in the figure.

Figure 3.11: Tournaments realizing alternating paths A_4, A_5, A_6.

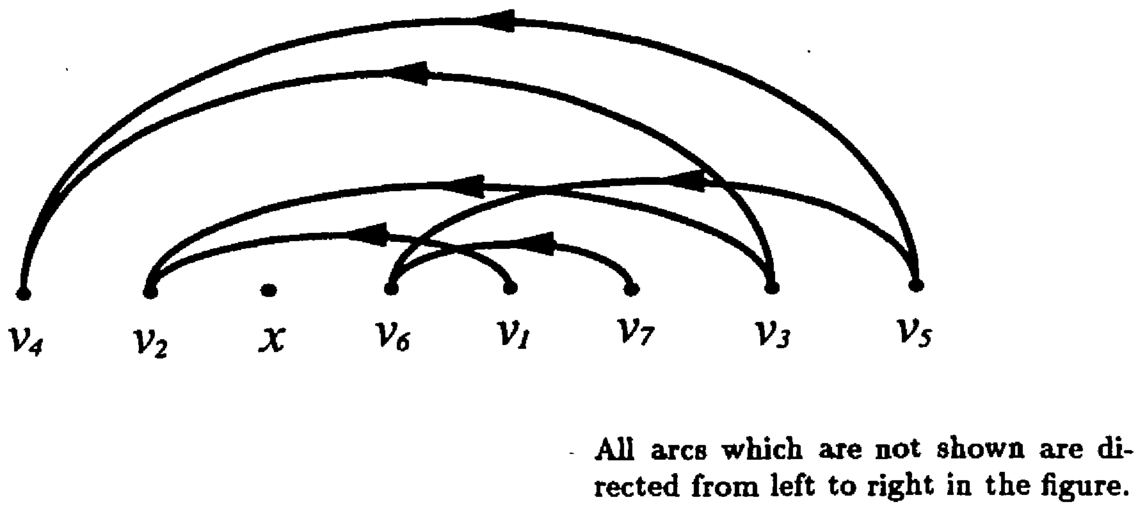


Figure 3.12: Tournament with A_7 as a minimum reversing set containing arc disjoint cycles (v_1, v_2, x, v_1), (v_3, v_2, v_7, v_3), (v_3, v_4, x, v_3), (v_5, v_4, v_2, v_5), (v_5, v_6, v_3, v_5), and (v_7, v_6, v_1, v_7).

Next we must show that A_7 is not a minimum reversing set of any tournament on 7 vertices. Suppose that there exists a tournament T^* on 7 vertices with A_7 as a minimum reversing set.

We first show that the outdegrees of T^* must be in $\{2, 3, 4\}$. If there were a vertex x in T^* with $d^+_{T^*}(x) = 5$ or 6 (respectively 0 or 1), then by reversing at most one arc, a tournament T with x as a source (respectively sink) is obtained. Recall the result of Bermond and Kodratoff [1976], used in Theorem 3.17, that m_6, the size of a largest minimum reversing set for a tournament on 6 vertices, is 4. Then $T|_{V(T^*)\setminus x}$ can be made acyclic with at most four reversals and by Lemma 3.2, the size of a minimum reversing set of T^* is at most five. Thus all outdegrees in T^* must be 2,3 or 4.

The outdegrees in T^* cannot all be 3, since in any reversing set the vertex which becomes the sink after reversal must be contained in three arcs which are reversed and there is no such vertex in A_7.

Thus, since the sum of the outdegrees of vertices in T^* is $\frac{n(n-1)}{2} = 21$, the multiset of outdegrees for T^* must be one of $\{2, 3, 3, 3, 3, 3, 4\}$, $\{2, 2, 3, 3, 3, 4, 4\}$, or

$\{2,2,2,3,4,4,4\}$. The outdegrees after reversal of the arcs in a minimum reversing set are $\{0,1,2,3,4,5,6\}$. Since the arcs of A_7 are those which are reversed in T^* to make the tournament acyclic, we see that the changes in outdegrees from T^* to $(T^* \setminus A_7) \cup A_7^R$ must be exactly three increases by two, two decreases by two, and two decreases by one. It is easy to see that these changes cannot transform the outdegrees $\{2,3,3,3,3,3,4\}$ into $\{0,1,2,3,4,5,6\}$. Thus $\{2,3,3,3,3,3,4\}$ cannot be the multiset of outdegrees.

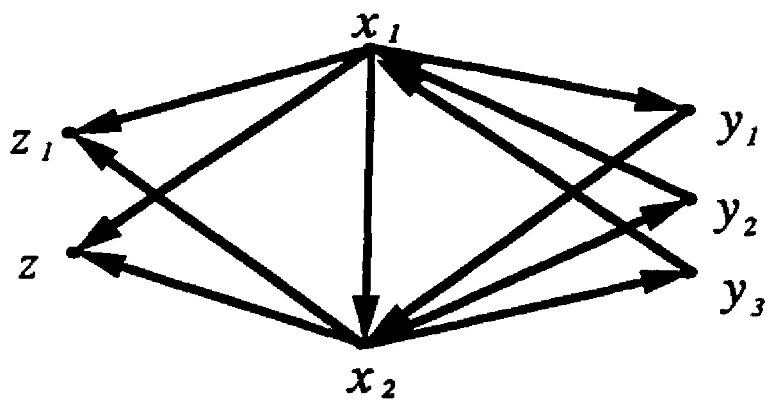
Consider next the case of $\{2,2,2,3,4,4,4\}$. Every tournament contains a Hamiltonian path (see for example Harary [1972]). Applying this observation to the subtournament of T^* induced by vertices of outdegree 4, we see that we can find x, y, z with $(x,y),(y,z) \in T^*$ and $d^+_{T^*}(x) = d^+_{T^*}(y) = d^+_{T^*}(z) = 4$. Consider an acyclic tournament in which x is a source, y is beaten only by x, and z is beaten only by x and y. That is the acyclic order for T has $\pi(x) < \pi(y) < \pi(z)$ and $\pi(z) < \pi(v)$ for all $v \neq x, y$. Here, two reversals in T^* are needed to make x a source. In T^*, y was beaten by two vertices, one of which was x, so one reversal is needed to put y in order. Also z was beaten by two vertices, y and another (possibly x), so at most one more reversal is needed to place z third in the acyclic order. Finally the remaining vertices form a tournament on four vertices; since $m_4 = 1$ at most one additional reversal is needed to make these acyclic. Thus an acyclic tournament T can be always obtained from T^* with at most five reversals, two for x, one for y, at most one for z, and at most one for the remaining vertices. Hence A_7 cannot be a minimum reversing set of a tournament with outdegrees $\{2,2,2,3,4,4,4\}$.

Finally, consider the outdegrees $\{2,2,3,3,3,4,4\}$. Denote by $X = \{x_1, x_2\}$ the vertices with outdegree four, $Y = \{y_1, y_2, y_3\}$ the vertices with outdegree three, and $Z = \{z_1, z_2\}$ the vertices with outdegree two. Since each vertex is contained in a total of six arcs, both x_1 and x_2 are the heads of two arcs. Assume without loss of generality that x_1 beats x_2, i.e., $(x_1, x_2) \in T^*$.

Consider first the case that there exists a vertex in Y (with outdegree three) which is beaten by both x_1 and x_2. Without loss of generality assume that this vertex is y_1.

Since $d^+_{T^*}(y_1) = 3$ and y_1 is beaten by both x_1 and x_2, y_1 must beat three of the four vertices $\{y_2, y_3, z_1, z_2\}$. The acyclic order with $\pi(x_1) < \pi(x_2) < \pi(y_1)$ and $\pi(y_1) < \pi(v)$ for all $v \in \{y_2, y_3, z_1, z_2\}$ can be obtained from T^* as follows: Two reversals for the arcs with x_1 as head, one reversal for the arc other than (x_1, x_2) with x_2 as head, one reversal for the arc from the one vertex in $\{y_2, y_3, z_1, z_2\}$ beating y_1, at most one reversal to put the four vertices $\{y_2, y_3, z_1, z_2\}$ in acyclic order. The last point follows since $m_4 = 1$. Thus an acyclic order is obtained from T^* with a total of at most five reversals, showing that A_7 is not a minimum reversing set of such a tournament.

Otherwise, there is no vertex in Y beaten by both x_1 and x_2. If this is the case, $(x_i, z_i) \in T^*$ for $i = 1, 2$. This follows since x_1 beats three of the vertices and x_2 beats four of the vertices among $Y \cup Z$ and no vertex in Y is beaten by both x_1 and x_2. Then x_2 must beat two of the vertices in Y and x_1 must beat one vertex in Y and these must be distinct. So we may assume that $(x_1, y_1), (x_2, y_2), (x_2, y_3) \in T^*$ (and that $(y_1, x_2), (y_2, x_1), (y_2, x_1) \in T^*$). Then T^* is as shown in Figure 3.13. Consider the



All arcs which are not shown may have any orientation.

Figure 3.13: T^*.

acyclic order with $\pi(x_1) < \pi(y_1) < \pi(x_2)$ and $\pi(x_2) < \pi(v)$ for $v \in \{y_2, y_3, z_1, z_2\}$. This is obtained from T^* by at most five reversals; two for reversing (y_2, x_1) and (y_3, x_1), two for the two arcs from a vertex in $\{y_2, y_3, z_1, z_2\}$ with y_1 as head (since $d^+_{T^*}(y_1) = 3$), and at most one to put $\{y_2, y_3, z_1, z_2\}$ in acyclic order. The last point follows since $m_4 = 1$. Thus A_7 is not a minimum reversing set of T^* in this case, completing the proof that T^* cannot have outdegrees $\{2,2,3,3,3,4,4\}$. This completes the proof that $r(A_7) \neq 0$. □

3.7 Further Research

1. We have shown that $2n - 4\log n \leq r(T_n) \leq 2n - 2$. It remains an open problem to find the exact value for $r(T_n)$.

2. We have determined exact values of the reversing number for alternating paths, directed stars, complete bipartite digraphs, and sets of disjoint arcs. Little is known about the reversing number of acyclic digraphs of other interesting classes. It would be interesting, for example to determine the reversing number for general bipartite digraphs or for digraphs representing weak orders or semiorders.

3. Little is known about $d(n, r)$, the largest arc size of a connected digraph on n vertices having reversing number r. It would be interesting to examine this value in more detail. In particular, we have not been able to show even that $d(n, r+1) > d(n, r)$ although it seems reasonable that this should hold.

4. Calculation of the reversing number in general seems difficult. (Note that calculating reversing number seems to require calculation of the size of a minimum reversing set and that this problem is NP-complete.) We currently do not know the complexity status of determining the reversing number. In fact we don't even have algorithms for determining reversing number for any class of digraphs. Again bipartite digraphs may be a good candidate for such algorithms.

5. Recall that for a minimum reversing set D of T, the acyclic order of $(T \setminus D) \cup D^R$ is a ranking of the vertices of T for which D is the set of arcs inconsistent with

the ranking. Another interesting question to consider would be to examine sets of arcs which arise as inconsistent arcs for other ranking procedures and to find the size of a smallest tournament giving rise to these sets of inconsistencies under the new ranking procedure. Such computations might provide another method to evaluate ranking procedures for round robin tournaments.

Along these lines, it might be interesting to examine the possibility of a concept similar to reversing number defined on weak orders arising from inconsistencies when ties are allowed in ranking.

Chapter 4

Bounded Discrete Representations of Interval Orders

4.1 Introduction

Finite interval orders are partial orders which can be represented by 'strictly greater than' on a set of closed real intervals. (Recall the definition given in the introductory chapter.) We examine classes of interval orders where the intervals are bounded and required to have integral endpoints.

Definition 4.1 *Let* $(A, \succ)$ *be a finite interval order and* $\boldsymbol{\alpha}, \boldsymbol{\beta} : A \to \mathbf{N}$, *non-negative integer constraints. An* $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ **bounded discrete representation** *of* $(A, \succ)$ *is a closed interval representation* $J : A \to \{[l, r] : l, r \in \mathbf{Z}\}$ *so that* $J(i) = [l_i, r_i]$ *with*

(1) $i \succ j \Leftrightarrow l_i > r_j$

and

(2) $\alpha(i) \geq r_i - l_i \geq \beta(i)$ *for all* $i \in A$.

We will use the non-bold notation $[\alpha, \beta]$ to indicate representations for which the upper and lower bounds are constants α and β.

It is also possible to define *open* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ bounded discrete representations for which the closed intervals $J(i) = [l_i, r_i]$ are replaced with open intervals $J(i) = (l_i, r_i)$ and *(1)* is replaced with

(1') $i \succ j \Leftrightarrow l_i \geq r_j$.

However, we will show in Remark 4.1 that these notions are essentially equivalent, and thus will consider only closed bounded discrete representations.

The following gives notation for interval orders which have bounded discrete representations.

Definition 4.2 *Let $(A,\succ)$ be a finite interval order. $(A,\succ) \in \mathcal{D}[\alpha,\beta]$ if and only if $(A,\succ)$ has an $[\alpha,\beta]$ bounded discrete representation.*

The notion of bounded discrete representations was introduced by Bogart and Stellpflug [1989] for semiorders. K.P. Bogart (personal communication) asked whether or not there is a polynomial algorithm to determine if $(A,\succ)$ is in $\mathcal{D}[\alpha,\beta]$ given the order $(A,\succ)$ and the bounds $[\alpha,\beta]$.

Fishburn [1983; 1985a, Chapter 8] studies bounded interval orders (with constant upper and lower bounds and no integrality constraint on the endpoints of intervals). He defines $\mathcal{P}[p,q]$ to be the set of finite interval orders which have a closed real representation in which all intervals have lengths between p and q. He notes that $\mathcal{P}[p,q] = \mathcal{P}[1,q/p]$. This follows from scaling the intervals. Using a set of linear inequalities which must be satisfied by a representation and Farkas' Lemma, he gives the following conditions for bounded interval orders. Recall from the introductory chapter, that the notation $x \succ^{\eta_1} \sim^{\eta_2} \cdots \succ^{\eta_k} y$ represents a sequence of relations from x to y with the first η_1 symbols $\succ$, the next η_2 symbols $\sim$, ..., the final η_k symbols $\succ$.

Theorem 4.1 (Fishburn 1983) *Suppose p and q are positive integers with $p \leq q$ that are relatively prime. Suppose also that $(A,\succ)$ is an interval order with A finite. Then $(A,\succ) \in \mathcal{P}[p,q]$ if and only if*

$$x \succ^{\zeta_1} \sim^{\xi_1} \cdots \succ^{\zeta_n} \sim^{\xi_n} y \Rightarrow x \succ y$$

and

$$x \sim^{\xi_n} \succ^{\zeta_n} \cdots \sim^{\xi_1} \succ^{\zeta_1} y \Rightarrow x \succ y$$

for $n = 1, \ldots, p$, $(\zeta_1, \xi_1, \ldots, \zeta_n, \xi_n) \geq (2,2,\ldots,2,1)$, $\sum_{i=1}^{n} \zeta_i = q + n$, and $\sum_{i=1}^{n} \xi_i = p + n - 1$.

We will show in Remark 4.4 that in fact only one of the sets of implications is needed in the statement. That is, $x \succ^{\zeta_1}\sim^{\epsilon_1} \cdots \succ^{\zeta_n}\sim^{\epsilon_n} y \Rightarrow x \succ y$ holds if and only if $x \sim^{\epsilon_n}\succ^{\zeta_n} \cdots \sim^{\epsilon_1}\succ^{\zeta_1} y \Rightarrow x \succ y$ holds for ζ_i as in the statement of Fishburn's Theorem.

An alternative presentation is to consider a list of forbidden interval orders such that $(A, \succ) \notin \mathcal{P}[1, r]$ if and only if there is a suborder of $(A, \succ)$ isomorphic to some forbidden order. Fishburn [1983] notes that there is a finite list of forbidden orders to $\mathcal{P}[1, r]$ if r is rational and there is no finite list of forbidden orders if r is irrational.

Bogart and Stellpflug [1989, 1990] study bounded discrete representations of semiorders. For each k, they construct a family of semiorders with the property that each order in the family has a $[k+1, k+1]$ bounded discrete representation and no $[k, k]$ bounded discrete representation. They show (personal communication) that the sizes of these families are the catalan numbers $\frac{1}{k+1}\binom{2k}{k}$.

In Section 4.2, we discuss interval orders, review some results from network theory used in later proofs, note a simple transformation between open and closed bounded discrete interval representations, and discuss briefly the analogous bounded discrete representations of interval graphs. In Section 4.3, we will make use of an integer linear programming formulation for a bounded discrete representation, total unimodularity, and Farkas' Lemma to reduce the problem of determining whether or not an interval order $(A, \succ)$ is in $\mathcal{D}[\alpha, \beta]$ to determining whether or not a corresponding digraph $D(A, \succ, \alpha, \beta)$ has a negative length cycle. If the digraph contains no negative cycles, then the lengths of shortest paths from certain vertices will provide endpoints for a representation. This formulation also provides an alternative algorithm to the use of linear programming implied in Fishburn [1983] for determining whether or not an interval order is in $\mathcal{P}[p, q]$.

In Section 4.4, we examine the structure of the digraphs $D(A, \succ, \alpha, \beta)$. The structure results will be used in the remaining sections of the chapter, where we consider constant upper bounds and constant lower bounds of 0 (*degenerate* intervals allowed) and 1 (*non-degenerate* intervals). In Section 4.5, we present necessary and sufficient conditions in the spirit of Fishburn's conditions for the cases of constant bounds. We also show in Section 4.5 that there is a finite list of forbidden orders to the family $\mathcal{D}[\alpha, 0]$, but there is no finite list for $\mathcal{D}[\alpha, 1]$.

In order to more carefully examine forbidden orders, we make the following definition for the family of minimal orders with no $[\alpha, \beta]$ representation.

Definition 4.3 *Let $(A, \succ)$ be a finite interval order. $(A, \succ) \in \mathcal{F}[\alpha, \beta]$ if and only if $(A, \succ)$ has no $[\alpha, \beta]$ bounded discrete representation and every proper suborder $(A', \succ)$ of $(A, \succ)$ has an $[\alpha, \beta]$ bounded discrete representation. That is, $(A, \succ) \notin \mathcal{D}[\alpha, \beta]$ and $(A', \succ) \in \mathcal{D}[\alpha, \beta]$ for all $A' \subset A$.*

Note that $(A, \succ) \notin \mathcal{D}[\alpha, \beta]$ if and only if some suborder of $(A, \succ)$ is isomorphic to an order in $\mathcal{F}[\alpha, \beta]$. In Section 4.6, we characterize $\mathcal{F}[\alpha, 0]$ and in Section 4.7 we investigate $\mathcal{F}[\alpha, 1]$. As noted above, $\mathcal{F}[\alpha, 0]$ is finite and $\mathcal{F}[\alpha, 1]$ is infinite. We will show in Section 4.5 that there are orders which are in both $\mathcal{F}[\alpha + 1, 1]$ and $\mathcal{F}[\alpha, 1]$. Such orders have no $[\alpha + 1, 1]$ bounded discrete representation and every proper suborder has an $[\alpha - 1, 1]$ bounded discrete representation.

4.2 Basic Results

In this section we briefly discuss interval orders, and review some basic results and notation which will be used later in the chapter. We will also note the relationship between closed and open interval representations and discuss which of the results on interval orders can apply to interval graphs.

We first briefly discuss interval orders. For more detailed discussions and reviews of related literature see Roberts [1979] or Fishburn [1985a] and see these books, Golumbic [1980], or Roberts [1978] for discussions of the related concepts of interval graphs, semiorders and indifference graphs (which are the co-comparability graphs of semiorders).

Interval orders arise in the study of temporal events. This observation was made as early as Wiener [1914] using the terminology 'relations of complete sequence'. (See Fishburrn and Monjardet [1990] for discussion of Weiner's early work on this subject). Each temporal event corresponds to some interval in time and event a occurs before event b if a ends before b begins. This is exactly the interval order model. Such a model can be used, for instance, in chronological dating in archaeology and paleontology and in production scheduling. In each case it seems reasonable to ask that the lengths of the intervals be bounded and that the endpoints be limited to a discrete set, providing motivation for the study of bounded discrete representations.

A related model, also noted by Weiner [1914], in which the interval order model arises is in the comparison of measured properties when measurement is subject to error. When measurement is subject to error, each measurement is represented by an interval of uncertainty rather than a point.

The term interval orders was introduced by Fishburn [1970] in the study of preference orderings which give rise to intransitive indifference. We can view $\succ$ in an order as representing preference and $\sim$ as representing indifference. Then semiorders and interval orders allow indifference which is intransitive. Many argue that human preferences often have intransitive indifference. (See for example Luce [1956] or Roberts [1979].) Luce [1956] first introduced semiorders in this context.

Finally, Luce and Suppes [1965] note that interval orders arise in the modelling of just noticeable differences in psychophysics (the study of the human perception of physical quantities such as length or sound). In this case $\succ$ represents a noticeable difference between the physical quantity being measured in two quantities and $\sim$ represents no noticeable difference. It is again desirable to allow $\sim$ to be intransitive, as in semiorders and interval orders.

Denote the row vector $(l_1, l_2, \ldots, l_n)$ by $\mathbf{l}$. Also let $\mathbf{1}$ denote the vector with each entry 1 and similarly for other real numbers. Finally, let $(\mathbf{l}, \mathbf{r})$ denote the concatenation of the vectors $\mathbf{l}$ and $\mathbf{r}$.

With this notation we can state Farkas' lemma in one of its forms (see e.g. Schrijver

[1986], pg. 89).

Lemma 4.2 (Farkas) *Let x and b be real row vectors, and M a real matrix of the appropriate size. Then, exactly one of the following holds, but not both:*
(a) *there exists* $\mathbf{x}$ *such that* $\mathbf{xM} \leq \mathbf{b}$
(b) *there exists* $\mathbf{c} \geq \mathbf{0}$ *such that* $\mathbf{Mc}^T = \mathbf{0}$ *and* $\mathbf{cb}^T < 0$.

A matrix is *totally unimodular* if each sub-determinant is 0,+1, or −1. It is well known that if $\mathbf{M}$ is totally unimodular and the entries of $\mathbf{b}$ are integral, then, if there is a solution $\mathbf{x}$ to $\mathbf{xM} \leq \mathbf{b}$, there is a solution $\mathbf{y}$ with each entry integral such that $\mathbf{yM} \leq \mathbf{b}$ (see e.g. Schrijver [1986]).

Given a digraph D, we can define its *vertex-arc* incidence matrix $\mathbf{M}$. The rows of $\mathbf{M}$ are indexed by the vertices of D, and the columns of $\mathbf{M}$ are indexed by the arcs of D. Denoting the arc corresponding to column j by (a, b), we have $m_{aj} = -1$, $m_{bj} = 1$ and $m_{rj} = 0$ for $r \neq a, b$. It is well known (see e.g. Lawler [1976]), that vertex-arc incidence matrices of digraphs are totally unimodular.

Given a digraph D, a *circulation* is a set of non-negative numbers (which we call *flows*) assigned to the arcs such that, for each vertex v, the sum of the flows over all arcs (w, v) 'entering' v is equal to the sum of the flows over all arcs (v, w) 'leaving' v. Let $c(x, y)$ be the flow on arc (x, y). Then, a circulation satisfies $\sum c(v, x) = \sum c(y, v)$ for all v, where the first sum is over all arcs (v, x) with v as the tail and the second sum is over all arcs (y, v) with v as the head. Thus, if the ordering of the columns of the vertex-arc incidence matrix $\mathbf{M}$ is the same as the ordering of the vector $\mathbf{c}$ of flows, a circulation satisfies $\mathbf{Mc}^T = \mathbf{0}$.

If lengths $\mathbf{k}$ are assigned to the arcs of a digraph D, the total flow in a circulation $\mathbf{c}$ is the inner product $\mathbf{ck}^T$. A circulation has negative total flow if this inner product is negative. If a digraph D admits a circulation with negative total flow, then it contains a negative length cycle C, i.e., a cycle C with $length(C) < 0$. To see this, note first that it is well known that a circulation can be decomposed as the sum of non-negative circulations on cycles in the digraph. That is, denoting by S the set of cycles in a digraph and $\mathbf{c}'_s$ the characteristic vector of a cycle $s \in S$ (a vector with entries 1 corresponding to arcs on the cycle and 0 elsewhere), we have $\mathbf{c} = \sum_{s \in S} t_s \mathbf{c}'_s$ for some non-negative t_s. Then $\mathbf{ck}^T = \sum_{s \in S} t_s (\mathbf{c}'_s \mathbf{k})^T$. Hence, since the inner product $\mathbf{c}'_s \mathbf{k}^T$ is the length of cycle s, one of these lengths must be negative if the total flow $\mathbf{ck}^T$ is negative.

A shortest path from x to y in a digraph D, with lengths on the arcs, is a path P from x to y such that $length(P)$ is less than or equal to the length of any other path from x to y. If there are no negative length cycles in D, a shortest chain is a path. If there are negative cycles, D contains chains with arbitrarliy small negative length (by including many traversals of a negative cycle). Thus we consider shortest paths to be defined only if there are no negative cycles in D. There are many well known polynomial algorithms which will either find the length of shortest paths between all pairs of vertices or determine that the digraph contains a negative cycle. See Lawler [1976] for more details.

Fix some root v and denote the length of a shortest path from v to w by s_w. Bellman's equations for shortest path lengths are $s_w = \min s_x + length(x, w)$, where the minimum is over all vertices x such that the arc (x, w) is in the digraph. In particular, Bellman's equations imply that for a digraph D with no negative cycles, if a vertex v is picked so that there is some path from v to every other vertex in D and if $\mathbf{s}$ is a vector representing shortest path lengths from v, then $\mathbf{s}$ well defined and satisfies $\mathbf{sM} \leq \mathbf{k}$. Here, as above, $\mathbf{M}$ is the vertex-arc incidence matrix of D and $\mathbf{k}$ is the vector of arc lengths. See Lawler [1976] for more details.

We have formulated the discrete representation problem in terms of closed intervals. The discrete open interval representation problem is equivalent to the the closed representation problem. Denote a representation with bounds $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ using open intervals by $(\boldsymbol{\alpha}, \boldsymbol{\beta})$.

Remark 4.1 An interval order $(A, \succ)$ has an open $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ discrete representation if and only if it has a closed $[\boldsymbol{\alpha} - \mathbf{1}, \boldsymbol{\beta} - \mathbf{1}]$ discrete representation. That is, there is an open interval representation if and only if there is a closed interval representation in which both upper and lower bounds are reduced by one. To see this, note that if l_i, r_i are integers for all i, $J' = \{(l_i, r_i) : i \in A\}$ satisfies the condition $i \succ j \Leftrightarrow l_i \geq r_j$ for an open interval representation if and only if $J = \{[l_i, r_i - 1] : i \in A\}$ satisfies the condition $i \succ j \Leftrightarrow l_i > r_j$ for a closed interval representation.

Finally, we briefly comment on the problem of finding bounded discrete interval representations for interval graphs. The solution to this problem does not follow immediately from the solution to the bounded discrete interval order representation problem. In general, there can be many different agreeing interval orders for which an interval graph G is the co-comparability graph. Thus, determining if a graph has a discrete interval graph representation satisfying given bounds $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ could require testing if $(A, \succ) \in \mathcal{D}[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ for each agreeing interval order.

Fishburn [1985a, Chapter 8; 1985b] sketches a proof showing that an interval graph G has a representation in which the intervals have lengths between p and q if and only if every interval order agreeing with G is in $\mathcal{P}[p, q]$. His proof also works for bounded discrete representations in the case that the lower bounds on interval lengths are a constant β. It suffices to note that the flips and permutations of intervals in Fishburn's proof can be accomplished preserving integrality. Thus, an interval graph G has a discrete interval graph representation satisfying the bounds $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ if and only if every agreeing interval order is in $\mathcal{D}[\boldsymbol{\alpha}, \boldsymbol{\beta}]$. We will not repeat Fishburn's proof because it requires extensive background material from Hanlon's [1982] analysis of agreeing interval orders. We note, however, that this relationship between representations for interval graphs and agreeing interval orders does not hold for variable lower bounds in the discrete case.

Let G be the interval graph shown in Figure 4.1(a). The discrete interval graph

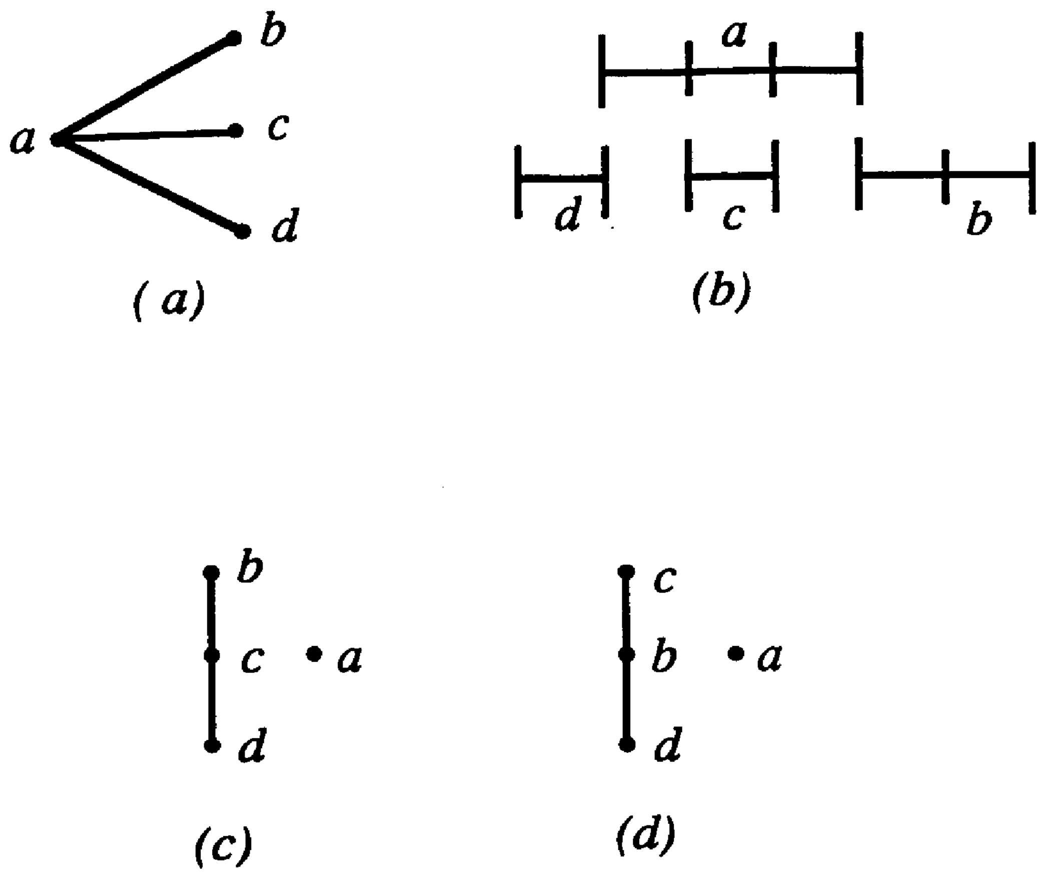


Figure 4.1: (a) An interval graph, (b) a discrete interval graph representation, and (c) and (d) two agreeing orders.

representation shown in Figure 4.1(b) satisfies $1 \leq |J(a)| \leq 3$, $2 \leq |J(b)| \leq 3$, $1 \leq |J(c)| \leq 3$, and $1 \leq |J(d)| \leq 3$. Here $|J(i)|$ indicates the length of interval $J(i)$. The agreeing interval order shown in Figure 4.1(c) (via its Hasse digram) also has the same representation satisfying the same bounds. However, the agreeing interval order shown in Figure 4.1(d) has no representation satisfying the bounds stated above. In this case, if the intervals for c and d both intersect the interval for a, there are at most two integers between the left endpoint of c and the right endpoint d. Then the interval for b, which must fall between these endpoints can have length at most 1, violating the lower bound for b.

4.3 Bounded Discrete Representations

Clearly, $(A, \succ) \in \mathcal{D}[\alpha, \beta]$ if and only if the following integer linear programming problem, which we will call ILP, has a solution.

$$\begin{array}{lll}
\forall i \in A & -l_i + r_i \leq \alpha(i) & \text{interval length is at most } \alpha(i) \\
\forall i \in A & l_i - r_i \leq -\beta(i) & \text{interval length is at least } \beta(i) \\
\forall i \succ j & -l_i + r_j \leq -1 & J(i) \text{ is greater than } J(j) \\
\forall i \sim j & l_i - r_j \leq 0 & J(i) \text{ is not greater than } J(j) \\
\forall i \in A & l_i, r_i \text{ integer} &
\end{array}$$

Note that the final inequality applied to $j \sim i$ insures also that interval $J(j)$ is not greater than $J(i)$, as is necessary for $i \sim j$. To see the third inequality, note that $i \succ j$ holds if and only if $l_i > r_j$. With the condition of integrality on the l_i and r_j, this is equivalent to $l_i \geq r_j + 1$.

Each row of the constraint matrix in ILP has exactly one -1 and one $+1$ entry. Thus, this matrix corresponds to the transpose of the vertex-arc incidence matrix of a certain directed graph.

We define the directed graph $D(A, \succ, \alpha, \beta)$ corresponding to an interval order $(A, \succ)$ and bounds $[\alpha, \beta]$ as follows. When there is no chance of confusion, we will refer to $D(A, \succ, \alpha, \beta)$ as D for simplicity. Let D have vertex set $L \cup R = \{l_1, \ldots, l_{|A|}\} \cup \{r_1, \ldots, r_{|A|}\}$ and arc set $U \cup V \cup W \cup Z$. The arc sets U, V, W, Z and the lengths on these arcs are

$$\begin{aligned}
U &= \{(l_i, r_i) : i = 1, \ldots, |A|\} && \text{with lengths } \alpha(i), \\
V &= \{(r_i, l_i) : i = 1, \ldots, |A|\} && \text{with lengths } -\beta(i), \\
W &= \{(l_i, r_j) : i \succ j\} && \text{with lengths } -1, \\
Z &= \{(r_j, l_i) : i \sim j\} && \text{with lengths } 0.
\end{aligned}$$

For convenience, we use the same notation for variables in ILP as for the vertices of D. There is a correspondence between constraint inequalities in ILP and arcs of D, with lengths of the arcs corresponding to the right hand side of the inequality. There are four types of inequalities and corresponding arcs; we shall refer to these as upper bounds on lengths (U), lower bounds on lengths (V), preference inequalities (W), and incomparability inequalities (Z). We will use the variables u_i, v_i, w_{ij} and z_{ij} to represent the dual variables corresponding to these inequalities. See Figure 4.2 for an example of an interval representation of an interval order and its corresponding digraph.

Note that D is *bipartite*; there are no arcs joining two vertices of L or two vertices of R. An arc from L to R must be in U or W and an arc from R to L must be in V or Z.

Construct the vertex-arc incidence matrix for the digraph described above with row j corresponding to l_j if $j \leq |A|$ and corresponding to $r_{j-|A|}$ if $j > |A|$. Also order the columns so that they are partitioned with the arcs in U appearing first, the arcs in V second, the arcs in W third and the arcs in Z last. Using this notation ILP becomes

$$(\mathbf{l}, \mathbf{r})\mathbf{M} \leq (\boldsymbol{\alpha}, -\boldsymbol{\beta}, -\mathbf{1}, \mathbf{0}) \tag{4.1}$$

$$l_i, r_i \text{ integer}$$

Note that since vertex-arc incidence matrices are totally unimodular and the right hand side is integral, if there is a feasible solution to (4.1), then there is a feasible integral solution. Thus we can drop the integrality constraints. This means that

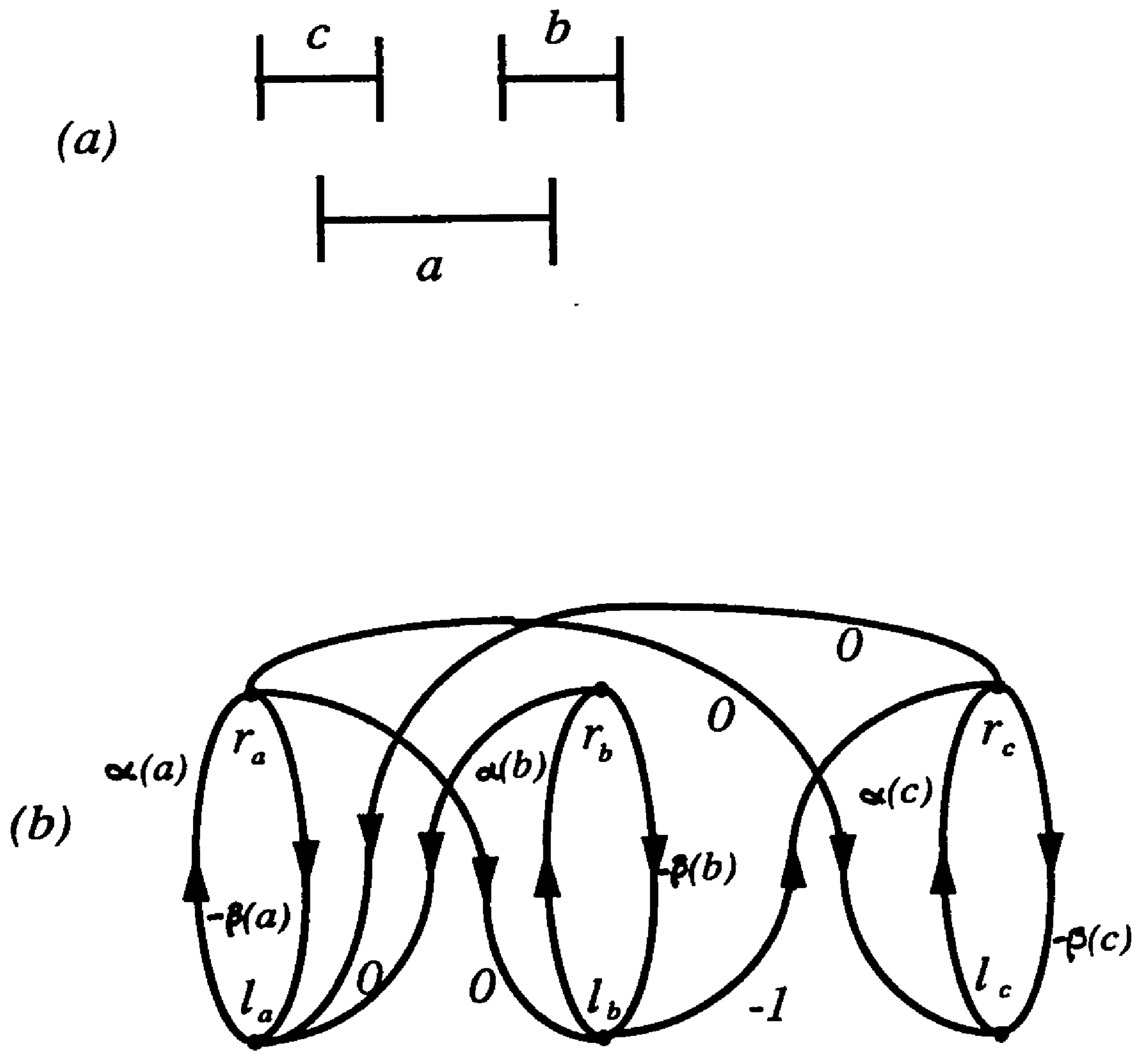


Figure 4.2: (a) A representation of an interval order and (b) its corresponding digraph $D(A, \succ, \alpha, \beta)$

regular linear programming can be used to solve the bounded discrete representation problem. However, making use of the digraph model provides a more efficient procedure to determine discrete representations and provides information on structures blocking such a representation.

Remark 4.2 We may use the ILP formulation with the cost function $\sum(r_i - l_i)$ to find a representation which minimizes the sums of the lengths. Other cost functions can also be minimized using linear programming (since total unimodularity insures integrality). However, by adding an extra element x to A such that $x \succ i$ for all remaining $i \in A$, requiring that the interval for x have length 0, and using the shortest path formulation which will be described in Corollary 4.4, we find a representation which minimizes the distance between the largest and smallest point covered by some interval without resorting to linear programming.

In order to find a more efficient procedure and to develop necessary and sufficient conditions for representability, we will apply Farkas Lemma to (4.1) to translate the problem of finding a bounded discrete representation for $(A, \succ)$ into the problem of finding shortest paths in D. First note that $\mathbf{M}(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z})^T = \mathbf{0}$ is the following set of equations.

$$-u_i + v_i - \sum_{j:i \succ j} w_{ij} + \sum_{j:i \sim j} z_{ji} = 0 \quad \forall i \in A \tag{4.2}$$

$$u_i - v_i + \sum_{j:j \succ i} w_{ji} - \sum_{j:i \sim j} z_{ij} = 0 \;\; \forall i \in A \tag{4.3}$$

Note also that if we view $(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z})$ as the vector representing flows on arcs in U, V, W, Z, then $(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z})$ is a circulation and (4.2) represents flow conservation at vertices l_i and (4.3) represents flow conservation at vertices r_i. Making use of these observations we get the following.

Theorem 4.3 *Let an interval order $(A, \succ)$ and bounds $[\alpha, \beta]$ be given. $(A, \succ) \in \mathcal{D}[\alpha, \beta]$ if and only if the digraph $D(A, \succ, \alpha, \beta)$ contains no negative cycles. Furthermore, if $D(A, \succ, \alpha, \beta)$ contains no negative cycles, pick any vertex $r_v \in D$ such* *that v is maximal with respect to $\succ$. Then the lengths of shortest paths from r_v to vertices l_i (respectively r_i) can be used as the left (respectively right) endpoints in a representation.*

Proof: $(A, \succ) \in \mathcal{D}[\alpha, \beta]$ if and only if ILP has a solution. By the definition of $D(A, \succ, \alpha, \beta)$, ILP has a solution if and only if (4.1) has a feasible solution. By total unimodularity and the assumption that the vector $(\boldsymbol{\alpha}, -\boldsymbol{\beta}, -\mathbf{1}, \mathbf{0})$ has integral entries, the integrality constraint on (4.1) can be dropped. That is, (4.1) without the integrality constraints has a solution if and only if it has an integral solution. By Farkas' Lemma, (4.1) without the integrality constraints has no solution if and only if there exists a $\mathbf{c} = (\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}) \geq \mathbf{0}$ such that $\mathbf{M}\mathbf{c}^T = \mathbf{0}$ and $\mathbf{c}(\boldsymbol{\alpha}, -\boldsymbol{\beta}, -\mathbf{1}, \mathbf{0})^T < 0$. Such a $\mathbf{c}$ represents a circulation in D by the constraint $\mathbf{M}\mathbf{c}^T = \mathbf{0}$. Note that $\mathbf{c}(\boldsymbol{\alpha}, -\boldsymbol{\beta}, -\mathbf{1}, \mathbf{0})^T = \sum \alpha(i)u_i + \sum -\beta(i)v_i - \sum w_{ij}$ is the total flow of the circulation, so there exists a $\mathbf{c} \geq \mathbf{0}$ with $\mathbf{c}(\boldsymbol{\alpha}, -\boldsymbol{\beta}, -\mathbf{1}, \mathbf{0})^T < 0$, and $\mathbf{M}\mathbf{c}^T = \mathbf{0}$ if and only if D admits a circulation with negative total flow. It is well known (recall the comments in Section 4.2), that if D admits a negative circulation, it contains a negative cycle. Clearly, if D contains a negative cycle, it admits a negative circulation. So D admits a negative circulation if and only if it contains a negative cycle. This proves the first part of the theorem.

Furthermore, if D contains no negative cycles, pick some $v \in A$ which is maximal with respect to $\succ$. Recall that this means that $v \succeq i$ for all $i \in A$. Then for all $i \in A$, $(l_v, r_i) \in D$ (if $v \succ i$) or $(r_v, l_i) \in D$ (if $v \sim i$). Also, $(r_v, l_v) \in D$. Thus there is a path in D from r_v to either the l vertex or the r vertex corresponding to each element. Since $(l_i, r_i) \in D$ and $(r_i, l_i) \in D$ for all i, there is a path in D from r_v to every other vertex. Thus shortest paths from r_v to every other vertex are defined. Letting l_i (respectively r_i) be the length of a shortest path from r_v to l_i (respectively r_i) yields an integral feasible solution to (4.1). That this is integral follows from the integrality of the arc lengths. The inequalities (4.1) hold since Bellman's equations for shortest paths hold. □

Corollary 4.4 *Let an interval order $(A, \succ)$ and bounds $[\alpha, \beta]$ be given. There is a polynomial procedure to determine if $(A, \succ) \in \mathcal{D}[\alpha, \beta]$. Moreover, the procedure produces an $[\alpha, \beta]$ discrete representation if one exists.*

Proof: Construct the corresponding digraph (in polynomial time) and use any all pairs shortest path algorithm on the digraph. If a negative cycle is detected, conclude that there is no representation. Otherwise, pick any vertex x such that the shortest paths from x to every other vertex are finite, i.e. some path exists. Such a vertex exists because, as noted in the proof of the theorem, this property holds for vertices r_v corresponding to maximal elements v in the interval order. Set $J(i) = [s_{l_i}, s_{r_i}]$ where s_w denotes the length of a shortest path from vertex x to vertex w in the digraph. This is the $[\alpha, \beta]$ discrete representation. $\square$

We note that with some modifications to ILP, this procedure works to determine representations when no integrality is expected, providing an alternative to a linear programming computation in that case.

Remark 4.3 For non-integral closed representations, the inequalities (W) for preference become $-l_i + r_j \leq -\epsilon$ for some small $\epsilon > 0$. This follows since a representation with $l_i > r_j$ satisfies $l_i \geq r_j + \epsilon$ for some $\epsilon > 0$. A digraph for this non-integral case can then be constructed putting a length of $-\epsilon$ on the arcs from W and the algorithm in the Corollary works.

By Remark 4.1, open interval representations can be transformed to closed representations in the discrete case. However, simple modifications allow direct solution in the open interval case and also allow mixes of closed, open and half open intervals. For open intervals (l_i, r_i) the condition for representation is $i \succ j \Leftrightarrow l_i \geq r_j$. Thus for discrete open intervals, the third inequality in ILP becomes $-l_i + r_j \leq 0$ for $i \succ j$ and the fourth inequality becomes $l_i - r_j \leq -1$. Similar modifications can be used if one interval is open and the other is closed.

We may also use the negative cycles in the corresponding digraph in the non-discrete case to remove one of the conditions in Fishburn's Theorem 4.1.

Remark 4.4 Consider the conditions in Theorem 4.1. Construct a digraph D using the non-discrete conditions as described in Remark 4.3. The implications $x \succ^{\zeta_1} \sim^{\xi_1} \cdots \succ^{\zeta_n} \sim^{\xi_n} y \Rightarrow x \succ y$ together with $y \sim x$ correspond to a negative cycle in D. If the implications are violated, there is such a negative cycle. Breaking the cycle with an arc $(r_y, l_x) \in Z$ such that the next arc in the cycle is $(l_x, r_x) \in U$ (such an arc will be shown to exist in Corollary 4.6) produces a set of implications $x \sim^{\xi_n} \succ^{\zeta_n} \cdots \sim^{\xi_1} \succ^{\zeta_1} y \Rightarrow x \succ y$ which are violated. The converse is shown in a similar manner, noting that any negative cycle must contain an arc from Z followed by an arc from W (where we break the cycle), since otherwise the cycle contains only arcs from U and V and is positive.

At this point we have a polynomial algorithm to recognize if an interval order $(A, \succ) \in \mathcal{D}[\alpha, \beta]$. This answers the original question posed by K.P. Bogart. However, the digraphs $D(A, \succ, \alpha, \beta)$ provide a good deal of information. We will continue to examine bounded discrete interval orders making use of these digraphs in order to obtain an analogue of Fishburn's Theorem 4.1 and furthermore to describe in detail the families of minimal orders $\mathcal{F}[\alpha, 0]$ and $\mathcal{F}[\alpha, 1]$.

4.4 Negative Cycles

In this section we will examine negative cycles in the digraphs $D(A, \succ, \alpha, 0)$ and $D(A, \succ, \alpha, 1)$. We will show that if there is a negative cycle in D, then there is one with certain minimal properties. We first prove a lemma about the relation between elements of A corresponding to vertices in paths in D that contain no length α arcs corresponding to the upper bounds (U). This lemma will not require any assumption of constant bounds. In fact, the lemma does not even require the assumption of integral endpoints.

Lemma 4.5 *If $P = u, \ldots, v$ is a path in $D(A, \succ, \alpha, \beta)$ containing no arcs from U then*

(a) $u = l_i$ *and* $v = r_j \Rightarrow i \succ j$

(b) $u = l_i$ *and* $v = l_j \Rightarrow i \succeq j$

(c) $u = r_i$ *and* $v = l_j$ *or* $v = r_j \Rightarrow i \succeq j$.

Proof: The proof will make use of a general interval representation $\tilde{J}$ on A so that $\tilde{J}(i) = [\tilde{l}_i, \tilde{r}_i]$ and $i \succ j \Leftrightarrow l_i > r_j$. It is well known that such a real representation exists if A is finite. We first show that the left endpoints of the intervals corresponding to l vertices in P form a decreasing sequence moving along the path.

Consider any path P that begins with a vertex from L. Denote P by $l_{\sigma(1)}, r_{\sigma(2)}, l_{\sigma(3)}, r_{\sigma(4)}, \ldots, r_{\sigma(2n)}, l_{\sigma(2n+1)}$. Since there are no arcs from U, the arcs $(l_{\sigma(2k-1)}, r_{\sigma(2k)})$ must be in W, so $\sigma(2k-1) \succ \sigma(2k)$. Thus, the right endpoint of the interval for $\sigma(2k)$ is less than the left endpoint of the interval for $\sigma(2k-1)$. That is, $\tilde{r}_{\sigma(2k)} < \tilde{l}_{\sigma(2k-1)}$. Also, the arc $(r_{\sigma(2k)}, l_{\sigma(2k+1)})$ must be from V or Z. If it is from V, $\sigma(2k) = \sigma(2k+1)$. If it is from Z, $\sigma(2k) \sim \sigma(2k+1)$. In either case, $\tilde{l}_{\sigma(2k+1)} \leq \tilde{r}_{\sigma(2k)}$ and thus $\tilde{l}_{\sigma(2k+1)} < \tilde{l}_{\sigma(2k-1)}$. From this decreasing sequence, $\tilde{r}_{\sigma(1)} \geq \tilde{l}_{\sigma(1)} > \tilde{l}_{\sigma(2n+1)}$ which implies (b).

To show (a), note that (since there are no arcs from U) $(l_{\sigma(2n-1)}, r_{\sigma(2n)}) \in W$, so $\sigma(2n-1) \succ \sigma(2n)$ and $\tilde{r}_{\sigma(2n)} < \tilde{l}_{\sigma(2n-1)}$. As in the proof of (b), $\tilde{l}_{\sigma(2n-1)} < \tilde{l}_{\sigma(1)}$ so $\tilde{r}_{\sigma(2n)} < \tilde{l}_{\sigma(1)}$. Thus $i \succ j$ and (a) holds.

Finally, for (c), consider $P = r_{\sigma(0)}, l_{\sigma(1)}, \ldots, r_{\sigma(2n)}, l_{\sigma(2n+1)}$. The arc $(r_{\sigma(0)}, l_{\sigma(1)})$ is from V or Z. In either case $\sigma(0) \sim \sigma(1)$, so $\tilde{r}_{\sigma(0)} \geq \tilde{l}_{\sigma(1)}$. From the proof of (b), $\tilde{l}_{\sigma(1)} > \tilde{l}_{\sigma(2n+1)}$. So $\tilde{r}_{\sigma(0)} > \tilde{l}_{\sigma(2n+1)}$, which yields (c) when $v = l_j$. From the proof of (b), $\tilde{l}_{\sigma(1)} > \tilde{r}_{\sigma(2n)}$. So $\tilde{r}_{\sigma(0)} > \tilde{r}_{\sigma(2n)} \geq \tilde{l}_{\sigma(2n)}$, which yields (c) when $v = r_j$. □

Corollary 4.6 *Every cycle in $D(A, \succ, \alpha, \beta)$ must contain an arc from U.*

Proof: Assume that some cycle C contains no arc from U and reach a contradiction. Since vertices of any cycle must alternate between r vertices and l vertices, C must contain a vertex $l_{\sigma(1)}$ from L. Breaking the cycle before this vertex, we denote C by

$C = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(2n)}, l_{\sigma(2n+1)}$ with $\sigma(2n+1) = \sigma(1)$. Then $(r_{\sigma(2n)}, l_{\sigma(1)})$ is an arc of C, so it is either from V or Z. If it is from V, then $\sigma(2n) = \sigma(1)$. If it is from Z, then $\sigma(1) \sim \sigma(2n)$. By part (a) of the Lemma, $\sigma(1) \succ \sigma(2n)$, a contradiction in both cases. □

The corollary shows that all cycles contain at least one arc from U corresponding to the upper bound. We will show that if there is a negative cycle in $D(A, \succ, \alpha, 1)$, there is one such that the arcs all appear 'consecutively' as a path alternating between Z arcs and U arcs. In the case that degenerate intervals are allowed, we have the following.

Lemma 4.7 *If $D(A, \succ, \alpha, 0)$ contains a negative cycle, then it contains a cycle C of length -1 that has exactly one arc from U.*

Proof: Let C be a negative cycle with more than one arc from U or length less than -1. We show that C can be reduced to a negative cycle C' such that C' has fewer arcs from U or C' contains exactly one arc from U and has length -1. When C already has exactly one arc from U, C' must have length -1 and one arc from U since reducing the number of arcs from U would in this case produce a negative cycle with no arcs from U, contradicting Corollary 4.6. Repeating the reduction yields the result.

Partition the cycle into paths containing exactly one U arc, with that arc appearing first in each path. Since C has negative length, one of these paths must have negative length. Pick any such negative length path $P = l_{\sigma(1)}, r_{\sigma(2)}, l_{\sigma(3)}, \ldots$ in the partition. Note that $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$ and $\sigma(1) = \sigma(2)$. For $i > 1$, consider the sum of arcs

$$S(i) = \sum_{h=2}^{i} length(x_{\sigma(h-1)}, x_{\sigma(h)}) \tag{4.4}$$

where x may be l or r. Other than the first arc $(l_{\sigma(1)}, r_{\sigma(2)})$ with length α, the arcs are from Z or V with length 0 or from W with length -1. So $S(1) = \alpha$ and for $i > 1$, $S(i) = S(i-1)$ or $S(i) = S(i-1) - 1$. S becomes negative since P has negative length.

Thus for some t, $S(t) = 0$ and $S(t+1) = -1$ with $(l_{\sigma(t)}, r_{\sigma(t+1)}) \in W$. From Lemma 4.5(c), $\sigma(1) = \sigma(2) \succeq \sigma(t+1)$. If $\sigma(1) \succ \sigma(t+1)$ then $(l_{\sigma(1)}, r_{\sigma(t+1)}) \in W$. In C, replace the subpath $l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(t+1)}$ of P with $l_{\sigma(1)}, r_{\sigma(t+1)}$ to get a new cycle C' with the same length as C and one less arc from U. The lengths are the same since $S(t+1) = -1$ and $length(l_{\sigma(1)}, r_{\sigma(t+1)}) = -1$. Alternatively, if $\sigma(1) \sim \sigma(t+1)$ then $(r_{\sigma(t+1)}, l_{\sigma(1)}) \in Z$ and $C' = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(t+1)}, l_{\sigma(1)}$ is a cycle in D. The length of C' is $S(t+1) = -1$ since $length(r_{\sigma(t+1)}, l_{\sigma(1)}) = 0$. Then C' is a cycle with length -1 and exactly one arc from U. □

In order to examine the case of non-degenerate $[\alpha, 1]$ representations, we make the following definition for a sequence of arcs alternating between U arcs and Z arcs.

Definition 4.4 *For $k \geq 1$, a path $P = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, l_{\sigma(2k-1)}, r_{\sigma(2k)}$ in $D(A, \succ, \alpha, \beta)$ is a* **UZ-Path** *if $\sigma(2i-1) = \sigma(2i)$ for $i = 1, \ldots, k$.*

As a consequence of the definition, a UZ-Path must contain arcs $(l_{\sigma(2i-1)}, r_{\sigma(2i)}) \in U$ for $i = 1, \ldots, k$. The arcs $(r_{\sigma(2i)}, l_{\sigma(2i+1)})$ for $i = 1, \ldots, k-1$ must be in Z, since otherwise, if they are in V, $\sigma(2i) = \sigma(2i+1) = \sigma(2(i+1))$, and the vertex $r_{\sigma(2i)} = r_{\sigma(2(i+1))}$ appears twice, contradicting the definition of a path. Thus $(r_{\sigma(2i)}, l_{\sigma(2i+1)}) \in Z$ for $i = 1, \ldots, k-1$ and it follows that $\sigma(2i) \sim \sigma(2i+1) = \sigma(2(i+1))$. The definition allows trivial UZ-Paths consisting of exactly one arc from U. We say that a subpath of a cycle (path) is a *maximal* UZ-Path if it is a UZ-Path and it is not included in a larger UZ-Path in the cycle (path).

In analogy to UZ-Paths, we introduce a path which alternates between arcs from W and arcs from V.

Definition 4.5 *For $k \geq 1$, a path $P = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, l_{\sigma(2k-1)}, r_{\sigma(2k)}$ in $D(A, \succ, \alpha, \beta)$ is a* **WV-Path** *if $\sigma(2i) = \sigma(2i+1)$ for $i = 1, \ldots, (k-1)$.*

As a consequence of the definition, a WV-Path must contain arcs $(r_{\sigma(2i)}, l_{\sigma(2i+1)}) \in V$ for $i = 1, \ldots, (k-1)$. The arcs $(l_{\sigma(2i-1)}, r_{\sigma(2i)})$ for $i = 1, \ldots, k-1$ must be in W

since otherwise, if they are in U, $\sigma(2i-1) = \sigma(2i) = \sigma(2i+1)$, and the vertex $l_{\sigma(2i+1)} = l_{\sigma(2i-1)}$ appears twice, contradicting the definition of a path. Similarly, $(l_{\sigma(2k-1)}, r_{\sigma(2k)}) \in W$, since otherwise, if it is in U, $\sigma(2(k-1)) = \sigma(2k-1) = \sigma(2k)$ and the vertex $r_{\sigma(2(k-1))} = r_{\sigma(2k)}$ appears twice, contradicting the definition of a path. So for $i = 1, \ldots, k$, $\sigma(2i-1) \succ \sigma(2i) = \sigma(2i+1)$. By transitivity, the elements in the order corresponding to a WV-Path satisfy $\sigma(1) \succ \sigma(3) \succ \cdots \succ \sigma(2k-1)$ and $\sigma(2k-1) \succ \sigma(2k)$. As with UZ-Paths, we say that a subpath is a maximal WV-Path if it is not included in a larger WV-path.

We can now state a lemma regarding negative cycles in the non-degenerate case.

Lemma 4.8 *If $D(A, \succ, \alpha, 1)$ contains a negative cycle, then it contains a cycle C of length -1 that has exactly one maximal UZ-Path, or α is odd and it contains a cycle of length -2 with exactly one arc from U and exactly one maximal WV-Path.*

Proof: Let C be a negative cycle. By Corollary 4.6, C has at least one arc in U and hence at least one UZ-Path (possibly a trivial one consisting of just this arc). Let X be the property that a cycle has length -1 or -2 and has exactly one maximal UZ-Path. We first show that there is a cycle satisfying property X. Let C be a negative cycle. We shall show the following.

(a) If C contains more than one maximal UZ-Path, then C can be reduced to a negative cycle C' with the property X or with fewer maximal UZ-Paths.

(b) If C contains exactly one maximal UZ-Path, but the length of C is not -1 or -2, then C can be reduced to a negative cycle C' with property X or with exactly one maximal UZ-Path and one less arc from U.

By continuing with (a), we eventually get a cycle satisfying X or we get to a situation where we can use (b). By continuing with (b) from that point on, we eventually get a cycle with property X. This follows since, by Corollary 4.6, the reduction can not produce a negative cycle containing no U arcs.

We prove both (a) and (b) simultaneously. Thus, start with a negative cycle C satisfying the hypothesis of (a) or (b). Partition C into paths containing exactly one maximal UZ-Path, with the maximal UZ-Path appearing first in the path. If C contains exactly one maximal UZ-path, then the partition consists of exactly one 'path' which in this case is the cycle C with the arc from Z which precedes the maximal UZ-Path deleted. Since C has negative length, one path in the partition must have negative length. Pick any such negative length path $P = l_{\sigma(1)}, r_{\sigma(2)}, l_{\sigma(3)}, \ldots$ in the partition. Denote the UZ-Path at the beginning of this path by $l_{\sigma(1)}, \ldots, r_{\sigma(2k)}$. Note that $\sigma(2i) = \sigma(2i-1)$ for $i = 1, \ldots, k$. As in equation (4.4) in the proof of Lemma 4.7, let $S(i)$ denote the sum of arc lengths up to the i^{th} vertex. The arcs in the UZ-Path are from U and Z and have lengths α and 0 respectively. So $S(2k) = \alpha k$. Consider $i > 2k$, that is, the part of P not containing the maximal UZ-Path. This part of the path contains no positive arcs from U since such an arc alone defines a UZ-Path. Thus, the arcs in the rest of P are from Z with length 0 and from V and W with length -1. So, for $i > 2k$, $S(i) = S(i-1)$ or $S(i) = S(i-1) - 1$. S becomes negative since P has negative length. Thus for some $t > 2k$, $S(t) = 0$ and $S(t+1) = -1$. There are two cases, depending on whether the arc causing the sum to become negative is from V or W.

Case 1: $(l_{\sigma(t)}, r_{\sigma(t+1)}) \in W$.

In this case, a W arc causes the sum to become negative. So $\sigma(t) \succ \sigma(t+1)$. There are three subcases depending on the relation between $\sigma(1)$ and $\sigma(t+1)$.

Subcase i: $\sigma(1) \succ \sigma(t+1)$.

In this case, replace $l_{\sigma(1)}, \ldots, r_{\sigma(t+1)}$ in C with $(l_{\sigma(1)}, r_{\sigma(t+1)}) \in W$ to get a new cycle C' with the same length as C. This follows since the arc $(l_{\sigma(1)}, r_{\sigma(t+1)})$ has length -1 and $l_{\sigma(1)}, \ldots, r_{\sigma(t+1)}$ has length $S(t+1) = -1$. C' has one less maximal UZ-Path than C.

Subcase ii: $\sigma(1) \sim \sigma(t+1)$.

In this case $(r_{\sigma(t+1)}, l_{\sigma(1)}) \in Z$ with length 0. Then, $C' = l_{\sigma(1)}, \ldots, r_{\sigma(t+1)}, l_{\sigma(1)}$ is a cycle with length $S(t+1) + length((r_{\sigma(t+1)}, l_{\sigma(1)}) = -1 + 0 = -1$ and exactly one maximal UZ-Path. So C' has property X.

Subcase iii: $\sigma(t+1) \succ \sigma(1) = \sigma(2)$.

In this case, $(l_{\sigma(t+1)}, r_{\sigma(2)}) \in W$. Also, by the definition of D, $(r_{\sigma(t+1)}, l_{\sigma(t+1)}) \in V$. Let $C' = r_{\sigma(2)}, l_{\sigma(3)}, \ldots, r_{\sigma(t+1)}, l_{\sigma(t+1)}, r_{\sigma(2)}$. Note that $P' = r_{\sigma(2)}, l_{\sigma(3)}, \ldots, r_{\sigma(t+1)}$ is a subpath of P and thus is itself a path (contains no repeated vertices). To show that C' is a cycle, we must show that $l_{\sigma(t+1)}$ does not appear in P'. Note first that $l_{\sigma(t+1)}$ is not part of the maximal UZ-Path in P since, if this were the case, then $r_{\sigma(t+1)}$ would also appear on the UZ-Path (by the definition of UZ-Path), contradicting $t+1 > 2k$. If $l_{\sigma(t+1)}$ appears on the part of P' not containing the maximal UZ-Path, say as $l_{\sigma(u)}$ for $2k < u < t+1$, then by Lemma 4.5(a) applied to $l_{\sigma(u)}, \ldots, r_{\sigma(t+1)}$, $\sigma(t+1) = \sigma(u) \succ \sigma(t+1)$, a contradiction. Thus, $l_{\sigma(t+1)}$ does not appear on P' and C' is indeed a cycle.

C' is formed from the part of P up to $\sigma(t+1)$ with the first arc $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$ deleted, shortening the path length by α. Also the two new arcs added to complete the cycle each have length -1, so the total length of C' is $S(t+1) - \alpha - 2 < 0$. C' contains exactly one maximal UZ-Path. Thus, it has fewer maximal UZ-Paths than C, unless C contained exactly one maximal UZ-Path; in the later case, C' has exactly one maximal UZ-Path and the maximal UZ-Path in C' contains one less arc from U.

Case 2: $(r_{\sigma(t)}, l_{\sigma(t+1)}) \in V$.

In this case, a V arc causes the sum to become negative. So $\sigma(t) = \sigma(t+1)$. Denote by $l_{\sigma(-1)}$ and $r_{\sigma(0)}$ the two vertices preceding $l_{\sigma(1)}$ in C. That is, the two arcs preceding $(r_{\sigma(1)}, l_{\sigma(2)})$ in C are $(l_{\sigma(-1)}, r_{\sigma(0)})$ and $(r_{\sigma(0)}, l_{\sigma(1)})$. Since C is a cycle, $r_{\sigma(0)} \neq r_{\sigma(2)}$ and $\sigma(0) \neq \sigma(1)$ (since $\sigma(1) = \sigma(2)$). Thus, $(r_{\sigma(0)}, l_{\sigma(1)}) \in Z$ as it is not in V. If $(l_{\sigma(-1)}, r_{\sigma(0)}) \in U$, then $l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}, \ldots, r_{\sigma(2k)}$ is a UZ-Path, contradicting the maximality of the UZ-Path $l_{\sigma(1)}, \ldots, r_{\sigma(2k)}$. So $(l_{\sigma(-1)}, r_{\sigma(0)}) \in W$. Then $length(l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}) = 0 - 1 = -1$.

Note that $l_{\sigma(-1)}$ and $r_{\sigma(0)}$ are not equal to any of the vertices appearing on $P' = l_{\sigma(1)}, \ldots, l_{\sigma(t+1)}$. This follows immediately from the definition of a path if C contains at least two maximal UZ-Paths (since no vertices are repeated in a path and since the last vertex of P' appears before the second maximal UZ-Path).

If C contains exactly one maximal UZ-Path, P is C with one arc from Z deleted, so $length(P) = length(C)$. Now, in this case, $l_{\sigma(-1)}, r_{\sigma(0)}$ appear as $l_{\sigma(u)}, r_{\sigma(u+1)}$ in P. Thus, since $length(l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}) = -1$, $length(C) = length(P) = S(u) + 1$. C satisfies the hypothesis of (b) (when C contains exactly one maximal UZ-Path), so $length(C) < -2$ and thus $S(u) < -3$ and $u + 1 > u > t + 1$. So $l_{\sigma(-1)}$ and $r_{\sigma(0)}$ are not equal to any of the vertices appearing on $P' = l_{\sigma(1)}, \ldots, l_{\sigma(t+1)}$ in the case that C contains exactly one maximal UZ-Path.

There are three possibilities for the relation between $\sigma(-1)$ and $\sigma(t)$.

Subcase i: $\sigma(-1) \succ \sigma(t)$.

In this case replace $l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}$ in C by $(l_{\sigma(-1)}, r_{\sigma(t)}) \in W$ to form a new cycle C' with one less maximal UZ-Path. The replaced path has length $S(t) + length(l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}) = 0 - 1 = -1$. The new arc $(l_{\sigma(-1)}, r_{\sigma(t)})$ also has length -1, so the length of C' is the same as the length of C.

Subcase ii: $\sigma(-1) \sim \sigma(t)$.

In this case, $(r_{\sigma(t)}, l_{\sigma(-1)}) \in Z$. Let $C' = l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}, l_{\sigma(-1)}$. Note that C' has exactly one maximal UZ-Path. The length of C' is $length(l_{\sigma(-1)}, r_{\sigma(0)}, l_{\sigma(1)}) + S(t) + length(r_{\sigma(t)}, l_{\sigma(-1)}) = -1 + 0 + 0 = -1$. So C' satisfies property X.

Subcase iii: $\sigma(t+1) = \sigma(t) \succ \sigma(-1)$.

Since $(l_{\sigma(-1)}, r_{\sigma(0)}) \in W$, $\sigma(-1) \succ \sigma(0)$. By transitivity of $\succ$, $\sigma(t+1) \succ \sigma(0)$ and $(l_{\sigma(t+1)}, r_{\sigma(0)}) \in W$. This arc has length -1 and $(r_{\sigma(0)}, l_{\sigma(1)})$ has length 0. Let $C' = r_{\sigma(0)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}, l_{\sigma(t+1)}, r_{\sigma(0)}$. Then C' has exactly one maximal UZ-Path. The length of C' is $length(r_{\sigma(0)}, l_{\sigma(1)}) + S(t+1) + length(l_{\sigma(t+1)}, r_{\sigma(0)}) = 0 - 1 - 1 = -2$. So C' satisfies property X.

This completes the proof that reductions (a) and (b) can be found, and thus that there is a cycle satisfying property X.

Finally, we show that if C has exactly one maximal UZ-Path and length -2, then: when α is even, C can be reduced to a C' such that C' has exactly one maximal UZ-Path and length -1; when α is odd, C can be reduced to a C' that has exactly one maximal UZ-Path and length -1, or C' contains exactly one arc from U, exactly one maximal WV-Path and has length -2.

Let the maximal UZ-Path contain γ arcs from U and let $C = P_0, P, l_{\pi(2\gamma)}$, with the maximal UZ-Path $P_0 = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots, l_{\pi(2)}, r_{\pi(1)}$ and $P = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(v)}$ containing no arcs from U. Note that for $j = 1, \ldots, \gamma$, $(l_{\pi(2j)}, r_{\pi(2j-1)}) \in U$ and for $j = 1, \ldots, \gamma - 1$, $(r_{\pi(2j+1)}, l_{\pi(2j)}) \in Z$. Then $length(P_0) = \alpha\gamma$. The arcs $(r_{\pi(1)}, l_{\sigma(1)})$ and $(r_{\sigma(v)}, l_{\pi(2\gamma)})$ joining P and P_0 and P_0 to $l_{\pi(2\gamma)}$ are from Z and have length 0. Then, $length(C) = length(P_0) + length(P)$ and since $length(C) = -2$, we have $length(P) = -\alpha\gamma - 2$.

The path P contains no arcs from U. If P contains an arc $(r_{\sigma(w)}, l_{\sigma(w+1)}) \in Z$, then $1 < w < w + 1 < v$ by the definition of P. So $P' = l_{\sigma(w-1)}, r_{\sigma(w)}, l_{\sigma(w+1)}, r_{\sigma(w+2)}$ is in P with $(l_{\sigma(w-1)}, r_{\sigma(w)}), (l_{\sigma(w+1)}, r_{\sigma(w+2)}) \in W$ (since there are no arcs from U in P). Then $length(P') = -1 + 0 + -1 = -2$. By Lemma 4.5(a) applied to P', $\sigma(w-1) \succ \sigma(w+2)$ and $(l_{\sigma(w-1)}, r_{\sigma(w+2)}) \in W$ with length -1. Replace P' in C with $l_{\sigma(w-1)}, r_{\sigma(w+2)}$ to obtain the cycle C'. The replaced path has length -2 and the new arc has length -1, so C' has length -1. Also, clearly, C' contains exactly one maximal UZ-Path P_0.

Thus, we may assume that P contains no arcs from Z. Since P also contains no arcs from U, it is a WV-Path. It is maximal since P_0 contains no arcs from W. So, for t odd $l_{\sigma(t)}$ appears in P and for t even $r_{\sigma(t)}$ appears in P. As in equation (4.4), let $S(i)$ denote the sum of the arcs along P, from $\sigma(1)$ to $\sigma(i)$. Since P is a WV-Path,

$S(i) = -i + 1$ and since $length(P) = -\alpha\gamma - 2$, $v = \alpha\gamma + 3$.

Case 1: α is even.

When α is even, note that $r_{\sigma(\alpha+2)}$ appears in P (since $\alpha+2$ is even and since $v = \alpha\gamma+3$). Let $P' = l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(\alpha+2)}$. Since $(l_{\pi(2)}, r_{\pi(1)}) \in U$, $\pi(2) = \pi(1)$. Then since C is a cycle, $\pi(1) \neq \sigma(1)$, and $(r_{\pi(1)}, l_{\sigma(1)}) \in Z$ with length 0. Then, $length(P') = \alpha + S(\alpha+2) = \alpha + -(\alpha+2)+1 = -1$. By Lemma 4.5(c), applied to $r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(\alpha+2)}$, $\pi(2) = \pi(1) \succeq \sigma(\alpha + 2)$. If $\pi(2) \succ \sigma(\alpha + 2)$, then $(l_{\pi(2)}, r_{\sigma(\alpha+2)}) \in W$. Replace P' in C by $l_{\pi(2)}, r_{\sigma(\alpha+2)}$ to obtain a cycle C' with the same length as C and one less arc from U. If $\pi(2) \sim \sigma(\alpha + 2)$, then $(r_{\sigma(\alpha+2)}, l_{\pi(2)}) \in Z$ with length 0 and $C' = P', l_{\pi(2)}$ is a cycle with exactly one arc from U (exactly one maximal UZ-Path) and $length(C') = length(P') = -1$. If C has exactly one arc from U, since C' must contain an arc from U by Corollary 4.6, the case $\pi(2) \succ \sigma(\alpha + 2)$ can not hold and it must be that $\pi(2) \sim \sigma(\alpha + 2)$.

Case 2: α is odd.

When α is odd, note that $r_{\sigma(\alpha+3)}$ appears in P (since $\alpha+3$ is even and since $v = \alpha\gamma+3$). Let $P' = l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(\alpha+3)}$. Since $(l_{\pi(2)}, r_{\pi(1)}) \in U$, $\pi(2) = \pi(1)$. Then since C is a cycle, $\pi(1) \neq \sigma(1)$, and $(r_{\pi(1)}, l_{\sigma(1)}) \in Z$ with length 0. Then, $length(P') = \alpha + S(\alpha+3) = \alpha + -(\alpha+3)+1 = -2$. By Lemma 4.5(c), applied to $r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(\alpha+3)}$, $\pi(2) = \pi(1) \succeq \sigma(\alpha + 3)$. If $\pi(2) \succ \sigma(\alpha + 3)$, then $(l_{\pi(2)}, r_{\sigma(\alpha+3)}) \in W$. Replace P' in C by $l_{\pi(2)}, r_{\sigma(\alpha+3)}$ to obtain a cycle C' with $length(C') = length(C) + 1 = -1$ and exactly one maximal UZ-Path. (The length is increased by one since $length(P') = -2$ and $length(l_{\pi(2)}, r_{\sigma(\alpha+3)}) = -1$.) If $\pi(2) \sim \sigma(\alpha + 3)$, then $(r_{\sigma(\alpha+3)}, l_{\pi(2)}) \in Z$ with length 0 and $C' = P', l_{\pi(2)}$ is a cycle with exactly one arc from U (exactly one maximal UZ-Path), exactly one WV-Path and $length(C') = length(P') = -2$. □

We shall note that negative cycles with exactly one maximal UZ-Path can be decomposed into maximal UZ-Paths and maximal WV-Paths, with the connections between these paths being arcs from Z. Consider any path P in D that contains no arcs from U. Removing all Z arcs from P produces a disconnected collection of paths alternating between arcs from W and arcs from V. Each of these subpaths (except possibly the first) must start with an arc a from W, since the arc preceding a is from Z and thus has a vertex from L as its head. So the tail of a must be in L. Since a is not in U, it must then be in W. Similarly, each of these subpaths (except possibly the last) must end in an arc a from W, since the arc following a is from Z and thus has a vertex from R as its tail. So the head of a must be in R. Since a is not in U, it must then be in W. Thus each of the subpaths except possibly the first and last is a WV-Path.

If C is a cycle containing exactly one maximal UZ-Path, the path P obtained by removing this path can be decomposed as described in the preceding paragraph. It is not difficult to see that the arc a in C following the last arc of the maximal UZ-Path and the arc b preceding the first arc of the maximal UZ-Path are from Z (since otherwise the UZ-Path would not be a path). Since a is the arc preceding the first arc of P and b is the arc following the last arc from P, in a manner similar to that in the previous paragraph, the first and last arcs in P must be in W. Thus P can be decomposed into WV-Paths and we have the following observation.

Remark 4.5 A cycle with exactly one maximal UZ-Path can be written as $C = P_0, P_1, \ldots, P_k$ where P_0 is the UZ-Path and for $i = 1, \ldots, k$, P_i is a WV-Path. The last vertex of P_i is connected to the first vertex of $P_{i+1}(\text{mod } k+1)$ by an arc from Z.

4.5 Necessary and Sufficient Conditions

In this section we study the negative cycles which, according to Theorem 4.3, block discrete representations. We use these cycles to obtain necessary and sufficient conditions for an order to be in $\mathcal{D}[\alpha, \beta]$. We then translate the existence of the negative cycles in $D(A, \succ, \alpha, 0)$ and $D(A, \succ, \alpha, 1)$ described in Lemmas 4.7 and 4.8 into a more compact set of conditions necessary and sufficient for membership in $\mathcal{D}[\alpha, 0]$ and $\mathcal{D}[\alpha, 1]$. We also discuss the cardinalities of the families $\mathcal{F}[\alpha, 0]$ and $\mathcal{F}[\alpha, 1]$ of minimal forbidden orders.

In $(A, \succ)$, a WV-Path $l_{\sigma(1)}, \ldots, r_{\sigma(2k)}$ corresponds to a chain $\sigma(1) \succ^k \sigma(2k)$. This follows since there are k arcs from W corresponding to $\succ$. The $k-1$ arcs from V simply correspond to $\sigma(2i-1) = \sigma(2i)$. Similarly, a UZ-Path $l_{\sigma(1)}, \ldots, r_{\sigma(2k)}$ corresponds to $\sigma(1) \sim^{k-1} \sigma(2k)$ since there are $(k-1)$ arcs from Z corresponding to $\sim$ and k arcs from U corresponding to $\sigma(2i-1) = \sigma(2i)$.

We now use the negative cycles described Theorem 4.3 to give necessary and sufficient conditions on the order.

Theorem 4.9 $(A, \succ) \in \mathcal{D}[\alpha, \beta]$ *if and only if*

$$x \sim^{\gamma_1} \succ^{\eta_1} \sim^{\gamma_2} \succ^{\eta_2} \ldots \sim^{\gamma_k} \succ^{\gamma_k} y \Rightarrow x \succ y \tag{4.5}$$

holds for all integral $\gamma_i, \eta_i, k \geq 1$ *such that*

$$\sum_{i=1}^{k} (\eta_i + \beta(\eta_i - 1)) > \left(\sum_{i=1}^{k} \alpha(\gamma_i - 1) \right) + \alpha. \tag{4.6}$$

Proof: By Theorem 4.3, it is enough to show that $D(A, \succ, \alpha, \beta)$ contains a negative cycle if and only if one of the conditions (4.5) is violated for η_i, γ_i, k satisfying (4.6). These conditions are simply translations of the relations implied by a negative cycle C in $D(A, \succ, \alpha, \beta)$ into chains of $\succ$ and $\sim$ in the order. In a manner similar to Remark 4.5, a negative cycle C can be decomposed as $C = P_1, P_2, \ldots, P_n, u$ where u is the first vertex of P_1 and the P_i are either UZ-Paths or WV-Paths. The last vertex of P_i is connected to the first vertex of P_{i+1} by an arc from Z and the last vertex of P_n is connected to the first vertex of P_1 by an arc from Z. By Corollary 4.6, C contains an arc from U and thus at least one of the paths, say P_1, is a UZ-Path. Furthermore, we may assume that no two consecutive P_i and P_{i+1}, are UZ-Paths since then P_i, P_{i+1} is itself a UZ-Path. The length of UZ-Paths is positive. Thus, if C is negative, there must be at least one WV-Path. So, we may assume that P_n is a WV-Path, since if not, we can combine P_n and P_1 into a larger UZ-Path. Let $u = l_x$ be the first vertex of P_1 and r_y be the last vertex of P_n. Then $(r_y, l_x) \in Z$ and $x \sim y$.

The sequence of paths $P_1, \ldots, P_n$ and the Z arcs joining the paths translate to $xR^1 \sim R^2 \sim \cdots \sim R^n y$ where R^i is $\sim^{k-1}$ if P_i is a UZ-Path with k arcs from U and R^i is $\succ^k$ if P_i is a WV-Path with k arcs from W. R^1 consists of $\sim$ terms and R^n consists of $\succ$ terms by the choice of P_1 and P_n. If R^i, $i \neq 1$, consists of $\sim$ terms, combine it with the $\sim$ term preceding it and the $\sim$ term following it. This can be done since no two consecutive P_i are UZ-Paths. Then, we can write $xR^1 \sim R^2 \sim \cdots \sim R^n y$ with $\sim$ and $\succ$ terms alternating. Thus, if C is negative, corresponding to $P_1, P_2, \ldots, P_n$ is the chain

$$x \sim^{\gamma_1} \succ^{\eta_1} \sim^{\gamma_2} \succ^{\eta_2} \ldots \sim^{\gamma_k} \succ^{\eta_k} y$$

in D. Here $\succ^{\eta_j}$ corresponds to a WV-Path with η_j arcs from W and $\eta_j - 1$ arcs from V. So the length of this path is $-\eta_j - \beta(\eta_j - 1)$. The $\sim^{\gamma_1}$ term is $\sim^{\gamma_1 - 1}\sim$ where the $\sim^{\gamma_1 - 1}$ corresponds to the first UZ-Path with γ_1 arcs from U followed by the $\sim$ term for the first Z arc linking P_1 to P_2. This subpath has length $\alpha\gamma_1$. For $j \neq 1$, the term $\sim^{\gamma_j}$ corresponds to a Z arc linking two WV-Paths if $\gamma_j = 1$. If $\gamma_j > 1$, the term is $\sim\sim^{\gamma_j - 2}\sim$. In this case, the first and last $\sim$ correspond to the linking Z arcs and $\sim^{\gamma_j - 2}$ corresponds to a UZ-Path with $\gamma_j - 1$ arcs from U and $\gamma_j - 2$ arcs from Z. Such a subpath has length $\alpha(\gamma_j - 1)$.

Summing the lengths for the P_i noted above, we get

$$\begin{aligned} & -\textstyle\sum_{i=1}^{k} (\eta_i + \beta(\eta_i - 1)) + \left(\sum_{i=2}^{k} \alpha(\gamma_i - 1)\right) + \alpha\gamma_1 \\ = \quad & -\textstyle\sum_{i=1}^{k} (\eta_i + \beta(\eta_i - 1)) + \left(\sum_{i=1}^{k} \alpha(\gamma_i - 1)\right) + \alpha. \end{aligned} \tag{4.7}$$

If $(A, \succ) \notin \mathcal{D}[\alpha, \beta]$, then there is a cycle C of negative length. It follows that $P_1, \ldots, P_n$ has negative length, so

$$-\sum_{i=1}^{k} (\eta_i + \beta(\eta_i - 1)) + \left(\sum_{i=1}^{k} \alpha(\gamma_i - 1)\right) + \alpha < 0,$$

so (4.6) holds. However, since P_n is joined to P_1 by a Z arc, we have $y \sim x$. Thus (4.5) fails.

Conversely, suppose that (4.6) holds but (4.5) fails for a sequence of relations. If (4.5) fails with $y \sim x$, then $x \sim^{\gamma_1} \succ^{\eta_1} \cdots \sim^{\gamma_k} \succ^{\eta_k} y \sim x$ gives a cycle whose length is given in equation (4.7). By (4.6), this length is less than zero. Thus, there is a negative cycle and $(A, \succ) \notin \mathcal{D}[\alpha, \beta]$. If (4.5) fails with $y \succ x$, then $(l_y, r_x) \in W$. This arc is in D and has length -1. Let P_i be the paths defined from $x \sim^{\gamma_1} \succ^{\eta_1} \cdots \sim^{\gamma_k} \succ^{\eta_k} y$ as above and $C = P_1, \ldots, P_n, u$ as above. Then (4.7) gives the length of C and by (4.6), this length is negative. Also, let P_1' be P_1 with the first arc (l_x, r_x), having length α, removed. So P_1' starts with vertex r_x. Then $length(P_1') < length(P_1)$. Recall that r_y is the last vertex in P_n. Then $C' = P_1', \ldots, P_n, l_y, r_x$ has negative length because $C = P_1, \ldots, P_n, u$ has negative length and $length(P_1') < length(P_1)$ and $length(l_y, r_x) = -1$. Either C' is a negative cycle or it contains a negative cycle (if the new vertex l_y also appears earlier in C). In either case, D contains a negative cycle and $(A, \succ) \notin \mathcal{D}[\alpha, \beta]$. □

The conditions in the previous theorem are not independent. In the cases that the lower bounds are 0 or 1, we can use the structure of the negative cycles in Lemmas 4.7 and 4.8 to state more concise conditions.

Theorem 4.10 $(A, \succ) \in \mathcal{D}[\alpha, 0]$ *if and only if*

$$x \sim\succ^{\eta_1} \sim\succ^{\eta_2} \cdots \sim\succ^{\eta_k} y \Rightarrow x \succ y \tag{4.8}$$

holds for all integral $\eta_i \geq 1$, *such that*

$$\sum_{i=1}^{k} \eta_i = \alpha + 1. \tag{4.9}$$

Proof: By Theorem 4.3, it is enough to show that $D(A, \succ, \alpha, 0)$ contains a negative cycle if and only if one of the conditions is violated. Suppose that there is a negative cycle C. We translate the relations implied by a negative cycle $C = l_x, r_x, P, l_x$ of the type described in Lemma 4.7 into chains in the order. C contains one arc (l_x, r_x) from U connected by an arc from Z to a sequence P of WV-Paths each also joined by an arc from Z. If y is the element corresponding to the last vertex r_y in P, then $x \sim\succ^{\eta_1} \sim\succ^{\eta_2} \cdots \sim\succ^{\eta_k} y$ holds in the order. Here, the η_i indicate the number of arcs from W in the WV-Paths. The WV-Paths are non-empty, so $\eta_i \geq 1$ for all i. The only arcs with non-zero length in C are those from W with length -1 and the one arc from U with length α. From Lemma 4.7, the cycle has length -1 and so there are $\alpha + 1$ arcs from W. Thus, $\sum_{i=1}^{k} \eta_i = \alpha + 1$ and (4.9) holds. Completing the negative cycle C is an arc $(r_y, l_x) \in Z$. This corresponds to $x \sim y$, violating (4.8).

Conversely, assume that (4.9) holds but (4.8) fails for some sequence of relations. By the correspondence between $x \sim\succ^{\eta_1} \cdots \sim\succ^{\eta_k} y$ and the path l_x, r_x, P in C, if (4.9) holds and (4.8) fails with $y \succ x$, we reach a contradiction. The contradiction is reached because by Lemma 4.5(c) applied to r_x, P, $x \succeq y$ since r_y is the last vertex in P. Thus it must be that (4.8) fails with $x \sim y$. Then D contains a negative cycle l_x, r_x, P, r_y, l_x. This is a cycle since $x \sim y$ implies that $(r_y, l_x) \in Z$. □

Theorem 4.11 *For*

$$x \sim^{\gamma} \succ^{\eta_1} \sim\succ^{\eta_2} \cdots \sim\succ^{\eta_k} y \Rightarrow x \succ y \tag{4.10}$$

and

$$\sum_{i=1}^{k} (2\eta_i - 1) = \gamma\alpha + 1, \tag{4.11}$$

we have the following.

(a) *When α is even: $(A, \succ) \in \mathcal{D}[\alpha, 1]$ if and only if (4.11) for integral $\eta_i \geq 1$, $\gamma \geq 1$ implies (4.10).*

(b) *When α is odd: $(A, \succ) \in \mathcal{D}[\alpha, 1]$ if and only if (4.11) for integral $\eta_i \geq 1$, $\gamma \geq 1$ implies (4.10) and*

$$x \sim\succ^{\frac{\alpha+1}{2}} y \Rightarrow x \succ y \tag{4.12}$$

holds.

Proof: By Theorem 4.3, it is enough to show that $D(A, \succ, \alpha, 1)$ contains a negative cycle if and only if one of the conditions is violated. Suppose these is a negative cycle.

We show that is α is even, (4.10) $\not\Rightarrow$ (4.11), and if α is odd (4.10) $\not\Rightarrow$ (4.11) or (4.12) fails. We translate the relations implied by a negative cycle of the type described in Lemma 4.8 into chains in the order. Such a negative cycle contains one maximal UZ-Path connected by an arc from Z to a sequence of maximal WV-Paths each also joined by an arc from Z. Denote this by $C = P_0, P, l_x$ where P_0 is the UZ-Path, P is a sequence of WV-Paths joined by arcs from Z, and l_x is the first vertex of P_0. Let y denote the element corresponding to the last vertex r_y in P. Then $x \sim^{\gamma-1} \sim\succ^{\eta_1} \sim\succ^{\eta_2} \cdots \sim\succ^{\eta_k} y$ holds in the order. Here, the η_i indicate the number of arcs from W in the WV-Paths and γ indicates the number of arcs from U in the UZ-Path.

The WV-Paths are non-empty, so $\eta_i \geq 1$ for all i. A WV-Path with η_i arcs from W has $\eta_i - 1$ arcs from V. Both of these arcs have length -1 in $D(A, \succ, \alpha, 1)$, so the WV-Paths have length $1 - 2\eta_i$. The UV-Path has length $\alpha\gamma$ since each U arc has length α and the arcs from Z have length 0. So

$$length(C) = \alpha\gamma + \sum_{i=1}^{k}(1 - 2\eta_i). \tag{4.13}$$

From Lemma 4.8, the cycle has length -1 or -2. If $length(C) = -1$ then, from (4.13), (4.11) holds. Completing the negative cycle C is an arc $(r_y, l_x) \in Z$. This corresponds to $x \sim y$, violating (4.10). In the case that $length(C) = -2$, by Lemma 4.8, C has exactly one arc from U and exactly one maximal WV-Path. So $\gamma = 1$ and $k = 1$ and $x \sim\succ^{\eta_1} y$ holds in the order. Also, (4.13) with $\gamma = k = 1$ and $length(C) = -2$ gives $\eta_1 = \frac{\alpha+3}{2}$. Then since $x \sim y$ (as arc $(r_y, l_x) \in Z$), (4.12) is violated.

Conversely, suppose that when α is even, (4.10) $\not\Rightarrow$ (4.11), or if α is odd (4.10) $\not\Rightarrow$ (4.11) or (4.12) fails. We show that $D(A, \succ, \alpha, 1)$ contains a negative cycle. By the correspondence between $x \sim^{\gamma} \succ^{\eta_1} \sim\succ^{\eta_2} \cdots \sim\succ^{\eta_k} y$ and the path P_0, P in C, if a condition (4.10) is violated or if (4.12) is violated with $x \sim y$, then D contains a negative cycle. The cycle can be formed since $x \sim y$ implies that $(r_y, l_x) \in Z$. The negativity follows from (4.11) or from $\eta_1 = \frac{\alpha+3}{2}$. In the case that (4.10) or (4.12) is violated with $y \succ x$, we proceed as in the proof of Theorem 4.9. If $y \succ x$, then $(l_y, r_x) \in W$. Let P_0' be P_0

with the first arc (l_x, r_x) removed. Then $length(P_0') < length(P_0)$ and $C' = P_0', P, l_y, r_x$ has negative length. Either C' is itself a negative cycle or it contains a negative cycle. In either case, D contains a negative cycle. □

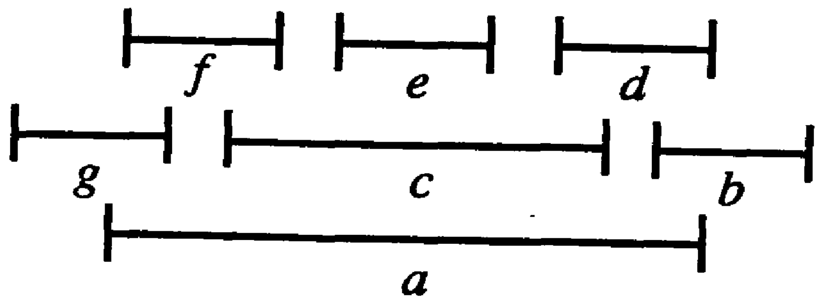


Figure 4.3: An order with a duplicated element in a negative cycle.

In both of the previous theorems, the elements of A appearing in the condition $x \sim\succ^{\eta_1}\sim \cdots \succ^{\eta_k} y$ are not necessarily distinct. For example, the order shown in Figure 4.3 contains $a \sim b \succ c \sim d \succ e \succ f \sim c \succ g$ with $a \sim g$. Thus $a \sim\succ\sim\succ^2\sim\succ g \Rightarrow a \succ g$ is violated. Since (4.9) holds for $\alpha = 3$, the order has no $[3, 0]$ representation. The element c appears twice in the chain, corresponding to appearances as r_c and l_c in a negative cycle in the digraph. It can be checked that there is no condition (4.8) satisfying (4.9) which is violated that does not contain a repeated element.

In the degenerate case ($\beta = 0$), there are a finite number of conditions (4.8) which must be satisfied since the η_i in (4.9) satisfy $1 \leq \eta_i \leq \alpha + 1$ and so k is at most $\alpha + 1$. It is not immediate that these conditions are independent. However, the description of the orders $(A, \succ) \in \mathcal{F}[\alpha, 0]$ given in Theorem 4.25 will imply that the set of conditions (4.8) satisfying (4.9) in Theorem 4.10 are independent. (Recall that $\mathcal{F}[\alpha, \beta]$ is defined in Defintion 4.3.) That is, for a given α and for each condition c defined by (4.8) and (4.9), there is an order which violates c but satisfies every other condition defined by (4.8) and (4.9).

In the non-degenerate case ($\beta = 1$), for each γ there are a finite number of conditions (4.10) as k is bounded by $\alpha\gamma + 1$ in (4.11). However the entire family of conditions described by (4.10) and (4.11) (and (4.12)) is infinite since γ may be any positive integer. The conditions in Theorem 4.11 are not independent. For example it can be shown that if (4.10) is violated for some η_i, $\gamma \geq 2$ and $k = 1$ satisfying (4.11) then a condition (4.10) satisfying (4.11) is violated with $\gamma = 1$ and $k = 1$ or (4.12) is violated. This is shown by a reduction of the corresponding cycles in the digraph. However, Theorem 4.13 and Theorem 4.29 show that an infinite set of independent conditions is necessary to describe membership in $\mathcal{D}[\alpha, 1]$.

We now state results concerning the cardinality of the minimal families. Theorem 4.13 is implied by Theorem 4.29, however, we present a short proof here which will give some insight into the general structure described in Theorem 4.29.

Theorem 4.12 *For a given $\alpha \geq 0$, $\mathcal{F}[\alpha, 0]$ is finite.*

Proof: If $(A, \succ) \in \mathcal{F}[\alpha, 0]$ then $(A, \succ)$ has no $[\alpha, 0]$ discrete representation. By Theorem 4.10, there is a subset of elements of A which violate (4.8) and satisfy (4.9) for some k and η_i. Also, as $(A, \succ)$ is minimal, each element of A must appear in the violated condition (4.8). Thus, the number of elements in A is bounded by $1 + \sum_{i=1}^{k}(\eta_i + 1) = \alpha + k + 2$, the number of elements appearing in a chain of the type in (4.8). Since $\sum_{i=1}^{k} \eta_i = \alpha + 1$ and the η_i are greater than or equal to one, k must satisfy $1 \leq k \leq \alpha + 1$. Thus, every order $(A, \succ)$ in $\mathcal{F}[\alpha, 0]$ satisfies $|A| \leq \alpha + k + 2 \leq 2\alpha + 3$. For a given α, there is a finite number of orders with at most $2\alpha + 3$ elements. So $\mathcal{F}[\alpha, 0]$ must be finite. $\square$

For non-degenerate discrete representations the situation is quite different. There is no finite list of forbidden suborders to an $[\alpha, 1]$ discrete representation. We give a simple example to prove this. The structure used in Theorem 4.29 will be similar to that described here.

Theorem 4.13 *For $\alpha \geq 2$, $\mathcal{F}[\alpha, 1]$ is infinite.*

Proof: For even α, construct an infinite family of minimal forbidden orders. A similar, slightly more complex construction, can be used for odd α. We will omit the construction for odd α since the result will follow from the more general construction of Theorem 4.29.

For α even, and any $\gamma \geq 3$, we will construct an order $(A^{\alpha,\gamma}, \succ)$ on $(\frac{\alpha}{2}) + 4 + \gamma$ elements using an interval representation for which every interval except one has length between 1 and α. The exceptional interval has length $\alpha + 1$. We then show that the corresponding digraph $D(A^{\alpha,\gamma}, \succ, \alpha, 1)$ contains a negative cycle, so there can be no $[\alpha, 1]$ discrete representation. Finally we show that $(A^{\alpha,\gamma}, \succ)$ is minimal by shifting the intervals to produce an $[\alpha, 1]$ discrete representation for any suborder obtained by removing one element from $A^{\alpha,\gamma}$.

Let $A^{\alpha,\gamma} = \{a'_1, a'_2, a'_3, a'_4\} \cup \{a_1, \ldots, a_{\frac{\alpha}{2}\gamma+1}\} \cup \{b_1, \ldots, b_\gamma\}$. Let an interval representation be given as follows.

$$J(a'_1) = [0, 1]$$
$$J(a'_2) = [2, 3]$$
$$J(a'_3) = [2(\tfrac{\alpha}{2}\gamma + 1) - 2, 2(\tfrac{\alpha}{2}\gamma + 1) - 1]$$
$$J(a'_4) = [2(\tfrac{\alpha}{2}\gamma + 1), 2(\tfrac{\alpha}{2}\gamma + 1) + 1]$$

$$J(a_1) = [1, 2]$$
$$J(a_2) = [3, 4]$$
$$\vdots$$
$$J(a_i) = [2i - 1, 2i]$$
$$\vdots$$
$$J(a_{\frac{\alpha}{2}\gamma+1}) = [2(\tfrac{\alpha}{2}\gamma + 1) - 1, 2(\tfrac{\alpha}{2}\gamma + 1)]$$

$$J(b_1) = [1, \alpha + 2]$$
$$J(b_2) = [\alpha + 2, 2\alpha + 2]$$
$$J(b_3) = [2\alpha + 2, 3\alpha + 2]$$
$$\vdots$$
$$J(b_i) = [(i-1)\alpha + 2, i\alpha + 2]$$
$$\vdots$$
$$J(b_\gamma) = [(\gamma - 1)\alpha + 2, \gamma\alpha + 2]$$

Note that $J(b_1)$ is the only interval which violates the constraints. Also, if $\alpha > 2$, the

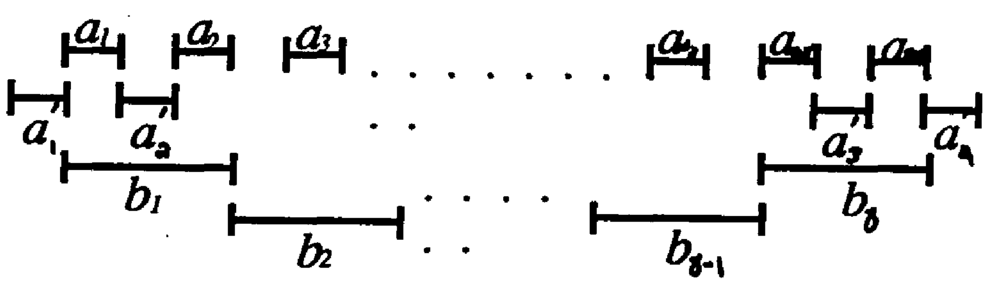


Figure 4.4: $(A^{\alpha,\gamma}, \succ)$

intervals $J(b_i)$ for $i \neq 1$ can be shifted one unit to the left (i.e. $J(b_i) = [(i-1)\alpha + 1, i\alpha + 1]$) without changing the relationship to the $J(a_j)$ intervals. When $\alpha = 2$, $J(b_{\gamma-1}) \cap J(a'_3) \neq \emptyset$, so shifting $J(b_{\gamma-1})$ one unit to the left will destroy this overlap unless the interval $J(a'_3)$ is also changed. We will acount for this special case when $\alpha = 2$ separately. See Figure 4.4 for a schematic representation of $(A^{\alpha,\gamma}, \succ)$. Note that when $\alpha = 2$, the intervals for $J(b_{\gamma-1})$ and $J(a'_3)$ should also overlap in this figure.

It is not difficult to check that the following are paths in $D(A^{\alpha,\gamma}, \succ, \alpha, 1)$.

$$P1 = l_{b_1}, r_{b_1}, l_{b_2}, r_{b_2}, \ldots, l_{b_\gamma}, r_{b_\gamma}$$

$$P2 = l_{a'_4}, r_{a'_3}$$

$$P3 = l_{a_{\frac{\alpha}{2}\gamma+1}}, r_{a_{\frac{\alpha}{2}\gamma}}, l_{a_{\frac{\alpha}{2}\gamma}}, r_{a_{\frac{\alpha}{2}\gamma-1}}, l_{a_{\frac{\alpha}{2}\gamma-1}}, \ldots, r_{a_2}, l_{a_2}, r_{a_1}$$

$$P4 = l_{a'_2}, r_{a'_1}, l_{b_1}$$

Additionally $C = P1, P2, P3, P4$ is a cycle with the links between each pair of paths having length 0. We also have $length(P1) = \alpha\gamma$, $length(P2) = -1$, $length(P3) = -(\alpha\gamma - 1) = -\alpha\gamma + 1$, and $length(P4) = -1$. So the total length of the cycle is -1. Thus, by Theorem 4.3, $(A^{\alpha,\gamma}, \succ) \notin \mathcal{D}[\alpha, 1]$.

To show that $(A^{\alpha,\gamma}, \succ)$ is minimal, i.e., that each proper suborder of $(A^{\alpha,\gamma}, \succ)$ is in $\mathcal{D}[\alpha, 1]$, we construct an $[\alpha, 1]$ representation for each suborder obtained by deleting one element from $(A^{\alpha,\gamma}, \succ)$. If b_1 is removed, the above representation suffices. If some other element is removed, we shift the intervals for some of the elements in order to shorten the interval for b_1 without changing any of the relations. We give the shifts below; in each case it is not difficult to check that no overlaps of intervals are created or destroyed.

(i) Remove b_j for some $1 < j \leq \gamma$:
shrink $J(b_1)$ and for $i < j$ shift $J(b_i)$ one unit to the left;

$$J(b_i) = [(i-1)\alpha + 1, i\alpha + 1] \text{ for } 1 < i < j \text{ and } J(b_1) = [1, \alpha + 1].$$

Additionally, when $\alpha = 2$ and $j = \gamma$, shift the left endpoint of $J(a'_3)$ one unit to the left;

$$J(a'_3) = [2(\frac{\alpha}{2}\gamma + 1) - 3, 2(\frac{\alpha}{2}\gamma + 1) - 1].$$

(ii) Remove a_j for some $3 \leq j \leq \frac{\alpha}{2}\gamma - 1$ (note that if $\alpha = 2$ and $\gamma = 3$ there is no such a_j):
shrink $J(b_1)$ and shift $J(a'_1)$, $J(a'_2)$ and $J(a_i)$ for $i < j$ one unit to the right;

$$J(a'_1) = [1, 2] \text{ and } J(a'_2) = [3, 4]$$

$$J(a_i) = [2i, 2i + 1] \text{ for } 1 \leq i < j \text{ and } J(b_1) = [2, \alpha + 2].$$

(iii) Remove a'_1:

shrink $J(b_1)$;

$$J(b_1) = [2, \alpha + 2].$$

(iv) Remove a'_2:

shrink $J(b_1)$ and shift $J(a'_1)$ to the right;

$$J(a'_1) = [1, 2] \text{ and } J(b_1) = [2, \alpha + 2].$$

(v) Remove a'_3;

shrink $J(b_1)$, move $J(b_i)$ one unit to the left for $i = 2, \ldots, \gamma$ and move $J(a'_4)$ one unit to the left;

$$J(b_i) = [(i-1)\alpha + 1, i\alpha + 1] \text{ for } i = 2, \ldots, \gamma \text{ and } J(b_1) = [1, \alpha + 1].$$

$$J(a'_4) = [2(\frac{\alpha}{2}\gamma + 1) - 1, 2(\frac{\alpha}{2}\gamma + 1)]$$

(vi) Remove a'_4;

shrink $J(b_1)$, and move $J(b_i)$ one unit to the left for $i = 2, \ldots, \gamma$;

$$J(b_i) = [(i-1)\alpha + 1, i\alpha + 1] \text{ for } i = 2, \ldots, \gamma \text{ and } J(b_1) = [1, \alpha + 1].$$

Additionally, in the case that $\alpha = 2$, move the left endpoint of $J(a'_3)$ one unit to the left;

$$J(a'_3) = [2(\frac{\alpha}{2}\gamma + 1) - 3, 2(\frac{\alpha}{2}\gamma + 1) - 1].$$

(vii) Remove a_1:

shrink $J(b_1)$, and shift $J(a'_1)$ and $J(a'_2)$ to the right;

$$J(a'_1) = [1, 2] \text{ and } J(a'_2) = [3, 4] \text{ and } J(b_1) = [2, \alpha + 2].$$

(viii) Remove a_2:

shrink $J(b_1)$ and move $J(a'_1)$, $J(a'_2)$, and $J(a_1)$ to the right;

$$J(a'_1) = [1, 2] \text{ and } J(a'_2) = [3, 4] \text{ and } J(a_1) = [2, 3] \text{ and } J(b_1) = [2, \alpha + 2].$$

(ix) Remove $a_{\frac{\alpha}{2}\gamma}$;

shrink $J(b_1)$, move $J(b_i)$ one unit to the left for $i = 2, \ldots, \gamma$, and move $J(a'_3)$, $J(a'_4)$ and $J(a_{\frac{\alpha}{2}\gamma+1})$ one unit to the left;

$$J(b_i) = [(i-1)\alpha + 1, i\alpha + 1] \text{ for } i = 2, \ldots, \gamma \text{ and } J(b_1) = [1, \alpha + 1]$$

$$J(a'_3) = [2(\frac{\alpha}{2}\gamma + 1) - 3, 2(\frac{\alpha}{2}\gamma + 1) - 2] \text{ and } J(a'_4) = [2(\frac{\alpha}{2}\gamma + 1) - 1, 2(\frac{\alpha}{2}\gamma + 1)]$$

$$J(a_{\frac{\alpha}{2}\gamma+1}) = [2(\frac{\alpha}{2}\gamma + 1) - 2, 2(\frac{\alpha}{2}\gamma + 1) - 1].$$

(x) Remove $a_{\frac{\alpha}{2}\gamma+1}$;

shrink $J(b_1)$, move $J(b_i)$ one unit to the left for $i = 2, \ldots, \gamma$, and move $J(a'_3)$, $J(a'_4)$ one unit to the left;

$$J(b_i) = [(i-1)\alpha + 1, i\alpha + 1] \text{ for } i = 2, \ldots, \gamma \text{ and } J(b_1) = [1, \alpha + 1]$$

$$J(a'_3) = [2(\frac{\alpha}{2}\gamma + 1) - 3, 2(\frac{\alpha}{2}\gamma + 1) - 2] \text{ and } J(a'_4) = [2(\frac{\alpha}{2}\gamma + 1) - 1, 2(\frac{\alpha}{2}\gamma + 1)].$$

□

Finally, we construct a special class of interval orders based on the violated condition (4.12) of Theorem 4.11.

Definition 4.6 *Given α odd $\alpha \geq 3$, the* **bi-minimal** *order $(A, \succ)$ with respect to α is such that the elements can be labeled $A = \{a_0, a_1, \ldots, a_{\frac{\alpha+5}{2}}\}$ with $\succ$ given by $a_1 \succ a_2 \succ \cdots \succ a_{\frac{\alpha+5}{2}}$ (and the relations implied by transitivity in this chain) and for $i = 1, \ldots, \frac{\alpha+5}{2}$, $a_0 \sim a_i$.*

Thus, the bi-minimal order with respect to α consists of a chain of $\frac{\alpha+5}{2}$ elements and a single element which is $\sim$ to every element in the chain. The bi-minimal order with respect to α has no $[\alpha+1, 1]$ discrete representation, but every proper suborder has an $[\alpha, 1]$ representation.

Theorem 4.14 *Given $\alpha \geq 3$, the bi-minimal order $(A, \succ)$ with respect to α satisfies $(A, \succ) \in \mathcal{F}[\alpha + 1, 1]$ and $(A, \succ) \in \mathcal{F}[\alpha, 1]$.*

Proof: Let $\zeta = \frac{\alpha+5}{2}$ and let the elements of the bi-minimal order $(A, \succ)$ be labeled as in Definition 4.6. It can be checked that $C = l_{a_0}, r_{a_0}, l_{a_1}, r_{a_2}, l_{a_2}, \ldots, r_{a_{\zeta-1}}, l_{a_{\zeta-1}}, r_{a_\zeta}, l_{a_0}$ is a cycle in $D(A, \succ, \alpha+1, 1)$ with length -1. (The cycle contains one arc from U, two arcs from Z, and a WV-Path with $\zeta - 1$ arcs from W and $\zeta - 2$ arcs from V.) So $(A, \succ) \notin \mathcal{D}[\alpha+1, 1]$ (and thus $(A, \succ) \notin \mathcal{D}[\alpha, 1]$).

The proof will be completed by showing that for all $a \in A$, $(A \setminus \{a\}, \succ) \in \mathcal{D}[\alpha, 1]$ (and thus $(A \setminus \{a\}, \succ) \in \mathcal{D}[\alpha+1, 1]$).

Consider $A \setminus \{a_0\}$. The set of intervals with length one given by $J(a_i) = [\alpha + 4 - 2i, \alpha + 5 - 2i]$ for $i = 1, \ldots, \frac{\alpha+5}{2}$ can easily be seen to represent $(A \setminus \{a_0\}, \succ)$.

Consider $A \setminus \{a_j\}$ for a given $j \in \{1, \ldots, \frac{\alpha+5}{2}\}$. The set of intervals given by $J(a_0) = [0, \alpha]$ and

$$J(a_i) = \begin{cases} [\alpha+2-2i, \alpha+3-2i] & \text{if } i < j \\ [\alpha+4-2i, \alpha+5-2i] & \text{if } i > j \end{cases}$$

can easily be seen to represent $(A \setminus \{a_j\}, \succ)$. □

Note that it is not difficult to construct an $[\alpha + 2, 1]$ discrete representation for the bi-minimal order $(A, \succ)$ with respect to α by using the representation given in the proof for the case $A \setminus \{a_0\}$ along with the interval $J(a_0) = [0, \alpha + 2]$. So $(A, \succ)$ has the property that it has an $[\alpha + 2, 1]$ discrete representation, but no $[\alpha + 1, 1]$ discrete representation, and every proper suborder has an $[\alpha, 1]$ discrete representation. So, by removing a single element, the length of the longest required interval is reduced by two.

4.6 Minimal Forbidden Orders — Degenerate Case

In this section we examine the family $\mathcal{F}[\alpha, 0]$ of minimal orders with no bounded discrete representation when degenerate intervals (of length 0) are allowed. We first introduce some more notation which will be useful in examining both $\mathcal{F}[\alpha, 1]$ and $\mathcal{F}[\alpha, 0]$.

Definition 4.7 *Given two chains* $C_a = a_1 \succ \cdots \succ a_{n-1} \succ a_n = a_1 \succ^{n-1} a_n$ *and* $C_b = b_1 \succ \cdots \succ b_{m-1} \succ b_m = b_1 \succ^{m-1} b_m$ *in an interval order* $(A, \succ)$*, the end of* C_a *is* **linked** *to the beginning of* C_b *if* $a_n \sim b_1$*,* $a_n \sim b_2$*,* $a_{n-1} \sim b_1$*, and* $a_{n-1} \succ b_2$*.*

We next define a structure in terms of the elements of the order which reflects part of the structure of the negative cycles in Lemma 4.7, i.e., the cycles of length -1 with exactly one arc from U. This alternative presentation to that in Theorem 4.10 will allow us to specify the remaining relations in the order.

Definition 4.8 *A* **sequence of linked chains** *in an interval order* $(A, \succ)$ *consists of chains*

$$\begin{aligned} C_1 &= a_{11} \succ a_{12} \succ \cdots \succ a_{1n_1} \\ &\vdots \\ C_i &= a_{i1} \succ a_{i2} \succ \cdots \succ a_{in_i} \\ &\vdots \\ C_k &= a_{k1} \succ a_{k2} \succ \cdots \succ a_{kn_k} \end{aligned} \tag{4.14}$$

such that the chains are non-trivial ($n_i \geq 2$) *and for* $i = 1, \ldots k-1$*, the end of* C_i *is linked to the beginning of* C_{i+1}*.*

We will call elements a_{ij} with $1 < j < n_i$ *inner.* The elements which appear as a first or last element a_{j1} or a_{jn_j} in some chain C_j are called *outer.* These terms are used with respect to a particular sequence of linked chains. We will show below that no vertex can be both inner and outer, inner vertices may appear only once, and outer vertices may appear once as the first element in a chain and once as the last element in another chain. See Figure 4.5 for an illustration of a sequence of linked chains.

The structure of a sequence of linked chains reflects the sequence of WV-Paths like those described in Remark 4.5.

Remark 4.6 If $(A, \succ)$ contains a sequence of linked chains (4.14), then in $D(A, \succ, \alpha, 0)$ corresponding to the sequence of linked chains, there is a path $P =$

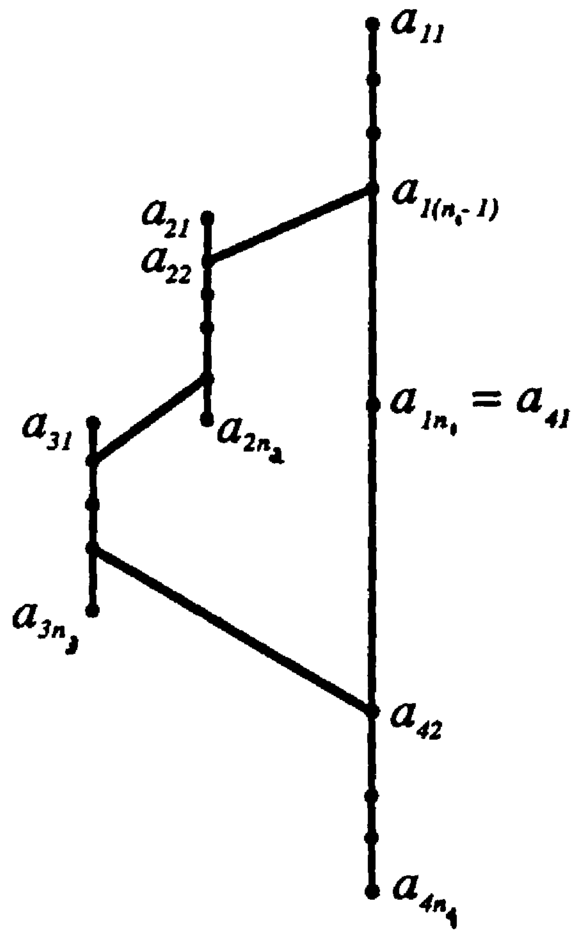


Figure 4.5: A sequence of linked chains.

$P_1, \ldots, P_k$ such that:

(a) for $i = 1, \ldots k$, $P_i = l_{a_{i1}}, r_{a_{i2}}, l_{a_{i2}}, r_{a_{i3}}, l_{a_{i3}}, \ldots, r_{a_{i(n_i-1)}}, l_{a_{i(n_i-1)}}, r_{a_{in_i}}$ is a WV-Path containing vertices corresponding to the elements of C_i,

and

(b) for $i = 1, \ldots k-1$, the last vertex of P_i is connected to the first vertex of P_{i+1} by an arc from Z.

Note that P has no positive length arcs from U. It is not difficult to check that $length(P) = -\sum_{j=1}^{k}(n_j - 1)$.

We can now apply Lemma 4.5 to the path P described above to derive some relations between elements in a sequence of linked chains.

Lemma 4.15 *In a sequence of linked chains (4.14), if $i = j$ and $s < t$ or if $i < j$, $s \neq n_i$, and $t \neq 1$, then $a_{is} \succ a_{jt}$. In particular, the inner elements form a linear order based on a lexicographic order on the subscripts. That is, inner elements a_{is}, $1 < s < n_i$ and a_{jt}, $1 < t < n_j$ form a linear order given by $a_{is} \succ a_{jt} \Leftrightarrow i < j$ or $i = j$ and $s < t$. Also, if $i < j$, $a_{in_i} \succeq a_{jt}$ for $1 \leq t \leq n_j$ and $a_{is} \succeq a_{j1}$ for $1 \leq s \leq n_i$.*

Proof: If $i = j$ and $s < t$, then $a_{is} \succ a_{jt}$ by transitivity of $\succ$ in the chain C_i. Consider the path P in D corresponding to the sequence of linked chains. As long as $s \neq n_i$, and $t \neq 1$, then for $i < j$, $P' = l_{a_{is}}, \ldots, r_{a_{jt}}$ is a subpath of P and Lemma 4.5(a) applies to show that $a_{is} \succ a_{jt}$. Since $s \neq n_i$ and $t \neq 1$ for inner elements a_{is} and a_{jt}, the linear order for inner elements follows immediately from the above implications.

Also, $r_{a_{in_i}}$ appears in P and for $i < j$, either $P' = r_{a_{in_i}}, \ldots, l_{a_{jt}}$ or $P' = r_{a_{in_i}}, \ldots, r_{a_{jt}}$ is a subpath of P. Apply Lemma 4.5(c) to get $a_{in_i} \succeq a_{jt}$. Similarly, for $i < j$, either $P' = l_{a_{is}}, \ldots, l_{a_{j1}}$ or $P' = r_{a_{is}}, \ldots, l_{a_{j1}}$ appears as a subpath of P. Apply Lemma 4.5(b) in the first case and Lemma 4.5(c) in the second to get $a_{is} \succeq a_{j1}$.
□

Lemma 4.15 provides a basis for examining potential multiple appearances of an element in a sequence of linked chains. We show that an element may appear twice in a sequence of linked chains only if it appears as the last element in C_i and the first element of some later chain C_j $(j > i)$. Otherwise the elements are distinct.

Lemma 4.16 *In a sequence of linked chains (4.14), no element appears more than twice. An element may appear twice, but only as $a_{in_i} = a_{j1}$ for $i < j$. Otherwise the elements are distinct.*

Proof: Assume that some element appears twice as $a_{is} = a_{jt}$ $(is \neq jt)$. Then (trivially) $a_{is} \sim a_{jt}$. Since $a_{is} \sim a_{jt}$, the appearances can not be as distinct elements in the same chain (by transitivity of $\succ$). So $i \neq j$. Without loss of generality assume $i < j$.

By Lemma 4.15, the only way for $a_{is} \sim a_{jt}$ to occur is if $s = n_i$ or if $t = 1$. Assume that $s = n_i$ and $t \neq 1$. Then $a_{(i+1)1} \succ a_{jt}$ by Lemma 4.15. (This uses the fact that by Definition 4.8, $n_i \geq 2$.) Also, $a_{(i+1)1} \sim a_{in_i}$ since the end of C_i is linked to the beginning of C_{i+1}. This contradicts $a_{in_i} = a_{jt}$. So $s = n_i \Rightarrow t = 1$. Similarly, assume $t = 1$ and $s \neq n_i$. Then $a_{is} \succ a_{(j-1)n_{j-1}}$ by Lemma 4.15 and $a_{j1} \sim a_{(j-1)n_{j-1}}$ since the end of C_{j-1} is linked to the beginning of C_j. This contradicts $a_{is} = a_{j1}$. So $t = 1 \Rightarrow s = n_i$. Thus, if there are two appearances they are as $a_{in_i} = a_{j1}$ for $i < j$.

Finally, note that an element can not have three appearances which are pairwise consistent with the above condition for two appearances. So an element has at most two appearances. □

This shows that with respect to a given sequence of linked chains an element is either inner or outer but not both. In some cases, an outer element a may replace another outer element a_{i1} in a sequence of linked chains. We will show later that each outer element will have at least one appearance for which it can not be replaced by another element in a sequence of linked chains corresponding to an order in $\mathcal{F}[\alpha, 0]$.

We now prove a technical lemma which will be useful in examining the structure of $\mathcal{F}[\alpha, 0]$.

Lemma 4.17 *Let C be a cycle in $D(A, \succ, \alpha, 0)$ containing exactly one arc from U. If $length(C) < -1$, then $(A, \succ) \notin \mathcal{F}[\alpha, 0]$.*

Proof: Denote C by $C = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(v)}, l_{\sigma(1)}$ with $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$. Thus $\sigma(1) = \sigma(2)$. Let $S(i)$ be the sum of the lengths of arcs from $l_{\sigma(1)}$ to the appearance of $\sigma(i)$ as in equation (4.4) in the proof of Lemma 4.7. Assume that C has length less than -1. $S(t) = 0$ and $S(t+1) = -1$ for some t with $(l_{\sigma(t)}, r_{\sigma(t+1)}) \in W$ (as these are the only negative length arcs). Since the length of C is less than -1, $v > t+1$.

From Lemma 4.5(c) applied to the path in C starting with $r_{\sigma(2)}$ we have $\sigma(2) \succeq \sigma(t+1)$. Furthermore, if $\sigma(1) = \sigma(2) \succ \sigma(t+1)$, then replacing $l_{\sigma(1)}, \ldots, r_{\sigma(t+1)}$ with $l_{\sigma(1)}, r_{\sigma(t+1)}$ yields a cycle with no arc from U, contradicting Corollary 4.6. Thus $\sigma(1) \succ \sigma(t+1)$ produces a contradiction and it must be that $\sigma(1) \sim \sigma(t+1)$.

Since $\sigma(1) \sim \sigma(t+1)$, $(r_{\sigma(t+1)}, l_{\sigma(1)}) \in Z$. Then $C' = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(t+1)}, l_{\sigma(1)}$ is a cycle in D with $length(C') = S(t+1) + length(r_{\sigma(t+1)}, l_{\sigma(1)}) = -1 + 0 = -1$.

Note that C' does not contain $r_{\sigma(v)}$ or $l_{\sigma(v)}$. To see this, first note that since $v > t+1$, the vertex $r_{\sigma(v)}$ is not in C'. By Lemma 4.5(a), $l_{\sigma(v)}$ cannot appear in $l_{\sigma(3)}, \ldots, r_{\sigma(v)}$. Also, $l_{\sigma(v)} \neq l_{\sigma(1)}$, for otherwise, since $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$, $\sigma(1) = \sigma(2) = \sigma(v)$ and $r_{\sigma(v)} = r_{\sigma(2)}$, contradicting the fact that C is a cycle. Thus, $l_{\sigma(v)}$ is not in C and so is not in C'. Then C' is a negative cycle on a digraph corresponding to a proper subset of A. This contradicts the minimality of $(A, \succ)$. So C must have length -1. □

Definition 4.9 *An interval order $(A, \succ)$ has a* **D-linked chain structure** *if there is an element $a_0 \in A$ such that $a_0 \sim a$ for all $a \in A$ and the elements of $A \setminus \{a_0\}$ form a sequence of linked chains $C_1, \ldots, C_k$ as in (4.14) such that $\sum_{j=1}^{k}(n_j - 1) = \alpha + 1$, where n_j denotes the number of elements in C_j.*

We now show that every minimal order with no $[\alpha, 0]$ discrete representation contains a D-linked chain structure.

Lemma 4.18 *If $(A, \succ) \in \mathcal{F}[\alpha, 0]$, then $(A, \succ)$ has a D-linked chain structure.*

Proof: If $(A, \succ) \in \mathcal{F}[\alpha, 0]$, then the corresponding digraph D contains a negative cycle with exactly one arc from U as in Lemma 4.7. As in Remark 4.5, any cycle with exactly one arc from U can be written as $C = l_{\sigma(1)}, r_{\sigma(2)}, P_1, \ldots, P_k, l_{\sigma(1)}$ with $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$, where the P_i are WV-paths and contain no arcs from Z. Pick a negative cycle C from D with exactly one arc from U such that the number k of WV-Paths is minimum.

Denote C by $C = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(v)}, l_{\sigma(1)}$ with $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$. Thus $\sigma(1) = \sigma(2)$. Let $S(i)$ be the sum of arcs from $l_{\sigma(1)}$ to the appearance of $\sigma(i)$ as in equation (4.4) in the proof of Lemma 4.7. Then $S(v) \leq -1$ since C has negative length and since $(r_{\sigma(v)}, l_{\sigma(1)}) \in Z$ and has length 0. From Lemma 4.5(c) applied to the path in C starting with $r_{\sigma(2)}$ we have $\sigma(2) \succeq \sigma(i)$ for $i = 3, \ldots, v$. Furthermore, if $\sigma(1) = \sigma(2) \succ \sigma(i)$ for some i, then replacing $l_{\sigma(1)}, \ldots, r_{\sigma(i)}$ with $l_{\sigma(1)}, r_{\sigma(i)}$ yields a cycle with no arc from U, contradicting Corollary 4.6. Replacing $l_{\sigma(1)}, \ldots, l_{\sigma(i)}$ with $l_{\sigma(1)}, r_{\sigma(i)}, l_{\sigma(i)}$ also yields a cycle with no arcs from U. Thus $\sigma(1) \succ \sigma(i)$ produces a contradiction and it must be that $\sigma(1) \sim \sigma(i)$. Since $(A, \succ)$ is minimal, each element $a \in A$ appears as $\sigma(i)$ for some i. Thus $a_0 = \sigma(1) \sim a$ for all $a \in A$.

By Lemma 4.17, $length(C) = -1$. If P_i contains η_i arcs from W, it corresponds to a chain $C_i = a_{i1} \succ^{\eta_i} a_{in_i}$ in $(A, \succ)$ with $n_i = \eta_i + 1$ elements. The length of C is α (from arc $(l_{\sigma(1)}, r_{\sigma(2)}) \in U$) plus $\sum \eta_i$ (η_i from WV-Path P_i). Thus, $-1 = \alpha + \sum \eta_i$, so $\sum \eta_i = \alpha + 1$ and $\sum(n_i - 1) = \alpha + 1$.

To show that the end of chain C_i is linked to the beginning of chain C_{i+1}, consider the corresponding arcs in the cycle. The last arc of path P_i is joined in C to the first arc of P_{i+1} by an arc from Z. Thus we have $P' = l_{a_{i(n_i-1)}}, r_{a_{in_i}}, l_{a_{(i+1)1}}, r_{a_{(i+1)2}}$ in C. For simplicity we will denote this by $P' = l_a, r_b, l_c, r_d$. Note that P' has length -2. To show that C_i is linked to C_{i+1} we need to show that $a \sim c$, $a \succ d$, $b \sim c$, and

$b \sim d$. Since $(r_b, l_c) \in Z$, $b \sim c$. By Lemma 4.5(a), $a \succ d$. Also by Lemma 4.5, $a \succeq c$ and $b \succeq d$. If $a \succ c$, then replace P' by l_a, r_c, l_c, r_d with length -2. If $b \succ d$ then replace P' by l_a, r_b, l_b, r_d with length -2. The new vertices, r_c in the first case and l_b in the second case, can not repeat vertices other than l_{a_0}, r_{a_0} in C since this would give a cycle containing no arcs from U, contradicting Corollary 4.6. Also, $r_c \neq r_{a_0}$ since otherwise $l_c = l_{a_0}$ in C, contradicting the fact that C is a cycle. Similarly, $l_b \neq l_{a_0}$. Thus, in both cases, replacing P' forms a new negative cycle with one less WV-Path, contradicting the choice of C with the minimum number of such paths. Thus $a \sim c$ and $b \sim d$, completing the proof that the C_i's are linked. □

Remark 4.7 If $(A, \succ) \in \mathcal{F}[\alpha, 0]$, then there is a D-linked chain structure on $(A, \succ)$. Then, as noted in Remark 4.6, there is a path P with length $-\sum_{j=1}^{k}(n_j - 1)$ corresponding to the sequence of linked chains (4.14) in the D-linked chain structure with $l_{a_{11}}$ as the first vertex in P and $r_{a_{kn_k}}$ as the last vertex. Since $a_0 \sim a$ for all $a \in A$, $(r_{a_0}, l_{a_{11}}), (r_{a_{kn_k}}, l_{a_0}) \in Z$. These arcs have length 0. Also, $(l_{a_0}, r_{a_0}) \in U$ with length α by the definition of $D(A, \succ, \alpha, 0)$. Then $l_{a_0}, r_{a_0}, P, l_{a_0}$ is a cycle in $D(A, \succ, \alpha, 0)$ with length $\alpha - \sum_{j=1}^{k}(n_j - 1) = -1$. The last equality follows from the definition of D-linked chain structure. This cycle has exactly one arc from U, like those described in Lemma 4.7.

In order to describe $\mathcal{F}[\alpha, 0]$, we will describe $[\alpha + 1, 0]$ representations of orders $(A, \succ)$ in the family. Shifting these representations will yield an $[\alpha, 0]$ representation of suborders, proving minimality. In order to prove that the representation that we will describe later can represent any order in $\mathcal{F}[\alpha, 0]$, we will first need to prove three technical lemmas.

Lemma 4.19 *Let $(A, \succ) \in \mathcal{F}[\alpha, 0]$. Let $C_1, \ldots, C_k$ as in (4.14) be the sequence of linked chains in a D-linked chain structure on $(A, \succ)$. Let $b = a_{in_i}$ and $c = a_{j1}$ be outer* *elements and $d = a_{fg}$, $1 < g < n_f$ an inner element with respect to the sequence of linked chains. If $d \sim b$ and $d \sim c$, then $b \sim c$.*

Proof: By Lemma 4.15, $i < f < j$ since $d \sim b$ and $d \sim c$. Also by Lemma 4.15, $b \succeq c$. Consider the negative cycle $C = l_{a_0}, r_{a_0}, P, l_{a_0}$ corresponding to the D-linked chain structure as in Remark 4.7. Since d is an inner element, the vertices r_d and l_d in D appear as arc $(r_d, l_d) \in V$ in C Note that this arc has length 0. The last vertex in P is $r_{a_{kn_k}}$. Assume that $b \succ c$ and reach a contradiction. If $b \succ c$, then $(l_b, r_c) \in W$ and has length -1. Also, since $d \sim b$ and $d \sim c$, the arcs (r_d, l_b) and (r_c, l_d) are in Z. In C, replace r_d, l_d having length 0 with r_d, l_b, r_c, l_d having length -1 to form C'. The new vertices r_c and l_b can not repeat vertices other than l_{a_0}, r_{a_0} in C since this would give a cycle containing no arcs from U, contradicting Corollary 4.6. Also, $r_c \neq r_{a_0}$ since otherwise $l_c = l_{a_0}$ in C contradicting the fact that C is a cycle. Similarly, $l_b \neq l_{a_0}$. Thus, C' is a cycle and its length is less than $length(C) = -1$. So $length(C') < -1$ and then Lemma 4.17 applies, since C' still has exactly one arc from U, contradicting $(A, \succ) \in \mathcal{F}[\alpha, 0]$. □

Lemma 4.20 *Let $(A, \succ) \in \mathcal{F}[\alpha, 0]$. Let $C_1, \ldots, C_k$ as in (4.14) be the sequence of linked chains in a D-linked chain structure on $(A, \succ)$. Let b and c be outer elements with respect to the sequence of linked chains Let $d = a_{ij}$ and $e = a_{i(j+1)}$ be two consecutive elements in some chain in the sequence of linked chains. Then it is not the case that $b \sim c$, $b \succ e$, and $d \succ c$ all hold.*

Proof: Let b, c, d and e be as in the statement of the lemma. Assume that all three relations hold and reach a contradiction. Let $C = l_{a_0}, r_{a_0}, P, l_{a_0}$ be as in Remark 4.7. The arc $(l_d, r_e) \in W$ with length -1 appears in P since d and e appear consecutively in a chain in $(A, \succ)$. By assumption, $(l_d, r_c) \in W$, $(r_c, l_b) \in Z$, and $(l_b, r_e) \in W$ are all in D. Replace l_d, r_e (with length -1) in P by l_d, r_c, l_b, r_e (with length -2) to form C'. As in the previous lemma, the new vertices r_c and l_b can not repeat vertices other than l_{a_0}, r_{a_0} in C since this would give a cycle containing no arcs from U, contradicting Corollary 4.6. Also, $r_c \neq r_{a_0}$ since otherwise $l_c = l_{a_0}$ in C, contradicting the fact that C is a cycle. Similarly, $l_b \neq l_{a_0}$. Thus, C' is a cycle and its length is less than -1. Then Lemma 4.17 provides a contradiction to $(A, \succ) \in \mathcal{F}[\alpha, 0]$. □

In order to state the third lemma, we examine the set of inner elements that are $\sim$ to a given outer element in a sequence of linked chains.

Definition 4.10 *Let y be an outer element with respect to a sequence of linked chains (4.14). Let $I' = \{a_{is} : 1 < s < n_i\} \cup \{a_{11}, a_{kn_k}\}$. (So the elements of I' are the inner elements in the sequence of linked chains along with a_{11} and a_{kn_k}.) The* **indifference interval** *$I(y)$ of an outer element y is given by*

$$I(y) = \{x \in A : y \sim x, x \in I'\}. \tag{4.15}$$

Note that the sets $I(y)$ are sub-intervals of the linear order consisting of the inner elements (as described in Lemma 4.15) along with maximal element a_{11} and minimal element a_{kn_k}. Thus the indifference intervals are linear orders and have a unique maximal and minimal element.

Lemma 4.21 *Let $(A, \succ) \in \mathcal{F}[\alpha, 0]$. Let $C_1, \ldots, C_k$ as in (4.14) be the sequence of linked chains in a D-linked chain structure on $(A, \succ)$. If a_{j1}, $j \neq 1$, and a_{in_i}, $i \neq k$, with $i < j$ are outer elements with respect to the sequence of linked chains, then $|I(a_{j1}) \cap I(a_{in_i})| \neq 1$.*

Proof: Assume that $|I(a_{j1}) \cap I(a_{in_i})| = 1$ and reach a contradiction. Let $C = l_{a_0}, r_{a_0}, P, l_{a_0}$ be as in Remark 4.7. Let a_{st}, $1 < t < n_s$, be the single element in $I(a_{j1}) \cap I(a_{in_i})$. By Lemma 4.15, $a_{j1} \sim a_{st} \Rightarrow s < j$. Also by Lemma 4.15, $a_{in_i} \sim a_{st} \Rightarrow i < s$. Thus $i < s < j$. So $a_{st} \neq a_{11}$ and $a_{st} \neq a_{kn_k}$ and a_{st} is an inner element with respect to the sequence of linked chains.

Note that $l_{a_{s(t-1)}}, r_{a_{st}}, l_{a_{st}}, r_{a_{s(t+1)}} = P'$ is a path that is part of P. Since a_{st} is an inner element, there exist $a_{s(t-1)} \succ a_{st} \succ a_{s(t+1)}$. By Lemma 4.15, $a_{in_i} \succeq a_{s(t+1)} \succeq a_{j1}$. If $a_{s(t+1)} \succ a_{j1}$, then transitivity of $\succ$ would imply $a_{st} \succ a_{j1}$, a contradiction. Thus, $a_{s(t+1)} \sim a_{j1}$ and since $I(a_{j1}) \cap I(a_{in_i}) = \{a_{st}\}$, it is not the case that $a_{in_i} \sim a_{s(t+1)}$. So $a_{in_i} \succ a_{s(t+1)}$. In a similar manner, it can be shown that $a_{s(t-1)} \succ a_{j1}$. Then, $(l_{a_{s(t-1)}}, r_{a_{j1}}) \in W$ and $(l_{a_{in_i}}, r_{a_{s(t+1)}}) \in W$. Also, by Lemma 4.19, $a_{in_i} \sim a_{j1}$ and thus $(r_{a_{j1}}, l_{a_{in_i}}) \in Z$. Replace P' in C with $P'' = l_{a_{s(t-1)}}, r_{a_{j1}}, l_{a_{in_i}}, r_{a_{s(t+1)}}$ to get C'. Note that $length(P') = length(P'') = -2$, so C' has negative length.

We now show that C' is indeed a cycle and that C' does not contain vertices $l_{a_{st}}$ and $r_{a_{st}}$ corresponding to a_{st}. Note that (as described in Remarks 4.7 and 4.6), $l_{a_{j1}}$ and $r_{a_{in_i}}$ appear in C. Since also (l_{a_0}, r_{a_0}) appears in C and there are no repeated vertices in C, $a_{j1} \neq a_0$, $a_{in_i} \neq a_0$, and thus $r_{a_{j1}} \neq r_{a_0}$ and $l_{a_{in_i}} \neq l_{a_0}$. Also, if $r_{a_{j1}}$ appears in P, then C' contains a cycle with no arcs from U, contradicting Corollary 4.6. This follows since $r_{a_{j1}}$ appears in P'', so if it also appears in P, then $r_{a_{j1}}$ appears twice in C', and the two appearences partition C' into two cycles, and C' contains exactly one arc from U. Similarly, $l_{a_{in_i}}$ does not appear in P. So $r_{a_{j1}}$ and $l_{a_{in_i}}$ appear only once in C' and C' is indeed a cycle. Finally, since replacing P' with P'' removes the vertices $l_{a_{st}}$ and $r_{a_{st}}$, C' contains no vertex which corresponds to a_{st}. Since C' is negative, this contradicts $(A, \succ) \in \mathcal{F}[\alpha, 0]$. □

Definition 4.11 *Let a sequence of linked chains $C_1, \ldots, C_k$ as in (4.14) be given. An outer element $x \neq a_{11}, a_{kn_k}$ with respect to the sequence of linked chains* **duplicates** *a_{j1}, $j > 1$, if x can replace a_{j1} in the sequence, that is, if $x \succ a_{j2}$, $x \sim a_{(j-1)n_{j-1}}$, and $x \sim a_{(j-1)(n_{j-1}-1)}$. Similarly, x* **duplicates** *a_{in_i}, $i < k$, if $a_{i(n_i-1)} \succ x$, $x \sim a_{(i+1)1}$, and $x \sim a_{(i+1)2}$. We will say that the appearance of an outer element y as a first element in a chain is* **duplicated** *if y is a_{j1} for some $1 < j \leq k$ and some x duplicates a_{j1} or if y duplicates a_{j1} for some $1 < j \leq k$. Similarly the appearance as a last element* *in a chain is* **duplicated** *if y is a_{in_i} for some $1 \le i < k$ and some x duplicates a_{in_i} or if y duplicates a_{in_i} for some $1 \le i < k$.*

Note that when x duplicates a_{j1} in a sequence of linked chains, we can replace a_{j1} with x without affecting the linkages in the sequence. In particular, both x and a_{j1} satisfy the relations given in Lemma 4.15. Thus, when there is no chance of confusion, we will use the notation $x = a_{j1}$ when x duplicates a_{j1}. Similarly, we will use $x = a_{in_i}$ when x duplicates a_{in_i}.

Lemma 4.22 *Let a sequence of linked chains $C_1, \ldots, C_k$ as in (4.14) be given. Let $a = a_{in_i}$, $i \neq k$, be an outer element with respect to the sequence of linked chains which does not duplicate a_{j1} for any j, and let $a' = a_{j1}$, $j \neq 1$, be an outer element with respect to the sequence of linked chains which does not duplicate a_{in_i} for any i. Let a_{st} be the minimal element in the indifference interval $I(a)$, and let $a_{s't'}$ be the maximal element in the indifference interval $I(a')$. Assume that $a_{st} \neq a_{kn_k}$ and $a_{s't'} \neq a_{11}$. Then,*

(a) $a \succ a_{s(t+1)}$,

(b) $a_{s'(t-1)} \succ a'$.

Proof: We prove (a). The proof of (b) is analogous, with a' replacing a, $s't'$ replacing st, switching the roles of $\prec$ and $\succ$ and switching the roles of $>$ and $<$.

Let $a = a_{in_i}$ and a_{st} be as in the statement of the lemma. Note that $a_{st} \neq a_{11}$ since by Lemma 4.15, $a_{11} \succ a_{in_i} = a$ and thus $a_{11} \notin I(a)$. Also, by assumption $a_{st} \neq a_{kn_k}$, so a_{st} is an inner element in the sequence of linked chains, i.e., $t < n_s$ and $a_{s(t+1)}$ exists. If $s \le i$, by Lemma 4.15 with $t < n_s$, we have $a_{st} \succ a_{in_i} = a$, a contradiction. So, $i < s$.

Clearly, $a \succeq a_{s(t+1)}$, since if $a_{s(t+1)} \succ a_{in_i} = a$, then by transitivity of $\succ$, and since $a_{st} \succ a_{s(t+1)}$, we have $a_{st} \succ a$, contradicting $a_{st} \in I(a)$. If $t+1 < n_s$, then $a_{s(t+1)}$ is an inner element, and $a \succ a_{s(t+1)}$ for otherwise $a \sim a_{s(t+1)}$, contradicting the minimality of a_{st} in $I(a)$. Thus, we may assume that $a_{s(t+1)}$ is not an inner element, i.e., that $t + 1 = n_s$.

Assume that $a \sim a_{sn_s} = a_{s(t+1)}$, and reach a contradiction. Let f be the smallest index among $s+1, s+2, \ldots, k$ such that $a \succ a_{f2}$. Such an f exists, because if a_{h2} is an inner element for some h among $s+1, s+2, \ldots, k$, then $a \succ a_{h2}$ since a_{st} is the minimal inner element in $I(a)$. If there is no such inner element, $a_{k2} = a_{kn_k}$ (since a_{k2} is not an inner element). But, $a_{kn_k} \notin I(a)$, since if it were, it would be the minimal element in $I(a)$ (by Lemma 4.15 and by the definition of $I(a)$). So $a \not\sim a_{k2}$. Then, since by Lemma 4.15, $a \succeq a_{k2} = a_{kn_k}$, it must be the case that $a \succ a_{k2}$. Thus, f exists.

If $f = s+1$, then $a \sim a_{sn_s} = a_{(f-1)n_{f-1}}$. Also, recall that $f-1 = s$ and $t+1 = n_s$, so $a_{(f-1)(n_{f-1}-1)} = a_{st}$ and $a \sim a_{st} = a_{(f-1)(n_{f-1}-1)}$. Now, we have $a \sim a_{(f-1)n_{f-1}}$, $a \sim a_{(f-1)(n_{f-1}-1)}$, and $a \succ a_{f2}$, so a duplicates a_{f1}, contradicting the assumption that a does not duplicate any a_{j1}.

Thus we may assume that $f > s+1$. In this case, $a \sim a_{(f-1)2}$ by the choice of f.

If $a_{(f-1)2}$ is an inner element, i.e., if $2 < n_{f-1}$, then since a_{st} is the minimal inner element in $I(a)$, and since $a_{st} \succ a_{(f-1)2}$ (by Lemma 4.15 with $s \leq f-1$), it is not the case that $a \sim a_{(f-1)2}$. Also, if $a_{(f-1)2} \succ a$, then by transitivity of $\succ$, $a_{st} \succ a$, contradicting $a_{st} \in I(a)$. So, $a \succ a_{(f-1)2}$, a contradiction.

So it must be the case that $n_{f-1} = 2$. By Lemma 4.15 with $i < s < f-1$, $a = a_{in_i} \succeq a_{(f-1)(n_{f-1}-1)} = a_{(f-1)1}$. It is not the case that $a \succ a_{(f-1)(n_{f-1}-1)} = a_{(f-1)1}$, since by transitivity of $\succ$ this would imply $a \succ a_{(f-1)2}$, a contradiction to the choice of f. So, $a \sim a_{(f-1)(n_{f-1}-1)} = a_{(f-1)1}$. Now, we have $a \sim a_{(f-1)2} = a_{(f-1)n_{f-1}}$, $a \sim a_{(f-1)(n_{f-1}-1)}$, and $a \succ a_{f2}$, so a duplicates a_{f1}, contradicting the assumption that a does not duplicate any a_{j1}. Thus we have reached a contradiction in the case $f > s+1$ also. So it must be that $a \succ a_{sn_s} = a_{s(t+1)}$. □

We now describe a set of intervals which will be shown to represent orders in $\mathcal{F}[\alpha, 0]$.

Definition 4.12 *Let $(A, \succ) \in \mathcal{F}[\alpha, 0]$. Let $C_1, \ldots, C_k$ as in (4.14) be the sequence of linked chains in a D-linked chain structure on $(A, \succ)$, and let indifference intervals for the outer elements in the sequence of linked chains be given as in (4.15). Define $J(A) : A \rightarrow \{[l, r] : l, r \in \mathbf{Z}\}$ (i.e., with l and r integral) as follows. In each case the sum will be considered to be 0 if upper limit is 0.*

$$J(a_0) = [0, \alpha + 1]. \tag{4.16}$$

$$J(a_{11}) = [\alpha + 1, \alpha + 1]. \tag{4.17}$$

$$J(a_{kn_k}) = [0, 0]. \tag{4.18}$$

For $1 < s < n_i$, define the point (length 0) intervals,

$$J(a_{is}) = [\alpha + 1 - \sum_{f=1}^{i-1}(n_f - 1) - (s - 1), \alpha + 1 - \sum_{f=1}^{i-1}(n_f - 1) - (s - 1)]. \tag{4.19}$$

For outer elements which appear as a_{in_i} $(i < k)$, such that the appearance as a_{in_i} is unduplicated and such the element has no appearance duplicating some a_{j1}, let a_{st} be the minimal element in indifference interval $I(a_{in_i})$. Then,

$$J(a_{in_i}) = [\alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t - 1), \alpha + 1 - \sum_{f=1}^{i}(n_f - 1)]. \tag{4.20}$$

For outer elements which appear as a_{j1} $(j > 1)$, such that the appearance as a_{j1} is unduplicated and such that the element has no appearance duplicating some a_{in_i}, let a_{st} be the maximal element in the indifference interval $I(a_{j1})$. Then,

$$J(a_{j1}) = [\alpha + 1 - \sum_{f=1}^{j-1}(n_f - 1), \alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t - 1)]. \tag{4.21}$$

For outer elements a which duplicate a_{j1} and a_{in_i}, for some $i < j$, let

$$J(a) = [\alpha + 1 - \sum_{f=1}^{j-1}(n_f - 1), \alpha + 1 - \sum_{f=1}^{i}(n_f - 1)]. \tag{4.22}$$

It is easy to check that the intervals given in the defintion of $J(A)$ do have integral endpoints and are indeed intervals (i.e., the right endpoint is greater than or equal to the left endpoint).

Note that the intervals given in (4.17), (4.18), and (4.19) all have length 0. Recall that by the defintion 4.9 of a D-linked chain structure, $\sum_{j=1}^{k}(n_j - 1) = \alpha + 1$ and that in a sequence of linked chains (4.14), $n_i \geq 2$ for $i = 1, \ldots, k$. Then, since $n_1 \geq 2$, the right endpoints of the intervals given in (4.20) and (4.22) are at most α and since $\sum_{j=1}^{k}(n_j - 1) = \alpha + 1$ the left endpoints of these intervals is at least 0. So the intervals in (4.20) and (4.22) have length at most α. Since $n_k \geq 2$, $j - 1 < k$, and $\sum_{j=1}^{k}(n_j - 1) = \alpha + 1$, the left endpoint of the intervals given in (4.21) is at least 1 and since the $n_i \geq 2$, the right endpoints of these intervals are at most $\alpha + 1$. So the intervals given in (4.21) have length at most α. Thus, the only interval in $J(A)$ that has length greater than α is $J(a_0)$, with length $\alpha + 1$.

We will show in Lemma 4.24 that if $(A, \succ) \in \mathcal{F}[\alpha, 0]$, then $J(A)$ is an interval representation of $(A, \succ)$. From the above comments, this representation is an $[\alpha, 0]$ discrete representation. We first prove a technical lemma.

Lemma 4.23 *For $i = 1, 2, \ldots, k$ let $n_i \geq 2$ be an integer. Also, let integers s, s', t, t' satisfy $1 \leq s, s' \leq k$, $1 \leq t \leq n_s$, and $1 \leq t' \leq n_{s'}$. Then, if either $t' > 1$ or if $t < n_s$,*

$$\sum_{f=1}^{s'-1}(n_f - 1) + (t' - 1) > \sum_{f=1}^{s-1}(n_f - 1) + (t - 1) \tag{4.23}$$

$$\Longleftrightarrow$$

$$s < s' \text{ or } [s = s' \text{ and } t < t']. \tag{4.24}$$

Proof: Assume that (4.24) holds. If $t = t'$ and $t < t'$, then clearly $\sum_{f=1}^{s'-1}(n_f - 1) + (t' - 1) > \sum_{f=1}^{s-1}(n_f - 1) + (t - 1)$. If $s < s'$,

$$\sum_{f=1}^{s-1}(n_f - 1) + (t - 1) \leq \sum_{f=1}^{s}(n_f - 1) \leq \sum_{f=1}^{s'-1}(n_f - 1) \leq \sum_{f=1}^{s'-1}(n_f - 1) + (t' - 1)$$

The first inequality folows since $t \leq n_s$. Since either $t < n_s$ or $t' > 1$, either the first or the third inequality is strict and (4.23) holds.

Conversely, if (4.24) fails, i.e., if $s' < s$ or if $[s = s'$ and $t' \leq t]$, we show that (4.23) fails. If $s = s'$ and $t' \leq t$ then clearly $\sum_{f=1}^{s'-1}(n_f - 1) + (t' - 1) \leq \sum_{f=1}^{s-1}(n_f - 1) + (t - 1)$

and (4.23) fails. If $s' < s$,

$$\sum_{f=1}^{s'-1}(n_f - 1) + (t' - 1) \le \sum_{f=1}^{s'}(n_f - 1) \le \sum_{f=1}^{s-1}(n_f - 1) \le \sum_{f=1}^{s-1}(n_f - 1) + (t - 1)$$

The first inequality hold since $t' \le n_{s'}$. Thus (4.23) fails. □

Lemma 4.24 *If $(A, \succ) \in \mathcal{F}[\alpha, 0]$, then $J(A)$ as in Definition 4.12 is an $[\alpha + 1, 0]$ bounded discrete representation of $(A, \succ)$.*

Proof: Let $(A, \succ) \in \mathcal{F}[\alpha, 0]$ and assume that $J(A)$ is given as in definition 4.12. We have noted above that if $J(A)$ represents $(A, \succ)$, then the representation is an $[\alpha, 0]$ discrete representation. So we must show that the intervals in $J(A)$ represent $(A, \succ)$.

For $a, a' \in A$, we must show that $a \succ a' \Leftrightarrow l_a > r_{a'}$. There are several cases to consider. In the presentation of the proof, we will make use of the fact that the left endpoints of intervals $J(a_{j1})$ can be written as in the formula for the left endpoints of intervals $J(a_{is})$ for inner elements a_{is}, $1 < s < n_s$, and the right endpoints of intervals $J(a_{jn_j})$ can be written as in the formula for right endpoints of intervals $J(a_{is})$ for inner elements a_{is}, $1 < i < n_s$.

Case 1: $a = a_{uv}$, $1 \le v \le n_u$ and $a' = a_{u'v'}$, $1 < v' < n'_{u'}$.

We first show that $a_{uv} \succ a_{u'v'} \Leftrightarrow u < u'$ or $[u = u'$ and $v < v']$. Since $v' \ne 1$ and $v \ne n_u$ then by Lemma 4.15, $u < u'$ or $[u = u'$ and $v < v'] \Rightarrow a_{uv} \succ a_{u'v'}$. Conversely, by Lemma 4.15 and by the definition of a chain, $u' > u$ or $[u = u'$ and $v' \ge v] \Rightarrow a_{u'v'} \succeq a_{uv}$ and thus not $a_{uv} \succ a_{u'v'}$. So, $a_{uv} \succ a_{u'v'} \Leftrightarrow u < u'$ or $[u = u'$ and $v < v']$.

We have $l_a = \alpha + 1 - \sum_{f=1}^{u-1}(n_f - 1) - (v - 1)$ and $r_{a'} = \alpha + 1 - \sum_{f=1}^{u'-1}(n_f - 1) - (v' - 1)$. Then,

$$l_a \quad > \quad r_{a'}$$

$$\Longleftrightarrow$$

$$\sum_{f=1}^{u'-1}(n_f - 1) + (v' - 1) \quad > \quad \sum_{f=1}^{u-1}(n_f - 1) + (v - 1)$$

$$\Longleftrightarrow$$

$$u < u' \text{ or } [u = u' \text{ and } v < v'].$$

The last $\Leftrightarrow$ follows from Lemma 4.23 since $v' > 1$. This completes the proof for case 1.

Case 2: $a = a_{un_u}$ for some u and $a' = a_{u'v'}$ for $1 < v' \leq n_{u'}$.

In this case, if $u' \leq u$, then, by Lemma 4.15, $a' \succ a$ and it is not the case that $a \succ a'$. Also, if $u' \leq u$, we have $r_{a'} = \alpha + 1 - \sum_{f=1}^{u'-1}(n_f - 1) - (v' - 1) \geq \alpha + 1 - \sum_{f=1}^{u}(n_f - 1) = r_a \geq l_a$, so it is not the case that $l_a > r_{a'}$. Thus, case 2 holds when $u' \leq u$.

If $u < u'$, note that $a \succeq a'$ by Lemma 4.15. If a duplicates a_{w1} for some w, then we are done by case 1 applied to $a = a_{w1}$ and a'. So we may assume that a does not duplicate any a_{w1}. Thus $l_a = \alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t - 1)$ where a_{st} is the minimal element in $I(a)$. Also, $r_{a'} = \alpha + 1 - \sum_{f=1}^{u'-1}(n_f - 1) - (v' - 1)$. Note that by Lemma 4.15, $a_{11} \succ a_{un_u}$ for all $1 \leq u \leq k$, so $a_{st} \neq a_{11}$. If $a' = a_{st}$, then $a' \sim a$ (since $a' = a_{st} \in I(a)$), and $l_a = r_{a'}$. So case 2 holds in this situation and we may assume that $a' \neq a_{st}$. We have,

$$\begin{array}{ccc} l_a & > & r_{a'} \\ & \Longleftrightarrow & \\ \sum_{f=1}^{u'-1}(n_f - 1) + (v' - 1) & > & \sum_{f=1}^{s-1}(n_f - 1) + (t - 1) \\ & \Longleftrightarrow & \end{array}$$

$$s < u' \text{ or } [s = u' \text{ and } t < v']. \tag{4.25}$$

The last $\Leftrightarrow$ follows from Lemma 4.23 since $v' > 1$. Thus we need to show that $a \succ a'$ if and only if (4.25) holds. There are two subcases.

Subcase i: $v' \neq n_{u'}$, i.e., a' is an inner element.

First assume that $a \succ a'$ and show that (4.25) holds. If $a \succ a'$, then $a' \notin I(a)$ by the definition (4.15) of $I(a)$. Either both a' and a_{st} are inner elements or a' is an inner element and $a_{st} = a_{kn_k}$. In either case, by Lemma 4.15, $a' \succ a_{st}$ or $a_{st} \succ a'$ (and not $a \sim a_{st}$). If $a' \succ a_{st}$, then by transitivity and since $a \succ a'$, we would have $a \succ a_{st}$, contradicting $a_{st} \in I(a)$. So $a_{st} \succ a' = a_{u'v'}$. Then, by Lemma 4.15, (4.25) holds.

Conversely, assume that (4.25) holds. Then $a_{st} \neq a_{kn_k}$ since $s < u$ or $t < v = n_u$. So a_{st} is an inner element and since both a' and a_{st} are inner elements (and $a' \neq a_{st}$), by Lemma 4.15, $a_{st} \succ a'$. Since a_{st} is the minimal element in $I(a)$, it is not the case that $a' \sim a$, and since $a \succeq a'$, $a \succ a'$.

Subcase ii: $v' = n_{u'}$. So $a' = a_{u'n_{u'}}$.

We first show that if (4.25) fails, then $a' \succeq a$. If (4.25) fails, then $u' < s$ or $[u' = s$ and $n_{u'} = v' \leq t]$. If $u' = s$ then also $t = n_{u'}$ (since t must be $\leq n_s = n_{u'}$). So then $a' = a_{u'n_{u'}} = a_{st} \sim a$ since $a_{st} \in I(a)$. Thus we may assume that $u' < v$. From Lemma 4.15, $a_{s1} \succ a_{st}$. (This uses the fact that either $a_{st} = a_{kn_k}$ or a_{st} is an inner element, so $t \neq 1$.) Assume that $a \succ a'$ and reach a contradiction. We have $a \succ a'$, $a_{s1} \succ a_{st}$, and $a \sim a_{st}$ (since $a_{st} \in I(a)$), so by the definition of an interval order, $a_{s1} \succ a' = a_{u'n_{u'}}$. But by Lemma 4.15 and since $u' < s$, $a' = a_{u'n_{u'}} \succeq a_{s1}$, a contradiction. Thus $a' \succ a$.

Conversely, we assume that (4.25) holds and show $a \succ a'$. If (4.25) holds, $a_{st} \neq a_{kn_k}$ since either $s < u'$ or $t < n_{u'}$. So, by Lemma 4.22, since we have assumed that a does not duplicate any a_{w1}, we have $a \succ a_{s(t+1)}$. Thus, if $t + 1 = n_s$, then $a \succ a_{sn_s}$. If $u' = s$, then $a \succ a_{sn_s} = a_{u'n_{u'}} = a'$ and we are done. So we may assume that $s < u'$. Also, if $t + 1 < n_s$, then by Lemma 4.15, $a_{s(t+1)} \succ a_{sn_s}$ and by transitivity of $\succ$, $a \succ a_{sn_s}$. Also $a_{u'1} \succ a_{u'n_{u'}} = a'$. Then, $a_{u'1} \succ a_{u'n_{u'}} = a'$ and $a \succ a_{sn_s}$ and by the definition of an interval order either $a \succ a'$ or $a_{u'1} \succ a_{sn_s}$. But, by Lemma 4.15, $a_{u'1} \succ a_{sn_s}$ can not hold since $s < u'$. Thus $a \succ a'$.

Case 3: $a = a_{uv}$ for $1 \leq v < n_u$ and $a' = a_{u'1}$ for some u'.

In this case, if $u' \leq u$, then by Lemma 4.15, $a' \succeq a$ and it is not the case that $a \succ a'$. Also, if $u' \leq u$, we have $r_{a'} \leq l_{a'} = \alpha+1-\sum_{f=1}^{u'-1}(n_f-1) \leq \alpha+1-\sum_{f=1}^{u-1}(n_f-1)-(v-1) = l_a$. So it is not the case that $l_a > r_{a'}$. Thus, case 3 holds when $u' \leq u$.

If $u < u'$, note that $a \succeq a'$ by Lemma 4.15. If a' duplicates a_{wn_w} for some w then we are done by case 1 applied to a and $a' = a_{wn_w}$. So, we may assume that a' does not duplicate a_{wn_w} for any w. Let $a_{s't'}$ be the maximal element in the indifference interval $I(a')$. Then we have $r_{a'} = \alpha + 1 - \sum_{f=1}^{s'-1}(n_f - 1) - (t' - 1)$. Also,

$l_a = \alpha + 1 - \sum_{f=1}^{u-1}(n_f - 1) - (v-1)$. Note that $a' \succ a_{kn_k}$ by 4.15, so $a_{s't'} \neq a_{kn_k}$. If $a = a_{s't'}$ then $a \sim a'$ (since $a = a_{s't'} \in I(a')$), and $l_a = r_{a'}$. So case 3 holds in this situation, and we may assume that $a \neq a_{s't'}$. We have

$$\begin{array}{ccc} l_a & > & r_{a'} \\ & \Longleftrightarrow & \\ \sum_{f=1}^{s'-1}(n_f - 1) + (t'-1) & > & \sum_{f=1}^{u-1}(n_f - 1) + (v-1) \\ & \Longleftrightarrow & \end{array}$$

$$u < s' \text{ or } [u = s' \text{ and } v < t']. \tag{4.26}$$

The last $\Leftrightarrow$ follows from Lemma 4.23 since $v < n_u$. Thus we need to show that $a \succ a'$ if and only if (4.26) holds. There are two subcases.

Subcase i: $v \neq 1$, i.e., a is an inner element.

First assume $a \succ a'$ and show that (4.26) holds. If $a \succ a'$ then $a \notin I(a')$. Note that a is an inner element and $a_{s't'}$ is a_{11} or an inner element, so, by Lemma 4.15, it is not the case that $a \sim a_{s't'}$ (since also $a \neq a_{s't'}$). If $a_{s't'} \succ a$, then by transitivity of $\succ$, $a_{s't'} \succ a'$, contradicting $a_{s't'} \in I(a')$. So, $a_{uv} = a \succ a_{s't'}$. Then, by Lemma 4.15, (4.26) holds.

Conversely, assume that (4.26) holds. Then $a = a_{uv} \succ a_{s't'}$ by Lemma 4.15 (since also $a \neq a_{s't'}$). If $a \sim a'$ then since a is an inner element, $a \in I(a')$, contradicting the maximality of $a_{s't'}$ in $I(a')$. Thus (since also $a \succeq a'$), $a \succ a'$.

Subcase ii: $v = 1$. So $a = a_{u1}$.

We first show that if (4.26) fails, then $a' \succeq a$. If (4.26) fails, then $u > s'$ or $[u = s'$ and $t' \leq v = 1]$. If $u = s'$ then also t' must be $= 1$. So then $a_{s't'} = a_{uv}$ and $a = a_{uv} = a_{s't'} \sim a'$ since $a_{s't'} \in I(a)$. Thus we may assume that $u > s'$. Also, by Lemma 4.15, $a_{s't'} \succ a_{s'n_{s'}}$ since either $a_{s't'}$ is an inner element, i.e., $t' < n_{s'}$ or $a_{s't'} = a_{11}$. Assume that $a \succ a'$ and reach a contradiction. We then have $a \succ a'$, $a_{s't'} \succ a_{s'n_{s'}}$, and $a' \sim a_{s't'}$, so, by the definition of an interval order, $a_{uv} = a \succ a_{s'n_{s'}}$. Then by Lemma 4.15, $u \leq s'$, a contradiction. So $a \sim a'$.

Conversely, assume that (4.26) holds and show that $a \succ a'$. Since $s' > u$ or $t' > v = 1$, $a_{s't'} \neq a_{11}$. Then by Lemma 4.22, since $a_{s't'} \neq a_{11}$, and since we have assumed that a' does not duplicate any a_{wn_w}, we have $a_{s'(t'-1)} \succ a'$. So $a_{s'1} \succ a'$. This follows immediately if $t'-1 = 1$ and by transitivity of $\succ$ if $t'-1 > 1$. Also, by Lemma 4.15 and since $u \leq s'$, $a_{un_u} \succeq a_{s'1}$. Note that $a = a_{u1} \succ a_{un_u}$. Then, if $a_{un_u} \succ a_{s'1}$, $a = a_{u1} \succ a_{un_u} \succ a_{s'1} \succ a'$, so by transitivity of $\succ$, $a \succ a'$. If $a_{un_u} \sim a_{s'1}$, then since also $a \succ a_{un_u}$ and $a_{s'1} \succ a'$, by the definition of an interval order $a \succ a'$.

Case 4: $a = a_{un_u}$ and $a' = a_{u'1}$.

In this case, if $u' \leq u$, then, by Lemma 4.15, $a' \succ a$ and it is not the case that $a \succ a'$. Also, if $u' \leq u$, we have $r_{a'} \geq l_{a'} = \alpha + 1 - \sum_{f=1}^{u'-1}(n_f - 1) > \alpha + 1 - \sum_{f=1}^{u}(n_f - 1) = r_a \geq l_a$, so it is not the case that $l_a > r_a$. Thus, case 4 holds when $u' \leq u$.

Finally, we consider $a = a_{un_u}$, $a' = a_{u'1}$, for $u < u'$. Since $u < u'$, by Lemma 4.15, $a \succeq a'$. If a duplicates a_{w1} for some w, we are done by case 3. If a' duplicates a_{wn_w} for some w, we are done by by case 2. Thus we may assume that a does not duplicate any a_{w1} and a' does not duplicate any a_{wn_w}. Let a_{st} be the minimal element in $I(a)$ and $a_{s't'}$ be the maximal element in $I(a')$. Then $l_a = \alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t-1)$ and $r_{a'} = \alpha + 1 - \sum_{f=1}^{s'-1}(n_f - 1) - (t'-1)$.

From the Definition 4.10 of indifference intervals, if $t' = 1$ then $a_{s't'} = a_{11}$, and if $t = n_s$, then $a_{st} = a_{kn_k}$. If $a_{s't'} = a_{11}$ and $a_{st} = a_{kn_k}$, then from the conditions for a D-linked chain structure, $l_a = 0$ and $r_{a'} = \alpha + 1$. So it is not the case that $l_a > r_{a'}$. Also, if $a \succ a'$, then, since $a' = a_{u'1} \succ a_{kn_k}$ by Lemma 4.15, we have by transitivity $a \succ a_{kn_k}$, contradicting the assumpton that $a_{kn_k} = a_{st} \in I(a)$. So it must be the case that $a \succ a'$ does not hold, and the proof is complete in the case that $t' = 1$ and $t = n_s$.

We may now asume that either $t' > 1$ or $t < n_s$. We have

$$l_a \quad > \quad r_{a'}$$

$$\Longleftrightarrow$$

$$\sum_{f=1}^{s'-1}(n_f - 1) + (t'-1) \quad > \quad \sum_{f=1}^{s-1}(n_f - 1) + (t-1)$$

$$\iff$$

$$s < s' \text{ or } [s = s' \text{ and } t < t']. \tag{4.27}$$

The last $\Leftrightarrow$ follows by Lemma 4.23 since either $t' > 1$ or $t < n_s$. Thus we need to show that $a \succ a'$ if and only if (4.27) holds.

We first show that if (4.27) fails, then $a' \succeq a$. Since $a \succeq a'$, we need to show $a \sim a'$. We have $l_a \leq r_{a'}$ if and only if $s > s'$ or $[s = s'$ and $t \geq t']$. Then by Lemma 4.15, $a_{s't'} \succ a_{st}$ or $a_{s't'} = a_{st}$. In the second case, $a_{st} \sim a$, $a_{st} \sim a'$ and thus by Lemma 4.19, $a \sim a'$. If $a_{s't'} \succ a_{st}$, then it is not the case that $a_{st} \succ a'$, since then, by transitivity of $\succ$, $a_{s't'} \succ a'$, contradicting $a_{s't'} \in I(a')$. If $a' \succ a_{st}$, then if $a \succ a'$, by transitivity of $\succ$, $a \succ a_{st}$, contradicting $a_{st} \in I(a)$. So if $a' \succ a_{st}$, then $a \sim a'$. Thus we may assume that $a' \sim a_{st}$. But then by Lemma 4.19 and since $a \sim a_{st}$, we have $a \sim a'$. This completes the proof that if (4.27) fails, then $a \sim a'$.

Conversely, assume that (4.27) holds and show that $a \succ a'$. By Lemma 4.15, if (4.27) holds, $a_{st} \succ a_{s't'}$. Also $a_{st} \neq a_{kn_k}$ since $s < s'$ or $t < t'$. Then, since we have assumed that a does not duplicate any a_{w1}, by lemma 4.22, we have $a \succ a_{s(t+1)}$. Clearly, $a_{st} \succeq a'$, since otherwise by transitivity of $\succ$, $a' \succ a_{s't'}$ contradicting $a_{s't'} \in I(a')$. Since a_{st} is an inner element, it is not the case that $a_{st} \sim a'$ because $a_{s't'}$ is the maximal element in $I(a')$. Thus $a_{st} \succ a'$. Then by Lemma 4.20 applied to $a = a_{un_u}$, $a' = a_{u'1}$, and the consecutive elements a_{st}, $a_{s(t+1)}$, $a \succ a'$. $\square$

We now can describe the structure of elements in $\mathcal{F}[\alpha, 0]$.

Theorem 4.25 *$(A, \succ) \in \mathcal{F}[\alpha, 0]$ if and only if there is a D-linked chain structure on $(A, \succ)$ with sequence of linked chains $C_1, \ldots, C_k$ as in (4.14) such that:*

(a) *Each outer element with respect to the sequence of linked chains has at least one unduplicated appearance.*

(b) *If a_{j1}, $j \neq 1$, and a_{in_i}, $i \neq k$ with $i < j$ are outer elements with respect to the sequence of linked chains, then $|I(a_{j1}) \cap I(a_{in_i})| \neq 1$.*

Proof: Assume that $(A, \succ) \in \mathcal{F}[\alpha, 0]$. By Lemma 4.18, there is a D-linked chain structure on $(A, \succ)$. By Lemma 4.21, condition (b) holds. If each appearance of an outer element a is duplicated, then replacing a with the duplicating elements produces a D-linked chain structure on $(A \setminus \{a\}, \succ)$. As in Remark 4.7, there is a negative cycle in $D(A \setminus \{a\}, \succ, \alpha, 0)$ corresponding to this D-linked chain structure. By Theorem 4.3, $(A \setminus \{a\}, \succ)$ has no $[\alpha, 0]$ discrete representation. This contradicts the minimality of $(A, \succ)$. Thus condition (a) must hold.

Conversely, let $(A, \succ)$ be an interval order satisfying the conditions. By Lemma 4.24, the intervals $J(A)$ defined in Definition 4.12 represent $(A, \succ)$. We must show that removing any element $a \in A$ results in an order which has an $[\alpha, 0]$ discrete representation. We will describe an $[\alpha, 0]$ discrete representation $J^*(A \setminus \{a\})$ of $(A \setminus \{a\}, \succ)$ based on the representation $J(A)$. Since $(A \setminus \{a\}, \succ)$ is a suborder of $(A, \succ)$ the structure contained in $J(A)$ will insure that J^* represents $(A \setminus \{a\}, \succ)$. Recall that the only interval in $J(A)$ with length greater than α is $J(a_0)$.

We consider the cases of removing different types of elements.

Case 1: remove a_0.

The only interval in J longer than α is $J(a_0)$, so J with this interval removed is an $[\alpha, 0]$ representation. That is, $J^*(b) = J(b)$ for $b \in A \setminus \{a_0\}$.

Case 2: remove an inner element a_{st}.

By condition (b), a_{st} can not be both a maximal element in some indifference interval $I(a_{j1})$ and a minimal element in some other indifference interval $I(a_{in_i})$.

Subcase i: a_{st} is not a minimal element in any $I(a_{in_i})$.

In this case, there is no interval $J(a_{in_i})$ with $p = \alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t - 1)$ as the left endpoint. (Note that p corresponds to the right and left endpoints of $J(a_{st})$.) This follows, since the only outer elements a_{in_i} which can have p as a left endpoint have a_{st} as a minimal element in $I(a_{in_i})$. Also, it is is easy to see that the left endpoints of outer elements $J(a_{i1})$, the left endpoints of inner elements (other than a_{st}) and the left endpoint of $J(a_{kn_k})$ can not be p. Define J^* by removing $J(a_{st})$ from $J(A)$ and shifting every endpoint greater than or equal to p one unit to the left. That is, for x either l or r, define J^* as follows (for $a \in A \setminus \{a_{st}\}$):

$$x_a^* = \begin{cases} x_a - 1 & \text{if } x_a \geq p \\ x_a & \text{if } x_a < p. \end{cases}$$

Note that the length $J^*(a_0)$ is α and the length of every other interval is the same or reduced by one (and no right endpoint is moved to the left of a right endpoint). From the definition of J^*, the only way for $l_a > r_{a'} \Leftrightarrow l_a^* > r_{a'}^*$ to be violated is if $l_a = p$ and $r_{a'} = p - 1$. However, we have already noted that since a_{st} is not the minimal element in any indifference interval, there is no $a \in A \setminus \{a_{st}\}$ with $l_a = p$. So $l_a > r_{a'} \Leftrightarrow l_a^* > r_{a'}^*$ holds and J^* is a representation of $(A \setminus \{a_{st}\}, \succ)$.

Subcase ii: a_{st} is not a maximal element in any $I(a_{j1})$.

This case is handled in a manner 'symmetric' to subcase (i). In a manner similar to subcase (i), we see that since there is no element a_{j1} with a_{st} as the maximal element in $I(a_{j1})$, there is no $a \in A \setminus \{a_{st}\}$ such that the right endpoint of $J(a)$ is $p = \alpha + 1 - \sum_{f=1}^{s-1}(n_f - 1) - (t - 1)$. Then we define J^* on $a \in A \setminus \{a_{st}\}$ as follows:

$$x_a^* = \begin{cases} x_a & \text{if } x_a > p \\ x_a + 1 & \text{if } x_a \leq p. \end{cases}$$

Once again it is not difficult to see that $l_a > r_{a'} \Leftrightarrow l_a^* > r_{a'}^*$ since there is no $a' \in A \setminus \{a_{st}\}$ with $r_{a'} = p$.

Case 3: remove an outer element a_{j1} or a_{in_i}.

Each outer element has at least one unduplicated end by condition (a). If an outer element has unduplicated appearances as both a_{j1} for some j and a_{in_i} for some i, then either subcase (i) or (ii) will suffice.

Subcase i: $a = a_{j1}$ (and there is no x which duplicates a_{j1}).

In this case, obtain J^* by removing the point $p = \alpha + 1 - \sum_{f=1}^{j-1}(n_f - 1)$ which is the left endpoint of the interval $J(a_{j1})$ and shift all endpoints greater than or equal to p

one unit to the left. That is, for x either l or r, define J^* on $A \setminus \{a_{j1}\}$ by

$$x_a^* = \begin{cases} x_a - 1 & \text{if } x_a \geq p \\ x_a & \text{if } x_a < p. \end{cases}$$

If $j \neq 1$, by the assumption that a_{j1} is unduplicated, there are no left endpoints equal to p and if $j = 1$, $p = \alpha + 1$, and by construction there are no left endpoints equal to p. Then, as in subcase (i) of case 2, J^* is an $[\alpha, 0]$ representation of $(A \setminus \{a_{j1}\}, \succ)$.

Subcase ii: $a = a_{in_i}$ (and there is no x which duplicates a_{in_i}).

In a similar manner, remove the point $p = \alpha + 1 - \sum_{f=1}^{i}(n_f - 1)$ which is the right endpoint of the interval $J(a_{in_i})$ and shift all endpoints less than or equal to p one unit to the right.

$$x_a^* = \begin{cases} x_a & \text{if } x_a > p \\ x_a + 1 & \text{if } x_a \leq p. \end{cases}$$

If $i \neq k$, by the assumption that a_{in_i} is unduplicated, there are no right endpoints equal to p and if $i = k$, $p = 0$, and there are no right endpoints equal to p. Then, as in subcase (ii) of case 2, J^* is an $[\alpha, 0]$ representation of $(A \setminus \{a_{in_i}\}, \succ)$. □

Remark 4.8 In the proof of Theorem 4.25 we have given $[\alpha + 1, 0]$ bounded discrete representations of each order in $\mathcal{F}[\alpha, 0]$. The D-linked chain structures representing orders in $\mathcal{F}[\alpha, 0]$ show that the conditions in Theorem 4.10 are independent, since for each condition there is a corresponding D-linked chain structure and order in $\mathcal{F}[\alpha, 0]$ completing that structure.

4.7 Minimal Forbidden Orders — Non-Degenerate Case

In this section we examine the family $\mathcal{F}[\alpha, 1]$ of minimal orders having no $[\alpha, 1]$ discrete representation. The structure of these orders will be similar to that described in Theorem 4.13 and shown in Figure 4.4. We will replace the linear order of a_i elements with a sequence of linked chains of the proper size. We will make use of the notation, definitions, and lemmas regarding sequences of linked chains from Section 4.6. The reductions on cycles in $D(A,\succ,\alpha,1)$ are similar to those on $D(A,\succ,\alpha,0)$ in Section 4.6 except that in the non-degenerate case arcs from V have non-zero length of -1.

Definition 4.13 *An interval order* $(A,\succ)$ *has an* **N-linked chain structure** *if there is a set of elements* $B=\{b_1,\ldots,b_\gamma\}$ *such that*

(1) $b_1 \sim b_2 \sim \cdots \sim b_\gamma$,

(2) $b_i \succ b_j$ *if* $j \geq i+2$.

(3) The elements of $A\setminus B$ *form a sequence of linked chains* $C_1,\ldots,C_k$ *such that*

$$\sum_{j=1}^{k}(2n_j-3)=\gamma\alpha+1$$

where n_j *denotes the number of elements in* C_j.

(4) These linked chains can be further decomposed as follows. For $s\neq 1$, *let*

$$\bar{S}(a_{is})=\left(\sum_{j=1}^{i-1}(2n_j-3)\right)+(2s-3)$$

and let

$$\bar{S}(a_{i1})=\sum_{j=1}^{i-1}(2n_j-3).$$

Let

$$A^j=\{a_{is}:(j-1)\alpha\leq \bar{S}(a_{is})\leq j\alpha\}.$$

Then, $A\setminus B=\cup_{j=1}^{\gamma}A^j\cup\{a_{kn_k}\}$ *(where the* A^i *are not necessarily disjoint),* $b_j\sim a\in A^j$ *for* $j=1,\ldots\gamma$ *and* $b_\gamma\sim a_{kn_k}$.

See Figure 4.6 for an example of an N-linked chain structure. Note that in the digraph corresponding to an N-linked chain structure, there is a cycle with exactly one maximal UZ-Path and with length -1 exactly like those with length -1 described in Lemma 4.8. The length -2 cycles in this lemma are described in the next definition. We will refer to the structure and the corresponding cycle interchangeably.

Recall Definition 4.6 of the bi-minimal order for α when α is odd. We now define a corresponding cyle.

indicates the interval of elements incomparable to b_i.

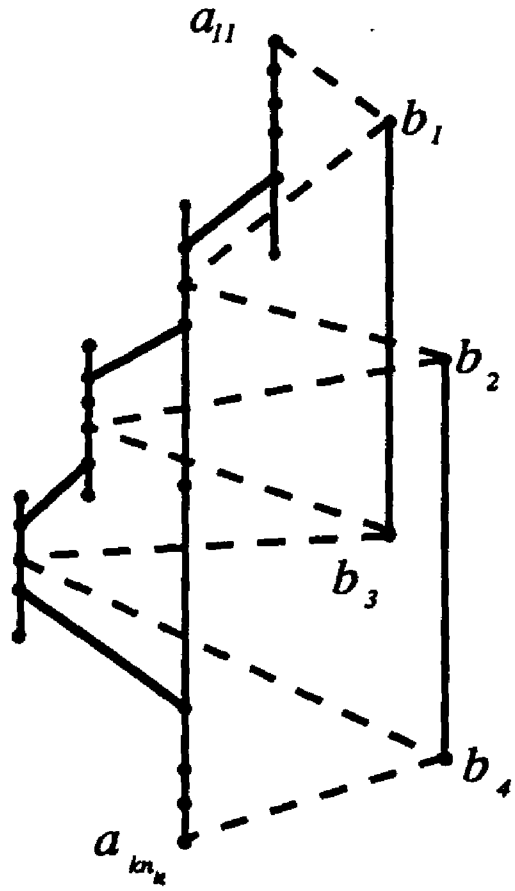


Figure 4.6: An N-linked chain structure.

Definition 4.14 *Given α odd, $\alpha \geq 3$, an α* **bi-minimal cycle** *is a cycle $C = l_{a_0}, r_{a_0}, P, l_{a_0}$ where P is a WV-Path containing $\alpha + 3$ vertices.*

Note that the length of an α bi-minimal cycle in $D(A, \succ, \alpha, 1)$ is -2. (This follows since a WV-Path containing $\alpha + 3$ vertices in has length $-\alpha - 2$ in $D(A, \succ, \alpha, 1)$, the arc (l_{a_0}, r_{a_0}) has length α and the arcs connecting r_{a_0} to P and P to l_{a_0} in C are from Z and have length 0.) Also, it can be shown that an α bi-minimal cycle such a cycle corresponds to the structure of a bi-minimal order for α for α odd.

Lemma 4.26 *Let C be a cycle in $D(A, \succ, \alpha, 1)$ with exactly one maximal UZ-Path. If $length(C) < -1$ and if C is not an α bi-minimal cycle, then $(A, \succ) \notin \mathcal{F}[\alpha, 1]$.*

Proof: Let $C = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(v)}, l_{\pi(2\gamma)}$ with the maximal UZ-Path $P_0 = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots l_{\pi(2)}, r_{\pi(1)}$. Then $P = l_{\sigma(1)}, \ldots, r_{\sigma(v)}$ consists of WV-Paths joined by arcs from Z and contains no arcs from U. Then, $length(C) = length(P_0) + length(P)$ since the arcs $(r_{\pi(1)}, l_{\sigma(1)})$ and $(r_{\sigma(v)}, l_{\pi(2\gamma)})$ must be in Z by the maximality of P_0. Analogous to $S(i)$ defined in (4.4) in Lemma 4.7, define $S'(i)$ to be the sum of the arcs along P from $l_{\sigma(1)}$ to $x_{\sigma(i)}$ where x can be either l or r.

If P contains an arc from Z, choose w so that $(r_{\sigma(w)}, l_{\sigma(w+1)}) \in Z$ is the first such arc in P, i.e., there is no w' such that $(r_{\sigma(w')}, l_{\sigma(w'+1)}) \in Z$. Note that from the definition of P, we have $1 < w < w + 1 < v$. Then $P' = l_{\sigma(w-1)}, r_{\sigma(w)}, l_{\sigma(w+1)}, r_{\sigma(w+2)}$ appears in P. Since P contains no arcs from U, $(l_{\sigma(w-1)}, r_{\sigma(w)})$ and $(l_{\sigma(w+1)}, r_{\sigma(w+2)})$ are in W. So $length(P') = -2$. Replace P' in C with the arc $(l_{\sigma(w-1)}, r_{\sigma(w+2)})$ to form the cycle C'. Then $length(C') = length(C) - length(P') + length(l_{\sigma(w-1)}, r_{\sigma(w+2)}) = length(C) + 1 \leq -1$. So C' is a negative cycle. The vertex $l_{\sigma(w+1)}$ does not appear in C', since it was removed with P'. The vertex $r_{\sigma(w+1)}$ can not appear in the UZ-Path P_0 since then $l_{\sigma(w+1)}$ appears in C in both P_0 and P, contradicting the definition of a cycle. If $r_{\sigma(w+1)}$ appears as $r_{\sigma(u)}$ for some $u \in \{1, \ldots, v\}$, then it must be that $u < w + 1$, since otherwise Lemma 4.5(a) applied to $l_{\sigma(w+1)}, \ldots, r_{\sigma(u)}$ produces the contradiction $\sigma(w + 1) \succ \sigma(u) = \sigma(w + 1)$. The arc $(r_{\sigma(w)}, l_{\sigma(w+1)})$ is in Z, so $u \neq w$. But then, consider the arc $(r_{\sigma(u)}, l_{\sigma(u+1)})$ in P. By the minimality of w, $(r_{\sigma(u)}, l_{\sigma(u+1)}) \in V$. So $\sigma(u+1) = \sigma(w+1)$ and $l_{\sigma(w+1)}$ appears twice in C, contradicting the definition of a cycle.

Thus, we may assume that P contains no arcs from Z, i.e., it is a WV-Path. Then, since C is not an α bi-minimal cycle, either $\gamma \geq 2$ or $v \neq \alpha + 3$ with α odd. Note also that since P is a WV-Path, v is even and $S'(v) = \mathit{length}(P) = -(v-1)$.

Consider the case $\gamma = 1$ and $v \neq \alpha + 3$ with α odd. In this case, P_0 consists of exactly one arc from U with length α. So, $\mathit{length}(C) = \mathit{length}(P_0) + \mathit{length}(P) = \alpha - v + 1$. Since $\mathit{length}(C) \leq -2$, and $v \neq \alpha + 3$, we have $\alpha - v + 1 \leq -2 \Rightarrow \alpha + 4 \leq v$. Then, since α is odd and v is even, we have $\alpha + 5 \leq v$. Let $C' = l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(v-3)}, l_{\pi(2)}$. We have $\mathit{length}(C') = \mathit{length}(l_{\pi(2)}, r_{\pi(1)}) + \mathit{length}(l_{\sigma(1)}, \ldots, r_{\sigma(v-3)})$ since the remaining arcs are from Z. So, $\mathit{length}(C') = \alpha + S'(v-3) = \alpha + -(v-4)$ (since P is a WV-Path) and thus $\mathit{length}(C') \leq \alpha - v + 4 \leq -1$. Also, C' contains no vertex corresponding to $\sigma(v)$. To see this, note that $r_{\sigma(v)}$ does not appear in C', since if it did, it would appear as $r_{\pi(1)}$ or as $r_{\sigma(w)}$ for $w \leq v-3$. In both cases, $r_{\sigma(v)}$ would appear twice in C, contradicting the assumption that C is a cycle. Since $\pi(1) = \pi(2)$, and $\pi(1) \neq \sigma(v)$, $l_{\pi(2)} \neq l_{\sigma(v)}$. So, if $l_{\sigma(v)}$ appears in C' it is as $l_{\sigma(w)}$ for $w \leq v-3$. Then by Lemma 4.5(a) applied to P we get $\sigma(v) = \sigma(w) \succ \sigma(v)$, a contradiction. Thus, C' is a negative cycle showing that $(A \setminus \{\sigma(v)\}, \succ)$ has no $[\alpha, 1]$ discrete representation, contradicting the minimality of $(A, \succ)$.

Thus, we may assume that $\gamma \geq 2$. Now, $\mathit{length}(P) = \mathit{length}(C) - \mathit{length}(P_0) \leq -\alpha - 2$ since P_0 has at least one arc from U and all arcs non-negative, so its length is at least α, and by assumption, $\mathit{length}(C) \leq -2$. Since P is a WV-Path, $S'(t) = -(t-1)$. The last vertex $r_{\sigma(v)}$ in P is an r vertex, and since $\mathit{length}(P) \leq \alpha - 2$, $S'(v) \leq -\alpha - 2$. Thus, since $S'(i)$ is non-decreasing in i, we can choose a smallest subscript $\sigma(t)$ of an r vertex such that $S'(t) < -\alpha$.

Since S' decreases by at most 1 for each arc and since D is bipartite, $S'(t) = -\alpha - 1$ or $-\alpha - 2$. By Lemma 4.5(c), $\pi(2) = \pi(1) \succeq \sigma(t)$.

If $\pi(2) \succ \sigma(t)$, then $(l_{\pi(2)}, r_{\sigma(t)}) \in W$. Thus, in D, there is a cycle $C' = l_{\pi(2\gamma)}, r_{\pi(\gamma-1)}, \ldots, l_{\pi(2)}, r_{\sigma(t)}, \ldots, r_{\sigma(v)}, l_{\pi(2\gamma)}$. C' is obtained by replacing the path $l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}$ with the arc $(l_{\pi(2)}, r_{\sigma(t)})$. The length of the replaced path is $\alpha + S'(t) = -1$ or -2, and the new arc has length -1. So the length of C' is at most the length of C plus 1. Thus C' is negative. Note that C' does not contain $l_{\sigma(1)}$. Furthermore, since $l_{\sigma(1)}$ is the first vertex in P, there is no r vertex in P corresponding to $\sigma(v)$. If there was, Lemma 4.5(a) would then imply $\sigma(v) \succ \sigma(v)$, a contradiction. Thus C' is a negative cycle containing no vertex corresponding to $\sigma(1)$. So $(A, \succ) \notin \mathcal{F}[\alpha, 1]$ if $\pi(2) \succ \sigma(t)$.

If $\pi(2) \sim \sigma(t)$, then $(r_{\sigma(t)}, l_{\pi(2)}) \in Z$ and $C' = l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}, l_{\pi(2)}$ is a cycle with length $\alpha + S'(t) < 0$ and exactly one arc from U. Since $\gamma \geq 2$, $\pi(4) = \pi(3) \in A$ and $l_{\pi(4)}$ and $r_{\pi(4)} = r_{\pi(3)}$ are not in C'. Thus $(A, \succ) \notin \mathcal{F}[\alpha, 1]$ in this case. □

Lemma 4.27 *If $(A, \succ) \in \mathcal{F}[\alpha, 1]$, then either $(A, \succ)$ is the bi-minimal order with respect to α, or $(A, \succ)$ has an N-linked chain structure.*

Proof: If $(A, \succ) \in \mathcal{F}[\alpha, 1]$, then by Lemma 4.8, the corresponding digraph D contains a negative cycle with exactly one maximal UZ-Path. As in Remark 4.5, a cycle containing exactly one maximal UZ-Path can be written as $P_0, P_1, \ldots, P_k$ where P_0 is a UZ-Path and P_i for $i = 1, \ldots, k$ is a maximal WV-Path. The last vertex of P_i is joined to the first vertex of P_{i+1} by an arc from Z, and the last vertex of P_k is joined to the first vertex of P_0 by an arc from Z. Pick a negative cycle in D containing exactly one maximal UZ-Path such that the number γ of arcs from U in P_0 is minimum and subject to that, such that the number k of maximal WV-Paths is minimum. Denote this cycle by $C = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots, l_{\pi(2)}, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(v)}, l_{\pi(2\gamma)}$ with the maximal UZ-Path $P_0 = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots, l_{\pi(2)}, r_{\pi(1)}$.

If $\gamma = 1$ and $k = 1$, then $C = l_{\pi(2)}, r_{\pi(1)}, P_1, l_{\pi(2)}$ where P_1 is a WV-Path and $(l_{\pi(2)}, r_{\pi(1)}) \in U$ with length α. So $\pi(2) = \pi(1)$. The WV-Path $P_1 = l_{\sigma(1)}, r_{\sigma(2)}, \ldots, r_{\sigma(v)}$ has $\frac{v}{2}$ arcs from W and $\frac{v}{2}-1$ arcs from V, for a length of $-v+1$. The arcs $(r_{\pi(1)}, l_{\sigma(1)})$ and $(r_{\sigma(v)}, l_{\pi(2)})$ are in Z with length 0. So $length(C) = length(l_{\pi(2)}, r_{\pi(1)}) + length(P_1) = \alpha - v + 1$, and since $length(C) = -2$, we have $v = \alpha + 3$. The elements corresponding to the WV-Path form a chain $C_1 = \sigma(1) \succ^{\frac{v}{2}} \sigma(v)$ containing $\frac{v+2}{2} = \frac{\alpha+5}{2}$ (distinct) elements. Also, since P_1 is a WV-Path there is a vertex $r_{\sigma(2t)}$ in P_1 corresponding to each of the elements except $\sigma(1)$ in the chain. By Lemma 4.5(c), $\pi(2) = \pi(1) \succeq \sigma(i)$ for $i = 1, \ldots, v$. If $\pi(1) \succ \sigma(2t)$ for some t, then $l_{\pi(2)}, r_{\sigma(2t)}, l_{\sigma(2t+1)}, \ldots, r_{\sigma(v)}, l_{\pi(2)}$ is a cycle with no arcs from U contradicting Corollary 4.6. So $\pi(2) \sim \sigma(2t)$. Also, since $(r_{\pi(1)}, l_{\sigma(1)}) \in Z$, $\pi(2) = \pi(1) \sim \sigma(1)$. Thus $\pi(2)$ is $\sim$ to every element in C_1. So the chain C_1 together with $\pi(2)$ are the bi-minimal order with respect to α. Since this order is contained in $(A, \succ)$ and since $(A, \succ) \in \mathcal{F}[\alpha, 1]$, $(A, \succ)$ is the bi-minimal order.

We may now assume that $\gamma > 1$, or $k > 1$. Since P_0 is a UZ-Path, $(l_{\pi(2i)}, r_{\pi(2i-1)}) \in U$ and $\pi(2i) = \pi(2i-1)$. Also, $(r_{\pi(2i-1)}, l_{\pi(2i-2)}) \in Z$ so $\pi(2i-1) \sim \pi(2i-2)$. For $i = 1, \ldots, \gamma$, let $b_i = \pi(2i) = \pi(2i-1)$. Then $b_i \sim b_{i-1}$. So $b_1 \sim b_2 \sim \cdots \sim b_\gamma$, which proves part *(1)* of Definition 4.13.

We next show by contradiction that $b_i \succ b_j$ for $j \geq i+2$. Assume first that $b_j \sim b_i$ for $j \geq i+2$. Then $(r_{\pi(2j-1)}, l_{\pi(2i)}) \in Z$. Replacing the positive length path $r_{\pi(2j-1)}, \ldots, l_{\pi(2i)}$ with the arc $(r_{\pi(2j-1)}, l_{\pi(2i)})$ having length 0 creates a new negative cycle with one maximal UZ-Path containing fewer arcs from U. The new cycle is $P_0', P_1, \ldots, P_k, l_{\pi(2\gamma)}$ where

$P_0' = l_{\pi(2\gamma)}, r_{\pi(2\gamma-1)}, \ldots, l_{\pi(2j)}, r_{\pi(2j-1)}, l_{\pi(2i)}, r_{\pi(2i-1)}, \ldots, l_{\pi(2)}, r_{\pi(1)}$. P_0' is a UZ-Path and it is maximal since there are no arcs from U among $P_1, \ldots, P_k$. This contradicts the choice of C. Next assume that for some i there is a $j \geq i+2$ such that $b_j \succ b_i$. Let j' be the largest index $\geq i+2$ such that $b_{j'} \succ b_i$. If $j' \neq \gamma$, then $b_{j'+1}$ exists and $b_i \succeq b_{j'+1}$. We have just shown that $b_i \sim b_{j'+1}$ can not hold so $b_i \succ b_{j'+1}$. But then by transitivity of $\succ$, $b_{j'} \succ b_{j'+1}$, contradicting $b_{j'} \sim b_{j'+1}$. So $j' = \gamma$. Then $(l_{\pi(2\gamma)}, r_{\pi(2i-1)}) \in W$. Replace the positive length path $l_{\pi(2\gamma)}, \ldots, r_{\pi(2i-1)}$ in C with the arc $(l_{\pi(2\gamma)}, r_{\pi(2i-1)})$ having length -1. This creates a new negative cycle $C' = l_{\pi(2\gamma)}, r_{\pi(2i-1)}, P_0', P_1, \ldots, P_k, l_{\pi(2\gamma)}$, where $P_0' = l_{\pi(2(i-1))}, r_{\pi(2(i-1)-1)}, \ldots, l_{\pi(2)}, r_{\pi(1)}$. P_0' is a UZ-Path and it is maximal since $P_1, \ldots, P_k, l_{\pi(2\gamma)}, r_{\pi(2i-1)}$ contains no arcs from U. Thus C' is a negative cycle with one maximal UZ-Path that contains fewer arcs from U than C, contradicting the choice of C. Thus $b_i \succ b_j$ for $j \geq i+2$ and we have shown that *(2)* in Definition 4.13 holds.

Since $(A, \succ) \in \mathcal{F}[\alpha, 1]$ and since C is a negative cycle, by Lemma 4.26, either C has length -1 or C is an α bi-minimal cycle, in which case $\gamma = 1$ and $k = 1$. The case $\gamma = 1$ and $k = 1$ has been accounted for, so C has length -1. P_0 has length $\gamma\alpha$. Thus $P_1, \ldots, P_k$ has length $-\gamma\alpha - 1$. If P_i contains η_i arcs from W, it corresponds to a chain $a_{i1} \succ^{\eta_i} a_{in_i}$ with $n_i = \eta_i + 1$ elements. Also, P_i has length $-(2\eta_i - 1)$. Thus, $\sum(2\eta_i - 1) = \alpha\gamma + 1$. So $\sum(2n_i - 3) = \alpha\gamma + 1$. This is part of the proof of *(3)* of Definition 4.13.

The proof that the end of chain C_i is linked to the beginning of chain C_{i+1} is almost identical to that in Lemma 4.18. The last vertex of path P_i is joined in C to the first vertex of P_{i+1} by an arc from Z. Denote P_i by
$P_i = l_{a_{i1}}, r_{a_{i2}}, l_{a_{i2}}, r_{a_{i3}}, l_{a_{i3}}, \ldots, r_{a_{i(n_i-1)}}, l_{a_{i(n_i-1)}}, r_{a_{in_i}}$. Then there is a subpath $P' = l_{a_{i(n_i-1)}}, r_{a_{in_i}}, l_{a_{(i+1)1}}, r_{a_{(i+1)2}}$ in C. For simplicity we will denote this by $P' = l_a, r_b, l_c, r_d$. Note that P' has length -2. To show that C_i is linked to C_{i+1} we need to show that $a \sim c$, $a \succ d$, $b \sim c$, and $b \sim d$. Since $(r_b, l_c) \in Z$, $b \sim c$. By Lemma 4.5(a), $a \succ d$. Also by Lemma 4.5, $a \succeq c$ and $b \succeq d$. If $a \succ c$, then replace P' by l_a, r_c, l_c, r_d with length -3. If $b \succ d$ then replace P' by l_a, r_b, l_b, r_d with length -3. The new vertices r_c in the first case and l_b in the second case can not repeat other vertices in C. If, say in the first case, r_c appears elsewhere in C, there is a cycle containing no arcs from U, contradicting Corollary 4.6 if r_c is on P_i for $i \geq 1$. Clearly r_c is not on P_0 since then l_c is also on P_0 and l_c appears twice in C, contradicting the definition of a cycle. Similarly, l_b can not repeat other vertices on C. Thus, in either case, we have found a

new negative cycle with one maximal UZ-Path, the same number of arcs from U and one less maximal WV-Path, contradicting the choice of C with the minimum number of such paths. Thus $a \sim c$ and $b \sim d$, completing the proof that the C_i's are linked. This completes the proof of part *(3)* of Definition 4.13

We now verify part *(4)* of Definition 4.13. Since $(r_{\sigma(v)}, l_{\pi(2\gamma)}) \in Z$, $b_\gamma = \pi(2\gamma) \sim \sigma(v) = a_{kn_k}$.

Consider the case $\gamma = 1$. As in the proof of Lemma 4.18, $\pi(2) = \pi(1) \sim \sigma(i)$ for $i = 1, \ldots v$. This holds since $\pi(1) \succeq \sigma(i)$ by Lemma 4.5(c), and if $\pi(2) \succ \sigma(i)$ then $l_{\pi(2)}, r_{\sigma(i)}, l_{\sigma(i+1)}, \ldots, r_{\sigma(v)}, l_{\pi(2)}$ (if i is even) and $l_{\pi(2)}, r_{\sigma(i)}, l_{\sigma(i)}, r_{\sigma(i+1)}, \ldots, r_{\sigma(v)}, l_{\pi(2)}$ (if i is odd) are cycles with no arc from U, contradicting Corollary 4.6. Since $\gamma = 1$, C is $l_{b_1}, r_{b_1}, l_{\sigma(1)}, \ldots, r_{\sigma(v)}, l_{b_1}$. If any $a \in A \setminus \{b_1\}$ is not equal to some $\sigma(i)$, then C is a negative cycle in $D(A \setminus \{\sigma(i)\}, \succ, \alpha, 1)$. So by Theorem 4.3, $(A \setminus \{\sigma(i)\}, \succ) \notin \mathcal{D}[\alpha, 1]$, contradicting $(A, \succ) \in \mathcal{F}[\alpha, 1]$. Thus, $b_1 \sim a \in A$, and $b_1 \sim a \in A^1 \subset A$.

Consider A^j in the case $\gamma \geq 2$. By an argument like the one just made, every element in $A \setminus B$ appears as some $\sigma(i)$. Thus, every element of $A \setminus B$ appears as some a_{is} in the sequence of linked chains. $\bar{S}$ is non-negative since the n_j in a sequence of linked chains are ≥ 2. The maximum $\bar{S}(a)$ for $a \neq a_{kn_k}$ is $\bar{S}(a_{k(n_k-1)}) = \sum_{j=1}^{k-1}(2n_j - 3) + (2(n_k - 1) - 3) = \sum_{j=1}^{k}(2n_j - 3) - 2 = \gamma\alpha - 1 \leq \gamma\alpha$. Thus each a_{is} other than a_{kn_k} is in some A^j and $A \setminus B = \cup_{j=1}^{\gamma} A^j \cup \{a_{kn_k}\}$.

It is left to show that $b_j \sim a \in A^j$ and $b_\gamma \sim a_{kn_k}$. Let a_{is} be an inner element in the sequence of linked chains corresponding to C. We know that $a_{is} = \sigma(t)$ for some t. The vertices $r_{a_{is}}$ and $l_{a_{is}}$ corresponding to a_{is} appear consecutively as $r_{\sigma(t)}, l_{\sigma(t+1)}$ (an arc from V) in $P_1, \ldots, P_k$. Then, if $S'(i)$ is defined to be the sum of arcs along P from $l_{\sigma(1)}$ to $x_{\sigma(i)}$ where x can be either l or r, as in the proof of Lemma 4.26, $\bar{S}(a_{is}) = -S'(t) = -S'(t+1) - 1$ since $r_{\sigma(t)}$ is the $(2s-2)^{nd}$ vertex in the i^{th} WV-Path. Similarly, $\bar{S}(a_{j1}) = -S'(t)$ when a_{j1} appears as $l_{\sigma(t)}$ in C, and $\bar{S}(a_{in_i}) = -S'(t)$ when a_{in_i} appears as $r_{\sigma(t)}$ in C.

Consider $a_{is} \in A^j$ for $s \neq 1$. Vertex $r_{\sigma(t)}$ for $\sigma(t) = a_{is}$ appears in $P_1, \ldots, P_k$. From the definition of A^j and since $\bar{S}(a_{is}) = -S'(t)$, $(j-1)\alpha \leq -S'(t) \leq j\alpha$. Recall also that $\pi(2j) = b_j$. If $\pi(2j) \succ \sigma(t)$, replace the path $l_{\pi(2j)}, \ldots, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}$ in C with the arc $(l_{\pi(2j)}, r_{\sigma(t)}) \in W$ to form a new cycle C'. The replaced path has length $j\alpha + S'(t) \geq 0$ and the new arc has length -1, so C' has negative length and it has fewer arcs from U than C. Since C' has one maximal UZ-Path, this contradicts the choice of C. So $\sigma(t) \succeq \pi(2j)$. If $\sigma(t) \succ \pi(2j) = \pi(2j-1)$, then $(l_{\sigma(t)}, l_{\pi(2j-1)}) \in W$. Let $C' = r_{\pi(2j-1)}, \ldots, r_{\pi(1)}, l_{\sigma(1)}, \ldots, r_{\sigma(t)}, l_{\sigma(t)}, r_{\pi(2j-1)}$. C' has length $(j-1)\alpha + S'(t) - 2 \leq -2$. (The -2 term comes from $r_{\sigma(t)}, l_{\sigma(t)}, r_{\pi(2j-1)}$.) Then C' is a negative cycle with one maximal UZ-Path and fewer arcs from U than C, contradicting the choice of C. Thus $b_j = \pi(2j) \sim \sigma(t) = a_{is}$.

Finally, we show that if $a_{i1} \in A^j$, then $b_j \sim a_{i1}$. Corresponding to a_{i1} is the first vertex $l_{\sigma(t)}$ of P_i. From the definition of A^j and since $\bar{S}(\sigma(t)) = -S'(t)$ for a_{i1} vertices, $(j-1)\alpha \leq -S'(t) \leq j\alpha$. Recall also that $\pi(2j) = b_j$. If $\pi(2j) \succ \sigma(t)$, then $(l_{\pi(2j)}, r_{\sigma(t)}) \in W$. Replace the path $l_{\pi(2j)}, \ldots, r_{\pi(1)}, l_{\sigma(1)}, \ldots, l_{\sigma(t)}$ in C with $l_{\pi(2j)}, r_{\sigma(t)}, l_{\sigma(t)}$ to form a new C' with fewer arcs from U than C and one maximal UZ-Path. The replaced path has length $j\alpha + S'(t) \geq 0$ and the new path has length -2. Thus C' has negative length and contradicts the choice of C. Thus $\sigma(t) \succeq \pi(2j) = \pi(2j-1)$. If $\sigma(t) \succ \pi(2j-1)$, then $(l_{\sigma(t)}, r_{\pi(2j-1)}) \in W$. Let $C' = r_{\pi(2j-1)}, \ldots, r_{\pi(1)}, l_{\sigma(1)}, \ldots, l_{\sigma(t)}, r_{\pi(2j-1)}$. Then C' has one maximal UZ-Path. The length of C' is $(j-1)\alpha + S'(t) - 1 \leq -1$. (The -1 comes from $(l_{\sigma(t)}, r_{\pi(2j-1)}) \in W$.) C' also has fewer arcs from U than C, a contradiction. Thus $b_j = \pi(2j) \sim \sigma(t) = a_{i1}$.

Finally, note that a_{kn_k} is the element corresponding to the last vertex of P_k, which is $r_{\sigma(v)}$. But $(r_{\sigma(v)}, l_{\pi(2\gamma)})$ is in Z, so $\sigma(v) \sim \pi(2\gamma) = b_\gamma$. Thus, $a_{kn_k} \sim b_\gamma$ and the proof of part *(4)* of Definition 4.13 is complete. □

The N-linked chain structure describes the basic structure contained in orders $(A, \succ) \in \mathcal{F}[\alpha, 1]$. This structure corresponds to a negative cycle in the digraph

$D(A, \succ, \alpha, 1)$. To completely characterize the entire family $\mathcal{F}[\alpha, 1]$, we would need to specify the possibilities for the relations between elements which are unspecified in the definition of an N-linked chain structure. That is, we have shown that it is necessary that $(A, \succ)$ have an N-linked chain structure if $(A, \succ) \in \mathcal{F}[\alpha, 1]$, and it remains to find necessary and sufficient conditions. This was done for the degenerate case ($\beta = 0$) in Theorem 4.25. A partial list of these conditions can be determined in a manner similar to those in Section 4.6, however, the conditions and the proofs are detailed and seem to provide little insight into orders in $\mathcal{F}[\alpha, 1]$ beyond the structure given in the N-linked chain structure. In Appendix 4.9 we give a construction of a family of orders based on the N-linked chain structure which is in the spirit of Definition 4.12 and a proof the the family is contained in $\mathcal{F}[\alpha, 1]$ in the spirit of Theorem 4.25. The construction of the family in Appendix 4.9 will also complete the proof of Theorem 4.13. While the construction is complex and does not provide a complete characterization of $\mathcal{F}[\alpha, 1]$, it seems that this construction will play an important role in in characterizing $\mathcal{F}[\alpha, 1]$ if the techniques of Section 4.6 are used to find such a characterization.

4.8 Further Research

1. We have noted that the algorithm for determining if there are bounded discrete representations of interval orders does not work for interval graphs when the lower bounds are variable. It would be interesting to determine the complexity of this problem.

2. Use the techniques of this chapter to provide an alternative proof of Fishburn's result stated in Theorem 4.1.

3. Find more compact necessary and sufficient conditions for membership in $\mathcal{D}[\alpha, \beta]$ and examine the families $\mathcal{F}[\alpha, \beta]$ when $\beta \geq 2$. In particular, this would provide an alternative approach to the results of Bogart and Stellpflug.

4. Determine $|\mathcal{F}[\alpha, 0]|$.

5. In this chapter we have examined interval orders on finite sets A. We might ask about representations of infinite interval orders. A simple example shows that it is not always the case that an infinite interval order with no $[\alpha, \beta]$ discrete representation contains a finite suborder with no representation. Take $\alpha = \beta = 0$ and $(A, \succ) = (\mathbf{Re}, >)$. The reals under $>$ are a linear order and thus an interval order. There can be no $[0,0]$ discrete representation, since if there were, the map from $\mathbf{Re}$ to the set of left endpoints would be an injection from the reals to the integers. However, every finite suborder $q_n > q_{n-1} > \cdots > q_1$ has a representation given by $J(q_i) = [i, i]$. In fact, $(\mathbf{Re}, >)$ has no real interval representation (with no integrality constraint or bounds on interval length). Given an interval order that has a real interval representation, but no $[\alpha, \beta]$ bounded discrete representation, is there always a finite suborder which also has no $[\alpha, \beta]$ bounded discrete representation?.

6. Examine values of partial order parameters, such as partial order dimension, on classes $\mathcal{D}[\alpha, \beta]$. (See for example Fishburn [1985a] for a definition of partial order dimension.)

4.9 Appendix: Construction of an Infinite Subfamily of $\mathcal{F}[\alpha, 1]$.

In this appendix we present interval representations of an infinite subfamily of $\mathcal{F}[\alpha, 1]$ based on the N-linked chain structure. This construction is in the spirit of the construction in Definition 4.12. The proof of Theorem 4.29 completes the proof of Theorem 4.13 (the case that α is odd). It is hoped that this construction, while it is somewhat complicated to describe will provide insight into a characterization of the family $\mathcal{F}[\alpha, 1]$ along the lines of the characterization of $\mathcal{F}[\alpha, 0]$ in Section 4.6. In particular, we expect that changing condition (d) in Definition 4.15 to allow intervals with length 1 under certain conditions will allow that the orders represented according to Definition 4.15 to be exactly those orders in $\mathcal{F}[\alpha, 1]$.

We will first describe interval representations of orders based on the N-linked chain structures as in Definition 4.13.

Definition 4.15 *For a given $\alpha \geq 2$, and $k, \gamma \geq 1$, let $n_f \geq 2$ for $f = 1, \ldots, k$, satisfy:*

(a)

$$\sum_{f=1}^{k}(2n_f - 3) = \gamma\alpha + 1.$$

(b) *For $i = 1, \ldots, \gamma$ and for $j = 2, \ldots, k$,*

$$\sum_{f=1}^{j-1}(2n_f - 3) \neq \alpha i \text{ or } \alpha i + 1.$$

(c) *For $i = 1, \ldots, \gamma$, and $j = 1, \ldots, k$, and $1 < s < n_j$,*

$$\sum_{f=1}^{j-1}(2n_f - 3) + (2s - 4) \neq \alpha i.$$

Also, let $A = \{a_{is} : 1 \leq i \leq k, 1 \leq s \leq n_i\} \cup \{b_1, \ldots, b_\gamma\}$. We assume that an element can appear twice in A as $a_{in_1} = a_{j1}$ for $i < j - 1$ if $2 \leq \sum_{f=1}^{j-1}(2n_f - 3) - \sum_{f=1}^{i}(2n_f - 3) \leq \alpha$. Otherwise, the elements of A are distinct. Define maps $J^h : A \to \{[l, r] : l, r \in \mathbf{Z}\}$ for $h = 1, \ldots, \gamma$. For a given h, specify δ^h and ζ^h by

$$\delta_j^h = \begin{cases} 1 & \text{if } 1 \leq j < h \\ 0 & \text{if } h \leq j \leq \gamma \end{cases}$$

$$\zeta_j^h = \begin{cases} 1 & \text{if } 1 \leq j \leq h \\ 0 & \text{if } h < j \leq \gamma \end{cases}$$

Note that when $h = 1$, the condition $1 < j \leq h$ is vacuous and when $h = \gamma$, the condition $h < j \leq \gamma$ is vacuous.

Then, the maps J^h are defined in terms of the n_f and conditions (d), (e), (f), and (g) as follows. Sums are considered to be 0 if the upper limit is 0.

For $i = 1, \ldots, \gamma$,

$$J^h(b_i) = [\alpha(\gamma - i) + \delta_i^h, \alpha(\gamma - i + 1) + \zeta_i^h].$$

Also,

$$J^h(a_{11}) = [\alpha\gamma + 1, \alpha\gamma + 2],$$

$$J^h(a_{kn_k}) = [-1, 0].$$

For $1 < s < n_i$,

$$J^h(a_{is}) = [\alpha\gamma - \sum_{f=1}^{i-1}(2n_f - 3) - (2s - 3), \alpha\gamma - \sum_{f=1}^{i-1}(2n_f - 3) - (2s - 4)].$$

If an element a *appears twice in* A *as* $a_{in_i} = a_{j1}$ *for some* $1 \leq i < j - 1 < j \leq k$, *then*

$$J^h(a) = [\alpha\gamma - \sum_{f=1}^{j-1}(2n_f - 3) + 1, \alpha\gamma - \sum_{f=1}^{i}(2n_f - 3) + 1].$$

Finally, for $1 < j \leq k$, *if* a_{j1} *appears once in* A,

$$J^h(a_{j1}) = [\alpha\gamma - \sum_{f=1}^{j-1}(2n_f - 3) + 1, r^h_{a_{j1}}],$$

and for $1 \leq i < k$, *if* a_{in_i} *appears once in* A,

$$J^h(a_{in_i}) = [l^h_{a_{in_i}}, \alpha\gamma - \sum_{f=1}^{i}(2n_f - 3) + 1].$$

For $1 < j, j' \leq k$, *and for* $1 \leq i, i' < k$, *and for all* $1 \leq h \leq \gamma$, *if* a_{j1}, $a_{j'1}$, $a_{i'n_{i'}}$, *and* a_{in_i} *all appear exactly once in* A, *then let* $r^h_{a_{j1}}$ *and* $l^h_{a_{in_i}}$ *and* $r^h_{a_{j'1}}$ *and* $l^h_{a_{i'n_{i'}}}$ *be such that (d), (e), (f) and (g) are satisfied.*

(d) *For* $1 < j \leq k$ *and* $1 \leq i < k$,

$$2 \leq |J^h(a_{j1})| \leq \alpha,$$

$$2 \leq |J^h(a_{in_i})| \leq \alpha$$

where $|J^h(x)|$ *is the length of the interval* $J^h(x)$.

(e) *Either* $r^h_{a_{j1}} = \alpha(\gamma - g) + \delta^h_g$ *for some* $g = 1, \ldots, \gamma - 1$, *or* $r^h_{a_{j1}} \neq \alpha(\gamma - g)$ *or* $\alpha(\gamma - g) + 1$ *for all* $g = 1, \ldots, \gamma - 1$ *and* $r^h_{a_{j1}}$ *is independent of* h.

Either $l^h_{a_{in_i}} = \alpha(\gamma - g) + \delta^h_g$ *for some* $g = 1, \ldots, \gamma - 1$, *or* $l^h_{a_{in_i}} \neq \alpha(\gamma - g)$ *or* $\alpha(\gamma - g) + 1$ *for all* $g = 1, \ldots, \gamma - 1$ *and* $l^h_{a_{in_i}}$ *is independent of* h.

(f) *For* $g = 2, \ldots, k$,

$$r_{a_{j1}} \neq \alpha\gamma - \sum_{f=1}^{g}(2n_f - 3) + 1.$$

For $g = 1, \ldots, k-1$,

$$l_{a_{in_i}} \neq \alpha\gamma - \sum_{f=1}^{g-1}(2n_f - 3) + 1.$$

(g) *There is no configuration as in Figure 4.7 for the endpoints which are independent of* h. *That is, for* $1 \leq s \leq k$ *and* $1 < t < n_s$, *if* $J(a_{st}) = [p, p+1]$ *then it is not the case that* $r_{a_{j1}} = l_{a_{in_i}} = p$ *and* $r_{a_{j'1}} = l_{a_{i'n_{i'}}} = p+1$.

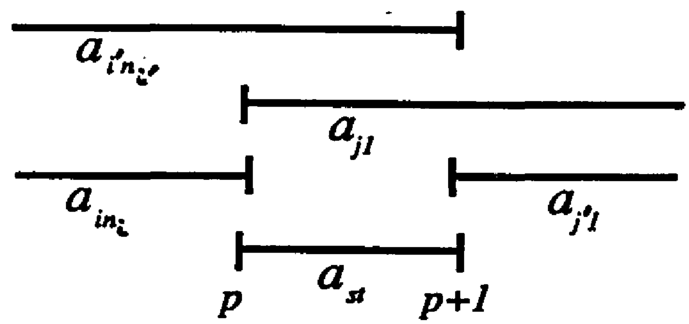


Figure 4.7: A forbidden configuration.

It can be checked that $h = 1, \ldots, \gamma$, the only interval in J^h that is longer than α is $J^h(b_h)$ with length $\alpha + 1$.

Lemma 4.28 *In Definiton 4.15, for given* n_f, $f = 1, \ldots, k$, *satisfying (a), (b), and (c) and for* $l_{a_{in_i}}$ *and* $r_{a_{j1}}$ *satisfying (d), (e), (f) and (g), for* $h = 1, \ldots, \gamma$, *the sets of intervals* J^h *represent the same interval order.*

Proof: By the definition of an interval representation, it is enough to show that for $x, y \in A$ and $1 \leq h, h' \leq \gamma$,

$$l_x^h > r_y^h \Leftrightarrow l_x^{h'} > r_y^{h'} \tag{4.28}$$

holds.

If r_y^h and l_x^h are independent of h, i.e., $r_y^{h'} = q$ for $h' = 1, \ldots, \gamma$ and $l_x = q'$, for $h' = 1, \ldots, \gamma$, then (4.28) holds trivially.

Note that for $h = 1, \ldots, \gamma$, and for $i = 1, \ldots, \gamma - 1$, $\alpha(\gamma - i) + \delta_i^h = \alpha(\gamma - i) + \zeta_{i+1}^h$. In particular, the right endpoint of $J(b_{i+1})$ agrees with the left endpoint of $J(b_i)$. For $i = 1, \ldots, \gamma - 1$, let $\alpha(\gamma - i) + \delta_i^h = p_i^h$, and let $p_i = \alpha(\gamma - i)$. Then, for a given $1 \leq h \leq \gamma$, we have $p_i^h = p_i$ or $p_i + 1$. From the definition of the intervals in J^h, the only endpoints which depend on h are p_i^h for some i. (This uses the fact that the right endpoint of $J(b_1)$ and the left endpoint of $J(b_\gamma)$ do not depend on h since, for $h = 1, \ldots, \gamma$, we have $\zeta_1^h = 1$ and $\delta_\gamma^h = 0$.) So, if r_y^h and l_x^h depend on h, we have $r_y^h = p_i^h$ for some i and $l_x^h = p_j^h$ for some j, and (since $\alpha \geq 2$ and $p_i^h = p_i$ or $p_i + 1$), (4.28) holds.

Consider the case that r_y^h is independent of h and l_x^h depends on h. Then, let $r_y^{h'} = q$ for $h' = 1, \ldots \gamma$ and $l_x^h = p_i^h$ for some $i = 1, \ldots, \gamma - 1$. Since $p_i^h = p_i$ or $p_i + 1$, the only way for (4.28) to fail is if $q = p_i$. By the definition of the intervals, $q \neq p_i$ if $y = b_1, a_{11}, a_{kn_k}$. By the definitions of the intervals, since r_y^h is independent of h, y is not b_j for $j = 2, \ldots, \gamma$. If $y = a_{f1}$ for some $1 < f \leq k$, then by condition (e), (and since l_y^h is independent of h), $q \neq p_i$. If $y = a_{fn_f}$ for some $1 \leq f < k$, then by condition (b), $q \neq p_i$. Finally, if $y = a_{fg}$, for $1 < g < n_f$, by condition (c), $q \neq p_i$.

Similarly, if $r_y^h = p_i^h$ depends on h and $l_x^{h'} = q$ for $h' = 1, \ldots, \gamma$ (is independent of h), then (4.28) fails only if $q = p_i + 1$. By the definitions of the intervals $q \neq p_i + 1$ if $x = b_\gamma, a_{11}, a_{kn_k}$. Also by the definitions of the intervals, x is not b_j for $j = 1, \ldots, \gamma - 1$ since l_x^h depend on h in these cases. If $x = a_{f1}$ for some $1 < f \leq k$, then by condition (b), $q \neq p_i + 1$. If $x = a_{fn_f}$ for some $1 \leq f < k$, then by condition (e), $q \neq p_i + 1$. Finally, if $x = a_{fg}$ for $1 \leq f \leq k$ and $1 < g < n_f$, then by (c), $q \neq p_i + 1$ since if $q = p_i + 1$, then $r_{f(g+1)} = p_i$ and (b) is violated if $g = n_f$ and (c) is violated if $g < n_f$. □

Remark 4.9 For each $\alpha \geq 3$ and for $\gamma \geq 1$, it is possible to find k, and n_f such that a representation satisfying the conditions of Definition 4.15 can be found. Thus, the families $\mathcal{F}[\alpha, 1]$ are indeed infinite. We briefly describe such an infinite family for α

odd. (Recall, that an infinite subfamily of $\mathcal{F}[\alpha,1]$ for α even was described in Theorem 4.13.) For α odd, $\alpha \geq 3$, and for $\gamma \geq 2$, let $k = \gamma - 1$ and let $n_1 = \alpha + 1$, $n_k = \frac{\alpha+5}{2}$ and for $i = 2, \ldots, k-1$, let $n_i = \frac{\alpha+3}{2}$. Also, for $i = 1, \ldots, k-2$, let $a_{in_i} = a_{(i+2)1}$. The only unspecified endpoints in this case are $r_{a_{21}}$ and $l_{a_{(k-1)n_{(k-1)}}}$. These must be picked to satisfy (d), (e), (f), and (g). Define these by $r_{a_{21}} = \alpha\gamma - (2n_1 - 3) + 4$ and $l_{a_{(k-1)n_{(k-1)}}} = \alpha\gamma - \sum_{j=1}^{k-1} -2$ (so that the corresponding intervals each have length three). It can be checked that the representation just described satisfies conditions (a), ..., (f).

Theorem 4.29 *Let $(A, \succ)$ be an interval order represented by the sets of intervals J^h as described in Definition 4.15. Then, $(A, \succ) \in \mathcal{F}[\alpha, 1]$.*

Proof: Let $(A, \succ)$ be an interval order represented by the sets of intervals J^h as described in Definition 4.15. Use the notation of the definition. Let

$$P_0 = l_{b_\gamma}, r_{b_\gamma}, l_{b_{\gamma-1}}, r_{b_{\gamma-1}}, \ldots, l_{b_1}, r_{b_1}.$$

For $i = 1, \ldots, k$, let

$$P_i = l_{a_{i1}}, r_{a_{i2}}, l_{a_{i2}}, r_{a_{i3}}, l_{a_{i3}}, \ldots, r_{a_{i(n_i-1)}}, l_{a_{i(n_i-1)}}, r_{a_{in_i}}.$$

It can be checked that in $D(A, \succ, \alpha, 1)$, P_0 is a UZ-Path with length $\alpha\gamma$, and that for $i = 1, \ldots, k$, P_i is a WV-Path with length $-2n_i + 3$. Also, there is an arc from Z in $D(A, \succ, \alpha, 1)$ connecting the last vertex of P_j to the first vetex of P_{j+1} (mod $k+1$). So $C = P_0, P_1, \ldots, P_k$ has length $\alpha\gamma - \sum_{j=1}^{k}(2n_j - 3)$ in in $D(A, \succ, \alpha, 1)$. Also, C is a cycle since if an element appears twice in A, then it is as $x = a_{j1}$ and $x = a_{in_i}$ for $i < j$ and then l_x and r_x each appear exactly once in C. By condition (a) in the definition, $length(C) = -1$. So by Theorem 4.3 $(A, \succ) \notin \mathcal{D}[\alpha, 1]$.

Conversely, we must show that an $[\alpha, 1]$ representation can be found for any proper suborder of $(A, \succ)$. We will describe such a representation based on one of the representations J^h. We consider cases of removing different types of elements.

For $x \in A$, we will describe an $[\alpha, 1]$ discrete representation $J^* : A \setminus \{x\} \to [l, r]$ of $(A \setminus \{x\}, \succ)$ based on one of the representations J^h of definition 4.15. Note that if an element appears twice in A, then we will describe two ways of removing it, one for each appearance. Either of these cases will suffice for removing such an element. To show that J^* represents $(A \setminus \{x\}, \succ)$ it is enough to show in each case that for all $y, z \in A \setminus \{x\}$

$$l_y^h > r_z^h \Leftrightarrow l_y^* > r_z^* \tag{4.29}$$

holds for some h.

Case 1: remove b_i for some $1 \leq i \leq \gamma$.
Start with the representation J^i. The only interval in this representation with length greater than α is $J(b_i)$. Remove it. That is, for $x \in A \setminus \{b_i\}$, let $J^*(x) = J^i(x)$. Then clearly (4.29) holds (with $h = i$).

Case 2: remove a_{11}.
Define J^* in terms of J^1. Let $q = \alpha\gamma + 1$. For x either l or r and for $a \in A \setminus \{a_{11}\}$, define J^* by

$$x_a^* = \begin{cases} x_a^1 - 1 & \text{if } x_a^1 \geq q \\ x_a^1 & \text{if } x_a^1 < q. \end{cases}$$

The only way for (4.29) to fail is if $l_y^1 = q$ and $r_x^1 = q - 1$ for some y, z. However, it is easy to see that there is no such y by the definitions of the intervals.

Since the intervals in J^1 all have length at least one, the only way for an interval in J^* to have length one is if for some z, $J^1(z) = [q - 1, q]$. There is no such interval by the definitons of the intervals and by condition (d). Also, the only interval in J^1 with length greater than α is $J(b_1)$. Since $r_{b_1} = q$, the length of $J^*(b_1)$ is α and J^* is an $[\alpha, 1]$ discrete representation.

Case 3: remove a_{kn_k}.
Define J^* in terms of J^k. Let $q = 0$. For x either l or r and for $a \in A \setminus \{a_{kn_k}\}$, define

J^* by

$$x_a^* = \begin{cases} x_a^k & \text{if } x_a^k > q \\ x_a^k + 1 & \text{if } x_a^k \le q. \end{cases}$$

The only way for (4.29) to fail is if $l_y^k = q+1$ and $r_z^k = q$ for some y, z. However, it is easy to see that there is no such z by the definitions of the intervals.

Since the intervals in J^k all have length at least one, the only way for an interval in J^* to have length one is if for some z, $J^k(z) = [q, q+1]$. There is no such interval by the definitions of the intervals and by condition (d). Also, The only interval in J^k with length greater than α is $J(b_k)$. Since $l_{b_k} = q$, the length of $J^*(b_k)$ is α and J^* is an $[\alpha, 1]$ discrete representation.

Case 4: remove a_{st} for $1 \le s \le k$ and $1 < t < n_s$.

Consider $J(a_{st}) = [\alpha\gamma - \sum_{j=1}^{s-1}(2n_s - 3) - (2t-3), \alpha\gamma - \sum_{j=1}^{s-1}(2n_s - 3) - (2t-4)]$ and denote this interval by $[q, q+1]$. By condition (c) in Definition 4.15, $q + 1 \neq \alpha i$ for $i = 1, \ldots, \gamma$. Then, for some $j = 1, \ldots, \gamma$, we have $\alpha(\gamma - j) + 1 < q \le \alpha(\gamma - j + 1)$. Use J^j as a basis for the representation J^*. Consider the overlaps of endpoints with the interval $[q, q+1]$. If $q \neq \alpha(\gamma - j + 1)$, then by condition (g), one of the following subcases holds:

(i) There is no z such that $l_z^j = q$.

(ii) There is no z such that $l_z^j = q + 1$.

(iii) There is no z such that $r_z^j = q$.

(iv) There is no z such that $r_z^j = q + 1$.

Otherwise, if $q = \alpha(\gamma - j + 1)$, then by conditions (b) and (e) and since $\zeta_j^j = 1$, there is no endpoint (other than $r_{a_{st}}$) which is $q + 1$. So when $q = \alpha(\gamma - j + 1)$, we can use subcase (i).

We give definitions for J^* in each of the subcases (i), (ii), (iii), (iv) seperately. In each case remove the interval $J(a_{st})$ and for endpoints x_a (with x either l or r), define the new endpoints $x^*(a)$ as follows:

(i)

$$x_a^* = \begin{cases} x_a^j - 1 & \text{if } x_a^j \geq q \\ x_a^j & \text{if } x_a^j < q. \end{cases}$$

(ii)

$$x_a^* = \begin{cases} x_a^j - 1 & \text{if } x_a^j \geq q+1 \\ x_a^j & \text{if } x_a^j < q. \end{cases}$$

(iii)

$$x_a^* = \begin{cases} x_a^j & \text{if } x_a^j > q \\ x_a^j + 1 & \text{if } x_a^j \leq q. \end{cases}$$

(iv)

$$x_a^* = \begin{cases} x_a^j & \text{if } x_a^j > q+1 \\ x_a^j + 1 & \text{if } x_a^j \leq q+1. \end{cases}$$

In (i), the only way for (4.29) to be violated is if $l_y^j = q$ and $r_x^j = q - 1$ for some y and x. However, there is no such y by the condition for subcase (i). Thus (4.29) holds. Similarly, the conditions for subcases (ii), (iii), and (iv) insure that (4.29) holds for the corresponding shifts.

The only interval in J^j which is longer than α is $J^j(b_j) = [\alpha(\gamma - j), \alpha(\gamma - j + 1) + 1]$ with length $\alpha + 1$. Since the right endpoint of this interval is greater than $q + 1$ and the left endpoint is smaller than q, the length of interval $J^*(b_j)$ will be α. Also, by condition (d) in the definition, and by the definition of the intervals, there is no interval (other than $J^j(a_{st})$) in J^j which is $[q-1, q]$, $[q, q+1]$ or $[q+1, q+2]$. Thus the shifts will not create any degenerate interval (with length 0). So the J^* are indeed $[\alpha, 1]$ discrete representations.

Case 5: remove a_{i1} for $i = 2, \ldots, k$.
Let $q = \alpha\gamma - \sum_{f=1}^{i-1}(2n_f - 3) + 1$, (the left endpoint of the interval $J(a_{i1})$). By condition (b), for some $1 \leq j \leq \gamma$, we have $\alpha(\gamma - j) < q < \alpha(\gamma - j + 1)$. Then for x either r or l and $a \in A \setminus \{a_{i1}\}$,

$$x_a^* = \begin{cases} x_a^j - 1 & \text{if } x_a^j \geq q \\ x_a^j & \text{if } x_a^j < q. \end{cases}$$

The only way for (4.29) to fail is with $l_y^j = q$ and $r_x^j = q - 1$ for some y, x. However, by condition (f) and by the construction of the intervals there is no such y.

Since the intervals in J^j all have length at least one, the only way for an interval in J^* to have length one is if for some z, $J^j(z) = [q-1, q]$. There is no such interval by the definitions of the intervals and by condition (d). Also, The only interval in J^j with length greater than α is $J(b_j)$. Since $l_{b_j}^j = \alpha(\gamma - j) < q < \alpha(\gamma - j + 1) < \alpha(\gamma - j + 1) + 1 = r_{b_j}^j$, the length of $J^*(b_j)$ is α and J^* is an $[\alpha, 1]$ discrete representation.

Case 6: remove $a = in_i$ for $i = 1, \ldots, k-1$.

Let $q = \alpha\gamma - \sum_{f=1}^{i}(2n_f - 3) + 1$, (the right endpoint of the interval $J(a_{in_i})$). By condition (b), for some $1 \leq j \leq \gamma$, we have $\alpha(\gamma - j) < q < \alpha(\gamma - j + 1)$. Then for x either r or l and $a \in A \setminus \{a_{in_i}\}$,

$$x_a^* = \begin{cases} x_a^j & \text{if } x_a^j > q \\ x_a^j + 1 & \text{if } x_a^j \leq q. \end{cases}$$

The only way for (4.29) to fail is with $l_y^j = q + 1$ and $r_x^j = q$ for some y, x. However, by condition (f) and by the construction of the intervals there is no such x.

Since the intervals in J^j all have length at least one, the only way for an interval in J^* to have length one is if for some z, $J^j(z) = [q, q+1]$. There is no such interval by the definitions of the intervals and by condition (d). Also, The only interval in J^j with length greater than α is $J(b_j)$. Since $l_{b_j}^j = \alpha(\gamma - j) < q < \alpha(\gamma - j + 1) < \alpha(\gamma - j + 1) + 1 = r_{b_j}^j$, the length of $J^*(b_j)$ is α and J^* is an $[\alpha, 1]$ discrete representation.

□

References

Alway, G.G., "Matrices and Sequences," Math. Gaz., **46** (1962), 208–213.

Aho, A.V., J.E. Hopcroft, and J.D. Ullman, *The Design and Analysis of Computer Algorithms*, Addision-Wesley, Reading MA, 1974.

Anstee, R.P., 'A Polynomial Algorithm for b-Matchings: An Alternative Approach,' Inf. Proc. Let., **24**, (1987), 153–157.

Baker, F.B. and L.J. Hubert, "Applications of Combinatorial Programming to Data Analysis: Seriation Using Asymmetric Proximity Measures," Br. J. Math. Statist. Psychol., **30** (1977), 154–164.

Barthelemy, J-P., O. Hudry, G. Isaak, F.S. Roberts, and B.Tesman, "The Reversing Number of a Digraph", manuscript (1990).

Berge, C., "Two Theorems in Graph Theory," Proc. Nat. Acad. Sci., USA, **43** (1957), 842.

Bermond, J.C., "Ordres a Distance Minimum d'un Tournoi et Graphes Partiels Sans Circuits Maximaux," Math. Sci. Hum., **37** (1972), 5–25.

Bermond, J.C., "The Circuit-Hypergraph of a Tournament," in A. Hajnal et al. eds., *Infinite and Finite Sets- Proc. Coll. Math. Janós Bolyai*, **10** (1975), 165–180.

Bermond, J.C. and Y. Kodratoff, "Une Heuristique Pour Le Calcul de l'Indice de Transitivite d'un Tournoi," R.A.I.R.O. Informatique Theorique, **10, no.3** (1976), 83–92.

Bermond, J.C. and C. Thomassen, "Cycles in Digraphs- a Survey," J. Graph Theory,

5 (1981), 1–43.

Bollabás, B., *Graph Theory: An Introductory Course*, Springer-Verlag, 1979.

Bogart, K.P. and K. Stellpflug, "Discrete Representation Theory for Semiorders I: Fundamental Examples," manuscript 1989.

Bogart, K.P. and K. Stellpflug, "Discrete Representation Theory for Semiorders II: Representation Theorems," manuscript 1990.

Cornuejols, G., D. Hartvigsen and W. Pulleyblank, "Packing Subgraphs in a Graph," OR Letters **1** (1982), 139–143.

Edmonds, J., "Paths, Trees and Flowers," Canad. J. Math **17** (1965), 449–467.

Edmonds, J. and E.L. Johnson, "Matching, Euler Tours, and the Chinese Postman," Math. Prog., **5** (1973), 88–124.

Erdős, P. and J.W. Moon, "On Sets of Consistent Arcs in a Tournament," Canad. Math. Bull., **8** (1965), 269–271.

Fishburn, P.C., "Intransitive Indifference and Unique Indifference Intervals," J. Math. Psych. **7** (1970), 144–149.

Fishburn, P.C., "Threshold-bounded Interval Orders and a Theory of Picycles," SIAM J. Alg. Disc. Meth. **4** (1983), 290-305.

Fishburn, P.C., *Interval Orders and Interval Graphs,* John Wiley and Sons, 1985a.

Fishburn, P.C., "Interval Graphs and Interval Orders," Disc. Math., **55** (1985b), 135–149.

Fishburn, P.C. and B. Monjardet, "Norbert Wiener on the Theory of Measurement," manuscript, 1990.

Gabow, H., and R.E. Tarjan, "A Linear-Time Algorithm for a Special Case of Disjoint Set Union," J. Comp. Sys. Sci. **30** (1985), 209–221.

Garey, M.R., and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman, NY, 1979.

Grötschel, M., M. Jünger and G. Reinelt, "A Cutting Plane Algorithm for the Linear Ordering Problem," Operations Research, **32** (1984), 1195–1220.

Grötschel, M., M. Jünger and G. Reinelt, "Facets of the Linear Ordering Polytope," Math. Programming, **33** (1985), 28–42.

Golumbic, M.C., *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, NY, 1980.

Hakimi, S.L., "On the Degrees of the Vertices of a Directed Graph," J. Franklin Inst., **279** (1965), 290–308.

Hanlon, P., "Counting Interval Graphs," Trans. AMS, **272** (1982), 383–426

Harary, F., *Graph Theory*, Addison-Wesley, London, 1972.

Harary, F., "On Minimal Feedback Vertex Sets of a Digraph," IEEE Trans. Circuit Theory, **22** (1975), 839–840.

Hell, P. and D.G. Kirkpatrick, "Packings by Cliques and by Finite Families of Graphs," Discrete Math **49** (1984), 45–59.

Hell, P. and D.G. Kirkpatrick, "Packings by Complete Bipartite Graphs," SIAM J. Alg. Disc. Meth. **7** (1986), 199–209.

Hubert, L., "Seriation Using Asymmetric Proximity Measures," Br. J. Math. Statist. Psychol., **29** (1976), 32–52.

Jünger, M., *Polyhedral Combinatorics and the Acyclic Subdigraph Problem*, Heldermann Verlag, Berlin, 1985.

Kamae, T., "Notes on a Minimum Feedback Arc Set," IEEE Trans. Circuit Theory, **14** (1967), 78–79.

Karp, R.M., "Reducibility Among Combinatorial Problems," in Raymond Miller and James Thatcher eds., *Complexity of Computer Computation*, Plenum Press 1972.

Kirkpatrick, D.G., and P. Hell, "On the Complexity of Graph Factor Problems," SIAM J. Comp. **12** (1983), 601–609.

Kotzig, A., "On the Maximal Order of Cyclicity of Antisymmetric Directed Graphs," Discrete Math., **12** (1975), 17–25.

Korte, B., "Approximative Algorithms for Discrete Optimization Problems," Annals of Disc. Math., **4** (1979), 85–120.

Lawler, E., "A Comment on Minimum Feedback Arc Sets," IEEE Trans. Circuit Theory, **11** (1964), 296–297.

Lawler, E., *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, NY, 1976.

Lempel, A. and I. Cederbaum, "Minimum Feedback Arc and Vertex Sets of a Directed Graph," IEEE Trans. Circuit Theory, **13** (1966), 399–403.

Lovász, L., and M.P. Plummer, *Matching Theory*, Akedémiai Kiadó, Budapest, 1986.

Luce, R.D., "Semiorders and a Theory of Utility Discrimination," Econometrica, **24** (1956), 178–191.

Luce, R.D., and P. Suppes, "Preference, Utility and Subjective Probability," in *Hanbook of Mathematical Psychology III*, R.D. Luce, R.R. Bush, and E. Galanter eds., Wiley, NY, 1965, 249–410.

Monjardet, B., "Tournois et Ordres Médians Pour une Opinion," Math. Sci. Hum., **43** (1973), 55–70.

Moon, J., *Topics on Tournaments*, Holt, Rinehart and Winston, New York, 1968.

Norman, R.Z., and M.O. Rabin, "An Algorithm for a Minimum Cover of a Graph," Proc. AMS, **10** (1959), 315–319.

Pulleyblank, W.R., and D. deWerra, "Odd Path Packings," European J. Comb., **10** (1989), 585–596.

Reid, K.B., "On Sets of Arcs Containing No Cycles in Tournaments," Canadian Math Bulletin, bf 12 (1969), 261–264.

Reid, K.B. and E.T. Parker, "Disproof of a Conjecture of Erdős and Moser on Tournaments," J. Comb. Theory, **9** (1970), 225–238.

Remage Jr., R. and W.A. Thompson Jr., "Maximum-likelihood Paired Comparison Rankings," Biometrika, **53** (1966), 143–149.

Roberts, F.S., *Graph Theory and its Applications to Problems of Society*, CBMS-NSF Series in Applied Mathematics 29, SIAM, Philadelphia 1978.

Roberts, F.S., *Measurement Theory*, Cambridge Univ. Press 1979.

Roberts, F.S., *Applied Combinatorics*, Prentice Hall, Englewood Cliffs, NJ, 1984.

Seshu, S. and M.B. Reed, *Linear Graphs and Electrical Networks*, Addison-Wesley, Reading, MA, 1961.

Schrijver, A., *Theory of Integer and Linear Programming*, John Wiley and Sons, 1986.

Slater, P., "Inconsistencies in a Schedule of Paired Comparisons," Biometrika, **48** (1961), 303–312.

Spencer, J., "Optimal Ranking of Tournaments," Networks, **1** (1971), 135–138.

Spencer, J., "Optimally Ranking Unrankable Tournaments," Periodica Math. Hungarica, **11** (1980), 131–144.

Tarjan, R.E., *Data Structures and Network Algorithms*, CBMS-NSF Series in Applied Mathematics 44, SIAM, Philadelphia 1983.

Thompson, W.A. Jr. and R. Remage Jr., "Rankings from Paired Comparisons," Ann. Math. Stat., **35** (1964), 739–747.

Tucker, A.W., "On Directed Graphs and Integer Programs," presented at 1960 Symposium on Combinatorial Problems, Princeton University, cited in Younger [1963].

de la Vega, W.F., "On the Maximum Cardinality of a Consistent Set of Arcs in a Random Tournament," J. Comb. Theory B, **35** (1983), 328–332.

Wiener, N., "A Contribution to the Theory of Relative Position," Trans. AMS, **5** (1914), 343–384.

deWerra, D., "Variation on a Theorem of König," Discrete Math. **51** (1984), 319–321.

deWerra, D., "Partitions Into Odd Chains," Mathematical Programming **37** (1987), 41–50.

deWerra, D., "Generalized Edge Packings," Mathematical Programming, **44** (1989), 235-245.

deWerra, D. and F.S. Roberts, "On the use of Augmenting Chains in Chain Packings," Discrete Applied Math, to appear (1990).

Yau, S.S., "Generation of all Hamiltonian Circuits, Paths, and Centers of a Graph, and Related Problems," IEEE Trans. Circuit Theory, **14** (1967), 79–81.

Younger, D.H., "Minimum Feedback Arc Sets for a Directed Graph," IEEE Trans. Circuit Theory, **10** (1963), 238–245.

Vita

Garth Isaak

1961 Born January 30 in Sellersville, Pa.

1979 Graduated from Pennridge High School, Perkasie, Pa.

1979–1980 Attended Bethel College, North Newton, Kan.

1980–1981 Work on low income housing and with homeless people, Mennonite Voluntary Service, Portland, Or.

1981–1984 Attended Bethel College, North Newton, Kan.

1984 B.A. in Chemistry, Mathematical Sciences, and Physics, Bethel College.

1984–1985 Mathematics Instructor, Bethel College.

1984–1985 Science Instructor, Hesston Junior College.

1985–1990 Graduate work in Operations Research, RUTCOR, Rutgers University.

1985–1986 Teaching Assistant, Department of Computer Science, Rutgers University.

1986–1989 Teaching Assistant, Department of Mathematics, Rutgers University.

1987–1990 Research Assistant, RUTCOR, Rutgers University.

1987–1990 Summers, Instructor, Department of Mathematics, Rutgers University.

1990– John Wesley Young Research Instructor, Department of Mathematics and Computer Science, Dartmouth College.

1990 Ph.D. in Operations Research, RUTCOR, Rutgers University.