Resenhas IME-USP 1998, Vol. 3, No. 3, 323 - 349.

The Complexity of Computing Medians of Relations[1]

Yoshiko Wakabayashi

Abstract: Let N be a finite set and $\mathcal{R}$ be the set of all binary relations on N. Consider $\mathcal{R}$ endowed with a metric d, the symmetric difference distance. For a given m-tuple $\Pi = (R_1, \ldots, R_m) \in \mathcal{R}^m$, a relation $R^* \in \mathcal{R}$ that minimizes the function $\sum_{k=1}^{m} d(R_k, R)$ is called a *median* relation of Π. In the social sciences, in qualitative data analysis and in multicriteria decision making, problems occur in which the m-tuple Π represents collected data (preferences, similarities, games) and the objective is that of finding a median relation of Π with some special feature (representing for example, consensus of preferences, clustering of similar objects, ranking of teams, etc.). In this paper we analyse the computational complexity of all such problems in which the median is required to satisfy one or more of the properties: reflexivity, symmetry, antisymmetry, transitivity and completeness. We prove that whenever transitivity is required (except when symmetry and completeness are also simultaneously required) then the corresponding median problem is $\mathcal{NP}$-hard. In some cases we prove that they remain $\mathcal{NP}$-hard even when the profile Π consists of one or two relations. We mention some applications and strategies that can be used to solve the median problems considered here.

Key words: Relations, clustering, complexity, median, order, transitivity

1. Introduction

In the social choice theory a classical problem that has been largely investigated and whose origin traces back to the late eighteenth century is the problem of aggregating individual preferences (linear orders) into a social preference (a linear order). The notion of consensus of preferences plays an important role in the social sciences, a reason why many efforts have been made to find realistic models to express it (*cf.* Leclerc [1988a], Day [1988]).

The first mathematical approaches on problems of aggregation of preferences are credited to Borda in 1770 and Condorcet in 1785, both concerned with the design of election procedures. Young [1990] discusses the model proposed by these two major figures of that time, gives some historical accounts and explains the Condorcet's theory of voting (see also Young and Levenglick [1978]).

The notion of median relation —a relation minimizing a "remoteness" function defined in terms of the symmetric difference distance— was introduced by Kemeny [1959], who investigated a method to aggregate individual preferences into a collective preference. His method, although being of metric nature, is in fact

[1]This work has been partially supported by CNPq Grant 304527/89-0, FAPESP Grant 96/04505-2 and ProNEx Project 107/97 (MCT/FINEP).

equivalent to the Condorcet's majority rule, according to which the winning collective preference should be the one supported by the largest number of votes (Young [1990], Barbut [1967], Fishburn [1977], Michaud [1987]). In cluster analysis a similar approach was proposed by Regnier [1965], then Mirkin [1974], for solving the problem of aggregating equivalence relations into an equivalence relation (see also Zahn [1964]).

The fact that the symmetric difference distance has been used in problems occurring in many different contexts is not a pure coincidence. Axiomatics supporting its use has been investigated in several cases, *cf.* Kemeny [1959], Monjardet [1978], Barthélemy [1979] and Barthélemy and Monjardet [1981]. However, the median approach, as any consensus procedure, has some defects as pointed out by Fishburn [1977], Leclerc [1988a], and Barthélemy and Monjardet [1988]. This last reference gives also an overview of the developments on the algorithmic approaches and extensions of the notion of median in other structures. The results concerning its algebraic definition that generalizes to any distributive lattice (*cf.* Barbut [1961], Monjardet [1980]), as well as more recent results on median semilattices, resp. (semi)modular (semi)lattices can be found in Monjardet [1987, 1988], resp. Barthélemy [1981] and Leclerc [1988b]. For a unified treatment of this subject the reader should refer to Barthélemy, Flament and Monjardet [1982]; Barthélemy, Leclerc and Monjardet [1986]; Barthélemy and Monjardet [1988] and Barthélemy [1988].

In this paper we analyse the computational complexity of a class of problems of finding medians with prescribed properties. This class includes those classical problems such as aggregation of preferences and clustering.

The material is organized as follows. In Section 2 we give the definitions and notation to be used and present the problems to be investigated. In Section 3 the main results on the computational complexity of these problems are presented, and in Section 4 we discuss special cases concerning restricted domains. In Section 5 we mention some applications and known strategies to solve median problems.

2. Definitions and Notation

Let N be a finite set with n objects and let $\mathcal{R}$ denote the set of all (binary) relations on N. Consider $\mathcal{R}$ endowed with a metric d, the symmetric difference distance, defined as

$$d(R,S) := |R \triangle S| := |R \cup S| - |R \cap S| \quad \text{for all } R, S \in \mathcal{R}.$$

A *profile* of relations in $\mathcal{R}$, or a profile in $\mathcal{R}^m$, is an m-tuple $\Pi = (R_1, \dots, R_m)$ where $R_k \in \mathcal{R}$ for $k = 1, \dots, m$. Given a profile $\Pi = (R_1, \dots, R_m)$ in $\mathcal{R}^m$, a relation $R^* \in \mathcal{R}$ that minimizes the function

$$D(\Pi, R) := \sum_{k=1}^{m} d(R_k, R)$$

is called a *median relation* of Π.

In this general form the problem of finding a median of a given profile is trivial and not interesting. However, if we require the median to satisfy certain properties the resulting problem becomes interesting and has nice applications. So, according to the desired properties of R^* we obtain different problems, and here we consider all those arising when the properties are chosen from the set

$$\mathbb{P} := \{\text{Reflexive, Symmetric, Antisymmetric, Transitive, Total}\}.$$

Let us recall some definitions. A relation $R \in \mathcal{R}$ is *reflexive* (REF) if $(i,i) \in R$ for all $i \in N$; R is *symmetric* (SYM) if $(i,j) \in R$ implies $(j,i) \in R$ for all $i,j \in N$; R is *antisymmetric* (ASY) if $(i,j) \in R$ and $(j,i) \in R$ imply $i=j$ for all $i,j \in N$; R is *transitive* (TRA) if $(i,j) \in R$ and $(j,k) \in R$ imply $(i,k) \in R$ for all $i,j,k \in N$; R is *total* (TOT) if $(i,j) \in R$ or $(j,i) \in R$ for all $i,j \in N, i \neq j$.

To simplify notation we use the abbreviated form of the *name* of the property (given in parentheses) to denote also the *set* of all relations having this property. Thus, for example, TRA denotes the set of all transitive relations in $\mathcal{R}$. Some relations having more than one of the properties in $\mathbb{P}$ are known by special names, not always standard in the literature. Here we adopt the following notation and terminology:

$\mathcal{C}$ denotes the set of all *complete preorders, i.e.* $\mathcal{C} = \text{TRA} \cap \text{TOT}$.
$\mathcal{T}$ denotes the set of all *tournaments, i.e.* $\mathcal{T} = \text{ASY} \cap \text{TOT}$.
$\mathcal{L}$ denotes the set of all *linear orders, i.e.* $\mathcal{L} = \text{ASY} \cap \text{TRA} \cap \text{TOT}$.
$\mathcal{O}$ denotes the set of all *partial orders, i.e.* $\mathcal{O} = \text{ASY} \cap \text{TRA}$.
$\mathcal{E}$ denotes the set of all *equivalence* relations, *i.e.* $\mathcal{E} = \text{REF} \cap \text{SYM} \cap \text{TRA}$.

For a subset $\mathcal{M} \subset \mathcal{R}$ the median problem relative to $\mathcal{M}$, denoted by $\text{MP}(\mathcal{R}, \mathcal{M})$, is defined as follows.

Median Problem relative to $\mathcal{M}$ — $\text{MP}(\mathcal{R}, \mathcal{M})$

Instance: Profile $\Pi = (R_1, \dots, R_m)$ of m relations in $\mathcal{R}$.
Objective: Find a relation $R^* \in \mathcal{M}$ such that $D(\Pi, R^*) = \min_{R \in \mathcal{M}} D(\Pi, R)$.

We expect the reader to be familiar with the basic concepts of graph theory and complexity theory. If not, the definitions not given here can be found in Bondy and Murty [1976], resp. Garey and Johnson [1979]. We present only the concepts we need to establish out notation.

A *graph* G with node set V and *edge* set E is denoted by $G = [V, E]$. A *digraph* (or directed graph) D with node set N and *arc* set A is denoted by $D = (N, A)$. A graph $G = [V, E]$, resp. digraph $D = (N, A)$, is called *complete* if $E = \{\{u,v\} : u,v \in V, u \neq v\}$, resp. $A = \{(u,v) : u,v \in N, u \neq v\}$. If $D = (N, A)$ is a digraph with $A = N \times N$ then D is called *l-complete* (*i.e.* complete with all loops). For a digraph $D = (N, A)$, we call the arcs in $(N \times N) \setminus A$ *missing arcs* (analogously, missing edges in case of a graph). A digraph is called *acyclic* if it does not contain any directed cycle. A *clique* of a graph is a complete subgraph of G. It needs not be maximal, as is sometimes assumed in the literature. A set of edges A in a graph $G = [V, E]$ is called a *clique partitioning* of G if there is a partition $V_1, \dots, V_k$ of V such that the subgraph induced by each V_i, $1 \le i \le k$, is a clique in G and A is the union of all edges in G with both endnodes in the same set of the partition. In this case, if for $1 \le i \le k$ the clique induced by V_i is denoted by Q_i, then we say that $\mathcal{C}(A) := \{Q_1, \dots, Q_k\}$ is the *clique set defined by* A.

3. Computational Complexity

We assume here that an instance of the median problem $\mathrm{MP}(\mathcal{R}, \mathcal{M})$ consisting of a profile $\Pi = (R_1, \dots, R_m)$ is given by an (n^2, m)-matrix $A = (a_{ek})$, where the rows correspond to the pairs $e \in N \times N$, the columns correspond to the relations $R_1, \dots, R_m$, and $a_{e,k} = 1$ if $e \in R_k$; $a_{e,k} = 0$ if $e \notin R_k$, $k = 1, \dots, m$. That is, each column k of A corresponds to the characteristic vector of the relation R_k. Clearly the size of such an instance is $\mathcal{O}(n^2 m)$.

It is well-known that the median problems we are considering can be formulated as 0/1 linear programs or optimization problems on weighted digraphs (see Grötschel and Wakabayashi [1988]). In fact, it is not difficult to prove that

$$D(\Pi, R) = \sum_{(i,j)} w_{ij} r_{ij} + \sum_{(i,j)} \alpha_{ij},$$

where

$$\alpha_{ij} := |\{k : (i,j) \in R_k\}|, \tag{3.1}$$

$$w_{ij} := m - 2\alpha_{ij} \quad \text{and} \tag{3.2}$$

$$r = (r_{ij}) \text{ is the characteristic vector of } R.$$

Thus, each given instance of $\mathrm{MP}(\mathcal{R}, \mathcal{M})$ can be formulated as the 0/1 linear program:

$$\begin{aligned} &\text{minimize} \sum_{(i,j)} w_{ij} r_{ij} \\ &\text{subject to: } r = (r_{ij}) \text{ is the characteristic vector of some relation } R \in \mathcal{M}. \end{aligned} \tag{3.3}$$

If the coefficients w_{ij} are interpreted as being weights associated with the arcs (i,j) of an l-complete digraph D_n on the node set N, then the problem becomes that of finding a minimum weighted subdigraph $D' = (N, R)$ of D_n, where $R \in \mathcal{M}$. For example, if $\mathcal{M} = \mathcal{L}$ the corresponding digraph problem is a special case of the weighted *feedback arc set problem* or *linear ordering problem*, and if $\mathcal{M} = \mathcal{E}$ we obtain the so-called *clique partitioning problem* (see Reinelt [1985], Grötschel, Jünger and Reinelt [1985], Barthélemy, Guenoche and Hudry [1988], resp. Wakabayashi [1986] and Grötschel and Wakabayashi [1988]).

From the above reduction one obtains immediately the following result (excluding some trivial non-interesting cases).

Proposition 3.4. *If* $\mathcal{M} \in \{\text{SYM, ASY, TOT, ASY} \cap \text{TOT}\}$ *then the median problem* $\text{MP}(\mathcal{R}, \mathcal{M})$ *is polynomially solvable.*

We can also make use of the given reduction, in a more specialized way, to show that $\text{MP}(\mathcal{R}, \mathcal{M})$ is $\mathcal{NP}$-hard for other subsets $\mathcal{M}$. Namely, we first note that the obtained digraph optimization problems are special in the sense that all of its weights w_{ij} are integers having the same parity. Furthermore, we observe that whenever we have such an l-complete weighted digraph $D_n = (N, A_n)$ with $m := \max_{e \in A_n} |w_e|$ we can construct a profile $\Pi = (R_1, \ldots, R_m)$ in $\mathcal{R}^m$ such that each $(i,j) \in N \times N$ belongs to precisely α_{ij} relations, where $\alpha_{ij} = (m - w_{ij})/2$ (see (3.1) and 3.2). In other words, these special digraph optimization problems are also reducible to $\text{MP}(\mathcal{R}, \mathcal{M})$.

In what follows we state more formally the results concerning the above reduction. Before, we introduce some notation. For each set $\mathcal{M} \subset \mathcal{R}$ we define a digraph optimization problem relative to $\mathcal{M}$ as follows.

Digraph Optimization Problem — $\text{DOP}(n, \mathcal{M}, m)$

Instance: l-complete digraph $D_n = (N, A_n)$; weights $w_e \in Z$ for each $e \in A_n$, all having the same parity and with $\max_e |w_e| = m$.

Objective: Find an arc set $A^* \subset A_n$ such that $A^* \in \mathcal{M}$ and $w(A^*) := \sum_{e \in A^*} w_e$ is minimum.

The reason to introduce these problems is justified by the following result.

Theorem 3.5. *Let* $\mathcal{M} \subset \mathcal{R}$. *If* $\text{DOP}(n, \mathcal{M}, m)$ *is* $\mathcal{NP}$*-hard and* m *is bounded by a polynomial in* n*, then* $\text{MP}(\mathcal{R}, \mathcal{M})$ *is* $\mathcal{NP}$*-hard.*

Proof. Let $D_n = (N, A_n)$, w and m be given as an instance $\mathcal{I}$ of $\text{DOP}(n, \mathcal{M}, m)$. The corresponding instance $\mathcal{I}'$ of $\text{MP}(\mathcal{R}, \mathcal{M})$ is constructed as follows. For each pair $(i,j) \in A_n$ we determine the number $\alpha_{ij} := (m - w_{ij})/2$ and set

$$R_k := \{(i,j) \in N \times N : \alpha_{ij} \geq k\}, \text{ for } k = 1, \ldots, m,$$

obtaining this way the profile $\Pi = (R_1, \ldots, R_m)$. In other words, we let (i,j) belong to the first α_{ij} relations $R_1, \ldots, R_{\alpha_{ij}}$.

The construction of the profile Π can be done in $\mathcal{O}(n^2 m)$ time. Thus, when m is bounded by a polynomial in n this construction is polynomial in the size of $\mathcal{I}$. The proof that an optimum solution of the instance $\mathcal{I}'$ gives an optimum solution of $\mathcal{I}$ is straightforward and will be omitted. $\square$

To prove the $\mathcal{NP}$-hardness of some problems, we consider the corresponding *decision version* of $\text{DOP}(n, \mathcal{M}, m)$ that will be denoted by $\text{DDP}(n, \mathcal{M}, m)$.

For technical reasons it will be convenient to consider a slight variation of the transitive relation, denoted by TRA*, defined as follows: if $(i,j) \in R$ and $(j,k) \in R$ then $(i,k) \in R$ for all $i,j,k \in N$, $i \neq j \neq k \neq i$. With this definition we can refer to the property TRA* on complete digraphs (instead of l-complete digraphs). For that, we define the corresponding digraph optimization (resp.

decision) problem DOP* (resp. DDP*), defined analogously as DOP (resp. DDP), except that the instance consists of a (loopless) complete digraph.

The next lemma shows that if we can prove an $\mathcal{NP}$-completeness result for DDP* with respect to TRA*, then we can derive an analogous result for DDP with respect to TRA (including or not the property REF). More precisely, the following holds.

Lemma 3.6. *Let $\mathcal{M}^* = S \cap \mathrm{TRA}^*$ for some relation S on N, and let*

$$\mathcal{M} \in \{S \cap \mathrm{TRA},\ S \cap \mathrm{TRA} \cap \mathrm{REF}\}.$$

If $\mathrm{DDP}^*(n, \mathcal{M}^*, m)$ is $\mathcal{NP}$-complete then $\mathrm{DDP}(n, \mathcal{M}, m)$ is $\mathcal{NP}$-complete.

Proof. Let $D_n = (N, A_n)$, w, m and k be an arbitrary instance of $\mathrm{DDP}^*(n, \mathcal{M}^*, m)$. The corresponding instance of $\mathrm{DDP}(n, \mathcal{M}, m)$, defined by D_n', w', m, k' is constructed as follows: $D_n' = (N, A_n')$ is the l-complete digraph obtained from D_n by adding to it all the missing loops, the weights w_e' are defined as:

$$w_e' := \begin{cases} w_e & \text{if} \quad e \in A_n, \\ 0 & \text{if} \quad e \notin A_n \text{ and } m \text{ is even}, \\ -1 & \text{if} \quad e \notin A_n \text{ and } m \text{ is odd}, \end{cases}$$

and

$$k' := \begin{cases} k & \text{if } m \text{ is even}, \\ k - n & \text{if } m \text{ is odd}. \end{cases}$$

We claim that D_n has an arc set B such that $B \in \mathcal{M}^*$ and $w(B) \leq k$ if and only if D_n' has an arc set B' with $B' \in \mathcal{M}$ and $w'(B') \leq k'$.

In fact, given $B \subset A_n$ take $B' := B \cup \{(i,i) : i \in N\}$; and conversely, given $B' \subset A_n'$ take $B := B' \setminus \{(i,i) : i \in N\}$. This proves the claim and establishes the $\mathcal{NP}$-completeness of $\mathrm{DDP}(n, \mathcal{M}, m)$. □

For the proof of the next theorem we need the fact that the following problem is $\mathcal{NP}$-complete (see Karp [1972]).

Acyclic Subdigraph Problem (ASP)

Instance: Digraph $D = (N, A)$ without loops; positive integer $k \leq |N|$.
Question: Is there a subset $B \subseteq A$ with $|B| \geq k$ such that $H = (N, B)$ is acyclic?

The next lemma (easy to be proved by induction) will be useful in theorem 3.8.

Lemma 3.7. *If $H = (N, B)$ is an acyclic digraph then there exists a graph $H' = (N, B')$ containing H, such that $B' \in \mathrm{ASY} \cap \mathrm{TRA} \cap \mathrm{TOT}$.*

In the subsequent $\mathcal{NP}$-completeness proofs we shall omit the straightforward verification that the considered problems are in the class $\mathcal{NP}$.

Theorem 3.8. *Let*

$$\mathcal{M}' = \mathrm{ASY} \cap \mathrm{TRA}^* \text{ and } \mathcal{M}'' = \mathrm{ASY} \cap \mathrm{TRA}^* \cap \mathrm{TOT}.$$

Then $\mathrm{DDP}^*(n, \mathcal{M}', m)$ *is* $\mathcal{NP}$*-complete for* $m \in \{2,3\}$, *and* $\mathrm{DDP}^*(n, \mathcal{M}'', m)$ *is* $\mathcal{NP}$*-complete for* $m \in \{1,2\}$.

Proof. [Transformation from the Acyclic Subdigraph Problem (ASP)]
(i) Assume first that $m \in \{2,3\}$ and let $\mathcal{M} \in \{\mathcal{M}', \mathcal{M}''\}$.
Suppose that $D = (N, A)$ and k are given as an instance of ASP.Then the corresponding instance of $\mathrm{DDP}^*(n, \mathcal{M}, m)$, defined by D_n, w, m and k', is obtained as follows: $D_n = (N, A_n)$ is the complete digraph obtained from D by adding to it all the missing arcs which are not loops; the weights w_e for $e \in A_n$ are defined as

$$w_e := \begin{cases} -m & \text{if } e \in A, \\ -(m-2) & \text{otherwise;} \end{cases}$$

and

$$k' := -2k - \binom{n}{2}(m-2).$$

We shall prove that D has an acyclic subdigraph $H = (N, B)$ with $|B| \geq k$ if and only if D_n has a subdigraph $H' = (N, B')$ with $B' \in \mathcal{M}$ and $w(B') \leq k'$.

a) Let $H = (N, B)$ be an acyclic subdigraph in D with $|B| \geq k$. Since H is also a subdigraph of D_n, then by Lemma (3.7) there exists in D_n a subdigraph $H' = (N, B')$ containing H such that $B' \in \mathcal{M}''$. Moreover,

$$\begin{aligned} w(B') &= w(B) + w(B' \setminus B) \\ &\leq |B|(-m) - \left(\binom{n}{2} - |B|\right)(m-2) \leq k'. \end{aligned}$$

b) Let $H' = (N, B')$ be a subdigraph in D_n such that $B' \in \mathcal{M}$ and $w(B') \leq k'$. Since H is acyclic, by Lemma (3.7) there exists in D_n a subdigraph $H'' = (N, B'')$ containing H' with $B'' \in \mathcal{M}''$. Note that B'' has at least k arcs with weight $-m$. Otherwise, if B'' has l arcs with weight $-m$, $l \leq k-1$, then

$$w(B') \geq w(B'') = l(-m) - \left(\binom{n}{2} - l\right)(m-2) > k'.$$

Thus, if we take $B := \{e \in B'' : w_e = -m\}$, clearly $H = (N, B)$ is an acyclic subdigraph of D with $|B| \geq k$.

(ii) If $m = 1$ then the above proof also holds for $\mathcal{M} = \mathcal{M}''$.

Since ASP is $\mathcal{NP}$-complete and the given transformation is polynomial, the result follows. □

We want to prove in the sequel that $\mathrm{DDP}^*(n,\mathcal{M},1)$ for $\mathcal{M} = \mathrm{SYM} \cap \mathrm{TRA}^*$ is $\mathcal{NP}$-complete. For that, we introduce the next problem which we prove later (see Theorem 3.13) to be $\mathcal{NP}$-complete.

Restricted Clique Partitioning Problem – RCCP

Instance: Complete graph $K_n = [V, E]$, weights $w_e \in \{-1, 0, 1\}$ for each $e \in E$, integer k.

Question: Is there a clique partitioning $A \subset E$ such that $w(A) \leq k$? (That is, is there a partition of the node set V_n such that the sum of the weights of all edges with both endnodes in the same set of the partition is less or equal to k?)

Theorem 3.9. *Let*

$$\mathcal{M} = \mathrm{SYM} \cap \mathrm{TRA}^*.$$

Then $\mathrm{DDP}^*(n,\mathcal{M},1)$ *is* $\mathcal{NP}$*-complete.*

Proof. [Transformation from RCPP] Note that it suffices to prove for $m = 1$. Let $K_n = [V_n, \mathcal{E}_n]$, w and k be an arbitrary instance of RCPP and assume that $V_n = \{1, 2, \ldots, n\}$. The corresponding instance of $\mathrm{DDP}^*(\mathcal{M}, 1)$ defined by D_n, w' and k', is constructed as follows: $D_n = (N, A_n)$ is a complete digraph with node set $N = V_n$, the weights w'_e for $e \in A_n$ are defined as

$$w'_{ij} := \begin{cases} 1 & \text{if } (w_{ij} = 1) \text{ or } (w_{ij} = 0 \text{ and } i < j), \\ -1 & \text{if } (w_{ij} = -1) \text{ or } (w_{ij} = 0 \text{ and } i > j); \end{cases}$$

and $k' := 2k$.

It is immediate that, if $K_n = [V_n, \mathcal{E}_n]$ has a clique partitioning A with $w(A) \leq k$, then $B := \{ij, ji : \{i,j\} \in A\}$, is an arc set in D_n such that $B \in \mathcal{M}$ and $w'(B) = 2w(A)$. Conversely, if D_n has an arc set $B \in \mathcal{M}$ with $w'(B) \leq k'$, then it is easy to see that the set $A := \{\{i,j\} : ij \in B\}$ is a clique partitioning of K_n with $2w(A) = w'(B)$. Since RCPP is $\mathcal{NP}$-complete (by Theorem 3.13), and the given transformation is polynomial, the result follows. □

It remains to analyse two more cases. Namely, when $\mathcal{M} \in \{\mathrm{TRA}^*, \mathrm{TRA}^* \cap \mathrm{TOT}\}$. This is done in the next two theorems.

Theorem 3.10. *Let*

$$\mathcal{M} = \mathrm{TRA}^* \cap \mathrm{TOT}.$$

Then the problem $\mathrm{DDP}^*(p,\mathcal{M},m)$, *where* m *is bounded by a polynomial in* p, *is* $\mathcal{NP}$*-complete.*

Proof. By Theorem 3.8, the problem $Q := \mathrm{DDP}^*(n,\mathcal{M}^*,2)$ with $\mathcal{M}^* = \mathrm{ASY} \cap \mathrm{TRA}^* \cap \mathrm{TOT}$ is $\mathcal{NP}$-complete. We want to prove that Q is polynomially reducible to $\widetilde{Q} := \mathrm{DDP}^*(p,\mathcal{M},m)$, where $m \leq p^4$. Let $D_n = (N, A_n)$, w and k be given as an instance of Q. Note that, we may assume that $k < n^2$, otherwise Q is trivially solvable. Suppose $N = \{1, 2, \ldots, n\}$, $n \geq 2$. The corresponding instance of $\widetilde{Q}$ defined by $\widetilde{D}_p$, $\widetilde{w}$ and $\tilde{k}$ is constructed as follows : $\widetilde{D}_p = (\widetilde{N}, \tilde{A})$ is the complete digraph of order $p = 2n$ with node set $\widetilde{N} := \{i_1, i_2 : i \in N\}$. To define the weights $\widetilde{w}_e$ for $e \in \tilde{A}$ we let

$$R := \bigcup_{1 \leq i < j \leq n} R_{ij} \text{ where } R_{ij} := \{(i_1, j_1), (j_1, i_2)\}$$

and set

$$\widetilde{w}_e = \begin{cases} 0 & \text{if } e = (i_2, i_1),\ i \in N \\ L & \text{if } e = (i_1, i_2),\ i \in N \\ w_{ij} & \text{if } e = (i_1, j_1),\ e \in R \\ w_{ji} & \text{if } e = (j_1, i_2),\ e \in R \\ M & \text{otherwise,} \end{cases}$$

where

$$M := 4n^2 \text{ and } L := 2n^4.$$

Observe that $|\widetilde{w}_e|$ is even and $|\widetilde{w}_e| \leq p^4$ for every $e \in \tilde{A}$.

The parameter $\tilde{k}$ is defined as

$$\tilde{k} := k + CM, \text{ where}$$
$$C := \binom{2n}{2} - n - \binom{n}{2} = 3\binom{n}{2}.$$

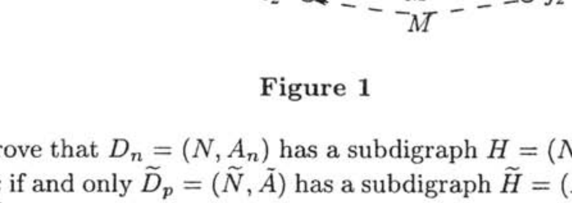


Figure 1

We shall prove that $D_n = (N, A_n)$ has a subdigraph $H = (N, B)$ with $B \in \mathcal{M}^*$ and $w(B) \leq k$ if and only $\widetilde{D}_p = (\widetilde{N}, \tilde{A})$ has a subdigraph $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ with $\widetilde{B} \in \mathcal{M}$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$.

It is clear that M, L and $\tilde{k}$ were chosen conveniently so that the above claim can be shown to hold. Before we give the proof, let us explain the idea behind the choice of the values for M, L and $\tilde{k}$. Note that for each pair i, j, $1 \leq i < j \leq n$, the arcs (i, j) and (j, i) in D_n correspond to the arcs (i_1, j_1) and (j_1, i_2) in $\widetilde{D}_p$, respectively, and that the assigned weights agree correspondingly. See Figure 1. Given a subdigraph $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ in $\widetilde{D}_p$ with $\widetilde{B} \in \mathcal{M}$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$, we want to construct a subdigraph $H = (N, B)$ in D_n with $B \in \mathcal{M}^*$ and $w(B) \leq k$. So we want $\widetilde{H}$ to have exactly one of the arcs $(i_1, j_1), (j_1, i_2)$ for each pair i, j, $1 \leq i < j \leq n$ (so that the corresponding arcs in D_n can be set into B). Thus we choose L conveniently (according to $\tilde{k}$) so that both of (i_1, j_1) and (j_1, i_2) cannot be in any transitive subdigraph $\widetilde{H}$ with $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$. This can be accomplished by choosing L so that whenever both of these arcs are chosen to be in a transitive subdigraph $\widetilde{H} = (\widetilde{N}, \widetilde{B})$, then the choice of (i_1, i_2) forced by the transitivity gives that $\widetilde{w}(\widetilde{B}) > \tilde{k}$. The values for $\tilde{k}$ and M are so chosen that $\widetilde{H}$ must be a subdigraph consisting of :

i) all arcs with weight 0;
ii) exactly one of the arcs (i_1, j_1), (j_1, i_2) for each pair i, j, $1 \leq i < j \leq n$;
iii) exactly $C := \binom{2n}{2} - n - \binom{n}{2}$ arcs with weight M.

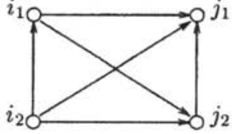
a) Given a subdigraph $H = (N, B)$ in D_n with $B \in \mathcal{M}^*$ and $w(B) \leq k$, construct $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ by setting:

$$\widetilde{B} := \widetilde{B}_1 \cup \widetilde{B}_2 \cup \widetilde{B}_3,$$

where

$$\begin{aligned}
\widetilde{B}_1 &:= \{(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) : \quad 1 \leq i < j \leq n \text{ and } (i, j) \in B\},\\
\widetilde{B}_2 &:= \{(j_1, i_1), (j_1, i_2), (j_2, i_1), (j_2, i_2) : \quad 1 \leq i < j \leq n \text{ and } (j, i) \in B\},\\
\widetilde{B}_3 &:= \{(i_2, i_1) : \quad i \in N\}.
\end{aligned}$$

Notice that $|\widetilde{B}| = \binom{2n}{2}$ and $\widetilde{B} \in \mathrm{TOT}$. For each pair $e = (i, j)$, $1 \leq i, j \leq n$, $i \neq j$, let S_e be the following *basic* subdigraph:



S_{ij}

Figure 2

Clearly, $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ is the union of all basic subdigraphs S_e each corresponding to an arc $e \in B$. By inspection, it is easy to see that these subdigraphs S_e are transitive. Thus it remains to be proved that if $e := (i_r, j_s)$ and $f := (j_s, l_t)$, with $r, s, t \in \{1, 2\}$, are arcs of $\widetilde{B}$ not in the same basic subdigraph, then $g := (i_r, l_t) \in \widetilde{B}$. Notice that when $i = j$ or $j = l$ then e and f are in a same basic subdigraph. Furthermore, $\widetilde{B}$ has no arcs such as e and f where $i = l$. Therefore we may assume that i, j, l are pairwise distinct.

Since $e := (i_r, j_s) \in \widetilde{B}$, then

$$if \begin{cases} i < j & \text{then } e \in \widetilde{B}_1 \text{ and } (i,j) \in B, \\ i > j & \text{then } e \in \widetilde{B}_2 \text{ and } (i,j) \in B. \end{cases}$$

Similarly, $f := (j_s, l_t) \in \widetilde{B}$ implies that

$$if \begin{cases} j < l & \text{then } f \in \widetilde{B}_1 \text{ and } (j,l) \in B, \\ j > l & \text{then } f \in \widetilde{B}_2 \text{ and } (j,l) \in B. \end{cases}$$

Thus, $(i,j) \in B$ and $(j,l) \in B$. Since $B \in$ TRA*, $(i,l) \in B$. If $i < l$ then $g \in \widetilde{B}_1$, otherwise $g \in \widetilde{B}_2$. Hence, $g \in \widetilde{B}$. This completes the proof that $\widetilde{B} \in$ TRA*.

Now let us prove that $\widetilde{w}(\widetilde{B}) \leq \widetilde{k}$. Notice that $\widetilde{B}_1, \widetilde{B}_2$ and $\widetilde{B}_3$ are pairwise disjoint, $R \cap \widetilde{B}_3 = \emptyset$ and $|R \cap \widetilde{B}| = \left|R \cap (\widetilde{B}_1 \cup \widetilde{B}_2)\right| = \frac{|R|}{2} = \binom{n}{2}$.

Thus,

$$\begin{aligned} \widetilde{w}(\widetilde{B}) &= \widetilde{w}(\widetilde{B} \cap R) + \widetilde{w}\big((\widetilde{B}_1 \cup \widetilde{B}_2) \setminus R\big) + \widetilde{w}(\widetilde{B}_3) \\ &= w(B) + \big(|\widetilde{B}| - |\widetilde{B}_3| - |\widetilde{B} \cap R|\big)M \\ &= w(B) + \left(\binom{2n}{2} - n - \binom{n}{2}\right)M \leq k + CM \quad = \quad \widetilde{k}. \end{aligned}$$

b) Let $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ be a subdigraph of $\widetilde{D}_p$ with $\widetilde{B} \in \mathcal{M}$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$. Then the following holds:

(b_1) $\widetilde{B}$ does not contain an arc e with $\widetilde{w}_e = L$.

Suppose $\widetilde{B}$ contains such an arc e. Then

$$\widetilde{w}(\widetilde{B}) \geq L - \sum_{e \in A_n} |w_e| + 2\binom{n}{2}M.$$

Since $\sum_{e \in A_n} |w_e| \leq 2n^2$, it follows that $\widetilde{w}(\widetilde{B}) \geq 6n^4 - 4n^3 - 2n^2$. On the other hand, $\tilde{k} = k + CM < n^2 + \left(\frac{3n(n-1)}{2}\right)4n^2 = 6n^4 - 6n^3 + n^2$ and therefore, $\widetilde{w}(\widetilde{B}) > \tilde{k}$, a contradiction.

(b_2) $\widetilde{B}$ contains all arcs e with $\widetilde{w}_e = 0$.

This follows immediately from (b_1) and the fact that $\widetilde{B} \in$ TOT.

(b_3) For every pair (i,j), $1 \leq i < j \leq n$, $|\widetilde{B} \cap R_{ij}| \leq 1$.

Suppose there is a pair (i, j) such that $\widetilde{B}$ contains both of (i_1, j_1) and (j_1, i_2). Since $\widetilde{B} \in \text{TRA}^*$, this implies that $(i_1, i_2) \in \widetilde{B}$; but as $\widetilde{w}(i_1, i_2) = L$, this contradicts ($b_1$).

(b_4) $\widetilde{B}$ contains exactly C arcs with weight M.

Suppose $\widetilde{B}$ has more than C arcs with weight M. Thus

$$\widetilde{w}(\widetilde{B}) \geq (C+1)M - \sum_{e \in A_n} |w_e| > \tilde{k},$$

a contradiction. So, $\widetilde{B}$ can have at most C arcs with weight M. On the other hand, since $|\widetilde{B}| \geq \binom{2n}{2}$ and $\widetilde{B}$ contains n arcs with weight 0 (by (b_2)), at most $\binom{n}{2}$ arcs of R (by (b_3)) and no arcs with weight L (by (b_1)), then $\widetilde{B}$ must contain at least $\binom{2n}{2} - n - \binom{n}{2} =: C$ arcs with weight M. Thus $\widetilde{B}$ contains exactly C arcs with weight M.

(b_5) For every pair (i, j), $1 \leq i < j \leq n$, $|\widetilde{B} \cap R_{ij}| = 1$.

Since $\binom{2n}{2} \leq |\widetilde{B}| = n + C + |\widetilde{B} \cap R|$, it follows that $|\widetilde{B} \cap R| \geq \binom{2n}{2} - n - C = \binom{n}{2}$. If for some pair (i, j), $1 \leq i < j \leq n$, $|\widetilde{B} \cap R_{ij}| < 1$ then by (b_3) $|\widetilde{B} \cap R| < \binom{n}{2}$, a contradiction. Thus, the statement is proved.

(b_6) $\widetilde{B}$ has no double arcs.

Immediate from (b_1), (b_4) and (b_5).

(b_7) $\widetilde{w}(\widetilde{B} \cap R) \leq k$.

Clearly, $\widetilde{w}(\widetilde{B}) = \widetilde{w}(\widetilde{B} \cap R) + CM$. Thus, $\widetilde{w}(\widetilde{B} \cap R) \leq \tilde{k} - CM = k$.
But $(i_1, j_2) \in \widetilde{B} \cap R$ and $(j_2, j_1) \in \widetilde{B}$ imply $(i_1, j_1) \in \widetilde{B}$. Thus, $(i_1, j_1) \in \widetilde{B}$. Analogously, analysing the cases $j < l$ and $j > l$ we conclude that $(j_1, l_1) \in \widetilde{B}$.

Since $\widetilde{B} \in \text{TRA}^*$, then $(i_1, j_1) \in \widetilde{B}$ and $(j_1, l_1) \in \widetilde{B}$ imply that $(i_1, l_1) \in \widetilde{B}$. Thus, if $i < l$ then $(i_1, l_1) \in \widetilde{B} \cap R$, and therefore $(i, l) \in B$. Suppose $i > l$. By (b_6) $(i_1, l_1) \in \widetilde{B}$ implies $(l_1, i_1) \notin \widetilde{B}$. By ($b_5$), if $l < i$ and $(l_1, i_1) \notin \widetilde{B} \cap R$ then $(i_1, l_2) \in \widetilde{B} \cap R$. But then, $(i, l) \in B$ and therefore $B \in \text{TRA}^*$.

Since the given transformation is polynomial, it follows that $\widetilde{Q}$ is $\mathcal{NP}$-complete. □

A construction similar to the one presented in the proof of Theorem 3.10 leads to the following result.

Theorem 3.11. *Let*

$$\mathcal{M} = \text{TRA}^*.$$

Then the problem $\text{DDP}^*(p, \mathcal{M}, m)$, *where* m *is bounded by a polynomial in* p, *is* $\mathcal{NP}$*-complete.*

Proof. Let $Q := \text{DDP}^*(n, \mathcal{M}^*, 2)$ with $\mathcal{M}^* = \text{ASY} \cap \text{TRA}^* \cap \text{TOT}$ be the $\mathcal{NP}$-complete problem considered in Theorem 3.8. Our aim is to prove that Q is polynomially transformable to $\widetilde{Q} := \mathrm{DDP}^*(p, \mathcal{M}, m)$, where $m \le p^6$. For that, let us assume that $D_n = (N, A_n)$, w and k, $k < n^2$, are given as an instance of Q, and let us construct the corresponding instance of $\widetilde{Q}$.

Let $\widetilde{D}_p = (\widetilde{N}, \tilde{A})$ be the complete digraph of order $p = 2n$ with node set $\widetilde{N} := \{i_1, i_2 : i \in N\}$.To define the weights $\widetilde{w}_e$, set

$$R := \bigcup_{1 \le i < j \le n} R_{ij}, \qquad \text{where } R_{ij} := \{(i_1, j_1), (j_1, i_2)\} ,$$
$$\overline{R} := \bigcup_{1 \le i < j \le n} \overline{R}_{ij}, \qquad \text{where } \overline{R}_{ij} := \{(j_1, i_1), (i_2, j_1)\} .$$

Let M be the *smallest even* integer such that

$$M > k + 2n^2 ,$$

and set

$$M^* := \left(\binom{n}{2} + 1\right) M ,$$
$$L := M + \binom{n}{2}(M^* + M) .$$

Now define $\widetilde{w}_e$ for each $e \in \tilde{A}$, as follows :

$$\widetilde{w}_e = \begin{cases} -M^* & \text{if } e = (i_2, i_1),\ i \in N \\ L & \text{if } e = (i_1, i_2),\ i \in N \\ w_{ij} - M^* & \text{if } e = (i_1, j_1),\ e \in R \\ w_{ji} - M^* & \text{if } e = (j_1, i_2),\ e \in R \\ M & \text{if } e \in \overline{R} \\ 0 & \text{otherwise .} \end{cases}$$

Observe that $|\widetilde{w}_e|$ is even and $|\widetilde{w}_e| \le p^6$ for every $e \in \tilde{A}$.

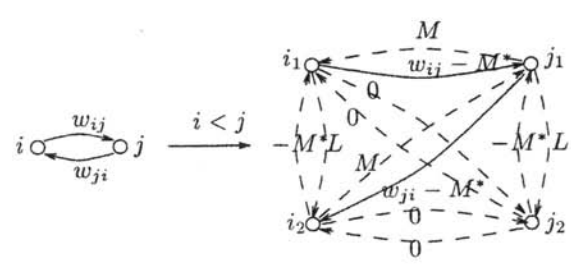


Figure 3

Let

$$\tilde{k} := k + \tbinom{n}{2} M - \left(\tbinom{n}{2} + n\right) M^* \ .$$

We claim that $D_n = (N, A_n)$ has a subdigraph $H = (N, B)$ with $B \in \mathcal{M}^*$ and $w(B) \leq k$ iff $\widetilde{D}_p = (\widetilde{N}, \tilde{A})$ has a subdigraph $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ with $\widetilde{B} \in \mathrm{TRA}^*$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$.

a) Given $H = (N, B)$ in D_n with $B \in \mathcal{M}^*$ and $w(B) \leq k$, let $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ be the subdigraph of $\widetilde{D}_p$ defined by :

$$\widetilde{B} := \widetilde{B}_1 \cup \widetilde{B}_2 \cup \widetilde{B}_3 \ ,$$

where

$$\begin{aligned}
\widetilde{B}_1 &:= \left\{(i_1, j_1), (i_1, j_2), (i_2, j_1), (i_2, j_2) : \quad 1 \leq i < j \leq n \text{ and } (i, j) \in B\right\} \ , \\
\widetilde{B}_2 &:= \left\{(j_1, i_1), (j_1, i_2), (j_2, i_1), (j_2, i_2) : \quad 1 \leq i < j \leq n \text{ and } (j, i) \in B\right\} \ , \\
\widetilde{B}_3 &:= \left\{(i_2, i_1) : \quad 1 \leq i \leq n\right\} \ .
\end{aligned}$$

Then

$$\begin{aligned}
\widetilde{w}(\widetilde{B}) &= \widetilde{w}(\widetilde{B} \cap R) + \widetilde{w}(\widetilde{B} \setminus R) \\
&= \widetilde{w}(\widetilde{B} \cap R) + \widetilde{w}(\widetilde{B}_1 \cup \widetilde{B}_2 \setminus R) + \widetilde{w}(\widetilde{B}_3) \\
&= w(B) - \binom{n}{2} M^* + \binom{n}{2} M - nM^* \\
&= w(B) - \left(\binom{n}{2} + n\right) M^* + \binom{n}{2} M \\
&\leq k - \left(\binom{n}{2} + n\right) M^* + \binom{n}{2} M \leq \tilde{k} \ .
\end{aligned}$$

Using the fact that $B \in \mathcal{M}^*$ it is not difficult to prove that $\widetilde{B} \in \mathrm{TRA}^*$. Indeed, the proof is analogous to the one present for Theorem 3.10, and therefore it will be omitted.

b) Let $\widetilde{H} = (\widetilde{N}, \widetilde{B})$ be a subdigraph of $\widetilde{D}_p$ with $\widetilde{B} \in \mathrm{TRA}^*$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$. Based on $\widetilde{H}$ we want to construct a transitive tournament $H = (N, B)$ in D_n with $w(B) \leq k$. For that, we first observe that $\widetilde{H}$ has the following properties :

(b_1) $\widetilde{B}$ does not contain an arc e with $\widetilde{w}_e = L$.

Suppose $\widetilde{B}$ contains such an arc e. Then

$$\begin{aligned}
\widetilde{w}(\widetilde{B}) &\geq L - nM^* - \sum_{e \in A} |w_e| - n(n-1)M^* \\
&\geq M + \binom{n}{2}(M^* + M) - nM^* - 2n^2 - n(n-1)M^* \\
&> k - \binom{n}{2} M^* - nM^* + \binom{n}{2} M = \tilde{k} \ ,
\end{aligned}$$

a contradiction.

(b_2) For every pair (i,j), $1 \leq i < j \leq n$, $\left|\widetilde{B} \cap R_{ij}\right| = 1$.

Suppose there is a pair (i,j) such that $\left|\widetilde{B} \cap R_{ij}\right| = 2$. In this case, since $\widetilde{B} \in \mathrm{TRA}^*$, it follows that $(i_1, i_2) \in \widetilde{B}$, contradicting ($b_1$). Thus, $\left|\widetilde{B} \cap R_{ij}\right| \leq 1$ for every pair (i,j), $1 \leq i < j \leq n$. Now suppose there is a pair (i,j) such that $\left|\widetilde{B} \cap R_{ij}\right| = 0$. Then

$$\widetilde{w}(\widetilde{B}) \geq -\left(\binom{n}{2} - 1\right) M^* - nM^* - \sum_{e \in A_n} |w_e|.$$

Using the fact that $\sum_{e \in A_n} |w_e| \leq 2n^2$ and making some substitutions we get $\widetilde{w}(\widetilde{B}) > \tilde{k}$. Since this contradicts our assumption, we conclude that (b_2) holds.

(b_3) For every i, $1 \leq i \leq n$, $(i_2, i_1) \in \widetilde{B}$.

Suppose for some i, $1 \leq i \leq n$, $(i_2, i_1) \notin \widetilde{B}$. Then

$$\begin{aligned}
\widetilde{w}(\widetilde{B}) &\geq -(n-1)M^* - \sum_{e \in A_n} |w_e| - \binom{n}{2} M^* \\
&\geq \left(\binom{n}{2} + 1\right) M - nM^* - 2n^2 - \binom{n}{2} M^* \\
&> \binom{n}{2} M + k - \left(\binom{n}{2} + n\right) M^* = \tilde{k}\ ,
\end{aligned}$$

a contradiction.

(b_4) For every pair (i,j), $1 \leq i, j \leq n$, $\left|\widetilde{B} \cap R_{ij}\right| = 1$.

By (b_2), for every pair (i,j), $1 \leq i < j \leq n$, exactly one of the arcs (i_1, j_1) or (j_1, i_2) is in $\widetilde{B}$. If $(i_1, j_1) \in \widetilde{B}$, since $(i_2, i_1) \in \widetilde{B}$, it follows that $(i_2, j_1) \in \widetilde{B}$. Analogously, if $(j_1, i_2) \in \widetilde{B}$ then $(j_1, i_1) \in \widetilde{B}$. Thus, $|\widetilde{B} \cap \overline{R}_{ij}| \geq 1$. Now suppose there is a pair (i,j), $1 \leq i < j \leq n$, such that $|\widetilde{B} \cap \overline{R}_{ij}| > 1$. This implies that $\widetilde{B}$ has more than $\binom{n}{2}$ arcs with weight M and therefore

$$\begin{aligned}
\widetilde{w}(\widetilde{B}) &\geq \left(\binom{n}{2} + 1\right) M - \sum_{e \in A_n} |w_e| - \binom{n}{2} M^* - nM^* \\
&> \binom{n}{2} M + k - \left(\binom{n}{2} + n\right) M^* = \tilde{k}\ ,
\end{aligned}$$

a contradiction. So, we have proved that (b_4) holds.

(b_5) $\widetilde{w}(\widetilde{B} \cap R) \leq k - \binom{n}{2} M^*$.

This follows from the fact that $\widetilde{w}(\widetilde{B} \cap R) = \widetilde{w}(\widetilde{B}) + nM^* - \binom{n}{2} M$ and $\widetilde{w}(\widetilde{B}) \leq \tilde{k}$.

Now let $H = (N, B)$ be the subdigraph of D_n with

$$B := \{(i,j) : (i_1, j_1) \in \widetilde{B} \cap R, \quad 1 \le i < j \le n\} \ \cup \\ \{(j,i) : (j_1, i_2) \in \widetilde{B} \cap R, \quad 1 \le i < j \le n\} .$$

We claim that $B \in \mathcal{M}^*$ and $w(B) \le k$. Note that $|B| = |\widetilde{B} \cap R| = \binom{n}{2}$. Furthermore, $w(B) = \widetilde{w}(\widetilde{B} \cap R) + \binom{n}{2} M^*$. Thus, by ($b_5$) it follows that $w(B) \le k$. The definition of B and fact (b_2) yield immediately that $B \in$ ASY$\cap$TOT. So it remains to be shown that $B \in$ TRA*. Let (i,j) and (j,l) be arcs of B, $i \ne j \ne l \ne i$. If $i < j$ and $(i,j) \in B$ then $(i_1, j_1) \in \widetilde{B} \cap R$. If $i > j$ and $(i,j) \in B$ then $(i_1, j_2) \in \widetilde{B} \cap R$. In the latter case, since $(j_2, j_1) \in \widetilde{B}$ and $\widetilde{B} \in$ TRA*, it follows that $(i_1, j_1) \in \widetilde{B}$. From an analogous analysis of cases $j < l$ and $j > l$, we conclude that $(j_1, l_1) \in \widetilde{B}$. Thus, $(i_1, j_1) \in \widetilde{B}$ and $(j_1, l_1) \in \widetilde{B}$, and therefore $(i_1, l_1) \in \widetilde{B}$. If $i < l$, then $(i_1, l_1) \in \widetilde{B} \cap R$, and hence $(i,l) \in B$. Suppose $l < i$. Then $(i_1, l_1) \in \widetilde{B} \cap \overline{R}$, and therefore by ($b_4$) it follows that $(l_2, i_1) \notin \widetilde{B}$. In this case, $(l_1, i_1) \notin \widetilde{B}$; otherwise $(l_2, l_1) \in \widetilde{B}$ and $(l_1, i_1) \in \widetilde{B}$ would imply $(l_2, i_1) \in \widetilde{B}$, a contradiction. But if $(l_1, i_1) \notin B$, by (b_2) we conclude that $(i_1, l_2) \in \widetilde{B}$. Thus $(i,l) \in B$, and this proves that $B \in$ TRA*.

Clearly, the transformation of Q to $\widetilde{Q}$ is polynomial and therefore $\widetilde{Q}$ is $\mathcal{NP}$-complete. □

The next result follows from theorems 3.8–3.11 together with Lemma 3.6 and Theorem 3.5.

Theorem 3.12. *Let*

$$\mathcal{M}' \in \{\text{TRA, ASY} \cap \text{TRA, SYM} \cap \text{TRA, TRA} \cap \text{TOT, ASY} \cap \text{TRA} \cap TOT\}.$$

Then MP$(\mathcal{R}, \mathcal{M})$ *is* $\mathcal{NP}$*-hard for* $\mathcal{M} \in \{\mathcal{M}', \mathcal{M}' \cap \text{REF}\}$. □

We can summarize the computational complexity results we have proved as follows.

a) The median problem MP$(\mathcal{R}, \mathcal{M})$ is easy whenever $\mathcal{M}$ results from any combination of the properties REF, SYM, ASY and TOT. [Proposition 3.4.]
b) Transitivity makes the median problem difficult. More precisely, if one of the properties required for $\mathcal{M}$ is TRA then the median problem MP$(\mathcal{R}, \mathcal{M})$ is $\mathcal{NP}$-hard except for the trivial combination TRA$\cap$SYM$\cap$TOT. The most interesting cases are included here: partial orders (TRA$\cap$ASY), linear orders (TRA $\cap$ ASY $\cap$ TOT) , complete preorders (TRA $\cap$ TOT) and equivalences REF $\cap$ TRA $\cap$ SYM. [Theorem 3.12.]
c) The property REF may be included or not, without changing the complexity status of the median problem. [Theorem 3.12.]
d) In some cases the median problem MP$(\mathcal{R}, \mathcal{M})$ is $\mathcal{NP}$-hard even when the profile Π consists of 1 or 2 relations. This is the case when $\mathcal{M}$ is a linear order and $\mathcal{M}$ is a partial order. [Theorem 3.8.]

To close this section we present the $\mathcal{NP}$-completeness proof of the Restricted Clique Partitioning Problem, needed to prove Theorem 3.9. We shall base our proof on the transformation from the *Simple Max-Cut Problem* (SMCP), known to be $\mathcal{NP}$-complete (Garey, Johnson and Stockmeyer [1976]). In this problem the instance consists of a graph $G = [V, E]$, and a positive integer k. The objective is to decide whether G has a cut of size at least k.

Theorem 3.13. *The Restricted Clique Partitioning Problem is $\mathcal{NP}$-complete.*

Proof. $\big[$ Transformation from the Simple Max-Cut Problem (SMCP) $\big]$Let $G = [V, E]$ and k be given as an instance of SMCP, and assume that $|V| = n$. Let $G' = [V', E']$ be a complete graph of order $3n$ obtained from G by adding to it $2n$ more nodes and completing it with all the missing edges, which are not loops. Assume that $V' = V \cup X \cup Y$, where $|X| = |Y| = n$. Assign weights w_e to each edge $e \in E'$ by setting

$$w_e := \begin{cases} 1 & \text{if} \quad e \in E \cup (X : Y) \ , \\ -1 & \text{if} \quad e \in (V : X \cup Y) \ , \\ 0 & \text{otherwise} \ ; \end{cases}$$

and let

$$k' := |E| - k - n^2 \ .$$

We claim that G has a cut C with $|C| \geq k$ iff G' has a clique partitioning A with $w(A) \leq k'$.

a) Assume that $C = E(V_1 : V \setminus V_1)$ is a cut in G with $|C| \geq k$, and let $V_2 := V \setminus V_1$. Then the edge set $A := E'(V_1 \cup X) \cup E'(V_2 \cup Y)$ is a clique partitioning of G' with

$$\begin{aligned} wA) &= w\big(E'(V_1 \cup X)\big) + w\big(E'(V_2 \cup Y)\big) \\ &= -|(V_1 : X)| - |(V_2 : Y)| + |E \setminus C| \\ &= -n^2 + |E| - |C| \ \leq \ -n^2 + |E| - k \ = \ k' \ . \end{aligned}$$

b) Assume that G' has a clique partitioning A with $w(A) \leq k'$. We want to prove that G has a cut C with $|C| \geq k$. Let us assume for the moment that the following holds:
Claim 1: G' has a clique partitioning A' with $w(A') \leq k'$ and $\mathcal{C}(A') = \{Q_1, Q_2\}$, where Q_1 and Q_2 are such that $VQ_1 = X \cup V_1$ and $VQ_2 = Y \cup V_2$ for some nonempty subsets V_1 and V_2 of V.

Note that, in this case, $w(A') = w(EQ_1) + w(EQ_2) = -n^2 + \big|E(V_1)\big| + \big|E(V_2)\big| = -n^2 + |E| - \big|E(V_1 : V_2)\big|$, and since $w(A') \leq k' = -n^2 + |E| - k$, it follows that $\big|E(V_1 : V_2)\big| \geq k$, and therefore $C := E(V_1 : V_2)$ defines the desired cut C in G.

Thus, in order to complete the proof it remains to be shown that Claim 1 holds. Before that, for notational convenience, let us give names to the different types of cliques we shall consider. According to its intersection with the sets V, X, Y, a clique $H = [VH, EH]$ may be of one of the following types (see Figure 4) :

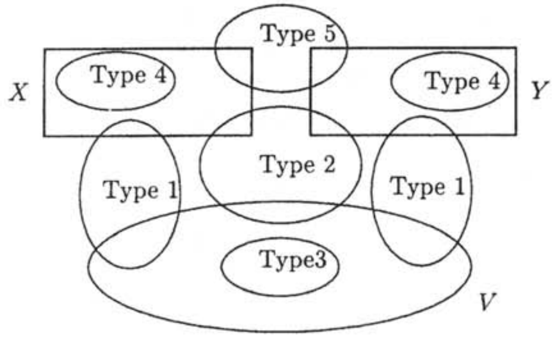


Figure 4

Type 1 :
(X, V)-intersecting (if $VH \cap X \neq \emptyset$, $VH \cap V \neq \emptyset$ and $VH \cap Y = \emptyset$) or
(Y, V)-intersecting (if $VH \cap Y \neq \emptyset$, $VH \cap V \neq \emptyset$ and $VH \cap X = \emptyset$) .
Type 2 :
(X, Y, V)-intersecting (if $VH \cap X \neq \emptyset$, $VH \cap Y \neq \emptyset$ and $VH \cap V \neq \emptyset$) .
Type 3 :
V-included (if $\emptyset \neq VH \subseteq V$) .
Type 4 :
X-included (if $\emptyset \neq VH \subseteq X$) or Y-included (if $\emptyset \neq VH \subseteq Y$) .
Type 5 :
(X, Y)-intersecting (if $VH \cap X \neq \emptyset$, $VH \cap Y \neq \emptyset$ and $VH \cap V = \emptyset$) .

Note that according to the given definitions the desired clique partitioning A' of Claim 1 must be such that $\mathcal{C}(A') = \{Q_1, Q_2\}$, where Q_1 and Q_2 are both of Type 1.

For simplicity, we say that a clique partitioning A_1 is *better* than a clique partitioning A_2 if either $w(A_1) < w(A_2)$, or $w(A_1) = w(A_2)$ and $|\mathcal{C}(A_1)| < |\mathcal{C}(A_2)|$.

Since each clique partitioning A is bijectively associated with the clique set $\mathcal{C}(A)$, when we refer to a clique partitioning B obtained from A by replacing some of the cliques in $\mathcal{C}(A)$ with others, we are in fact defining how $\mathcal{C}(B)$ is constructed and therefore defining in this way the arc set B.

Now consider the following Claim 2 to be used in the proof of Claim 1.

Claim 2: Let Q be a clique partitioning of G' and assume that $Q_1, \ldots, Q_l$, $l \geq 2$, are cliques in $\mathcal{C}(Q)$ all of which are (W, V)-intersecting, where $W = X$ or $W = Y$. Let $\alpha := \sum_{i=1}^{l} |VQ_i \cap W|$ and $\beta := \sum_{i=1}^{l} |VQ_i \cap V|$. If $\alpha \geq \beta$, then the clique partitioning Q' obtained from Q by replacing the cliques $Q_1, \ldots, Q_l$ with the clique $G'[VQ_1 \cup \ldots \cup VQ_l]$ is such that $w(Q') < w(Q)$.

The proof of Claim 2 will be omitted as it can be obtained without any difficulty by induction on l. (For $l \geq 3$ prove that there exist two cliques Q_i and Q_j, $1 \leq i < j \leq l$, such that $|(VQ_i \cup VQ_j) \cap W| \geq |(VQ_i \cup VQ_j) \cap V|$.)

Proof of Claim 1

Let $\mathcal{A}$ be the set of the clique partitionings $\tilde{A}$ of G' with $w(\tilde{A}) \leq k'$ and such that $\mathcal{C}(\tilde{A})$ contains the smallest number possible p of cliques of Type 2. Clearly, $\mathcal{A} \neq \emptyset$ and $p \geq 0$. Our aim is to prove first that $p = 0$, and then show the existence of the desired clique partitioning A'.

Let us start by assuming that $p \geq 1$. Now let $\hat{A}$ be a *best* clique partitioning in $\mathcal{A}$, and let $H_1, \ldots, H_p$ be the cliques of Type 2 contained in $\mathcal{C}(\hat{A})$.

For $1 \leq i \leq p$ let

$$X_i := VH_i \cap X,\ x_i := |X_i|\ ,$$
$$Y_i := VH_i \cap Y,\ y_i := |Y_i|\ ,$$
$$V_i := VH_i \cap V,\ v_i := |V_i|\ .$$

Suppose $\mathcal{C}(\hat{A})$ contains a clique H_i, $1 \leq i \leq p$, such that $v_i \leq x_i$. Then we can split H_i into a clique of Type 1, $G'[X_i \cup V_i]$, and a clique of Type 4, $G'[Y_i]$, obtaining this way a clique partitioning $\tilde{A}$ with $w(\tilde{A}) \leq w(\hat{A})$ and with $\mathcal{C}(\tilde{A})$ containing $p-1$ cliques ot Type 2. Since this contradicts the choice of $\hat{A}$, we conclude that $v_i > x_i$ for $i = 1, \ldots, p$. By symmetry, we also conclude that $v_i > y_i$ for $i = 1, \ldots, p$.

It is immediate that $\mathcal{C}(\hat{A})$ contains no (X,Y)-intersecting cliques. Otherwise, a better clique partitioning could be obtained from $\hat{A}$ by replacing each (X,Y)-intersecting clique with 2 cliques, one being X-included and the other Y-included.

It is also easy to see that $\mathcal{C}(\hat{A})$ contains no X-included and no Y-included cliques. For, if H were an X-included clique in $\mathcal{C}(\hat{A})$ then by replacing the clique H_1 with the clique $H_1 \cup H$ we could obtain a better clique partitioning in $\mathcal{A}$ (since $v_1 > y_1$). By symmetry, the same holds with respect to Y-included cliques.

Since $\sum_{i=1}^{p} x_i < \sum_{i=1}^{p} v_i \leq n = |X|$ and $\sum_{i=1}^{p} y_i < |Y|$, then $\mathcal{C}(\hat{A})$ must contain (X,V)-intersecting cliques, say $\widetilde{H}_1, \ldots, \widetilde{H}_h$, $h \geq 1$, and (Y,V)-intersecting cliques, say $\widetilde{Q}_1, \ldots, \widetilde{Q}_q$, $q \geq 1$.

Let $\alpha := \bigcup_{i=1}^{h} (V\widetilde{H}_i \cap X)$ and $\beta := \bigcup_{i=1}^{h} (V\widetilde{H}_i \cap V)$.

Since $\alpha \geq \beta$, if $h \geq 2$ then by Claim 2 the cliques $\widetilde{H}_1, \ldots, \widetilde{H}_h$ can be replaced with the clique $\bigcup_{i=1}^{h} \widetilde{H}_i$ yielding this way a better clique partitioning in $\mathcal{A}$. By symmetry, if $q \geq 2$ then a better clique partitioning can also be obtained. Thus, we conclude that $h = 1$ and $q = 1$, and for simplicity we let $H := \widetilde{H}_1$ and $Q := \widetilde{Q}_1$.

If $\mathcal{C}(\hat{A})$ contains V-included cliques, say $H'_1, \ldots, H'_l$, $l \geq 1$, then it is easy to see that these cliques can be combined with the clique H giving this way a better clique partitioning. It suffices to note that $|VH \cap X| > |VH \cap V| + |\bigcup_{i=1}^{l} VH_i'|$.

Thus, we conclude that $\mathcal{C}(\hat{A}) = \{H_1, \ldots, H_p\} \cup \{H, Q\}$.

Let $H_X := VH \cap X$, $h_X := |H_X|$; $H_V := VH \cap V$, $h_V := |H_V|$; $Q_Y := VQ \cap Y$, $q_Y := |Q_Y|$; $Q_V := VQ \cap V$, $q_V := |Q_V|$. Note that $h_X > h_V$ and $q_Y > q_V$.

Let us now focus our attention on the cliques H, Q and H_1.

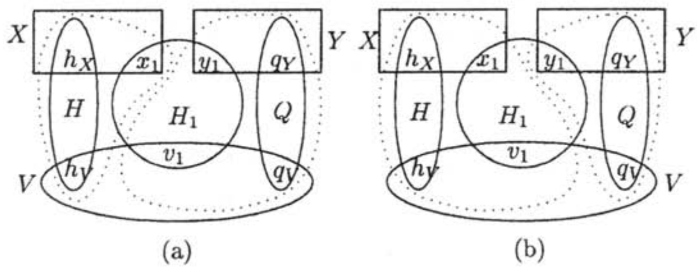


Figure 5

Suppose $v_1 \leq h_V + y_1$.

In this case, let $\tilde{A}$ be the clique partitioning obtained from $\hat{A}$ by splitting C_1 into the new cliques $G'[VH \cup X_1]$ and $G'[VQ \cup Y_1 \cup V_1]$, and preserving the (old) cliques $H_2, \ldots, H_p$ (see Figure 5.a). Thus, $\mathcal{C}(\tilde{A})$ contains $p-1$ cliques of Type 2 and $w(\tilde{A}) = w(\hat{A}) - x_1 h_V - x_1 y_1 + x_1 v_1 - y_1 q_V - v_1 q_Y + |E(V_1 : Q_V)| \leq w(\hat{A}) - x_1(h_V + y_1 - v_1) - y_1 q_V - v_1(q_Y - q_V)$.

Since $h_V + y_1 - v_1 \geq 0$ and $q_Y - q_V > 0$ we conclude that $w(\tilde{A}) < w(\hat{A}) \leq k'$, and therefore we have a contradiction to the choice of $\hat{A}$.

Assume now that $v_1 > h_V + y_1$.

In this case, let $\tilde{A}$ be the clique partitioning obtained from $\hat{A}$ by performing a splitting of C_1, symmetric to the previous one. That is, $\tilde{A}$ consists of the new cliques $G'[VH \cup X_1 \cup V_1]$ and $G'[VQ \cup Y_1]$, and the cliques $H_2, \ldots, H_p$ (see Figure 5.b). Thus $\mathcal{C}(\tilde{A})$ contains $p-1$ cliques of Type 2 and

$$\begin{aligned} w(\tilde{A}) &\leq w(\hat{A}) - y_1 q_V - x_1 y_1 + y_1 v_1 - x_1 h_V - v_1 h_X + v_1 h_V \\ &= w(\hat{A}) - y_1 q_V + y_1(v_1 - x_1) + h_V(v_1 - x_1) - v_1 h_X \\ &= w(\hat{A}) - y_1 q_V + (y_1 + h_V)(v_1 - x_1) - v_1 h_X \ . \end{aligned}$$

Since $y_1 + h_V < v_1$ and $v_1 > x_1$, it follows that

$$w(\tilde{A}) < w(\hat{A}) - y_1 q_V + v_1(v_1 - x_1) - v_1 h_X \ .$$

Now using the fact that $h_X = n - (x_1 + \dots x_p) = (h_V + q_V + v_1 + \dots v_p) - (x_1 + \dots x_p) > v_1 - x_1$, we obtain that $w(\tilde{A}) < w(\hat{A})$, again a contradiction to the choice of $\hat{A}$.

This completes the proof that $p = 0$.

Now let us assume that A' is a best clique partitioning in $\mathcal{A}$ and that $\mathcal{C}(A')$ contains no cliques of Type 2. It is immediate that $\mathcal{C}(A')$ must contain at least a clique of Type 1; otherwise we would have $w(A') \geq 0$ and therefore $w(A') > k'$, a contradiction. Let Q_1 be a clique of Type 1 contained in $\mathcal{C}(A')$, and assume without loss of generality that Q_1 is (X, V)-intersecting.

Clearly, $\mathcal{C}(A')$ contains no (X, Y)-intersecting cliques. It is also immediate that $\mathcal{C}(A')$ contains no X-included cliques, since they could all be combined with Q_1 giving this way a better clique partitioning.

If the clique set $\mathcal{C}(A')$ contains other (X, V)-intersecting cliques different from Q_1, say $H_1, \dots, H_h$, then by Claim 2, if we set $Q_1 := Q_1 \cup H_1 \cup \dots \cup H_h$, then we obtain a better clique partitioning. Thus, we conclude that $\mathcal{C}(A')$ contains a unique (X, V)-intersecting.

If $\mathcal{C}(A')$ contains a V-included clique, say H, then (since $n = |VQ_1 \cap X| > |VQ_1 \cap V|$) we can set $Q_1 := Q_1 \cup H$ and obtain a better clique partitioning.If $\mathcal{C}(A')$ contains no (Y, V)-intersecting cliques, then it consists of the clique $Q_1 = G'[X \cup V]$ and some Y-included cliques, and therefore $w(A') = -n^2 + |E| > k'$, a contradiction. Thus, let Q_2 be a (Y, V)-intersecting clique in $\mathcal{C}(A')$. If $\mathcal{C}(A')$ contains Y-included cliques and/or other (Y, V)-intersecting cliques, by performing analogous transformations to the ones we defined with respect to Q_1, we can construct a better clique partitioning. Hence, we conclude that Q_2 is the unique (Y, V)-intersecting clique in $\mathcal{C}(A')$ and therefore, $\mathcal{C}(A') = \{Q_1, Q_2\}$ with Q_1 and Q_2 both of Type 1.

Thus, we have proved that Claim 1 holds, and therefore we have completed the proof of the theorem. □

4. The case of restricted domains

In the preceding section we have proved the $\mathcal{NP}$-hardness of MP$(\mathcal{R}, \mathcal{M})$ for certain subsets $\mathcal{M} \subset \mathcal{R}$. One may now ask whether the following special cases of these problems have also the same complexity: instead of $\mathcal{R}$ (an unrestricted domain of the relations in the profile Π), we may have the information that the given relations are endowed with some properties from $\mathbb{P}$. In this case, instead of $\mathcal{R}$, we have a subset $\mathcal{M}' \subset \mathcal{R}$ and we are lead to the problem MP$(\mathcal{M}', \mathcal{M})$ defined analogously. In other words, when we consider that the domain is $\mathcal{R}$, this means that we have no information about the properties of the input relations, and when we specify a subset $\mathcal{M}' \subset \mathcal{R}$ this means that the input relations are known to be in $\mathcal{M}'$ (they belong to a restricted domain).

We have shown that in some cases the problem MP$(\mathcal{M}', \mathcal{M})$ is $\mathcal{NP}$-hard even when the profile Π consists of a fixed number m of relations. Let us denote by MP$(\mathcal{M}', \mathcal{M}, m)$ this latter problem.

When $\mathcal{M}' = \mathcal{M} = \mathcal{L}$ (the given relations are linear orders and the objective relation is also a linear order), Orlin [1981] (in an unpublished manuscript) and Bartholdi, Tovey and Trick [1988] proved that $\mathrm{MP}(\mathcal{L},\mathcal{L})$ is $\mathcal{NP}$-hard. Note that this implies that the more general problem $\mathrm{MP}(\mathcal{R},\mathcal{L})$ is $\mathcal{NP}$-hard, but not that $\mathrm{MP}(\mathcal{R},\mathcal{L},1)$ is $\mathcal{NP}$-hard — the result we have shown.

Let us turn now to the case of equivalence relations. Křivánek and Morávek [1986] proved that MP(SYM,$\mathcal{E}$,1) is $\mathcal{NP}$-hard. This result yields as a corollary the fact that $\mathrm{MP}(\mathcal{R},\mathcal{E},1)$ is $\mathcal{NP}$-hard (Theorem 3.9). Furthermore, from it one can also derive that $\mathrm{MP}(\mathcal{E},\mathcal{E})$ is $\mathcal{NP}$-hard. The reduction given by Křivánek and Morávek is from a problem on hierarchical-tree clustering, whose $\mathcal{NP}$-completeness proof is very laborious. For the sake of completeness of the class of results covered in this paper we have included our weaker result. We should observe however, that in this case rather than the $\mathcal{NP}$-hardness of $\mathrm{MP}(\mathcal{R},\mathcal{E},1)$, the interest lies more on the Theorem 3.13 from which the result could be derived.

There remains a number of open problems concerning the computational complexity status of $\mathrm{MP}(\mathcal{M}',\mathcal{M},m)$ for some combinations of $\mathcal{M}$ and $\mathcal{M}'$. We recall that the problems we know to be $\mathcal{NP}$-hard are $\mathrm{MP}(\mathcal{R},\mathcal{L},1)$, $\mathrm{MP}(\mathcal{R},\mathcal{O},2)$ and MP(SYM,$\mathcal{E}$,1). It would be interesting to establish the complexity of the problems $\mathrm{MP}(\mathcal{T},\mathcal{L},m)$ and $\mathrm{MP}(\mathcal{L},\mathcal{L},m)$ for small m.

5. Applications and strategies to solve some median problems

In qualitative data analysis, social choice theory, and paired comparison methods there are many problems that can be modelled as median problems. In these contexts, the data arise from the measurement of a number of characteristics (or attributes) associated with each object of a given set and the objective is that of finding a linear order, or a partial order or a clustering of the objects that 'best represents' the given data.

For example, the data may arise by collecting the preferences of m voters with respect to a set of objects (candidates, teams) and the objective is to find a ranking of the objects that best represents the given preferences. Note that the preferences of the m voters may be seen as m relations (eventually linear orders) on the object set, and the ranking of the objects may be seen as a linear order that best represents the given relations. So here we have the median problem $\mathrm{MP}(\mathcal{R},\mathcal{M})$, where $\mathcal{M}$ is a linear order. This is the problem of aggregation of preferences we have mentioned in the introductory section.

Instead of *preferences* the relations may indicate *dominances* (or *hierarchies*) on the object set. These applications occur in behavioral sciences in the study of the dominance relationship in a group of animals (see Marcotorchino and Michaud [1979]). Applications in marketing are mentioned by Slater [1961] and Reinelt [1985], on the design of publicity campaign for products based on voting upon different types of advertisements.

When the relations represent *similarities* (that can be deduced by considering the attributes of the objects), and the aim is that of finding a best partition of the object set into 'homogeneous' disjoint classes (or clusters), we have the so-called clustering problem. In this case, we are given a set of m relations (each relation indicating the similarities of the objects with respect to one attribute) and we are looking for an equivalence relation. Thus here we have the median problem MP($\mathcal{R}, \mathcal{M}$), where $\mathcal{M}$ is an equivalence relation.

Clustering problems occur in many areas: zoology, botanics, sociology, politics and economics. A number of real problems that were modelled as clustering problems coming from these different areas can be found in Grötschel and Wakabayashi [1989].

As we have mentioned, the median problems that we have considered can be reduced to optimization problems on weighted digraphs. Some of these digraph problems have been largely investigated, in special, the linear ordering problem and the clique partitioning problem. The latter solves the problem of finding a median that is an equivalence relation.

Thus, the strategies that have been developed to solve these problems can be used to solve the corresponding median problem. These strategies go from simple heuristics to sophisticated branch-and-cut algorithms.

For the linear ordering problem, Reinelt [1985] developed a branch-and-cut algorithm and reported very good computational results obtained by solving a number of problems in economics, in special triangulation problems for input-output tables (see also Grötschel, Jünger and Reinelt [1984a, 1984b]). This algorithm combines the cutting plane approach and branch-and-bound techniques. Marcotorchino and Michaud [1979] developed heuristics and reported computations on some ranking problems. Other problems that are equivalent to the linear ordering problem are the acyclic subdigraph problem (Jünger [1985]) and the permutation problem (Young [1978]). Exact methods were developed by de Cani [1972], Kaas [1981], Tüshaus [1983], and others.

For the clique partitioning problem, Wakabayashi [1986] developed a branch-and-cut algorithm (see also Grötschel and Wakabayashi [1988] for the computational results on the performance of the proposed algorithm). Results on the polyhedral investigations of the clique partitioning polytope that were used to develop this algorithm can be found in Grötschel and Wakabayashi [1990a, 1990b]. Among the heuristic methods that were proposed for the clustering problem we mention Marcotorchino and Michaud [1981], and Schader [1981]. Exact methods were developed by Tüshaus [1983]; Schader and Tüshaus [1985] also proposed an approach that combines heuristics with a subgradient method.

Polyhedral results for the partial order polytope were obtained by Müller [1996]; Gurgel [1992] investigated the facial structure of the complete preorder polytope (see also Gurgel and Wakabayashi [1993]). These results can be used in the design of a cutting plane based algorithm to solve the partial ordering problem and the complete preorder problem. Exact methods for these problems were investigated by Tüshaus [1981, 1983].

6. References

Barbut, M. [1961] Médiane, distributivité, éloignement. *Mathématiques et Sciences Humaines 70, 5-31 (reproduced in 1980).*

Barbut, M. [1967] Médianes, Condorcet et Kendall. *Mathématiques et Sciences*

Humaines 69, 5-13 (reproduced in 1980).

BARTHÉLEMY, J.P. [1979] Caractérisations axiomatiques de la distance de la différence symétrique entre des relations binaires. *Mathématiques et Sciences Humaines 67, 85-113.*

BARTHÉLEMY, J.P. [1981] Trois propriètés de la médiane dans un treillis modulare. *Mathématiques et Sciences Humaines 75, 83-91.*

BARTHÉLEMY, J.P. [1988] Social welfare and aggregation procedures: combinatorial and algorithmic aspects. *in: Applications of Combinatorics and Graph Theory to the Biological and Social Sciences.*

BARTHÉLEMY, J.P; FLAMENT, C. AND MONJARDET, B. [1982] Ordered Sets and Social Sciences. *in: Ordered Sets (I. Rival, ed.), D. Reidl, Dordrech, 721-758.*

BARTHÉLEMY, J.P.; GUENOCHE, A. AND HUDRY, O. [1988] Median linear orders: heuristics and a branch-and-bound algorithm. *manuscript in preparation.*

BARTHÉLEMY, J.P. AND MONJARDET, B. [1981] The median procedure in cluster analysis and social choice theory. *Mathematical Social Sciences 1, 235-267.*

BARTHÉLEMY, J.P. AND MONJARDET, B. [1988] The median procedure in data analysis: new results and open problems. *in: Classification and Related Methods of Data Analysis (H.H. Bock ed.) North-Holland, Amsterdam, 309-316.*

BARTHÉLEMY, J.P.; LECLERC, B. AND MONJARDET, B. [1986] On the use of ordered sets in problems of comparison and consensus classifications. *Journal of Classification 3, 187-224.*

BARTHOLDI, J.J. III; TOVEY, C.A. AND TRICK, M.A. [1988] Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare.*

BONDY, J.A. AND MURTY, U.S.R. [1976] *Graph Theory with Applications, Macmillan, London.*

DE CANI, J.S. [1972] A branch and bound algorithm for maximum likelihood paired comparison ranking. *Biometrika 59, 131-135.*

DAY, W.H.E. [1988] Consensus methods as tools for data analysis. *in: Classification and Related Methods of Data Analysis (H.H. Bock ed.) North-Holland, Amsterdam.*

FISHBURN, P.C. [1977] Condorcet social choice functions. *SIAM Journal of Applied Mathematics 33, 469-489.*

GAREY, M.R. AND JOHNSON, D.S. [1979] Computers and Intractability: A Guide to the Theory of NP-completeness. *Freeeman, San Francisco, 1979.*

GAREY, M.R.; JOHNSON, D.S. AND STOCKMEYER, L. [1976] Some simplified NP-complete graph problems. *Theoretical Computer Science 1, 237-267.*

GRÖTSCHEL, M., JÜNGER, M. AND REINELT, G. [1984a] A cutting plane algorithm for the linear ordering problem. *Operations Research, vol. 32 (6), 1195-1220.*

GRÖTSCHEL, M., JÜNGER, M. AND REINELT, G. [1984b] Acyclic subdigraphs and linear orderings: polytopes, facets, and a cutting plane algorithm. *Graphs and order (Banff, Alta., 1984), 217–264, NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., 147, Reidel, Dordrecht-Boston, Mass.*

GRÖTSCHEL, M. AND WAKABAYASHI, Y. [1989] A cutting plane algorithm for a clustering problem. *Mathematical Programming 45 (1), 59-96.*

GRÖTSCHEL, M. AND WAKABAYASHI, Y. [1990a] Facets of the clique partitioning polytope. *Mathematical Programming 47, Series A, 367-387.*

GRÖTSCHEL, M. AND WAKABAYASHI, Y. [1990b] Composition of facets of the clique partitioning polytope. *Topics in combinatorics and graph theory (Oberwolfach, 1990), 271–284, Physica, Heidelberg.*

GURGEL, M.A.M.C. [1992] Poliedros de Grafos Transitivos. *Thesis (in Portuguese).*

GURGEL, M.A.M.C. AND WAKABAYASHI, Y. [1993] The complete preorder polytope: facets and separation problem. *Technical Report RT-MAC-9303, IME-USP, Brazil.*

JÜNGER, M. [1985] Polyhedral Combinatorics and the Acyclic Subdigraph Problem. *Research and Expositions in Mathematics, 7. Heldermann Verlag, Berlin.*

KAAS, R. [1981] A branch and bound algorithm for the acyclic subgraph problem. *European Journal of Operations Research 8, 335-362.*

KARP, R.M. [1972] Reducibility among combinatorial problems. *in: Complexity of Computer Computations (R.E. Miller and J.W. Thatcher eds.) Plenum Press, N.Y., 85-103.*

KEMENY, J.G. [1959] Mathematics without numbers. *Daedalus, 88, 577-591.*

KŘIVÁNEK, M. AND MORÁVEK, J. [1986] NP-hard problems in hierarchical-tree clustering. *Acta Informatica 23, 311-323.*

LECLERC, B. [1988a] Consensus Applications in the Social Sciences. *in: Classification and Related Methods of Data Analysis (H.H. Bock ed.) North-Holland, Amsterdam.*

LECLERC, B. [1988b] Medians and majorities in semimodular lattices. *Research Report C.A.M.S.- P.031, CAMS-EHESS, Paris.*

MARCOTORCHINO, J.F. AND MICHAUD, P. [1979] Optimisation en Analyse Ordinale des Donns. *Masson, Paris.*

MARCOTORCHINO, J.F. AND MICHAUD, P. [1981] Heuristic approach to the similarity aggregation problem. *Methods of Operations Research, Vol. 43, 395-404.*

MICHAUD, P. [1987] Condorcet — a man of the avant-garde. *Applied stochastic models and data analysis, Vol.3, 173-189.*

MIRKIN, B.G. [1974] The problems of approximation in space of relations and qualitative data analysis. *Automatika i Telemechanica, translated in : Information and Remote Control 35 (9), 1424-1431.*

MONJARDET, B. [1978] An axiomatic theory of tournament aggregation. *Math. Oper. Research 3 (4), 334-351.*

MONJARDET, B. [1980] Théorie et applications de la médiane dans les treillis distributifs finis. *Annals of Discrete Math. 9, 87-91.*

MONJARDET, B. [1987] Arrowian characterization of latticial federation consensus functions. *Research report C.A.M.S. - P.030, CAMS-EHESS, Paris.*

MÜLLER, R. [1996] On the partial order polytope. *Mathematical Programming 73 , Series A, 31-49.*

ORLIN, J.B. [1988] *Unpublihed manuscript.*

RÉGNIER, S. [1965] Sur quelques aspects mathématiques des problèmes de classification automatique. *I.C.C. Bulletin 4, 175-191, reprod. 1983 in Math. Sci. Hum. 82, 13-29.*

REINELT, G. [1985] The linear ordering problem: algorithms and applications. *Research and Exposition in Mathematics 8, Heldermann Verlag, Berlin.*

SCHADER, M. [1981] Scharfe und unscharfe Klassification qualitativer Daten. *Athenäum, Königstein.*

SCHADER, M. AND TÜSHAUS [1985] Ein Subgradientenverfahren zur Klassification qualitativer Daten. *O.R. Spektrum 7, 1-5.*

SLATER, P. [1961] Inconsistencies in schedule of paired comparisons. *Biometrika 48, 202-312.*

TÜSHAUS, U. [1981] Approximation of binary relations by preorders. *Methods of Operations Research. Vol. 43, 445-454.*

TÜSHAUS, U. [1983] Aggregation binärer Relationen in der qualitativen Datenanalyse. *Mathematical Systems in Economics 82..*

WAKABAYASHI, Y. [1986] Aggregation of binary relations: algorithmic and polyhedral investigations. *Thesis, Universität Augsburg, Germany.*

YOUNG, H.P. [1978] On permutations and permutation polytopes. *Mathematical Programming Study 8, 129-140.*

YOUNG, H.P. [1990] Condorcet's theory of voting. *Math. Inform. Sci. Humaines, no. 111, 45-59.*

YOUNG, H.P. AND LEVENGLICK, A. [1978] A consistent extension of Condorcet's election principle. *SIAM Journal of Applied Mathematics 35, 285-300.*

ZAHN, C.T. [1964] Approximating symmetric relations by equivalence relations. *SIAM Journal of Applied Mathematics 12, 840-847.*

Yoshiko Wakabayashi
Instituto de Matemática e Estatística,
Universidade de São Paulo,
Rua do Matão 1010 – Cidade Universitária
05508-900, São Paulo SP
yw@ime.usp.br
Brazil