**DIMACS Technical Report 95-08**
**April 1995**

# Optimization under Ordinal Scales: When is a Greedy Solution Optimal?

by

Aleksandar Pekeč[1,2]
Deptartment of Mathematics
Rutgers University
New Brunswick, New Jersey 08903

[1]Affiliated Graduate Student Member, DIMACS Graduate Assistant
[2]The author gratefully acknowledges the support of the National Science Foundation under grant number SES-9211492 to Rutgers University.

---

DIMACS is a cooperative project of Rutgers University, Princeton University, AT&T Bell Laboratories and Bellcore.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

**ABSTRACT**

Mathematical formulation of an optimization problem often depends on data which can be measured in more than one acceptable way. If the conclusion of optimality depends on the choice of measure, then we should question whether it is meaningful to ask for an optimal solution. If meaningful optimal solution exists and the objective function depends on data measured on an ordinal scale of measurement, then the greedy algorithm will give such a solution for a wide range of objective functions.
**Keywords**: Optimization, Measurement Theory, Greedy Algorithms

# 1 Introduction

Efficent algorithms for solving optimization problems are rarely based on the formalization of the problem and usually rely on some properties characteristic to the particular problem. For example, both the traveling salesman problem (TSP) and the minimal weight spanning tree problem can be formulated as 0-1 programming problems where we need to minimize the objective function $f(\mathbf{x}) = \sum_{e \in E(G)} w_e x_e = \mathbf{w}^T \mathbf{x}$, where $w_e$ are weights on the edges of a graph $G$ and $\mathbf{x}$ is an incidence vector of the subgraph in question (Hamiltonian cycle for TSP, spanning tree for the minimal weight spanning tree problem) and we are optimizing $f$ over the set of all feasible incidence vectors $\mathbf{x}$ (all Hamiltonian cycles in the former case and all spanning trees in the latter). However, TSP is the most celebrated example of an NP-complete problem [1], while it is hard to imagine an easier algorithm than Kruskal's greedy algorithm [5] for finding an optimal spanning tree. It is clear that in order to fully understand the problem one cannot just formalize it and forget about any of its specifics. In this paper we will try to analyze optimization problems which involve objective functions which depend on scales of measurement and how we can use information about scales of measurement (in particular ordinal scales) to develop efficient optimization algorithms.

Usually there is more than one acceptable way to measure data. To keep up with our example, we can think of $w_e$ as a price we need to pay for using edge $e$. Obviously, we can express that price in dollars or in cents or in any other currency. We should expect that the optimal solution will remain optimal regardless of which currency we use to calculate the total cost. More generally, the weights can measure preferences or priorities. For concreteness, we shall sometimes refer to the weights $w_e$ as priorities.

Let us briefly introduce the language of measurement theory. There exsists an extensive literature on this subject (e.g. [4, 14, 6, 9, 10, 11]), and we will follow Roberts [11]. A scale of measurement is a mapping which assigns (positive) real numbers to objects being measured in such a way that certain empirical relations are preserved. An **admissible transformation** of scale is a transformation $f$ of scale values which leads to another acceptable mapping, i.e., one which again preserves the empirical relations. For example, all exchange rates are of the form $f(x) = \alpha x, \alpha > 0$. Similarly, we can measure mass in pounds or kilograms and here we have $p(x) = 2.2x$ where $x$ is the amount in kg. and p(x) is the amount in lb. The set of all admissible transformations of a particular scale of measurement defines a **scale type**. For example, if the set of all admissible transformations of the scale of measurement is $\{f(x) = \alpha x : \alpha > 0\}$ we are talking about a **ratio scale** ("objects are measured on a ratio scale"). Measurement of mass, currency amounts, time (seconds, minutes, hours) are examples of ratio scales. An **interval scale** is a scale where the set of admissable transformations is $\{f(x) = \alpha x + \beta : \alpha > 0, \beta \in \mathbf{R}\}$. Temperature is measured on an interval scale: for example, $f(x) = \frac{9}{5}x + 32$ is a transformation from $^oF$ to $^oC$. If we want to make sure that our graph optimization problem doesn't change an optimal solution if we use an exchange rate and pay a flat fee for each transaction, then this is the same as to say that weights can be measured on an interval scale (TSP is an example of such a problem). In many instances very precise data is not available and the most that is given is a relationship among data. For example, we can only have ordering or a preference list of all the weights (priorities): $w_1 \geq w_2 \geq \ldots \geq w_n$ (where we know if the inequality is strict or we have an equality) and any set of $n$ numbers satisfying this relationship is an acceptable way to measure data. In such a case any increasing function $f$ is an admissible transformation. We say that data is measured on an **ordinal scale** if $\{f : f \text{ is increasing}\}$ is the set of the admissible transformations.

If we can conclude from the nature of the problem (or data available) that data are measured on some particular scale of measurement, then we should expect that an optimal solution remains optimal if we apply an admissible transformation of the scale to all of the data. In other words, we expect a solution to be optimal regardless of the choice of an acceptable way to measure data. For example, it would be meaningless to conclude that a solution is optimal if we measure our cost in dollars but it is not optimal if we measure it in cents. In measurement theory, we call a statement using scales **meaningful** if its truth or falsity is unchanged after applying admissible transformations to all of the scales in the statement.

The meaningfulness of the conclusion of optimality for combinatorial optimization problems was studied by Roberts in [12, 13]. In this paper we analyze the problems where data is assigned positive real numbers and is measured on an ordinal scale

In the next section we will formulate our problem and introduce additional terminology. In Section 3 we analyze optimization under an ordinal scale and show that one can have far reaching conclusions about the optimization just from knowing that some data is measured on an ordinal scale of measurement. We will also try to show how we can use information about scales of measurement to develop an efficient algorithm, showing that ideas from measurement theory might be valuable in developing algorithms for solving optimization problems.

## 2 Formulation of the Problem

Throughout $\mathbf{x}$ will denote a vector $(x_1, \ldots, x_n)$, and $\mathbf{R}^n_+$ will denote a set of all $n$-dimensional vectors with nonnegative coordinates, i.e. $\mathbf{R}^n_+ = \{\mathbf{w} \in \mathbf{R}^n : w_i \geq 0;\ 1 \leq i \leq n\}$. We will also use the notation $f(\mathbf{w}) := (f(w_1), \ldots, f(w_n))$ where $\mathbf{w} \in \mathbf{R}^n$ and $f : \mathbf{R} \to \mathbf{R}$.

We consider a penalty function $P$ of the form

$$P : X \times \mathbf{R}^n_+ \to \mathbf{R}$$

where $X = X_1 \times X_2 \times \ldots \times X_n$ and each $X_i$ is a metric space with a metric $d_i$. We will analyze the optimization problem

$$\min_{\mathbf{x} \in F} P(\mathbf{x}; \mathbf{w}) \tag{1}$$

where $w \in \mathbf{R}^n_+$ is fixed and $F \subset X$ is a set of **feasible solutions**.

Let us suppose that $w_1, w_2, \ldots, w_n$ are measured on some scale of measurement. Then we say that $\mathbf{x} \in F$ is a **meaningful optimal solution** for the problem (1) if $x$ is an optimal solution for the problem (1) and the conclusion of optimality is a meaningful statement. More formally, $x \in F$ is a meaningful optimal solution for the problem (1) if and only if

$$P(\mathbf{x}; \mathbf{w}) \leq P(\mathbf{z}; \mathbf{w}) \Leftrightarrow P(\mathbf{x}; f(\mathbf{w})) \leq P(\mathbf{z}; f(\mathbf{w}))$$

for every $z \in F$, and every admissible transformation $f$ of the scale of measurement of $w_1, w_2, \ldots, w_n$.

We will also say that $w_1, w_2, \ldots, w_n$ from the problem (1) **can be measured** on some scale of measurement $\mathcal{S}$ if every optimal solution $\mathbf{x}$ is a meaningful optimal solution as long as $w_1, w_2, \ldots, w_n$ are measured on a scale $\mathcal{S}$.

Throughout we will assume that $P$ has the following property: For every $i$ there exist $a_i \in X_i$ such that

$$\forall \mathbf{x} \in X, \forall \mathbf{w} \in \mathbf{R}^n_+ : \quad P(x_1, \ldots, x_{i-1}, a_i, x_{i+1}, \ldots, x_n; \mathbf{w}) \leq P(\mathbf{x}; \mathbf{w}) \tag{2}$$

We call $a_i$ a **desirable value** for variable $i$.

Let us make some additional "reasonable" assumptions about the function $P$. We will say that $\mathbf{x} =_i \mathbf{y}$ if $x, y \in X$ and $x_j = y_j$ for all $j \neq i$.

First we assume that $P$ is increasing as the variable $i$ is moving away from its desired value $a_i$. In other words, for every $i$:

$$\forall \mathbf{x}, \mathbf{y} \in X, \mathbf{x} =_i \mathbf{y}, \forall \mathbf{w} \in \mathbf{R}^n_+ : \quad P(\mathbf{x}; \mathbf{w}) \leq P(\mathbf{y}; \mathbf{w}) \Leftrightarrow d_i(x_i, a_i) \leq d_i(y_i, a_i). \tag{3}$$

Note that this condition is equivalent to saying that $P(\mathbf{x}; \mathbf{w})$ is

$$P(d_1(x_1, a_1), \ldots, d_n(x_n, a_n); \mathbf{w})$$

where $P$ is nondecreasing in each of the first $n$ coordinates.

Obviously, if $\mathbf{a} = (a_1, \ldots, a_n) \in F$ then the minimum is attained at $\mathbf{a}$ and we may assume that $P(\mathbf{a}; \mathbf{w}) = 0$ (otherwise we just define $P'(\mathbf{x}; \mathbf{w}) := P(\mathbf{x}; \mathbf{w}) - P(\mathbf{a}; \mathbf{w})$) and, therefore, $P \geq 0$.

We will also assume that $P$ is nondecreasing if any of the $w_i$ is increased:

$$\forall \epsilon > 0, \forall \mathbf{x} \in X, \forall \mathbf{w} \in \mathbf{R}_+^n : \quad P(\mathbf{x}, \mathbf{w}) \leq P(\mathbf{x}, \mathbf{w} + \epsilon \mathbf{e_i}) \tag{4}$$

where $\mathbf{e_i} = (0, \ldots, 0, 1, 0, \ldots, 0)$ (1 is on $i$th position).

We say that $P(\mathbf{x}; \mathbf{w})$ is a **regular penalty function** if $\exists a_1 \in X_1, a_2 \in X_2, \ldots, a_n \in X_n$ such that (2),(3), and (4) hold for every $i = 1, 2, \ldots, n$

The whole setup might look artificial but a wide range of combinatorial optimization problems can be formulated in this way.

**Example:** Many graph optimization problems (shortest path, optimal spanning tree, maximal matching, TSP, minimal cut, etc.) can be formulated as 0-1 programming problems where we need to optimize an objective function $f(x) = \mathbf{w}^T\mathbf{x}$, where $\mathbf{x}$ is an incidence vector of the edge set and $\mathbf{w}$ is the vector of corresponding weights of the edges. If we define $P(\mathbf{x}; \mathbf{w}) = \mathbf{w}^T\mathbf{x}$, all the assumptions are satisfied with $X = \{0, 1\}^n$ (where $n$ is the number of edges in the graph) and $a_i = 0$ for $w_i \geq 0$ and $a_i = 1$ for $w_i \leq 0$ in the case of a minimization problem (Note that if there are negative weights we can't assume $P \geq 0$.). The set of feasible solutions $F$ is the set of all incidence vectors of feasible structures (paths, trees, matchings, hamiltonian cycles, cuts, etc.).

**Example:** Note that any linear (integer, 0-1) programming problem can fit in the above setup. Moreover, this can be generalized for nonlinear objective functions. For example, every objective function of the form

$$\sum c_{i_1, i_2, \ldots, i_k}(\mathbf{w}) f_{i_1, i_2, \ldots, i_k}(x_{i_1}, \ldots, x_{i_k}),$$

where $c_{i_1, i_2, \ldots, i_k}(\mathbf{w}) : \mathbf{R}_+^n \to \mathbf{R}_+$ and $f_{i_1, i_2, \ldots, i_k}(x_{i_1}, \ldots, x_{i_k}) : \mathbf{R}_+^k \to \mathbf{R}_+$ are increasing functions in every coordinate, satisfies all the assumptions for $P$.

# 3 Optimization

As mentioned in the previous section, if $\mathbf{a} = (a_1, \ldots, a_n) \in F$, then the problem is trivial since P achieves a global minimum at $\mathbf{a}$. In general, any optimal solution will have the property that $d_i(x_i, a_i)$ is "as small as possible" and variables with higher priority "will tend to be closer" to their desired value than variables with lower priority. One may try to attack this problem using a greedy heuristic:

**Greedy Algorithm:**

1. Rename all the variables such that $w_1 \geq w_2 \geq \ldots \geq w_n$. Let $i = 1$.
2. Find $x_i$ such that
   $d_i(x_i, a_i) = \min\{d_i(y_i, a_i) : \exists y_i, \ldots, y_n \text{ } s.t. \text{ } (x_1, \ldots, x_{i-1}, y_i, \ldots, y_n) \in F\}$.
   If this set is empty then STOP (there is no feasible solution).
3. If $i < n$ then let $i := i + 1$ and goto Step 2.
4. STOP ($\mathbf{x} = (x_1, \ldots, x_n)$ is a greedy solution.)

Note that this algorithm will produce a feasible solution if $F \neq \emptyset$. The algorithm is not necessarily fast; in some cases determining $x_i$ might be an easy task but it might be as hard as the original problem if the set $F$ is complicated. Moreover, it is clear that there is no guarantee that the algorithm will ever produce an optimal solution.

We say that $\mathbf{x} \in F$ is a **greedy solution** if it can be produced by the greedy algorithm. A greedy solution is not uniquely determined in general. If there are more variables having the same priority then there is more than one way how to rename them which might affect later decisions. Similarly, there can be infinitely many choices for $x_i$ and any particular choice might also affect later decisions. On the other hand, in many instances a greedy solution will be unique.

**Proposition 1** *Suppose that $w_i \neq w_j$ for all $i \neq j$ and suppose that for every $k$ and every $c \geq 0$ the equation $d_k(y_k, a_k) = c$ has at most one solution $y_k \in X_k$. Then there can be at most one greedy solution.*

**Proof:** Note that Step 1 is uniquely determined since all $w_i$'s are different. Also $x_k$ is uniquely determined if it exists. ∎

Whenever $w_1, \ldots, w_n$ are measured on an ordinal scale, something more can be said about greedy solution(s). However, we will need to impose new restrictions on $P$.

We say that $P(\mathbf{x};\mathbf{w})$ is **w-unbounded** if for every $i = 1, \ldots, n$:

$$\forall \mathbf{w} \in \mathbf{R}^n_+, \forall \mathbf{x}, \mathbf{y} \in X, \mathbf{x} \neq \mathbf{y}, x =_i \mathbf{y} : \lim_{\mathbf{w}_i \to \infty} |P(\mathbf{x};\mathbf{w}) - P(\mathbf{y};\mathbf{w})| = \infty \tag{5}$$

Roughly speaking, $P$ is $\mathbf{w}$-unbounded if there can be an unbounded increase of $P$ if we move variable $i$ away from its desired value. Note that for regular penalty functions this limit always exists.

$P(\mathbf{x};\mathbf{w})$ is **controllable on** $X$ if moving a variable with small priority cannot result in big changes in penalty. More precisely, for every $i = 1, \ldots, n$:

$$\forall \mathbf{x}, \mathbf{y} \in X, \mathbf{x} =_i \mathbf{y}, \exists M_i(\mathbf{x}, \mathbf{y}), \forall \mathbf{w} : \lim_{\mathbf{w}_i \to 0} |P(\mathbf{x};\mathbf{w}) - P(\mathbf{y}, \mathbf{w})| \leq M_i \tag{6}$$

Now we can state and prove the main result:

**Theorem 2** *Consider optimization problem (1) where $P$ is regular, $\mathbf{w}$-unbounded and controllable on $X$. If $w_1, \ldots, w_n$ are measured on an ordinal scale then any meaningful optimal solution can be produced by the greedy algorithm.*

**Proof:** We first rename all the indices such that

$$w_1 \geq w_2 \geq \ldots \geq w_n \geq 0. \tag{7}$$

Note that this can be done uniquely if and only if all $w_i$'s are different. We say that indexing is **proper** if (7) holds. Let $\mathbf{z}$ be a meaningful optimal solution. Let $i$ be the smallest index such that for any proper indexing there is no greedy solution $\mathbf{x}$ with $x_k = z_k$ for all $k = 1, \ldots, i$. If there is no such $i$ then $\mathbf{z}$ is one greedy solution. Note that $i \neq n$ since otherwise (3) and Step 2 of the algorithm would give $P(\mathbf{x}) < P(\mathbf{z})$. Therefore, we may assume $i < n$. Let $\mathbf{x}$ be a greedy solution for which $x_k = z_k$ for all $k = 1, \ldots, i-1$. Now,

$$\begin{aligned} P(\mathbf{z};\mathbf{w}) - P(\mathbf{x};\mathbf{w}) \;\; = \;\; & P(\mathbf{z};\mathbf{w}) - P(z_1, \ldots, z_{n-1}, x_n;\mathbf{w}) \\ & + P(z_1, \ldots, z_{n-1}, x_n;\mathbf{w}) - P(z_1, \ldots, z_{n-2}, x_{n-1}, x_n;\mathbf{w}) \\ & \vdots \\ & + P(z_1, \ldots, z_i, x_i + 1, \ldots, x_n;\mathbf{w}) - P(\mathbf{x};\mathbf{w}). \end{aligned}$$

Since $w_i$ are measured on an ordinal scale we can choose any $\mathbf{w}$ satisfying (7) and maintain the optimality of $z$. We choose $w_k$, $k = i+1, \ldots, n$, such that (7) holds and such that

$$|P(z_1, \ldots, z_{k-1}, z_k, x_{k+1}, \ldots, x_n;\mathbf{w}) - P(z_1, \ldots, z_{k-1}, x_k, x_{k+1}, \ldots, x_n;\mathbf{w})| \leq M_k.$$

We can do the latter because $P$ is controllable on $X$. Now we have:

$$\begin{aligned} P(\mathbf{z};\mathbf{w}) - P(\mathbf{x};\mathbf{w}) \;\; \geq \;\; & -M_n - \ldots - M_{i+2} - M_{i+1} \\ & + P(z_1, \ldots, z_i, x_i + 1, \ldots, x_n;\mathbf{w}) - P(\mathbf{x};\mathbf{w}). \end{aligned}$$

Since $P$ is $\mathbf{w}$-unbounded and since $x_k = z_k$ for $k = 1, \ldots, i-1$, we can choose $w_i$ such that

$$P(z_1, \ldots, z_i, x_i + 1, \ldots, x_n;\mathbf{w}) - P(\mathbf{x};\mathbf{w}) > M_{i+1} + M_{i+2} + \ldots + M_n$$

and all other $w_k$ ($k = 1, \ldots, i-1$) to satisfy (7). Then we have $P(\mathbf{z};\mathbf{w}) - P(\mathbf{x};\mathbf{w}) > 0$ which is a contradiction of the optimality of $\mathbf{z}$ ∎

**Corollary 3** *Under the conditions of Theorem 2 and Proposition 1 the following are equivalent:*

1. *$w_1, \ldots, w_n$ can be measured on an ordinal scale of measurement*
2. *Every meaningful optimal solution is greedy.*

**Proof:** Obviously, Theorem 2 gives 1. $\Rightarrow$ 2.. The converse is trivial since there is a unique greedy solution by Proposition 1 and the greedy solution is completely determined by the ordering of $w_i$'s. ▮

**Corollary 4** *Let $\mathcal{F} \subset 2^X$ be a matroid. Consider the problem (1) where $F$ is the set of all incidence vestors of independent sets in $\mathcal{F}$ and $P(\mathbf{x};\mathbf{w}) = \mathbf{w}^T\mathbf{x}$. Then $w_1,\ldots,w_n$ can be measured on an ordinal scale.*

**Proof:** It is well known that the greedy algorithm gives all optimal solutions ([3]) and every greedy solution is completely determined by the ordering of $w_i$'s. ▮

In order to conclude that every meaningful optimal solution (if such a solution exsists) is greedy, no information about the set of feasible solutions $F$ is needed and the penalty function $P$ doesn't need to be linear. Therefore, we can't hope for any kind of underlying structure on $X$ just because $w_1,\ldots,w_n$ are measured on an ordinal scale. However, if we have an additional information about the problem (as in the matroid case) it might be much easier to prove that the greedy algorithm gives an optimal solution than to prove that $w_1,\ldots,w_n$ can be measured on an ordinal scale.

Note that we can't hope that Theorem 2 will be true under much weaker conditions on $P$. For example, we need $\mathbf{w}$-unboundedness:

**Example:** Let $X = \{0,1\}^3$. Let $F = \{(1,0,0),(0,1,1),(1,0,1),(1,1,0),(1,1,1)\}$, $P = \mathbf{c}^T\mathbf{x}$ where $c_i = 2 - 1/(1+w_i)$, and let $w_1 > w_2 > w_3 \geq 0$. Note that $P$ is regular and controllable on $X$ but it is not $\mathbf{w}$-unbounded. Then it is easy to check that the greedy solution $(0,1,1)$ is never optimal and the optimal one is $(1,0,0)$. Note that this is a meaningful optimal solution as long as we measure priorities on an ordinal scale of measurement.

Unfortunately, a solution produced by the Greedy Algorithm might not be unique. Rules for resolving situations where there are few ways to continue execution of the algorithm may vary from problem to problem. However, as Proposition 1 shows, a greedy solution will be unique for a large class of optimization problems. For example, for any 0-1 programming problem

$$\min_{\mathbf{x}\in F}\sum_{i=1}^{n} w_i x_i; \quad F \subset \{0,1\}^n$$

the greedy solution will be unique provided that $w_i$ are all different. Clearly, $P(\mathbf{x};\mathbf{w}) = \mathbf{w}^T\mathbf{x}$ is regular, $\mathbf{w}$-unbounded and controllable on $\{0,1\}^n$.

In some cases ties among $w_i$ can be resolved arbitrarily. An example of such a case is the minimal weight spanning tree problem. Since all the spanning trees are maximal independent sets in a specific matroid structure on $E(G)$, we know that the greedy algorithm will produce all optimal solutions ([3]). However, it can be shown directly that if weights $w_i$ are measured on an ordinal scale then every optimal solution is meaningful. More precisely, if we define $F$ to be the set of all incidence vectors of spanning trees in $G$, then it can be shown that for every $\mathbf{x},\mathbf{y} \in F \subset \{0,1\}^n$ and every increasing $f$:

$$\sum_{i=1}^{n} w_i x_i \leq \sum_{i=1}^{n} w_i y_i \Leftrightarrow \sum_{i=1}^{n} f(w_i) x_i \leq \sum_{i=1}^{n} f(w_i) y_i.$$

Therefore, by Theorem 2, any optimal solution is a greedy one. Note that the Greedy Algorithm in this case is exactly Kruskal's algorithm. If there are ties among $w_i$ there will be more than one greedy solution but it is not hard to show that every greedy solution has the same value. Thus, every greedy solution is optimal and possible ties can be resolved arbitrarily.

Another example of this type is a well known scheduling problem denoted in standard notation by $1|p_j|\sum C_j$ (e.g. see [2]): we are given $n$ jobs with processing times $p_j$ which need to be scheduled on a single machine (one at a time) so that the sum of their completion times (or avarage completion time) is minimal. Without loss of generality we may assume that $p_1 \geq p_2 \geq \ldots \geq p_n$. Note that this problem can be formulated as the optimization problem

$$\min_{\mathbf{x}\in F}\mathbf{p}^T\mathbf{x}$$

where $X = \{1, \ldots, n\}^n$ and $F \subset X$ is the set of all $\mathbf{x} \in X$ having all coordinates different. $F$ is in obvious one to one correspondence with all permutations on $n$ elements. Any $\mathbf{x} \in F$ represents a schedule where $x_i = k$ if and only if the $i$-th job is scheduled $(n+1-k)$-th ($k$-th from the end). It is not *a priori* clear that the conclusion of optimality is meaningful if we measure priorities on an ordinal scale (for example if we don't know exact processing times but, for all $i$ and $j$, we know if job $i$ has longer processing time than job $j$). If $\mathbf{x}$ is an optimal solution then $p_i < p_j \Leftrightarrow x_i > x_j$ (otherwise we switch $x_i$ and $x_j$). Since $p_i < p_j \Leftrightarrow f(p_i) < f(p_j)$ for every increasing $f$, we conclude that $\mathbf{x} = (1, 2, \ldots, n)$ is a meaningful optimal solution and, by Theorem 2, a greedy one also. Since every two greedy solutions have the same value (ties among $p_i$ can be resolved arbitrarily) we conclude that every greedy solution is an optimal solution. Note that greedy solutions schedule jobs in the shortest processing time (SPT) order.

Optimal solutions for both minimal weight spanning tree, and the $1|p_j|\sum C_j$ problem can be found easily without observing that the conclusion of optimality is meaningful if priorities are measured on an ordinal scale. Sometimes it is straightforward to realize if the Greedy Algorithm gives an optimal solution. However, for most of the non-trivial problems this is not the case. Theorem 2 suggests the new approach which might be useful in some cases. If conditions of Theorem 2 hold, search for an optimal solution can be reduced to greedy solutions only. In many cases, only a slight modification of the Greeedy Algorithm will eliminate greedy solutions which are not optimal. Therefore, information that priorities can be measured on an ordinal scale is valuable and can be the main step towards finding an efficient optimization algorithm.

Mahadev, Pekec, and Roberts [7, 8] consider a specific scheduling problem arising from the problems at the Air Mobility Command of the Unites States Air Force: $n$ items need to be assigned arrival dates in such a way that no two items have the same arrival date. Every item $i$ has a priority $t_i$ and a desired arrival date $a_i$. Every schedule can be represented as $\mathbf{x} \in X = \mathbf{N}^n$ where $x_i$ represents the scheduled arrival time of the item $i$. Mahadev, Pekec and Roberts analyze the meaningfulness of the conclusion of optimality for a variety of Penalty functions. For example they show in [8] that for the problem where $a_1 = a_2 = \ldots = a_{n-1} < a_n$ and where the penalty for a schedule $\mathbf{x}$ is given by

$$P(\mathbf{x}; \mathbf{t}) = \sum_{i=1}^{n} h(t_i)|x_i - a_i|$$

for some increasing function $h$, the conclusion of optimality is meaningful if priorities are measured on an ordinal scale. Since $P$ is a regular (with desired values $a_i$), $\mathbf{t}$-increasing function which is controllable on $X$, by Theorem 2, we conclude that every optimal solution is given by the Greedy Algorithm. It is not hard to show that ties among $t_i$ can be resolved arbitrarily. However, if there is a choice for $x_i$ (in Step 2 of the algorithm), we cannot pick $x_i$ arbitrarily. For example, if $a_1 = a_2 = 2, a_3 = 3$, $t_1 > t_2 > t_3$ and $h(t) = t$, then both $\mathbf{x} = (2, 3, 4)$ and $\mathbf{y} = (2, 1, 3)$ are greedy solutions but $\mathbf{x}$ is not an optimal solution. Note that, for any $a_i$ and $c$, the equation $|x_i - a_i| = c$ can have at most two solutions. If we always pick the smaller of at most two candidates in Step 2, a greedy solution will be an optimal solution.

Theorem 2 can be interpreted as a "negative" result also. If we consider a problem where the greedy algorithm doesn't in general give any optimal solution (almost every problem is like that, e.g. TSP) and $w_i$ are measured on an ordinal scale, then we conclude that there can't be a meaningful optimal solution. In other words, it doesn't make much sense to ask for an optimal solution. Therefore, for every non-trivial problem (non-trivial in the sense that the greedy solution is not necessarily an optimal solution) where $w_i$ are measured on an ordinal scale, the conclusion of optimality is a meaningless conclusion.

# 4 Closing Remarks

It should be noted that the conditions of Theorem 2 are satisfied for a wide range of functions $P$, and that there are no conditions whatsoever on the structure of the set of feasible solutions $F \subset X$. The main reason that every optimal solution is a greedy one is in the fact that the $w_i$ are measured on an ordinal scale. Our main result has nothing to do with the well known result that

$$\max_{S\in\mathcal{F}} \sum_{e\in S} w_e$$

can be maximized by the greedy algorithm whenever all the feasible solutions $S \subset X$ are independent sets of some matroid structure on $X$ ([3]). In that case, the greedy algorithm produces an optimal solution because of the special form of the objective function together with a special structure of the set of feasible solutions.

The purpose of this paper was not to give explicit algorithms for solving certain optimization problems but to show that the information about a problem itself, such as the way of measuring data, which appears to be lost after the problem is formalized, can be crucial for understanding of the problem. We also try to exhibit the link between ordinal scales and greedy algorithms, i.e., connections between two seemingly unrelated areas: measurement theory and combinatorial optimization.

Our conclusions have limitations due to the setup of the optimization problem and it would be interesting to see where are the limits for the "equivalence" of greedy algorithms and ordinal scales. It would be interesting to know if anything can be said about the execution of the greedy algorithm in the cases when the algorithm doesn't determine a unique solution: maybe some additional rule would determine a unique greedy solution for a wide range of optimization problems.

The results presented seem to have more implications for practice than for theory since the main message is that if the relation among the data is as vague as preference lists, the only meaningful approach is a greedy approach regardless of any additional information about the problem. On the other hand, real life preference lists still have some fixed relations in the sense that it is not true that any ordered set of numbers can present it (i.e. $1 < 2 < 5$ might be acceptable but $10^{-10} < 1 < 10^{10}$ might not be an acceptable representation of the preference list $w_1 < w_2 < w_3$.). This problem remains to be understood.

**Acknowledgement**

The author thanks Dr. Fred S. Roberts for his helpful comments.

# References

[1] M.R. Garey and D.S. Johnson, Computers and intractability: a guide to the theory of NP-completeness. *W.H. Freeman, San Francisco* (1979).

[2] R.L. Graham, E.L. Lawler, J.K. Lenstra, and A.H.G Rinooy Kan, Optimization and approximation in deterministic sequencing and scheduling: a survey. *Annals of Discr. Math.* **5** (1979), 287-326.

[3] B. Korte, L. Lovász, and R. Schrader, Greedoids. *Springer-Verlag, NY* (1991).

[4] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky, Foundations of Measurement, Vol. I. *Academic Press, NY* (1971).

[5] J.B. Kruskal, On the shortest spanning tree of a graph and the traveling salesman problem. *Proc. Amer.Math.Soc.* **7** (1956), 48-50.

[6] R.D. Luce, D.H. Krantz, P. Suppes, and A. Tversky, Foundations of Measurement, Vol. III. *Academic Press, NY* (1990).

[7] N.V.R. Mahadev, A. Pekec, and F.S. Roberts, On the Scheduling Problem with Priorities and Lateness/Earliness Penalties: When is an Optimal Solution Not Optimal? *Research Report, Rutgers Center for Operations Research* **RRR 19-94** (1994).

[8] N.V.R. Mahadev, A. Pekec, and F.S. Roberts, On the Scheduling Problem with Priorities and Lateness/Earliness Penalties: Some Meaningful Conclusions with Nonconstant Arrival Times. *Research Report, Rutgers Center for Operations Research* **RRR 20-94** (1994).

[9] L. Narens, Abstract Measurement Theory. *MIT Press, Cambridge, MA* (1985).

[10] J. Pfanzagl, Theory of Measurement *Wiley, NY* (1968).

[11] F.S. Roberts, Measurement Theory. *Addison-Wesley, Reading, MA* (1979).

[12] F.S. Roberts, Meaningfulness of conclusions from combinatorial optimization. *Discr. Appl. Math.* **29** (1990), 221-241.

[13] F.S. Roberts, Limitations of conclusions using scales of measurement. *Handbooks in OR & MS, S.M. Pollock, M.H. Rothkopf, and A. Barnett,Eds.* **6** (1994), 621-671

[14] P. Suppes, D.H. Krantz, R.D. Luce, and A. Tversky, Foundations of Measurement, Vol. II. *Academic Press, NY* (1989).