


8-2005

Greedy-Type-Resistance of Combinatorial Problems

Gareth Bendall
University of Kentucky

François Margot
Carnegie Mellon University, fmargot@andrew.cmu.edu

Follow this and additional works at: <http://repository.cmu.edu/tepper>

 Part of the [Economic Policy Commons](#), and the [Industrial Organization Commons](#)

Published In

Discrete Optimization, 3, 4, 288-298.

This Article is brought to you for free and open access by Research Showcase @ CMU. It has been accepted for inclusion in Tepper School of Business by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Greedy-type-resistance of Combinatorial Problems

Gareth Bendall¹ François Margot²

April 2004; Revised July 2005

Abstract

This paper gives a sufficient condition for a combinatorial problem to be greedy-type-resistant, i.e. such that, on some instances of the problem, any greedy-type algorithm will output the unique worst possible solution. The condition is used to show that the Equipartition, the k -Clique, the Asymmetric Traveling Salesman, the Hamiltonian Path, the Min-Max Matching, and the Assignment Problems are all greedy-type-resistant.

Keywords: Greedy algorithms, greedy-resistance, domination analysis.

¹Department of Mathematics, University of Kentucky, 715 POT, Lexington, KY 40506.
Present address: Dept. of Mathematics, Physics, and Computer Science, Georgetown College, Georgetown KY 40324. Email: Gareth.Bendall@georgetowncollege.edu .

²Department of Mathematics, University of Kentucky, 715 POT, Lexington, KY 40506.
Present address: Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213. Email: fmargot@andrew.cmu.edu .

1 Introduction

The *greedy algorithm* is a simple algorithm for constructing feasible solutions of combinatorial optimization problems. It proceeds, starting from the partial solution $X = \emptyset$, by repeatedly adding to X an element x from the ground set E . Moreover x is the element with minimum weight in a set $I(X)$, where the set $I(X)$ is the set of all $x \in E - X$ such that $X \cup x$ can be completed to a feasible solution of the problem. *Greedy-type* algorithms are similar algorithms where $I(X)$ is replaced by one of its subsets. Well-known examples of greedy-type algorithms are the *nearest neighbor* heuristic for the Traveling Salesman Problem [6, 12] or *Prim's* Algorithm for the Minimum Spanning Tree Problem [11].

Natural questions are the assessment of the quality of solutions obtained by specific greedy-type algorithms and, more generally, to decide if there exists a greedy-type algorithm with some solution quality guarantee for solving a problem. For general algorithms, several measures of solution quality have been used. The most widely used measure is probably the *performance guarantee* (defined as the ratio of the values of the solution returned by the heuristic algorithm and of the optimal solution), yielding the study of *approximation algorithms*. In this paper, a different measure of solution quality is used, namely the *domination number* of the algorithm. This concept was introduced by Glover and Punnen [5] and has been the topic of several recent papers [2, 7, 8, 9, 10, 14]. Loosely speaking (precise definitions are given in the next section), this number is a lower bound on the number of solutions that are worse than the solution returned by the algorithm, as a function of the size of the problem. Of course, the larger the domination number is, the better the algorithm. In this paper, we study conditions on problems implying that the domination number of any greedy-type algorithm is 1. Such problems are said to be *greedy-type-resistant*.

Problems that are resistant to the greedy algorithm are studied in [8, 9]. In [9] a sufficiency condition is introduced and used to show that the Asymmetric Traveling Salesman Problem and the Assignment Problem are greedy-resistant. In [2] a complete characterization of problems where the weight function is restricted to the values $\{1, 2, \dots, r\}$ and for which the greedy algorithm has a domination number of 1 is given. This result is used to show that the Traveling Salesman Problem and the Equipartition Problem are greedy-resistant on instances with this restriction. This characterization as well as the sufficiency conditions for greedy-resistance set forth in [9] seem difficult to apply. Both sets of conditions require multiple comparisons of sums of sizes of intersections. This paper extends the analysis to greedy-type

algorithms, while relaxing the condition on the weight function. Our main result is Theorem 2 which gives a sufficient condition for a problem to be greedy-type resistant. While this does not yield a complete characterization of greedy-type resistance, the conditions are not too difficult to check and not too restrictive as most known interesting cases of greedy-type resistance satisfy them. As illustration, Theorem 2 is used to prove that six classical combinatorial optimization problems are greedy-type resistant.

2 Greedy-type Algorithms

An *independence system* is a pair consisting of a set of elements E and a family of subsets \mathcal{F} such that

- (I1) $\emptyset \in \mathcal{F}$;
- (I2) If $X \in \mathcal{F}$ and $Y \subseteq X$, then $Y \in \mathcal{F}$.

A maximal set in \mathcal{F} is called a *base*.

Many combinatorial optimization problems can be formulated as finding a set (or a base) in \mathcal{F} with minimum weight, where each element $e \in E$ has a weight $c(e)$ and the weight of a set $S \subseteq E$ is defined as

$$c(S) = \sum_{e \in S} c(e) .$$

We refer to these combinatorial optimization problems as independence optimization problems (*IOP*). An instance of such a problem is a triplet (E, \mathcal{F}, c) , where (E, \mathcal{F}) defines the independence system, and c is a weight function from E to \mathbb{R} .

For a given instance of an IOP (E, \mathcal{F}, c) and an algorithm \mathcal{H} which finds a base B in (E, \mathcal{F}) , it is possible to assess the quality of B as a solution to the IOP by counting the number of bases $S \in \mathcal{F}$ with $c(S) \geq c(B)$. This value $D_{\mathcal{H}}(E, \mathcal{F}, c)$ is called the *domination number* of \mathcal{H} on the instance (E, \mathcal{F}, c) . Considering all instances of the IOP (E, \mathcal{F}) , the domination number \mathcal{H} is

$$D_{\mathcal{H}}(E, \mathcal{F}) = \min_{c: E \rightarrow \mathbb{R}} \{D_{\mathcal{H}}(E, \mathcal{F}, c)\} .$$

We consider the algorithm \mathcal{H} to fail the IOP when \mathcal{H} returns the unique worst solution for some instance of the IOP. That is, \mathcal{H} fails the IOP (E, \mathcal{F}) if $D_{\mathcal{H}}(E, \mathcal{F}) = 1$; in this case we say the IOP is *\mathcal{H} -resistant*.

Given an independence system (E, \mathcal{F}) , if $X \in \mathcal{F}$, define $I(X)$ to be all the elements $x \in E - X$ such that $X \cup x$ is an independent set; that is

$$I(X) := \{x \in E - X \mid (X \cup x) \in \mathcal{F}\} .$$

For simplicity, we write $I(x_1, x_2, \dots, x_k)$ instead of $I(\{x_1, x_2, \dots, x_k\})$.

The *greedy algorithm* to solve the IOP (E, \mathcal{F}, c) is the algorithm that starts with $X = \emptyset$ and then repeatedly adds to X an element of minimum weight in $I(X)$ until X is a base of \mathcal{F} . (Ties are broken arbitrarily.) It is well-known that the greedy algorithm returns an optimal solution to the IOP for all functions $c \geq 0$ if and only if (E, \mathcal{F}) is a matroid [13]. When (E, \mathcal{F}) is not a matroid, however, the quality of the solution returned by the greedy algorithm may vary. In [9], the concept of *antimatroids* was introduced. Antimatroids are independence systems (E, \mathcal{F}) for which there exists a weight function $c \geq 0$ such that a unique base $B \in \mathcal{F}$ has maximum weight and B is obtained by the greedy algorithm on (E, \mathcal{F}, c) . Thus, on an antimatroid, the greedy algorithm may return the unique worst possible solution. Hence, when (E, \mathcal{F}) is a matroid, the domination number of the greedy algorithm is the number of bases in \mathcal{F} . When (E, \mathcal{F}) is an antimatroid, the domination number is 1, and so antimatroids are *greedy-resistant*.

In this paper we focus on *greedy-type* algorithms, introduced in [8]. Such an algorithm \mathcal{H} is similar to the greedy algorithm: start with the partial solution $X = \emptyset$ and then repeatedly add to X an element of minimum weight in $I_{\mathcal{H}}(X)$ (ties are broken arbitrarily) until X is a base of \mathcal{F} where $I_{\mathcal{H}}(X)$ is a subset of $I(X)$ that does not depend on the cost function c , but only on the independence system (E, \mathcal{F}) and the set X . Moreover, $I_{\mathcal{H}}(X)$ is non empty if $I(X)$ is non empty, a condition that guarantees that \mathcal{H} always outputs a base.

In the next section, we derive our main result, Theorem 2, which allows us to prove that the following classical IOPs are \mathcal{H} -resistant, where \mathcal{H} is any greedy-type algorithm:

- (1) The Equipartition Problem: partition the nodes of a complete graph into two equal size sets so that the sum of the weights of the edges between the sides of the partition is minimum.
- (2) The k -Clique Problem: find a set of k nodes in a complete graph so that the sum of the weights of the edges between them is minimum.
- (3) The Asymmetric Traveling Salesman Problem: find a directed cycle using all the nodes of a complete directed graph so that the sum of the weights of the edges in the cycle is minimum.

- (4) The Hamiltonian Path Problem: find a path containing all the nodes of a complete graph so that the sum of the weights of the edges in the path is minimum.
- (5) The Min-Max Matching Subgraph Problem: find a maximal (with respect to inclusion) matching so that the sum of the weights of the edges in the matching is minimum.
- (6) The Assignment Problem: find a perfect matching so that the sum of the weights of the edges in the matching is minimum.

For a more complete description of these problems, the reader is referred to [4].

3 Greedy-Type Resistance

Let (E, \mathcal{F}) be an independence system and \mathcal{H} be a greedy-type algorithm for the IOP (E, \mathcal{F}, c) . Let B' be a base of (E, \mathcal{F}) and order its elements as $(x_1, x_2, \dots, x_{|B'|})$. If there exists a weight function c such that \mathcal{H} outputs B' when applied on (E, \mathcal{F}, c) , adding element x_i at step i , then B' with ordering $(x_1, x_2, \dots, x_{|B'|})$ is \mathcal{H} -allowed.

In essence, a greedy-type algorithm \mathcal{H} examines at each step a subset of the elements and selects the element of least weight. It is reasonable to assume that, at some point during the running of \mathcal{H} , some elements must be chosen to be in the base. Consider the k -Clique Problem with edge set E and for $k \geq 3$; once two edges that share an endpoint are chosen, the third edge that completes the triangle is in every base with these two edges, and so must be chosen at some point by the algorithm. Which elements must be chosen depends not only on which elements \mathcal{H} chooses, but also on the order in which they are selected.

Let X be an independent set that could be selected by \mathcal{H} during the first $|X|$ steps of the algorithm in the order $(x_1, x_2, \dots, x_{|X|})$. Then $e \in E - X$ is in the $X\mathcal{H}$ -completion, denoted $comp(X, \mathcal{H})$, if and only if e is in every base that might be constructed by \mathcal{H} once the elements $x_1, x_2, \dots, x_{|X|}$ have been selected (in that order). For $x_j \in X$, x_j is in $comp(X, \mathcal{H})$ if and only if x_j is in every base that might be constructed by \mathcal{H} once the elements x_1, x_2, \dots, x_{j-1} have been selected (in that order). For example, if the set $I_{\mathcal{H}}(x_1, \dots, x_{t-1})$ contains only the element x_t , then $x_t \in comp(X, \mathcal{H})$.

In most cases, we only concern ourselves with the completion of bases. As such, we define the elements of the ordered base B' that are not in the $B'\mathcal{H}$ -completion as the $B'\mathcal{H}$ -decision elements, denoted $decn(B', \mathcal{H})$. For both the

$B'\mathcal{H}$ -completion and the $B'\mathcal{H}$ -decision, when \mathcal{H} and B' are understood, they are dropped from the notation.

The idea of the completion is important when looking at greedy-type algorithms. Consider the following lemma.

Lemma 1 *Let \mathcal{P} be an IOP on an independence system where all bases have the same cardinality, and let \mathcal{H} be a greedy-type algorithm for \mathcal{P} . If every \mathcal{H} -allowed base B' has an empty $B'\mathcal{H}$ -completion, then \mathcal{P} is not \mathcal{H} -resistant.*

Proof: Choose some instance of \mathcal{P} where all the elements of the ground set are given a positive weight. Let B' be the base chosen by \mathcal{H} in the order $x_1, \dots, x_{|B'|}$. Since $x_{|B'|}$ is not in the completion, there exists an element $y \neq x_{|B'|}$ such that

$$y \in I_{\mathcal{H}}(x_1, \dots, x_{|B'|-1}) .$$

Hence, the set $B = \{x_1, \dots, x_{|B'|-1}, y\}$ is independent, and so is a base. The algorithm \mathcal{H} selects the element $x_{|B'|}$ over y , implying that $c(x_{|B'|}) \leq c(y)$. Since all elements have a positive weight, we have $c(B') \leq c(B)$. \square

We use the completion to define a particular base and an order for it. Consider a base constructed under the following conditions: Let every element of E have weight zero and apply the greedy-type algorithm \mathcal{H} with the added condition that after the set X is chosen by \mathcal{H} , \mathcal{H} will only choose an element of $comp(X, \mathcal{H})$ when no other element is available. In this way, a base is found whose completion elements are chosen as late as possible. More formally, given a greedy-type algorithm \mathcal{H} , define a new greedy-type algorithm to be \mathcal{H}' where

$$I_{\mathcal{H}'}(X) = \begin{cases} I_{\mathcal{H}}(X) - comp(X, \mathcal{H}) & \text{when this is nonempty} \\ I_{\mathcal{H}}(X) & \text{otherwise.} \end{cases}$$

An \mathcal{H}' -allowed base B' with order $(x_1, x_2, \dots, x_{|B'|})$ is said to be \mathcal{H} -delayed.

Along with the idea of elements that \mathcal{H} must include in the base is the idea of elements that can no longer be selected. Notice that for the greedy algorithm, when an element e can not be chosen at step i , then e can no longer be chosen at any subsequent step. This is not necessarily true of all greedy-type algorithms. Since it is possible for an element to become ineligible to be chosen by \mathcal{H} and then later become eligible again, we concern ourselves only with elements that are eligible to be selected at step i of \mathcal{H} , but cannot be selected at any step after step i .

Let B' be an \mathcal{H} -allowed base B' with ordering $(x_1, x_2, \dots, x_{|B'|})$. For any $1 \leq r \leq |B'| + 1$, we define the (\mathcal{H}, B', r) -critical class as:

$$\kappa(\mathcal{H}, B', r) = \{y \in I_{\mathcal{H}}(x_1, \dots, x_{r-1}) : y \notin I_{\mathcal{H}}(x_1, \dots, x_s), \forall s \text{ with } |B'| \geq s \geq r\} .$$

The set $\kappa(\mathcal{H}, B', r)$ is the set of elements that could be added by \mathcal{H} at step r while constructing the base, but can never be added after x_r has been selected. This basic idea is the heart of Theorem 2. Observe that $x_r \in \kappa(\mathcal{H}, B', r)$, that the critical classes are disjoint and form a partition of the elements that are contained in at least one of the bases. Notice also that when B' is \mathcal{H} -delayed, if $x_i \in \text{comp}(B', \mathcal{H})$ then $\mathcal{I}_{\mathcal{H}}(x_1, \dots, x_{i-1})$ contains only elements from $\text{comp}(B', \mathcal{H}) - \{x_1, \dots, x_{i-1}\}$, since x_i is selected.

Theorem 2 makes use of the following condition concerning critical classes: for a base B and an \mathcal{H} -allowed base B' , if

$$t := \max\{i = 1, \dots, |B'| : |\kappa(\mathcal{H}, B', i) \cap B| \neq 1\}$$

is well defined, and if we have

$$|\kappa(\mathcal{H}, B', t) \cap B| = 0,$$

then we say that B satisfies Condition 1 for B' .

Theorem 2 *Let \mathcal{P} be an IOP on (E, \mathcal{F}) , and \mathcal{H} a greedy-type algorithm for \mathcal{P} . Let B' be an \mathcal{H} -delayed base with ordering $(x_1, x_2, \dots, x_{|B'|})$ such that every base $B \neq B'$ containing $\text{comp}(B', \mathcal{H})$ satisfies Condition 1 for B' . Then \mathcal{P} is \mathcal{H} -resistant.*

Proof: The proof proceeds by demonstrating a weight function which gives the \mathcal{H} -delayed base B' the unique worst weight.

For $e \in E$ define a weight function:

$$c(e) = \begin{cases} |E|^{|E|} + i & \text{if } e = x_i, \text{ for } x_i \in \text{comp}(B', \mathcal{H}) \\ |E|^i - 1 & \text{if } e = x_i, \text{ for } x_i \in \text{decn}(B', \mathcal{H}) \\ |E|^j & \text{if } e \in (\kappa(\mathcal{H}, B', j) - x_j) \\ -1 & \text{otherwise} \end{cases}$$

Note that as $(\text{decn}(B', \mathcal{H}), \text{comp}(B', \mathcal{H}))$ is a partition of B' , as the critical classes are disjoint, and as $(\kappa(\mathcal{H}, B', j) - x_j) \cap B' = \emptyset$ for all j , each element receives a unique weight.

Since B' is \mathcal{H} -delayed, the weight function c forces \mathcal{H} to select the base B' in the order $x_1, x_2, \dots, x_{|B'|}$. The base B' contains $|\text{comp}(B', \mathcal{H})|$ elements of weight larger than $|E|^{|E|}$ and one element of weight $|E|^i - 1$ for each $x_i \in \text{decn}(B', \mathcal{H})$.

Choose any base $B \neq B'$. If B does not contain the set $\text{comp}(B', \mathcal{H})$, then because B' contains at least one more element of weight larger than $|E|^{|E|}$,

$c(B') > c(B)$. If B does contain $\text{comp}(B', \mathcal{H})$, then, by hypothesis, B satisfies Condition 1 for B' . Hence, there exists a t such that B contains no elements of $\kappa(\mathcal{H}, B', t)$ and exactly one element from $\kappa(\mathcal{H}, B', i)$, for $i = t+1, t+2, \dots, |B'|$.

If $t = 1$, then B consists of exactly one element from each (\mathcal{H}, B', r) -critical class for $r = 2, 3, \dots, |B'|$ and has no elements in $(\mathcal{H}, B', 1)$ -critical class. As B is contained in the union of all the critical classes, we have that $c(B') > c(B)$.

If $t > 1$ then let $\mathcal{D} = \text{decn}(B', \mathcal{H})$. Since B contains at most $|E| - (|B'| - (t - 1))$ elements from

$$\cup_{i \in \mathcal{D}, i \leq t-1} \kappa(\mathcal{H}, B', i) ,$$

we have

$$c(B) \leq (|E| - |B'| + t - 1)|E|^{t-1} + a$$

where

$$a = |E|^{|E|} |\text{comp}(B', \mathcal{H})| + \sum_{i: x_i \in \text{comp}(B', \mathcal{H})} i + \sum_{i: x_i \in \mathcal{D}, i \geq t+1} |E|^i .$$

On the other hand,

$$\begin{aligned} c(B') &= |E|^{|E|} |\text{comp}(B', \mathcal{H})| + \sum_{i: x_i \in \text{comp}(B', \mathcal{H})} i + \sum_{i: x_i \in \mathcal{D}} (|E|^i - 1) \\ &= |E|^{|E|} |\text{comp}(B', \mathcal{H})| + \sum_{i: x_i \in \text{comp}(B', \mathcal{H})} i + \sum_{j: x_j \in \mathcal{D}, j \leq t} (|E|^j - 1) + \\ &\quad \sum_{i: x_i \in \mathcal{D}, i \geq t+1} (|E|^i - 1) \\ &\geq \sum_{j: x_j \in \mathcal{D}, j \leq t} (|E|^j - 1) + a - (|B'| - t) \end{aligned}$$

Hence, we have

$$\begin{aligned} c(B') - c(B) &\geq \sum_{j: x_j \in \mathcal{D}, j \leq t} (|E|^j - 1) + a - (|B'| - t) - (|E| - |B'| + t - 1)|E|^{t-1} - a \\ &\geq |E|^t - 1 + \sum_{j: x_j \in \mathcal{D}, j \leq t-1} (|E|^j - 1) - |B'| + t - |E|^t + (|B'| - t + 1)|E|^{t-1} \\ &> 0. \end{aligned}$$

The last inequality holds as $t > 1$ and $|B'| - t + 1 \geq 1$ and so $c(B') > c(B)$.
□

It is not difficult to show that, when \mathcal{H} is the greedy algorithm, Theorem 2 implies the sufficiency condition for greedy-resistance of [2] when applied to problems with a weight function with values in $\{1, 2, \dots, |E|^{|E|} + |E|\}$. On the other hand, we do not know if the converse holds or not.

4 Applications of Theorem 2

An immediate consequence of Theorem 2 is the following useful corollary:

Corollary 3 *Let \mathcal{P} be an IOP and B' be an \mathcal{H} -delayed base. If B' is the unique base containing the $B'\mathcal{H}$ -completion, then \mathcal{P} is \mathcal{H} -resistant.*

Proof: In order to satisfy the hypothesis of Theorem 2, we need to verify for every base $B \neq B'$ which contains the completion that Condition 1 holds for B' . But, by hypothesis, there are no such bases. Hence, the equation is vacuously true, and by Theorem 2, \mathcal{P} is \mathcal{H} -resistant. \square

4.1 Equipartition Problem

We use Corollary 3 extensively to show that the Equipartition Problem \mathcal{E} is \mathcal{H} -resistant for all greedy-type algorithms. We present here only the proof for when the problem is on eight or more nodes. The proofs for the smaller cases can be found in [3].

We define \mathcal{E} in the following way: for a weighted complete graph G having $2n$ nodes, find a partition of the nodes of G into two equal sized sets (V, W) such that the sum of the weights of the edges with one end point in V and the other in W is minimum. The partition (V, W) is called an *equipartition*, and the set of edges with one end point in V and one end point in W is called a *balanced cut*. Hence, \mathcal{E} is sometimes referred to as the Minimum Weight Balanced Cut Problem. When we view \mathcal{E} as an IOP (E, \mathcal{F}) , we consider the edges of G to be E and the balanced cuts to be the bases of \mathcal{F} .

Before applying Corollary 3 in the case of \mathcal{E} , we need the following lemmas.

Lemma 4 *Let \mathcal{E} be the equipartition problem, \mathcal{H} be a greedy-type algorithm for \mathcal{E} , and B' be some \mathcal{H} -delayed base. If R is the subgraph induced by the $B'\mathcal{H}$ -decision edges, then R contains no cycles.*

Proof: Note that since the edges in B' form a complete bipartite graph, the only cycles that could occur in R are even cycles of length at least 4. Let $P = (v_1, e_1, v_2, e_2, \dots, v_{2j})$ be a path in R , where $j \geq 2$. Each edge of P is in

the balanced cut B' , and so the endpoints of each edge are in opposite sides of the partition. This means that one partition contains the odd indexed nodes of P and the other the even indexed nodes. Hence, every base that contains the edges of P must contain the edge (v_1, v_{2j}) , and so this edge is in $\text{comp}(B', \mathcal{H})$. \square

Lemma 5 *Let \mathcal{E} be the equipartition problem with partition size $n \geq 3$, and \mathcal{H} be a greedy-type algorithm for \mathcal{E} . Suppose that there exists an \mathcal{H} -delayed base B' such that, for some $(u, w) \in B'$, the only edges of B' with endpoint w are decision edges and exactly one decision edge of B' has endpoint u . Then \mathcal{E} is \mathcal{H} -resistant.*

Proof: Let B' be an \mathcal{H} -delayed base with corresponding equipartition (V', W') constructed by \mathcal{H} where $u \in V'$ and $w \in W'$. Also, let R be the subgraph induced by the decision edges, and let B be a base with corresponding equipartition (V, W) containing the $B'\mathcal{H}$ -completion edges.

The node u has degree 1 in R , and so every edge between u and a node of $W' - w$ must be a completion edge. Hence, without loss of generality, $u \in V$ and $W' - w \subset W$. Therefore, if $B \neq B'$, there must be a node $v \in V'$ such that $V = (V' - v) \cup w$ and $W = (W' - w) \cup v$. Since B contains all the $B'\mathcal{H}$ -completion edges, this means that every edge of B' with endpoint v is a $B'\mathcal{H}$ -decision edge. By hypothesis $|V| \geq 3$, and so let $z \in V$ where $z \neq u, v$. By Lemma 4, R contains no cycles, and so z cannot be adjacent to a node of $W' - w$ in R (if it were, then we would have a cycle using this node of $W' - w$, and the nodes z, w and v). From this we conclude that every node of $V' - v$ has degree 1 in R .

If the edge (v, w) is the last chosen edge in R by \mathcal{H} , then since $n \geq 3$, $R - (v, w)$ is the disjoint union of two vertex stars each containing $n \geq 3$ nodes; hence, every balanced cut containing $R - (v, w)$ also contains (v, w) , which implies this is a completion edge, a contradiction because $(v, w) \in R$. Similarly, if $e = (w, \bar{v})$ (respectively $e = (v, \bar{w})$) is the last edge chosen where $\bar{v} \in V' - v$ (resp. $\bar{w} \in W' - w$), then $R - e$ contains a vertex star containing n nodes with v (resp. w) as its center. A vertex star is contained in exactly one base, and so e is a completion edge; contradiction with $e \in R$. Therefore $B = B'$; that is the $B'\mathcal{H}$ -completion edges are contained in a unique base. By Corollary 3 the result holds. \square

Theorem 6 *Let \mathcal{H} be a greedy-type algorithm for \mathcal{E} with partition size n where $n \geq 4$. Then \mathcal{E} is \mathcal{H} -resistant.*

Proof: Let B' be an \mathcal{H} -delayed base with corresponding equipartition (V', W') constructed by \mathcal{H} with ordering $(x_1, x_2, \dots, x_{|B'|})$, and let R be the subgraph induced by its decision edges. Also let B be a base with corresponding equipartition (V, W) containing the completion edges of B' .

By Lemma 4, R contains no cycles. Hence, R is a forest, and so R contains at least one node u of degree 1 in R . Without loss of generality let $u \in V'$. Let w be the neighbor of u in R , and so $w \in W'$. The node u has degree 1 in R , and so every edge between u and a node of $W' - w$ must be a completion edge. Hence, without loss of generality, $u \in V$ and $W' - w \subset W$. Moreover, we can assume that $w \notin W$ as otherwise $W = W'$ and $B = B'$.

We distinguish four cases, depending on the number of completion edges adjacent to w :

Case 1: If w is incident to no edges of $\text{comp}(B', \mathcal{H})$, then Lemma 5 applies.

Case 2: If w is incident to two edges of $\text{comp}(B', \mathcal{H})$, then both these neighbors must be in W , a contradiction because W already contains $|W - w| = (n - 1)$ nodes. This means B' is the only base containing $\text{comp}(B', \mathcal{H})$, and the result follows from Corollary 3.

Case 3: If w is incident to exactly one edge (v, w) of $\text{comp}(B', \mathcal{H})$, and if v is incident to two edges of $\text{comp}(B', \mathcal{H})$, then v must be connected to a node in $W' - w$ by an edge of $\text{comp}(B', \mathcal{H})$. Since w is also a neighbor of v , w must be in the same side of the partition as $W' - w$, a contradiction with $w \notin W$.

Case 4: Finally, assume that w is incident to exactly one edge (v, w) of $\text{comp}(B', \mathcal{H})$ and v is incident to exactly one edge (v, w) of $\text{comp}(B', \mathcal{H})$. Let e be that last edge chosen by \mathcal{H} that is in R . Note that e cannot be an edge of R connecting a node of $V' - v$ to a node of $W' - w$, because all the edges of R incident to v plus all the edges of R incident to w are enough to determine the base B' , and so no edge chosen after these edges can be a decision edge. Hence, e is incident to v or w .

Without loss of generality, let v be an endpoint of e . Let Q be the set of edges in $R - e$ and incident to v . Note that $|Q| = (n - 2)$ since e is incident to v and v is adjacent to exactly one completion edge. Before \mathcal{H} chooses e , all edges of Q are chosen, and so $|Q|$ nodes are fixed on some side of the partition. Also before \mathcal{H} chooses e , all decision edges adjacent to w are chosen, and so $(n - 1)$ nodes are fixed in the partition not containing w . If $n > 3$, then $|Q| > 1$ and $(n - 1) + |Q| > n$. Hence, there is no base containing in the same side of the partition the nodes connected to w by decision elements and the nodes connected to v by edges of Q . Therefore, before e is chosen by \mathcal{H} , the

equipartition is determined, a contradiction since $e \in \text{decn}(B', \mathcal{H})$. Hence this case is impossible. \square

4.2 Asymmetric TSP

Notice that in order to prove that \mathcal{E} was \mathcal{H} -resistant for all greedy-type algorithms \mathcal{H} , we made no use of Condition 1 or even any use of critical classes. This is not only an example of the utility of Corollary 3, but it leads us to wonder about the necessity of the more complex Condition 1. The fact that Condition 1 brings more than the simple Corollary 3 is demonstrated by its use on the case of the Asymmetric Traveling Salesman Problem (ATSP).

We formulate the ATSP as an IOP by viewing the ground set E as the set of edges of the complete directed graph \vec{K}_n . An independent set is any subset of edges that are in some directed Hamiltonian tour; that is, a set is independent if and only if it is a collection of node disjoint directed paths. A base is a set of edges that form a directed Hamiltonian tour and so has size n .

Theorem 7 *Let \mathcal{H} be a greedy-type algorithm for the ATSP. Then the ATSP is \mathcal{H} -resistant.*

Proof: Let B' be an \mathcal{H} -delayed base with order (x_1, x_2, \dots, x_n) . Notice that because the graph is directed, any independent set of size $(n - 2)$ is contained in only one base. This means that the last two elements selected by \mathcal{H} are in the $B'\mathcal{H}$ -completion.

We show that every base $B \neq B'$ containing x_{n-1} and x_n satisfies Condition 1 for B' . Choose such a base B , and let t be that largest index such that $x_t \in \text{decn}(B', \mathcal{H})$ and

$$|\kappa(\mathcal{H}, B', t) \cap B| \neq 1 .$$

We need to show that such a t exists and that

$$|\kappa(\mathcal{H}, B', t) \cap B| = 0$$

in order to use Theorem 2.

We first assume that t exists, and we show, by reverse induction on the index, that

$$\{x_{t+1}, x_{t+2}, \dots, x_n\} \subset B .$$

Since B contains $\text{comp}(B'\mathcal{H})$, $\{x_j, \dots, x_n\} \subset B$ for $j = (n - 1), n$. Hence, for some j' between t and $n - 1$ assume this holds for all $j > j'$. We want to know what edges could be in both the j' -critical class and B . That is, what edges satisfy at least the following conditions:

- (i) $e \notin I_{\mathcal{H}}(x_1, x_2, \dots, x_k)$ for all $j' \leq k \leq n$.
- (ii) $e \in I_{\mathcal{H}}(x_1, x_2, \dots, x_{j'-1})$
- (iii) $\{e, x_{j'+1}, x_{j'+2}, \dots, x_n\}$ is an independent set,

where condition (iii) comes from the fact that $e \in B$, and by induction, $\{x_{j'+1}, x_{j'+2}, \dots, x_n\} \subset B$. Notice that a consequence of condition (i) is that $e \notin \{x_{j'+1}, x_{j'+2}, \dots, x_n\}$. Edges satisfying these conditions are not guaranteed to be in the j' -critical class, but we show that the only edge that satisfies them is $x_{j'}$.

Let $\delta^-(e)$ and $\delta^+(e)$ denote the tail and head respectively of an edge e . Then condition (ii) is the same as saying that e is an edge such that

$$\begin{aligned} \delta^-(e) &\in \{\delta^-(x_{j'}), \delta^-(x_{j'+1}), \dots, \delta^-(x_n)\} && \text{and} \\ \delta^+(e) &\in \{\delta^+(x_{j'}), \delta^+(x_{j'+1}), \dots, \delta^+(x_n)\} . \end{aligned}$$

Adding this to condition (iii), we see the only edges satisfying (ii) and that are independent from $\{x_{j'+1}, x_{j'+2}, \dots, x_n\}$ are the edges $\{x_{j'}, x_{j'+1}, \dots, x_n\}$. By condition (i), e cannot be $\{x_{j'+1}, x_{j'+2}, \dots, x_n\}$, and thus $e = x_{j'}$.

From this, we see that our assumption that t exists is valid. Otherwise,

$$|\kappa(\mathcal{H}, B', i) \cap B| = 1$$

for all i , and by the above, this would mean $B = B'$. Since we have chosen $B \neq B'$, t exists.

Of course, the same three conditions must also hold for the t -critical class. The above reasoning shows that $|\kappa(\mathcal{H}, B', t) \cap B| \leq 1$, and since the definition of t implies that $|\kappa(\mathcal{H}, B', t) \cap B| \neq 1$, we have that

$$|\kappa(\mathcal{H}, B', t) \cap B| = 0$$

and applying Theorem 2 completes the proof. □

The same analysis yields an analogous result for the Hamiltonian Path Problem (given two nodes, find a minimum weight Hamiltonian Path connecting them). Also a similar proof can be applied to the Symmetric Traveling Salesman Problem.

4.3 k -Clique Problem

The k -Clique Problem \mathcal{K} can be stated as the following IOP (E, \mathcal{F}) : The ground set E is the set of edges of a complete weighted graph K_n , and \mathcal{F} is any subset of edges incident to k or fewer nodes. Hence, a base is a maximum sized set of edges incident to exactly k nodes, which is a k -clique.

The proof for showing that \mathcal{K} is resistant to all greedy-type algorithms also makes use of verifying Condition 1. However, the proof is not as straightforward as it was for the ATSP, and we need to break it into two cases; the first corresponding to a base being unique in containing its completion, the second where this is not so. We also need the following lemma.

Lemma 8 *Let \mathcal{H} be a greedy-type algorithm for \mathcal{K} and let B' be an \mathcal{H} -delayed base with order $(x_1, x_2, \dots, x_{|B'|})$. If*

$$x_j \in \text{decn}(B', \mathcal{H})$$

then at least one of the endpoints of x_j is not the endpoint of some edge in $\{x_1, \dots, x_{j-1}\}$.

Proof: Assume both endpoints of x_j are endpoints of edges in $\{x_1, \dots, x_{j-1}\}$. Then, any k -clique C that contains the edges $\{x_1, \dots, x_{j-1}\}$ contains the endpoints of x_j , which means C contains the edge x_j , i.e. x_j is a completion edge, a contradiction. \square

Theorem 9 *Let \mathcal{H} be a greedy-type algorithm for \mathcal{K} where $k \geq 4$. Then \mathcal{K} is \mathcal{H} -resistant.*

Proof: Let B' be an \mathcal{H} -delayed base with order $(x_1, x_2, \dots, x_{|B'|})$, and so B' is a set of edges of a k -clique; for simplicity we refer to the subgraph determined by the edges of B' as B' .

Also let R be the subgraph determined by the edge set $\text{decn}(B', \mathcal{H})$. From Lemma 8, we can conclude that R contains no cycle. We can also conclude from this lemma that any set of edges spanning the vertices in B' is contained uniquely in B' .

The proof breaks into two cases.

Case 1: Assume that the edges of $\text{comp}(B', \mathcal{H})$ are incident to every node of B' .

In this case, as we have noted, $\text{comp}(B', \mathcal{H})$ is contained in exactly one base B' . Hence, by Corollary 3, the theorem holds.

Case 2: Assume that the edges of $\text{comp}(B', \mathcal{H})$ are not incident to every node of B' .

In this case, there must be a node $u \in V(B')$ that is not the endpoint of any edge in $\text{comp}(B', \mathcal{H})$, and so every edge in B' incident to u must be in R . Lemma 8 then implies that R is a vertex star with u as the center.

Let $B \neq B'$ be a base that contains $\text{comp}(B', \mathcal{H})$. Then, since $\text{comp}(B', \mathcal{H}) = B' - E(R)$ and so contains edges incident to every node in B' but u ,

$$V(B) = (V(B') - u) \cup w$$

for some node $w \notin V(B')$. Let $x_t = (u, v)$ be the last edge of R chosen by \mathcal{H} , and let $N = V(R) - \{u, v\}$. Also let Q be the set of edges between w and the nodes of N . See Figure 1. Since no edge of R has both endpoints in N , we

Figure 1: The bases B and B' .

know that every edge connecting nodes in $N \cup v$ is in $\text{comp}(B', \mathcal{H})$.

If $Q \cap \kappa(B', t) = \emptyset$ then

$$B \cap \kappa(B', t) \subseteq \{(w, v)\}$$

since the only edge of B not in $\text{comp}(B', \mathcal{H})$ or Q is (w, v) . However,

$$(w, v) \notin \kappa(B', t)$$

because \mathcal{H} selects B' in the order $x_1, x_2, \dots, x_{|B'|}$ with x_t the last $B'\mathcal{H}$ -decision element, and the edges of $\{x_1, x_2, \dots, x_{t-1}, (w, v)\}$ are incident to $(k+1)$ nodes; hence this set of edges is not independent. Therefore

$$|B \cap \kappa(B', t)| = 0$$

and so this is the t that satisfies Condition 1 for B' . By Theorem 2, \mathcal{K} is \mathcal{H} -resistant.

If $Q \cap \kappa(B', t) \neq \emptyset$, then there exists an edge $q = (w, z) \in Q$ such that

$$q \in I_{\mathcal{H}}(x_1, x_2, \dots, x_{t-1}) .$$

This implies there exists a base \hat{B} where

$$\{x_1, x_2, \dots, x_{t-1}\} \cup q \subseteq \hat{B} .$$

The set of nodes that $\{x_1, x_2, \dots, x_{t-1}\} \cup q$ is incident to is exactly $N \cup \{u, w\}$, which is a set of k nodes.

Hence, \hat{B} with initial ordering $(x_1, x_2, \dots, x_{t-1}, q)$ is \mathcal{H} -allowed. Further,

$$\text{decn}(\hat{B}, \mathcal{H}) - q = \text{decn}(B', \mathcal{H}) - x_t$$

because B' and \hat{B} are indistinguishable in the first $(t-1)$ steps of \mathcal{H} , and whether x_j is a decision element or not depends only on the elements that have been selected by \mathcal{H} in the first $(j-1)$ steps of \mathcal{H} . Therefore every edge of the form (u, z) (where $z \in N$) is in $\text{decn}(\hat{B}, \mathcal{H})$, and by Lemma 8, no edge connecting two nodes of N is in $\text{decn}(\hat{B}, \mathcal{H})$. Since $k \geq 4$, this implies that every node in N is the endpoint of some edge in $\text{comp}(\hat{B}, \mathcal{H})$. Thus, since the set $Q - q$ and the edge (w, u) are also in $\text{comp}(\hat{B}, \mathcal{H})$, the edges of $\text{comp}(\hat{B}, \mathcal{H})$ are incident to every node of \hat{B} . Hence, \hat{B} is \mathcal{H} -allowed. It is an easy matter to see that \hat{B} is \mathcal{H} -delayed. Applying Corollary 3 to \hat{B} shows that the theorem holds. □

4.4 Min-Max Matching Subgraph Problem

The Min-Max Matching Subgraph Problem can be stated as follows: given a weighted graph G (not necessarily complete), what is a minimum weight, maximal (with respect to inclusion) matching? Hence, in terms of an IOP, the ground set E is the set of edges of G , the independence sets are matchings, and the bases are maximal matchings. For this problem, we can use Theorem 2 to characterize which instances are greedy-resistant.

Theorem 10 *Let \mathcal{P} be the Min-Max Matching Subgraph Problem on the graph G , and let \mathcal{H} be a greedy-type algorithm for \mathcal{P} . Also, let G have no isolated nodes. If G has a perfect matching, then \mathcal{P} is \mathcal{H} -resistant. If G does not have a perfect matching and \mathcal{H} is the greedy algorithm, then \mathcal{P} is not \mathcal{H} -resistant.*

Proof: We apply Theorem 2 using an \mathcal{H} -delayed perfect matching as the base B' . The proof that Condition 1 holds for B' is similar to the ATSP case.

Let $b = \frac{|V(G)|}{2}$, the size of a perfect matching. Label the decision edges of B' in the order they are chosen by \mathcal{H} as (x_1, x_2, \dots, x_r) . We know that $r < b$.

Choose a base $B \neq B'$ containing the $B'\mathcal{H}$ -completion edges, and let t be the largest index such that

$$|\kappa(B', t) \cap B| \neq 1 .$$

We need to show that such a t exists and that

$$|\kappa(B', t) \cap B| = 0$$

in order to use Theorem 2.

We first assume that t exists and show, by reverse induction on the index, that $\{x_{t+1}, x_{t+2}, \dots, x_r\} \subset B$. First, we want to know what edges could be in both the (B', \mathcal{H}, r) -critical class and B . That is, what edges satisfy the following conditions:

- (i) $e \notin I_{\mathcal{H}}(x_1, x_2, \dots, x_k)$ for all $r \leq k \leq b$.
- (ii) $e \in I_{\mathcal{H}}(x_1, x_2, \dots, x_{r-1})$
- (iii) $e \cup \text{comp}(B'\mathcal{H})$ is an independent set

where condition (iii) comes from the fact that $e \in B$ and $\text{comp}(B'\mathcal{H}) \subseteq B$. Condition (i) implies that e shares an endpoint with one of the edges of $\{x_1, x_2, \dots, x_r\}$. Condition (ii) implies that e does not share an endpoint with any of the edges $\{x_1, x_2, \dots, x_{r-1}\}$. Condition (iii) implies that e doesn't share an endpoint with any edge in the completion of B' . By assumption, B' is a perfect matching, and so the only edge that satisfies all of these conditions is the edge x_r . We repeat this argument for $j = r - 1, \dots, t + 1$ by modifying condition (iii) to read

- (iii) $\{e, x_{j+1}, x_{j+2}, \dots, x_r\} \cup \text{comp}(B'\mathcal{H})$ is an independent set.

Hence, we have $\{x_{t+1}, x_{t+2}, \dots, x_r\} \subset B$.

From this, we see that our assumption that t exists is valid. Otherwise,

$$|\kappa(\mathcal{H}, B', i) \cap B| = 1$$

for all i , and by the above, this would mean $B = B'$. Since we have chosen $B \neq B'$, t exists.

A similar three conditions have to hold for any edge in the t -critical class and in B . As the above reasoning shows that $|\kappa(B', t) \cap B| \leq 1$ and as the choice of t implies $|\kappa(B', t) \cap B| \neq 1$, we have

$$|\kappa(B', t) \cap B| = 0$$

and Condition 1 is satisfied for B' . Theorem 2 completes the proof

Assume that G has no perfect matching and that \mathcal{H} is the greedy algorithm. By way of contradiction, assume that \mathcal{H} does return the unique worst solution.

Let B' be the base returned by \mathcal{H} in the order $x_1, x_2, \dots, x_{|B'|}$. By assumption, B' is not a perfect matching, and so there exists a node v left unmatched by B' . The graph G has no isolated nodes, so there exists some edge e adjacent to v . Clearly $e \notin B'$. Also the other endpoint of e is matched by B' (otherwise, e could be added to B' and B' would not be maximal). Let the other endpoint of e be v_k , where v_k is matched by the edge x_k to the node w_k .

For all $j < k$, we know that $\{x_1, x_2, \dots, x_j, e\}$ is a matching, because both v and v_k are unmatched by x_1, x_2, \dots, x_j . Hence, at the k th iteration of the greedy algorithm, x_k was chosen over e , and so $c(x_k) \leq c(e)$. This implies that $(B' - \{x_k\}) \cup \{e\}$ has weight greater than or equal to B' . But B' is the unique worst base. Hence, $(B' - \{x_k\}) \cup \{e\}$ is not a base.

Let $B_1 = (B' - \{x_k\}) \cup \{e\}$. We know that B_1 is a matching and that it is not a base. Hence B_1 must not be a maximal matching. Let e_1 be an edge between two nodes unmatched by B_1 . Clearly these two nodes are not unmatched by B' , or we could add e_1 to B' . The only node unmatched by B_1 and matched by B' is w_k , and so e_1 has exactly one endpoint adjacent to a node matched by B' and one endpoint that is not. Hence, the same argument that implied that $c(x_k) \leq c(e)$ implies that $c(x_k) \leq c(e_1)$. But then any base containing B_1 would have to have weight greater than or equal to B' . This contradicts the assumption that B' is the unique worst base. \square

4.5 Further Research

It has yet to be established whether the conditions set forth in Theorem 2 herein are both necessary and sufficient for greedy-type-resistance, or merely

sufficient. Though it seems unlikely these conditions are necessary, a counter-example is elusive, at least when examining IOPs where every \mathcal{H} -allowed base has a nonempty completion. In fact, even restricting our attention to only the greedy algorithm, it is unknown whether the conditions in Theorem 2 are necessary for greedy-resistance.

References

- [1] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela and M. Protasi, *Complexity and Approximation*, Springer, Berlin, 1999.
- [2] J. Bang-Jensen, G. Gutin, A. Yeo, “When the greedy algorithm fails”, *Discrete Optimization* 1 (2004), no 2, 121–127.
- [3] G. Bendall, *Domination Analysis Beyond the Traveling Salesman Problem*, Ph.D. Thesis (2004), Department of Mathematics, University of Kentucky.
- [4] M. Garey, D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, New York, 1979.
- [5] Glover F., Punnen A.P., “The Traveling Salesman Problem: New Solvable Cases and Linkages with the Development of Approximation Algorithms”, *Journal of the Operational Research Society* 48 (1997), 502-510.
- [6] Gutin G. and Punnen A.P., *The traveling salesman problem and its variations*, Kluwer Academic Publishers, Boston, 2002.
- [7] G. Gutin, A. Vaishtein, A. Yeo, “Domination Analysis of Combinatorial Optimization Problems”, *Discrete Applied Mathematics* 129 (2003), no 2-3, 513–520.
- [8] G. Gutin, A. Vainshtein, A. Yeo, “When Greedy-Type Algorithms Fail”, unpublished manuscript.
- [9] G. Gutin, A. Yeo, “Anti-Matroids”, *Operations Research Letters* 30 (2002), no, 2, 97–99.
- [10] G. Gutin, A. Yeo, A. Zverovich, “Traveling Salesman Should Not be Greedy: Domination Analysis of Greedy-Type Heuristics for the TSP”, *Discrete Applied Mathematics* 117 (2002), no. 1-3, 81–86.

- [11] B. Korte, J. Vygen, *Combinatorial Optimization, Theory and Algorithms*, Springer (2002).
- [12] Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G., Shmoys D.B., *The Traveling Salesman Problem*, Wiley, New York, (1985).
- [13] J. Oxley, *Matroid Theory*, Oxford University Press, Oxford.
- [14] A. Punnen, F. Margot, S. Kabadi “TSP Heuristics: Domination Analysis and Complexity”, *Algorithmica* 35 (2003), no. 2, 111–127.