Computation of median orders: complexity results

Olivier Hudry[*]

Abstract

Given a set of individual preferences defined on a same finite set of candidates, we consider the problem of aggregating them into a collective preference minimizing the number of disagreements with respect to the given set and verifying some structural properties like transitivity. We study the complexity of this problem when the individual preferences as well as the collective one must verify different properties, and we show that the aggregation problem is NP-hard for different types of collective preferences, even when the individual preferences are linear orders.

Key words: Complexity, partially ordered relations, median relations, aggregation of preferences.

1 Introduction

The problem that we deal with in this paper can be stated as follows: given a set (called a *profile*) $\Pi = (R_1, R_2, \ldots, R_m)$ of m binary relations R_i ($1 \leq i \leq m$) defined on the same finite set X, find a binary relation R^* defined on X verifying certain properties like transitivity and summarizing Π as accurately as possible. This problem occurs in different fields, for instance in the social sciences, in electrical engineering, in agronomy or in mathematics (see for example L. Hubert (1976), J.-P. Barthélemy *et alii* (1981, 1986, 1988, 1989, 1995), M. Jünger (1985), G. Reinelt (1985), A. Guénoche *et*

---

[*] Olivier Hudry, École nationale supérieure des télécommunications, 46, rue Barrault, 75634 Paris cedex 13, France. Olivier.Hudry@enst.fr

alii (1994)). For example, in voting theory, X can be considered as a set of candidates, Π as a profile of individual preferences expressed by voters and R^* as the collective preference that we look for. Though the problem occurs in different fields, as said above, we shall keep this illustration from voting theory in the following.

The aim of this paper is to study the complexity of finding R^* (for the theory of complexity, see for instance M.R. Garey and D.S. Johnson (1979) or J.-P. Barthélemy *et alii* (1996)). We consider different types of ordered relations for the individual preferences of Π as well as for R^* and we show that for most cases, the computation of R^* is an NP-hard problem. This problem has been already studied in some special cases, namely for the aggregation of a profile of linear orders into a linear order by J.B. Orlin (1988) and by J.J. Bartholdi III, C.A. Tovey and M.A. Trick (1989), and for the aggregation of a profile of binary relations into a linear order, a partial order, a complete preorder or a preorder (see below for the definitions of these structures) by Y. Wakabashi (1986 and 1998). The results displayed in this paper generalize the previous ones by extending them to other cases. They slightly strengthen and sometimes generalize the ones presented in O. Hudry (1989).

In the following, the relations to aggregate are assumed to represent preferences, and thus will not be symmetric. Anyway, the aggregation of symmetric relations has also been studied: M. Krivanek and J. Moravek (1986) showed that the approximation of a symmetric relation by an equivalence relation (a reflexive, symmetric, and transitive relation) is NP-hard. This case corresponds with the aggregation of a profile reduced to only one symmetric relation while R^* is assumed to be an equivalence relation. From this, we may derive that the aggregation of several symmetric relations or of equivalence relations into one equivalence relation is also NP-hard (see J.-P. Barthélemy and B. Leclerc (1995)). On contrary, the aggregation of symmetric relations or of equivalence relations into a symmetric relations is trivially polynomial.

The paper is organized as follows. Section 2 recalls the definitions of the ordered relations that we take into account. In Section 3, we show how the aggregation problems can be formulated in graph theoretical terms. Then we prove our complexity results upon these aggregation problems in Section 4. The conclusions take place in Section 5 and summarize the main results got in Section 4.

2 The ordered relations

Given a finite set X, a binary relation R defined on X is a subset of $X \times X = \{(x, y) : x \in X \text{ and } y \in X\}$. We note n the number of elements of X and we suppose that n is great enough (typically, at least equal to 4). We note xRy instead of

$(x, y) \in R$ and $x\overline{R}y$ instead of $(x, y) \notin R$. The following properties that a binary relation R can satisfy are basic:

- *reflexive*: $\forall\ x \in X, xRx$;
- *irreflexive*: $\forall\ x \in X,\ x\overline{R}x$;
- *antisymmetric*: $\forall\ (x, y) \in X^2$, (xRy and $x \neq y$) $\Rightarrow\ y\overline{R}x$;
- *asymmetric*: $\forall\ (x, y) \in X^2, xRy \Rightarrow\ y\overline{R}x$;
- *transitive*: $\forall\ (x, y, z) \in X^3$, ($xRy$ and yRz) $\Rightarrow xRz$;
- *complete*: $\forall\ (x, y) \in X^2$ with $x \neq y$, xRy or (inclusively) yRx.

From a binary relation R, we may define an asymmetric relation R^a (called the *asymmetric part* of R) by: $xR^ay \Leftrightarrow$ (xRy and $y\overline{R}x$).

By combining the above properties, we may define different types of binary relations (see for instance J.-P. Barthélemy and B. Monjardet (1981) or P.C. Fishburn (1985)). As a binary relation R defined on X is the same as the oriented graph $G = (X, R)$ (*i.e.* (x, y) is an arc of G if and only if we have xRy), we illustrate these types with graph theoretic examples:

- a *partial order* is an asymmetric and transitive binary relation; $\mathcal{O}$ will denote the set of the partial orders defined on X;

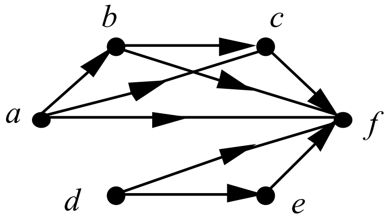

Figure 1. A partial order.

- a *linear order* is a complete partial order; $\mathcal{L}$ will denote the set of the linear orders defined on X;

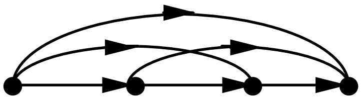

Figure 2. A linear order. The partial order of Figure 1 is not a linear order, for instance because the vertices a and d are not compared.

- a *tournament* is a complete and asymmetric binary relation; $\mathcal{T}$ will denote the set of the tournaments defined on X; notice that a transitive tournament is a linear order and conversely;

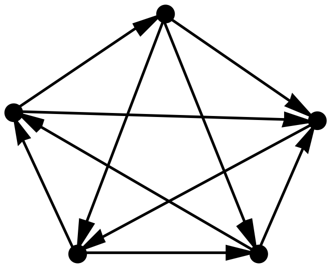

Figure 3. A tournament.

- a *preorder* is a reflexive and transitive binary relation; $\mathcal{P}$ will denote the set of the preorders defined on X;

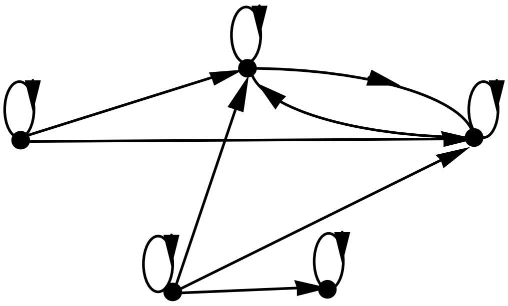

Figure 4. A preorder.

- a *complete preorder* is a reflexive, transitive and complete binary relation; $\mathcal{C}$ will denote the set of the complete preorders defined on X;

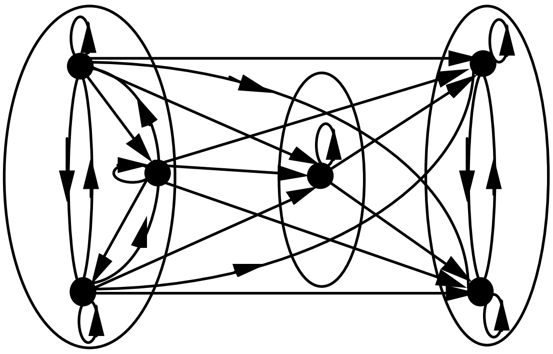

Figure 5. A complete preorder.

• a *weak order* is the asymmetric part of a complete preorder; $\mathcal{W}$ will denote the set of the weak orders defined on X;

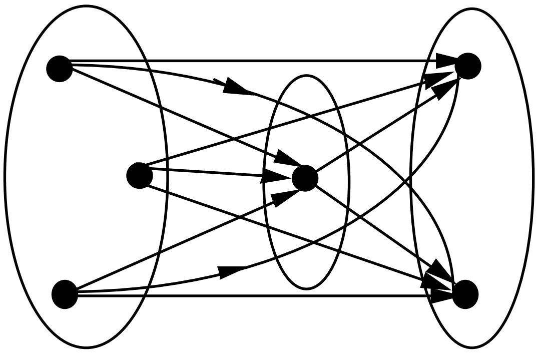

Figure 6. A weak order (namely, the asymmetric part of the complete preorder of Figure 5).

• an *interval order* is a partial order R satisfying: $\forall$ $(x, y, z, t) \in X^4$, (xRy and zRt) $\Rightarrow$ {xRt or (inclusively) zRy}; $\mathcal{I}$ will denote the set of interval orders defined on X (the name *interval order* comes from the fact that we may represent such an order by intervals spread on the real axis and associated with each element x of X: then xRy means that the interval associated with x is completely on the left of the one associated with y, while $x\overline{R}y$ and $y\overline{R}x$ mean that the intervals associated with x and y overlap; the above condition means that if x is on the left of y and z on the left of t, then the intervals associated with x and t on one hand and the ones associated with z and y on the other hand cannot overlap simultaneously).

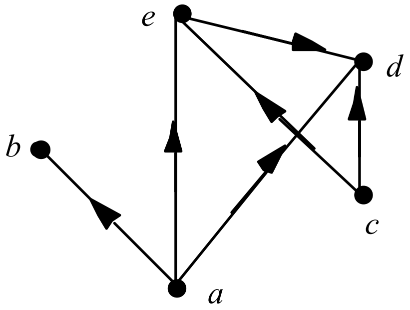

Figure 7. An interval order .The partial order of Figure 1 is not an interval order for instance because of the vertices b, c, d, e: bRc and dRe but we have not bRe nor dRc.

- a *semiorder* is an interval order R satisfying: $\forall$ $(x, y, z, t) \in X^4$, (xRy and yRz) $\Rightarrow$ {xRt or (inclusive) tRz}; $\mathcal{S}$ will denote the set of interval orders defined on X (with respect to the representation as intervals, an interval orders is a semiorder if we may associate intervals with the same length to all the elements of X).

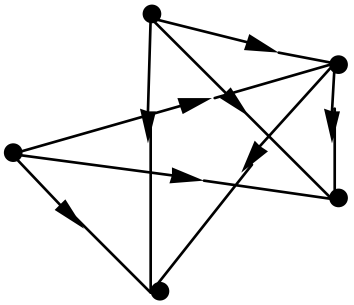

Figure 8. A semiorder .The interval order of Figure 7 is not a semiorder for instance because of the vertices b, c, d, e: cRe and eRd but we have not cRb nor bRd.

- a *quasi-order* is a reflexive and complete relation of which the asymmetric part is a semiorder; $\mathcal{Q}$ will denote the set of the quasi-orders defined on X;

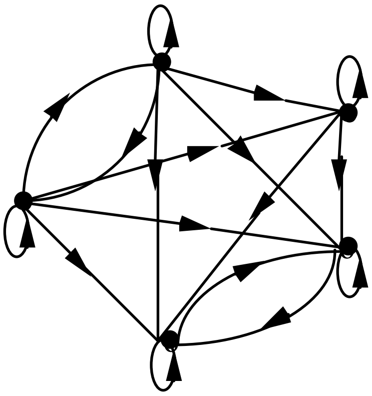

Figure 9. A quasi-order. Its asymmetric part is the semiorder of Figure 8.

- an *acyclic relation* is a relation R without directed cycle (*circuit*), *i.e.* verifying: $\forall$ $1 \leq k \leq n$, ($x_i R x_{i+1}$ for $1 \leq i \leq k - 1$) $\Rightarrow$ $x_k \overline{R} x_1$; $\mathcal{A}$ will denote the set of acyclic relations defined on X.

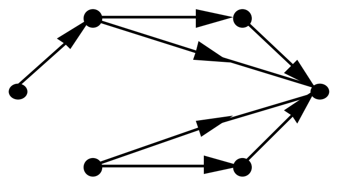

Figure 9. An acyclic relation. Its transitive closure is the partial order of Figure 1.

Checking that a given relation (or a given graph) fulfils the requirements of these structures can be done in polynomial time with respect to n. From this remark, it will follow that the problems considered below all belong to NP.

It is possible to get other structures by adding or by removing reflexivity or irreflexivity from the above definition (and by changing asymmetry by antisymmetry when necessary). In fact, the distinction between reflexive and irreflexive relations is not relevant for our study, as we shall see below: the complexity results will remain the same. Thus, in the following, we do not take reflexivity or irreflexivity into account (for instance, we will consider that a linear order is also a preorder).

These types include the most studied and used partially ordered relations. We will also consider generic binary relations, without any particular property. The set of the binary relations will be noted $\mathcal{R}$. We may notice several inclusions between these sets, especially the following one: $\forall\ \mathcal{Z} \in \{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{W}\}$, $\mathcal{L} \subseteq \mathcal{Z}$; in other words, a linear order can be considered as a special case of any one of the other types.

3 Formulations of the aggregation problem

In order to get an optimization problem to deal with, it is necessary to explicit what we mean when we say that R^* must summarize Π "as accurately as possible". To do so, we consider the symmetric difference distance δ: given two binary relations R and S defined on the same set X, we have

$$\delta(R,S) = \left|\left\{(x,y) \in X^2 : \left[xRy \text{ and } x\bar{S}y\right] \text{ or } \left[x\bar{R}y \text{ and } xSy\right]\right\}\right|$$

This quantity $\delta(R,S)$ measures the number of disagreements between R and S. Though some authors consider sometimes another distance, δ is used widely and is appropriate for many applications. J.-P. Barthélemy (1979) shows that δ satisfies a number of naturally desirable properties and J.-P. Barthélemy and B. Monjardet (1981) recall that $\delta(R,S)$ is the Hamming distance between the characteristic vectors of R and S and point out the links between δ and the L_1 metric or the square of the Euclidean distance between these vectors (see also K.P. Bogart (1973 and 1975) and B. Monjardet (1979 and 1990)).

Then, for a profile $\Pi = (R_1, R_2, \ldots, R_m)$ of m relations, we can define the *remoteness* $\Delta(\Pi, R)$ (J.-P. Barthélemy and B. Monjardet (1981)) between a relation R and the profile Π by:

$$\Delta(\Pi, R) = \sum_{i=1}^{m} \delta(R, R_i).$$

The remoteness $\Delta(\Pi, R)$ measures the total number of disagreements between Π and R.

Our aggregation problem can be seen now as a combinatorial problem: given a profile Π, determine a binary relation R^* minimizing Δ over one of the sets $\mathcal{A}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{L}$, $\mathcal{O}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, $\mathcal{W}$. Such a relation R^* will be called a *median relation* of Π (J.-P. Barthélemy and B. Monjardet (1981)). According to the number m of relations of the profile and to the properties assumed for the relations belonging to Π or required from the median relation, we get many combinatorial problems. They are too numerous to state them explicitly; so we note them as follows:

Problems $P_f(\mathcal{Y}, \mathcal{Z})$. For $\mathcal{Y}$ belonging to $\{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{W}\}$ and $\mathcal{Z}$ belonging also to $\{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{W}\}$, for a function f defined from the set **N** of integers to **N**, $P_f(\mathcal{Y}, \mathcal{Z})$ denotes the following problem: given a finite set X of n elements, given a profile Π of $m = f(n)$ binary relations all belonging to $\mathcal{Y}$, find a relation R^* belonging to $\mathcal{Z}$ with: $\Delta(\Pi, R^*) = \underset{R \in \mathcal{Z}}{\mathrm{Min}}\, \Delta(\Pi, R)$.

An interesting case is the one for which f is a constant m, i.e. the particular case for which the number m of relations is fixed outside the instance. We will denote this problem by $P_m(\mathcal{Y}, \mathcal{Z})$. For instance, $P_2(\mathcal{R}, \mathcal{L})$ will denote the aggregation of 2 binary relations into a linear order. We will see that the parity of m will play a role in the following results. Anyway, it will be easy to see from the following computations that, if $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard for some function f and some sets $\mathcal{Y}$ and $\mathcal{Z}$, then $P_{f+2}(\mathcal{Y}, \mathcal{Z})$ will also be NP-hard: it will be sufficient to add any linear order and its reverse order to the considered profile to get this result (since the linear orders are special cases of any one of the other types).

We do not explicit the statements of the decision problems associated with the problems $P_f(\mathcal{Y}, \mathcal{Z})$, because they are obvious. Similarly, it is obvious to show that these decision problems belong to NP. Thus, we deal with the NP-hardness of $P_f(\mathcal{Y}, \mathcal{Z})$, but we could deal with the NP-completeness of the decision problems associated with $P_f(\mathcal{Y}, \mathcal{Z})$.

To study the complexity of $P_f(\mathcal{Y}, \mathcal{Z})$, we develop the expression of Δ. For this, consider the characteristic vectors $r^i = \left(r^i_{xy}\right)_{(x,y)\in X^2}$ of the relations R_i $(1 \leq i \leq m)$ defined by $r^i_{xy} = 1$ if xR_iy and $r^i_{xy} = 0$ otherwise, and similarly the characteristic vector $r = \left(r_{xy}\right)_{(x,y)\in X^2}$ of any binary relation R. Then, it is easy to get a linear expression of $\Delta(\Pi, R)$:

$$\delta(R, R_i) = \sum_{(x,y)\in X^2} \left| r_{xy} - r^i_{xy} \right| = \sum_{(x,y)\in X^2} \left| r_{xy} - r^i_{xy} \right|^2 = \sum_{(x,y)\in X^2} \left[r_{xy}\left(1 - 2r^i_{xy}\right) + r^i_{xy} \right]$$

hence $$\Delta(\Pi, R) = \sum_{i=1}^{m} \sum_{(x,y)\in X^2} \left| r_{xy} - r^i_{xy} \right|$$

and, after simplifications: $$\Delta(\Pi, R) = C - \sum_{(x,y)\in X^2} m_{xy} \, . \, r_{xy}$$

with $C = \sum_{i=1}^{m} \sum_{(x,y)\in X^2} r^i_{xy}$ and $m_{xy} = \sum_{i=1}^{m} \left(2r^i_{xy} - 1\right) = 2\sum_{i=1}^{m} r^i_{xy} - m$.

Notice that the quantities m_{xy} can be non-positive or non-negative, and that they all have the same parity (the one of m). Notice also that, from this expression of $\Delta(\Pi, R)$, it is easy to get a 0-1 linear programming formulation of the problems $P_f(\mathcal{Y}, \mathcal{Z})$ by adding the 0-1 linear constraints associated with each type of median relation (but it will not be the way that we are going to follow in the sequel). For example, the transitivity of R can be written: $\forall (x, y, z) \in X^3$, $r_{xy} + r_{yz} - r_{xz} \leq 1$ (see for instance Y. Wakabayashi (1986) or O. Hudry (1989) for details). Such a 0-1 linear programming formulation was applied as soon as 1960 (A.W. Tucker (1960); see also D.H. Younger (1963), J.S. de Cani (1969), D. Arditti (1984), and more generally J.-P. Barthélemy and B. Monjardet (1981) for references).

Before going further, the following lemma shows that reflexivity or irreflexivity of the median relation do not change the complexity of the problems $P_f(\mathcal{Y}, \mathcal{Z})$.

Lemma 1. For any set $\mathcal{Z}$ of median relations, let $\mathcal{Z}_r$ (resp. $\mathcal{Z}_i$) be the set of median relations got from the elements of $\mathcal{Z}$ by adding the reflexivity (resp. irreflexivity) property. Then, for any set $\mathcal{Y}$ and any function f, $P_f(\mathcal{Y}, \mathcal{Z})$, $P_f(\mathcal{Y}, \mathcal{Z}_r)$, and $P_f(\mathcal{Y}, \mathcal{Z}_i)$ have the same complexity.

Proof. To show this result, consider any profile Π of m $(= f(n))$ relations belonging to $\mathcal{Y}$ and any relation Z belonging to $\mathcal{Z}$. Let Z_r (resp. Z_i) be the reflexive (resp. irreflexive) relation got from Z by adding the reflexivity (resp. irreflexivity) property. Then it is easy to state the following relations:

$$\Delta(\Pi, Z_r) = \Delta(\Pi, Z) + \sum_{x:(x,x)\notin Z} m_{xx} \text{ and } \Delta(\Pi, Z_i) = \Delta(\Pi, Z) - \sum_{x:(x,x)\in Z} m_{xx}.$$

Hence the result, since the computation of $\sum_{x:(x,x)\notin Z} m_{xx}$ and of $\sum_{x:(x,x)\in Z} m_{xx}$ can trivially be done in polynomial time w.r.t. the size of the considered instance. ❑

Because of Lemma 1, we shall not pay attention from now on to reflexivity or irreflexivity: all the complexity results remain the same if we add or remove reflexivity or irreflexivity.

In the following, we will not consider the previous 0-1 linear programming formulation to study the complexity of the problems $P_f(\mathcal{Y}, \mathcal{Z})$, but a graph theoretic representation. Indeed, we may associate a complete, symmetric, weighted, oriented graph $G = (X, U)$ to any profile Π: the vertex set of G is X and G owns all the arcs (i.e. oriented edges) that a simple graph can own; in other words, we have: $U = X \times X - \{(x,x) \text{ for } x \in X\}$ (remember that reflexivity does not matter now on). In the following, we will write U_X to denote the set $X \times X - \{(x,x) \text{ for } x \in X\}$ and the graph associated with Π is thus $G = (X, U_X)$. The arcs (x, y) of G (with $x \in X$, $y \in X$ and $x \neq y$) are weighted by m_{xy}. Then minimizing $\Delta(\Pi, Z)$ for Z belonging to one of the sets $\mathcal{A}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{L}$, $\mathcal{O}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, $\mathcal{W}$ is exactly the same as extracting a partial graph $H = (X, Z)$ from G in order to maximize $\sum_{(x,y)\in Z} m_{xy}$ while the kept arcs describe the structure that Z must respect (H must belong to $\mathcal{A}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{L}$, $\mathcal{O}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, $\mathcal{W}$, where these sets are seen as sets of graphs).

Then the question arises: which weighted graphs $G = (X, U_X)$ can be associated to a profile Π? By combining results by P. Erdös, L. Moser (1964) and by B. Debord (1987) (see also D. McGarvey (1953) and R. Stearns (1959)), we get such characterisations, which depend on the nature of the relations of Π. For the statements of the following theorems, let M denote the highest absolute value of the weights of G: $M = \underset{(x,y) \in U_X}{\text{Max}} \left| m_{xy} \right|$.

Theorem 2. The graph $G = (X, U_X)$ weighted by the (non-positive or non-negative) integers m_{xy} represents a profile Π of m binary relations if the following conditions are fulfilled:

1. all the weights m_{xy} have the same parity;
2. m has the same parity as the weights m_{xy};
3. $m \geq M$.

Theorem 3. The graph $G = (X, U_X)$ weighted by the (non-positive or non-negative) integers m_{xy} represents a profile Π of m tournaments if the following conditions are fulfilled:

1. all the weights m_{xy} have the same parity;
2. m has the same parity as the weights m_{xy};
3. $m \geq M$;
4. $\forall \ (x,y) \in U_X, \ m_{xy} = -m_{yx}$.

Theorem 4. The graph $G = (X, U_X)$ weighted by the (non-positive or non-negative) integers m_{xy} represents a profile Π of m linear orders if the following conditions are fulfilled:

1. all the weights m_{xy} have the same parity;
2. m has the same parity as the weights m_{xy};
3. $m \geq \dfrac{c.n.M}{\log n}$ where c is a constant;
4. $\forall \ (x,y) \in U_X, \ m_{xy} = -m_{yx}$.

Notice that, for Theorems 2 and 3, M is the lowest possible value of m. For Theorem 4, it is sometimes possible to find a profile of m linear orders associated with

G with m less than $\frac{c.n.M}{\log n}$. Anyway, in all these cases, there exists a profile Π with $m = M$ binary relations, or $m = M$ tournaments, or $m = \left|\frac{c.n.M}{\log n}\right|$ or $m = \left|\frac{c.n.M}{\log n}\right| + 1$ (depending on the parity of $\left|\frac{c.n.M}{\log n}\right|$ and of the weights of G) linear orders. Moreover, if we assume that M is upper-bounded by a polynomial in n (as it will be the case further), then the construction of Π can be done in polynomial time with respect to the size of G. Indeed, as any binary relation R defined on X can be described by $O(n^2)$ bits, it is possible to code $\Pi = (R_1, R_2, \ldots, R_M)$ with $O\left(M.n^2\right)$ bits, while the size of G is at least n^2 (at least 1 bit for the weight of each arc of G); hence the result. Notice also that, if M is upper-bounded by a constant, it is possible to fix the number m of relations of Π in Theorems 2 and 3; in this case, we may associate to G an instance of the problems $P_m(\mathcal{R}, \mathcal{Z})$ (Theorem 2) or the problems $P_m(\mathcal{T}, \mathcal{Z})$ (Theorem 3) for an appropriate set $\mathcal{Z}$.

From this polynomial link between the problems $P_f(\mathcal{Y}, \mathcal{Z})$ and their graph theoretic representations, it appears that we may study the complexity of the problems $P_f(\mathcal{Y}, \mathcal{Z})$ with the help of weighted graphs. It is what we do below. More precisely, we are going to study the following decision problems, stated as graph theoretic problems:

Problems $Q_0(\mathcal{Y}, \mathcal{Z})$ with $\mathcal{Y} \in \{\mathcal{L}, \mathcal{R}, \mathcal{T}\}$ and $\mathcal{Z} \in \{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{S}, \mathcal{W}\}$

Instance: a graph $G = (X, U_X)$ weighted by (non-positive or non-negative) even integers m_{xy} and which represents a profile of relations belonging to $\mathcal{Y}$; an integer K;

Question: does there exist a partial graph (X, U) of G belonging to $\mathcal{Z}$ with $\sum_{(x,y) \in U} m_{xy} \geq K$?

Problems $Q_1(\mathcal{Y}, \mathcal{Z})$ with $\mathcal{Y} \in \{\mathcal{L}, \mathcal{R}, \mathcal{T}\}$ and $\mathcal{Z} \in \{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{S}, \mathcal{W}\}$

Instance: a graph $G = (X, U_X)$ weighted by (non-positive or non-negative) odd integers m_{xy} and which represents a profile of relations belonging to $\mathcal{Y}$; an integer K;

Question: does there exist a partial graph (X, U) of G belonging to $\mathcal{Z}$ with $\sum_{(x,y) \in U} m_{xy} \geq K$?

4 The complexity results

As for the problems $P_f(\mathcal{Y}, \mathcal{Z})$, the problems $Q_0(\mathcal{Y}, \mathcal{Z})$ and $Q_1(\mathcal{Y}, \mathcal{Z})$ obviously belong to NP for $\mathcal{Y} \in \{\mathcal{L}, \mathcal{R}, \mathcal{T}\}$ and $\mathcal{Z} \in \{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{S}, \mathcal{W}\}$. To show that they are NP-complete, we use the well-known Feedback Arcset Problem (see M.R. Garey, D.S. Johnson (1979)):

Instance: a directed, asymmetric graph $H = (X, W)$; an integer h;

Question: does there exist $W' \subset W$ with $|W'| \leq h$ and such that W' contains at least one arc of each circuit (directed cycle) of H ?

R. Karp (1972) showed that this problem is NP-complete. We may also state it as follows:

Problem FAS

Instance: a directed, asymmetric graph $H = (X, W)$; an integer h;

Question: does there exist $W' \subset W$ with $|W'| \leq h$ and such that removing the elements of W' from H leaves a graph $(X, W - W')$ without any circuit (such a set W' is called a *feedback arc set* of H of cardinality at most h) ?

In the following, we use this latter formulation. We use also the following (obvious) lemma:

Lemma 5.

a. Any partial graph of a graph without circuit is itself without circuit.

b. Any graph without circuit can be completed into a linear order by adding appropriate arcs.

We now pay attention to the complexity of $Q_0(\mathcal{R}, \mathcal{Z})$ and $Q_1(\mathcal{R}, \mathcal{Z})$ (i.e. when the graph represents a profile of any binary relations) for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$ (i.e. when we look for an acyclic relation, an interval order, a linear order, a partial order, a semiorder or a weak order).

Theorem 6. The problems $Q_0(\mathcal{R}, \mathcal{Z})$ are NP-complete for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$.

Proof. We polynomially transform any instance $H = (X, W)$ and h of FAS into an instance G and K of $Q_0(\mathcal{R}, \mathcal{Z})$. For this, we set the vertex set of G as being X; hence the set of arcs of G: U_X. We define the weights m_{xy} of the arcs (x, y) of G and K as follows:

* if $(x, y) \in W$, $m_{xy} = 2$

* if $(x, y) \notin W$, $m_{xy} = 0$.

Then we set $K = 2(|W| - h)$.

This transformation is obviously polynomial. Let us show that it keeps the answer « yes » or « no ».

Indeed, assume that there exists a subset W' of W with $|W'| \leq h$ and such that removing the elements of W' from H leaves a graph $(X, W - W')$ without any circuit. If we consider it as a partial graph of G, its weight is $2(|W| - |W'|)$, which is greater than or equal to $2(|W| - h) = K$. If $\mathcal{Z} = \mathcal{A}$, we are done. Otherwise, thanks to Lemma 5 b, it is possible to complete $(X, W - W')$ into a linear order by adding extra arcs. As all the weights are non-negative, we get a linear order (X, L), that we may consider as a partial order if $\mathcal{Z} = \mathcal{O}$, or as an interval order if $\mathcal{Z} = \mathcal{I}$, or as a semiorder if $\mathcal{Z} = \mathcal{S}$, or as a weak order if $\mathcal{Z} = \mathcal{W}$, with $\sum_{(x,y) \in L} m_{xy} \geq K$.

Conversely, assume that the instance (G, K) of $Q_0(\mathcal{R}, \mathcal{Z})$ admits the answer « yes »: there exists a partial graph (X, U) of $G = (X, U_X)$ which is without circuit if $\mathcal{Z} = \mathcal{A}$, or a linear order if $\mathcal{Z} = \mathcal{L}$, or a partial order if $\mathcal{Z} = \mathcal{O}$, or an interval order if $\mathcal{Z} = \mathcal{I}$, or a semiorder if $\mathcal{Z} = \mathcal{S}$, or a weak order if $\mathcal{Z} = \mathcal{W}$, with $\sum_{(x,y) \in U} m_{xy} \geq K$. In every case, (X, U) is without circuit. By definition of G, we have $\sum_{(x,y) \in U_X} m_{xy} = 2|W|$. Let W' be the subset of W defined by $W' = W - W \cap U$. Then we have $|W'| = |W| - |W \cap U|$. Since the elements of U which do not belong to W have a weight equal to 0, and since the other arcs have a weight equal to 2, we have then $|W'| = |W| - \frac{1}{2} \sum_{(x,y) \in U} m_{xy} \leq |W| - \frac{1}{2} K = h$.

Moreover, the graph $(X, W - W')$ is equal to $(X, W \cap U)$, which is without circuit by Lemma 5 a and because the graph (X, U) is without circuit.

In conclusion, the answer is kept by the transformation, and hence the problems $Q_0(\mathcal{R}, \mathcal{Z})$ are NP-complete for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$. ❑

Corollary 7. For any even integer $m \geq 2$, the problems $P_m(\mathcal{R}, \mathcal{Z})$ are NP-hard for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$.

Proof. It follows from the fact that, in the proof of Theorem 6, it is possible to upper bound the weights of the graph by 2. ❑

Theorem 8. The problem $Q_1(\mathcal{R}, \mathcal{L})$ is NP-complete.

Proof. We apply the same transformation as for Theorem 6 (and thus we keep the same notations), but with the following weights:

* if $(x, y) \in W$, $m_{xy} = 1$ and $m_{yx} = -1$

* if $(x, y) \notin W$ and $(y, x) \notin W$, $m_{xy} = 1$ and $m_{yx} = 1$

(notice that the weights m_{xy} are well-defined, because H is assumed to be asymmetric),

and with $K = \frac{n(n-1)}{2} - 2h$.

This transformation is obviously polynomial. Let us show that it keeps the answer « yes » or « no ». The proof is quite similar as the one of Theorem 6.

Indeed, consider a minimum-sized subset W' of W such that removing the elements of W' from H gives a graph $(X, W - W')$ without any circuit, and assume that we have $|W'| \leq h$. If we consider $(X, W - W')$ as a partial graph of G, its weight is $|W| - |W'|$, which is greater than or equal to $|W| - h$. Thanks to Lemma 5 b, it is possible to complete $(X, W - W')$ into a linear order (X, L) by adding extra arcs. As we are looking for a linear order, it is necessary to add the arcs (x, y) such that (y, x) belongs to W' (because of the completeness of a linear order); there are $|W'|$ such arcs, and their weights are equal to –1. Because of the asymmetry of a linear order, the other extra arcs

(x, y) cannot belong to W' and are such that (y, x) neither belong to W'; there are $\frac{n(n-1)}{2} - |W|$ such arcs, and their weights are equal to 1. Thus we get:

$$\begin{aligned}\sum_{(x,y)\in L} m_{xy} &= \sum_{(x,y)\in W-W'} m_{xy} + \sum_{\substack{(x,y)\in L\\(y,x)\in W'}} m_{xy} + \sum_{\substack{(x,y)\in L-(W-W')\\(y,x)\notin W'}} m_{xy}\\ &= |W| - |W'| - |W'| + \frac{n(n-1)}{2} - |W|\\ &= \frac{n(n-1)}{2} - 2|W'|.\end{aligned}$$

From $|W'| \le h$, we get $\sum_{(x,y)\in L} m_{xy} \ge \frac{n(n-1)}{2} - 2h = K$: the answer admitted by the instance (G, K) of $Q_1(\mathcal{R}, \mathcal{L})$ is also « yes ».

Conversely, assume that the instance (G, K) of $Q_1(\mathcal{R}, \mathcal{L})$ admits the answer « yes »: there exists a partial graph (X, L) of $G = (X, U_X)$ which represents a linear order, with $\sum_{(x,y)\in L} m_{xy} \ge K$. Notice that ($X$, L) is without circuit. Let W' be the subset of W defined by $W' = W - W \cap L$. If an arc (x, y) belongs to W' (and thus to W), it does not belong to L; then (y, x) belongs to L (completeness of L) but not to W (asymmetry of H), and so its weight is equal to –1. Conversely, let (x, y) be an arc of L with a weight equal to –1; then it does not belong to W but is such that (y, x) does belong to W and not to L: (y, x) belongs to W'. So, the number of arcs of L with a weight equal to –1 is equal to $|W'|$. The other elements of L (there are $\frac{n(n-1)}{2} - |W'|$ such arcs) have a weight equal to 1. Hence the relation: $\sum_{(x,y)\in L} m_{xy} = \frac{n(n-1)}{2} - 2|W'|$. From the inequality $\sum_{(x,y)\in L} m_{xy} \ge K = \frac{n(n-1)}{2} - 2h$, we draw $|W'| \le h$. Moreover, the graph (X, $W - W'$) is equal to (X, $W \cap L$), which is without circuit by Lemma 5 a and because the graph (X, L) is without circuit.

In conclusion, the answer is kept by the transformation, and hence the problem $Q_1(\mathcal{R}, \mathcal{L})$ is NP-complete. ❑

Corollary 9. For any odd integer $m \ge 1$, the problems $P_m(\mathcal{R}, \mathcal{L})$ are NP-hard.

Proof. It follows from the fact that, in the proof of Theorem 8, it is possible to upper bound the weights of the graph by 1. ❑

Theorem 10. $Q_1(\mathcal{R}, \mathcal{A})$ is NP-complete.

Proof. The proof is similar to the one of Theorem 8, and we do not detail it here. The construction is the following, with the same notations as above:

* if $(x, y) \in W$, $m_{xy} = 1$ and $m_{yx} = -1$

* if $(x, y) \notin W$ $m_{xy} = 1$

* $K = \frac{n(n-1)}{2} - h$.

With respect to the proof of Theorem 8, instead of considering the linear order called (X, L) above, we consider the same set of arcs L without the arcs with a weight equal to –1. Details are left to the reader. ❑

Corollary 11. For any odd integer $m \geq 1$, the problems $P_m(\mathcal{R}, \mathcal{A})$ are NP-hard.

Proof. It follows from the fact that, in the proof of Theorem 10, it is possible to upper bound the weights of the graph by 1. ❑

Theorem 12. The problems $Q_1(\mathcal{R}, \mathcal{Z})$ are NP-complete for $\mathcal{Z} \in \{\mathcal{I}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$.

Proof. We apply the same transformation as for Theorem 8 (and thus we keep the same notations), but with the following weights:

* if $(x, y) \in W$, $m_{xy} = 3$

* if $(x, y) \notin W$ $m_{xy} = 1$ and $m_{yx} = 1$

and with $K = \frac{n(n-1)}{2} + 2|W| - 2h$.

This transformation is obviously polynomial. Let us show that it keeps the answer « yes » or « no ». The proof is quite similar as those of Theorems 6 and 8.

Indeed, assume that there exists a subset W' of W with $|W'| \leq h$ and such that removing the elements of W' from H leaves a graph (X, $W - W'$) without any circuit. If we consider it as a partial graph of G, its weight is $3(|W| - |W'|)$. Thanks to Lemma 5 b, it is possible to complete (X, $W - W'$) into a linear order (X, L) by adding $\frac{n(n-1)}{2} - (|W| - |W'|)$ extra arcs. As the weights are all greater than or equal to 1, we get a linear order (X, L), that we may consider as a partial order if $\mathcal{Z} = \mathcal{O}$, or as an interval order if $\mathcal{Z} = \mathcal{I}$, or as a semiorder if $\mathcal{Z} = \mathcal{S}$, or as a weak order if $\mathcal{Z} = \mathcal{W}$, with:

$$\sum_{(x,y)\in L} m_{xy} \geq 3(|W| - |W'|) + \frac{n(n-1)}{2} - (|W| - |W'|) = \frac{n(n-1)}{2} + 2|W| - 2|W'| \geq K,$$

which shows that the answer of the instance (G, K) of $Q_1(\mathcal{R}, \mathcal{Z})$ is « yes ».

Conversely, assume that the instance (G, K) of $Q_1(\mathcal{R}, \mathcal{Z})$ admits the answer « yes »: there exists a partial graph (X, U) of $G = (X, U_X)$ which represents an element of $\mathcal{Z}$, with $\sum_{(x,y)\in U} m_{xy} \geq K$. Notice that ($X$, U) is without circuit. It is then possible, by Lemma 5 b, to complete U into a linear order L by adding extra arcs. As all the weights are positive, we get a linear order (X, L) with $\sum_{(x,y)\in L} m_{xy} \geq K$. Let W' be the subset of W defined by $W' = W - W \cap L$. The graph (X, $W - W'$) is equal to (X, $W \cap L$), which is without circuit by Lemma 5 a. Let us now compute $\sum_{(x,y)\in L} m_{xy}$:

$$\begin{aligned}\sum_{(x,y)\in L} m_{xy} &= \sum_{(x,y)\in L\cap W} m_{xy} + \sum_{(x,y)\in L-W} m_{xy}\\ &= 3|L \cap W| + |L - W|\\ &= 2|L \cap W| + |L|\\ &= 2|W - W'| + \frac{n(n-1)}{2}\\ &= 2|W| - 2|W'| + \frac{n(n-1)}{2}.\end{aligned}$$

Hence, from $\sum_{(x,y)\in L} m_{xy} \geq K = \frac{n(n-1)}{2} + 2|W| - 2h$, we get $|W'| \leq h$. The set W' shows that the instance (H, h) of FAS admits the answer « yes ».

In conclusion, the answer is kept by the transformation, and hence the problems $Q_1(\mathcal{R}, \mathcal{Z})$ is NP-complete for $\mathcal{Z} \in \{\mathcal{I}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$. ❑

Corollary 13. For any odd integer $m \geq 3$ and for $\mathcal{Z} \in \{\mathcal{I}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$, the problems $P_m(\mathcal{R}, \mathcal{Z})$ are NP-hard.

Proof. It follows from the fact that, in the proof of Theorem 12, it is possible to upper bound the weights of the graph by 3. ❑

Remarks.

In fact, the above proof can more generally be applied to any set $\mathcal{Z}$ with $\mathcal{L} \subseteq \mathcal{Z} \subseteq \mathcal{A}$, what is the case for the above sets.

We may notice that the weights of the graph G of Theorem 12 are chosen to be positive, so that an optimal solution is in fact a linear order. The « price » of this trick is that we need some weights to be greater than 1. Because of this, the complexities of the problems $P_1(\mathcal{R}, \mathcal{Z})$ for $\mathcal{Z} \in \{\mathcal{I}, \mathcal{O}, \mathcal{S}, \mathcal{W}\}$ remain open.

We now consider a profile Π of linear orders, i.e. the problems $Q_0(\mathcal{L}, \mathcal{Z})$ and $Q_1(\mathcal{L}, \mathcal{Z})$ for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{C}, \mathcal{I}, \mathcal{L}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}$, or $\mathcal{W}\}$. The study of the complexity is more difficult because the graphs associated with Π are more constrained. Another consequence is that we cannot fix the number m of relations of Π any longer (because of the reconstruction of Π from the graph; see above) though it will be possible to upper bound m by a polynomial of n. To study the complexities of $Q_0(\mathcal{L}, \mathcal{Z})$ and $Q_1(\mathcal{L}, \mathcal{Z})$, we use the NP-completeness of two more constrained versions of FAS, that we call BFAS and BFAS′ because they deal with bipartite graphs.

Problem BFAS

Instance: a directed, asymmetric, and bipartite graph $H = (Y \cup Z, W_1 \cup W_2)$ where $Y = \{y_i\colon 1 \leq i \leq |Y|\}$ and $Z = \{z_i\colon 1 \leq i \leq |Z| = |Y|\}$ give the two classes of H and with $W_1 = \{(z_i, y_i)$ for $1 \leq i \leq |Y|\}$ and $W_2 \subseteq \{(y_i, z_j)$ for $1 \leq i \leq |Y|$ and $1 \leq j \leq |Y|\}$; an integer h;

Question: does there exist $W' \subset W_1 \cup W_2$ with $|W'| \leq h$ and such that removing the elements of W' from H leaves a graph $(Y \cup Z, (W_1 \cup W_2) - W')$ without any circuit ?

Problem BFAS′

Instance: the same as for BFAS;

Question: does there exist $W' \subset W_1$ with $|W'| \leq h$ and such that removing the elements of W' from H leaves a graph $(Y \cup Z, (W_1 \cup W_2) - W')$ without any circuit ?

Figure 10 shows how such a graph looks like. So the only difference between BFAS and BFAS′ is that, in BFAS′ and with respect to the drawing of Figure 10, W' is only made of horizontal arcs.

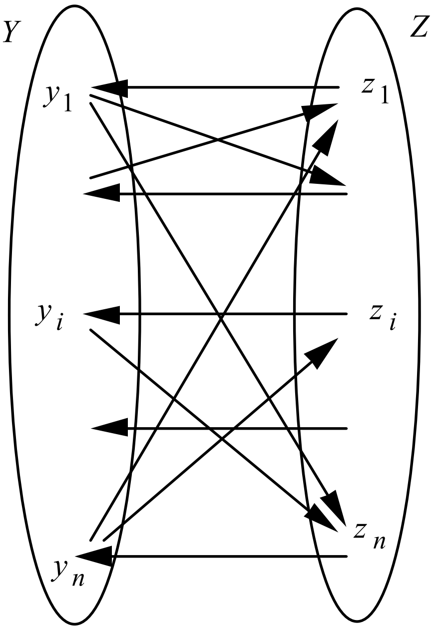

Figure 10: an instance of BFAS.

Theorem 14. BFAS and BFAS′ are NP-complete.

Proof. It is easy to show that BFAS and BFAS′ belong to NP (details are left to the reader). To prove that they are NP-complete, we transform the following problem, called Vertex Cover, into BFAS or BFAS′:

Problem Vertex Cover (VC)

Instance: an undirected graph $G = (X, U)$; an integer g;

Question: does there exist $X' \subset X$ with $|X'| \leq g$ and verifying the following property: $\forall\ \{x, y\} \in U$, $x \in X'$ or $y \in X'$ (X' is then a *vertex cover* of G of cardinality at most g) ?

It is known that VC is NP-complete (see R. Karp [1972]). Let (G, g) be any instance of VC. We define an instance (H, h) of BFAS or of BFAS′ as follows:

* for any vertex $x_i \in X$ ($1 \leq i \leq |X| = n$), we create two vertices of H: y_i and z_i ($1 \leq i \leq n$) and we set $Y = \{y_i, 1 \leq i \leq n\}$ and $Z = \{z_i, 1 \leq i \leq n\}$;

* for any edge $\{x_i, x_j\}$ of H, we create two arcs of H: (y_i, z_j) and (y_j, z_i); W_2 will denote the set of these arcs: $W_2 = \{(y_i, z_j), (y_j, z_i)$, for i and j such that $\{x_i, x_j\}$ belongs to $U\}$;

* we complete H by adding all the arcs of the form (z_i, y_i) for $1 \leq i \leq n$; they constitute the set W_1: $W_1 = \{(z_i, y_i)$ for $1 \leq i \leq n\}$;

* we set $g = h$.

Then we claim that G admits a vertex cover of cardinality at most g if and only if H admits a feedback arc set included into W_1 of cardinality at most h.

Indeed, assume that there exists a vertex cover X' of G with $|X'| \leq g$. Then let W' be defined by: $W' = \{(z_i, y_i)$ for $x_i \in X'\}$. We clearly have $|W'| = |X'| \leq g = h$. Moreover, assume that there exists a circuit in the graph $(Y \cup Z, (W_1 \cup W_2) - W')$. Then this circuit necessarily goes through an arc (z_i, y_i) for some i such that x_i does not belong to X' and then goes through an arc (y_i, z_j) for an appropriate j ($1 \leq j \leq n$); the only way to go on the circuit is to follow the arc (z_j, y_j) (it is the only arc with z_j as its tail),which involves that x_j does not belong to X'. But, as the arc (y_i, z_j) exists in H, $\{x_i, x_j\}$ must be an edge of G, and this edge is not covered by X', a contradiction.

Conversely, assume that (H, h) admits a subset W' of $W_1 \cup W_2$ which is a feedback arc set of cardinality at most h. Then there exists a subset W'' of W_1 which is a feedback arc set of H of cardinality at most h (for BFAS′, W' is necessarily such a set; so the following is useful only for BFAS). Indeed, for any arc (y_i, z_j) of $W' \cap W_2$, remove (y_i, z_j) from W' and replace it in W'' by the arc (z_i, y_i). We get thus a subset of W_1 with at most h elements. To see that W'' is a feedback arc set of H, it is enough to notice that any circuit of H going through (y_i, z_j) goes also through (z_i, y_i), since (z_i, y_i) is the only arc with y_i as its head. So define X' as the set of vertices of G associated with the elements of W'': $X' = \{x_i$ for $(z_i, y_i) \in W''\}$. Then obviously: $|X'| \leq g$. Moreover, assume that X' is not a VC of G. It means that there exists an edge $\{x_i, x_j\}$ of G with $x_i \notin X'$ and $x_j \notin X'$. So, similarly, we have in H: $(z_i, y_i) \notin W''$ and $(z_j, y_j) \notin W''$. But in these conditions, the arcs (z_i, y_i), (y_i, z_j), (z_j, y_j), and (y_j, z_i) (these four arcs do exist in H) define a circuit in the graph $(Y \cup Z, W_1 \cup W_2 - W'')$, and W'' is not a feedback arc set, a contradiction.

So the proposed transformation keeps the answer. As it is trivially polynomial with respect to the size of the transformed instance (G, g), BFAS and BFAS′ are NP-complete. ❑

Now we study the complexity of the problem $Q_0(\mathcal{L}, \mathcal{L})$, and then the one of $Q_1(\mathcal{L}, \mathcal{L})$:

Theorem 15. $Q_0(\mathcal{L}, \mathcal{L})$ is NP-complete.

Proof. As noticed above, $Q_0(\mathcal{L}, \mathcal{L})$ belongs to NP. We transform BFAS into $Q_0(\mathcal{L}, \mathcal{L})$. Let $(H = (Y \cup Z, W_1 \cup W_2), h)$ be any instance of BFAS, with the same notations as above. Let (G, K) be the instance of $Q_0(\mathcal{L}, \mathcal{L})$ defined by:

- the vertex set of G is $X = Y \cup Z$;

- the arc set of G is U_X;

- for any arc (x, y) of G, the weight m_{xy} of (x, y) is equal to: 2 if $(x, y) \in W_1 \cup W_2$, -2 if $(y, x) \in W_1 \cup W_2$, 0 otherwise;

- $K = 2|W_1 \cup W_2| - 4h$.

Notice that G is well defined since H is asymmetric. Moreover, the transformation is clearly polynomial with respect to the size of the instance (H, h) of BFAS.

Now, assume that the instance (H, h) of BFAS admits the answer « yes »: there exists $W' \subset W_1 \cup W_2$ with $|W'| \leq h$ and such that the graph $(X, (W_1 \cup W_2) - W')$ is without any circuit. If we consider $(X, (W_1 \cup W_2) - W')$ as a partial graph of G, its weight is $2|(W_1 \cup W_2) - W'|$, i.e. $2|W_1 \cup W_2| - 2|W'|$. Then, by Lemma 5 b, we may complete $(X, (W_1 \cup W_2) - W')$ into a linear order (X, L) by adding appropriate arcs. Among these extra arcs, there are at most $|W'|$ arcs (x, y) such that (y, x) belongs to $W_1 \cup W_2$, i.e. at most $|W'|$ arcs with a weight equal to -2. More precisely, the weights of the arcs of L belonging to $W_1 \cup W_2$ are equal to 2, and there are at least $|(W_1 \cup W_2) - W'|$ such arcs;

the weights of the arcs (x, y) of L such that (y, x) belongs to $W_1 \cup W_2$ are equal to –2, and there are at most $|W'|$ such arcs; the other arcs of L have a weight equal to 0. Hence:

$$\sum_{(x,y)\in L} m_{xy} \geq 2|W_1 \cup W_2| - 2|W'| - 2|W'| = 2|W_1 \cup W_2| - 4|W'| \geq 2|W_1 \cup W_2| - 4h = K.$$

So (G, K) admits also the answer « yes ».

Conversely, assume that (G, K) admits the answer « yes »: there exists $L \subset U_X$ with $\sum_{(x,y)\in L} m_{xy} \geq K$ and such that (X, L) is a linear order. Thus consider the set $W' = (W_1 \cup W_2) - (W_1 \cup W_2) \cap L$. As (X, L) is a linear order, (X, L) is without any circuit. Thus, by Lemma 5 a, $(W_1 \cup W_2) \cap L = (W_1 \cup W_2) - W'$ is also without any circuit. Let (x, y) be an element of W' (and thus of $W_1 \cup W_2$); then the arc (y, x) belongs to L (because L is complete and (x, y) does not belong to L) and its weight is –2. So, suppose that we have $|W'| > h$. Then there are at most $|(W_1 \cup W_2) - W'|$ arcs which belong to $W_1 \cup W_2$ and to L. So we get:

$$\sum_{(x,y)\in L} m_{xy} \leq 2|(W_1 \cup W_2) - W'| - 2|W'| = 2|W_1 \cup W_2| - 4|W'| < 2|W_1 \cup W_2| - 4h = K,$$

a contradiction. It means that W' satisfies all the conditions and the answer admitted by (X, L) is « yes ».

All these considerations show that $Q_0(\mathcal{L}, \mathcal{L})$ is NP-complete. ❑

Corollary 16. For $f = \Omega(n/\log n)$, with f taking even values, $P_f(\mathcal{L}, \mathcal{L})$ is NP-hard and, for any even integer $m \geq 2$, $P_m(\mathcal{T}, \mathcal{L})$ is NP-hard.

Theorem 17. $Q_1(\mathcal{L}, \mathcal{L})$ is NP-complete.

Proof. It is the same proof as for $Q_0(\mathcal{L}, \mathcal{L})$, but with the weights 1, n, and –n instead of 0, 2, and –2 respectively, and with $K = n|W_1 \cup W_2| - 2n.h$ instead of $2|W_1 \cup W_2| - 4h$. Details are left to the reader. ❑

Corollary 18. For $f = \Omega(n^2/\log n)$, with f taking odd values, $P_f(\mathcal{L}, \mathcal{L})$ is NP-hard and, for $f = \Omega(n)$ with f taking odd values, $P_f(\mathcal{T}, \mathcal{L})$ is NP-hard.

Remark. An interesting case is the problem $P_1(\mathcal{T}, \mathcal{L})$, set by P. Slater (1961) (also known under other names; see for instance I. Charon *et alii* (1997) or O. Hudry *et alii* (2005) for references), i.e. the approximation of a tournament by a linear order. Unfortunately, the proof of Corollary 18 does not allow to know the complexity of Slater's problem, which remains an open problem.

Proofs similar to the previous ones (and not given here) lead to the following results:

Theorem 19. $Q_0(\mathcal{L}, \mathcal{A})$, $Q_1(\mathcal{L}, \mathcal{A})$, $Q_0(\mathcal{L}, \mathcal{O})$, and $Q_1(\mathcal{L}, \mathcal{O})$ are NP-complete. For $f = \Omega(n/\log n)$, with f taking even values, $P_f(\mathcal{L}, \mathcal{A})$ and $P_f(\mathcal{L}, \mathcal{O})$ are NP-hard. For any even integer $m \geq 2$, $P_m(\mathcal{T}, \mathcal{A})$ and $P_m(\mathcal{T}, \mathcal{O})$ are NP-hard. For $f = \Omega(n^2/\log n)$, with f taking odd values, $P_f(\mathcal{L}, \mathcal{A})$ and $P_f(\mathcal{L}, \mathcal{O})$ are NP-hard. For $f = \Omega(n)$ with f taking odd values, $P_f(\mathcal{T}, \mathcal{A})$ and $P_f(\mathcal{T}, \mathcal{A})$ are NP-hard.

Now we study the complexity of the problems $Q_0(\mathcal{L}, \mathcal{Z})$ for $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$.

Theorem 20. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$, $Q_0(\mathcal{L}, \mathcal{Z})$ is NP-complete.

Proof. Let $\mathcal{Z}$ belong to $\{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$. As for the other problems above, $Q_0(\mathcal{L}, \mathcal{Z})$ obviously belongs to NP. We transform BFAS′ into $Q_0(\mathcal{L}, \mathcal{Z})$. Let $(H = (Y \cup Z, W_1 \cup W_2), h)$ be any instance of BFAS′, with the same notations as above. Let (G, K) be the instance of $Q_0(\mathcal{L}, \mathcal{Z})$ defined by:

- the vertex set of G is $X = Y \cup Z$;

- the arc set of G is U_X;

- for any arc (x, y) of G, the weight m_{xy} of (x, y) is equal to: 2 if $(x, y) \in W_1$, -2 if $(y, x) \in W_1$, $4n - 2$ if $(x, y) \in W_2$, $-(4n - 2)$ if $(y, x) \in W_2$, 0 otherwise;

- $K = (4n-2)|W_2| + 2n - 4h$.

Notice that G is well defined since H is asymmetric. Moreover, the transformation is clearly polynomial with respect to the size of the instance (H, h) of BFAS′.

Now, assume that the instance (H, h) of BFAS′ admits the answer « yes »: there exists $W' \subset W_1$ with $|W'| \leq h$ and such that the graph $(X, (W_1 \cup W_2) - W')$ is without any circuit. We prove that the instance (G, K) admits also the answer « yes » as in

Theorem 15. If we consider (X, $(W_1 \cup W_2) - W'$) as a partial graph of G, its weight is $2|W_1 - W'| + (4n+2)|W_2|$. By Lemma 5 b, we may complete (X, $(W_1 \cup W_2) - W'$) into a linear order (X, L) (that we shall consider as an interval order if Z is I, or as a semiorder if Z is S, or as a complete preorder if Z is C) by adding appropriate arcs. Among these extra arcs, there are at most $|W'|$ arcs (x, y) such that (y, x) belongs to W_1, i.e. at most $|W'|$ arcs with a weight equal to –2, while the other extra arcs all belong to $U_X - (W_1 \cup W_2)$ and have a weight equal to 0. More precisely, the weights of the arcs of L belonging to W_1 are equal to 2; there are at least $|W_1 - W'|$ such arcs. The weights of the $|W_2|$ arcs of L belonging to W_2 are equal to $4n - 2$. The weights of the arcs (x, y) of L such that (y, x) belongs to W_1 are equal to –2; there are at most $|W'|$ such arcs. The other arcs of L have a weight equal to 0. Hence, since $|W_1| = n$ and $W' \subset W_1$:

$$\sum_{(x,y)\in L} m_{xy} \geq 2|W_1 - W'| + (4n-2)|W_2| - 2|W'| = 2n + (4n-2)|W_2| - 4|W'|$$

and so, since $|W'| \leq h$: $\sum_{(x,y)\in L} m_{xy} \geq 2n + (4n-2)|W_2| - 4h = K.$

So (G, K) admits also the answer « yes ».

Conversely, assume that (G, K) admits the answer « yes ». We consider two main subcases: $Z \in \{I, S\}$ or $Z = C$.

- 1rst subcase: $Z \in \{I, S\}$

Since a semiorder is an interval order, there exists $I \subset U_X$ with $\sum_{(x,y)\in I} m_{xy} \geq K$ and such that (X, I) is an interval order. We want to show that then the instance (H, h) of BFAS′ also admits the answer « yes ». Notice that if h is greater than n, the answer of (H, h) is trivially « yes »; so, assume that we have $h \leq n - 1$. Let us show that I contains all the arcs of W_2. Assume the contrary. The arcs of I with a non-negative weight would be at most the n elements of W_1 (with a weight equal to 2) and at most $|W_2| - 1$ arcs of W_2 (with a weight equal to $4n - 2$). So we would get: $\sum_{(x,y)\in I} m_{xy} \leq 2n + (4n-2)(|W_2| - 1)$. On the other hand, we are supposed to have: $\sum_{(x,y)\in I} m_{xy} \geq K = (4n-2)|W_2| + 2n - 4h$, from which we draw $4h \geq 4n - 2$, which is incompatible with $h \leq n - 1$. Hence: $W_2 \subset I$ and, because of the antisymmetry of I, there is no arc in I of the form (z_i, y_j) with $i \neq j$. We prove now that we may construct a linear order L with $\sum_{(x,y)\in L} m_{xy} \geq K$ from I. For this, set $J = I - \{(x, y) \in I \text{ with } m_{xy} = 0\}$, and gather the vertices of X into the following three sets:

- $X_1 = \{y_k \in Y,\ z_k \in Z$ such that $(z_k, y_k) \in J\}$
- $X_2 = \{y_k \in Y,\ z_k \in Z$ such that $(y_k, z_k) \notin J$ and $(z_k, y_k) \notin J\}$
- $X_3 = \{y_k \in Y,\ z_k \in Z$ such that $(y_k, z_k) \in J\}$.

The situation is illustrated by Figure 11. We are going to show that the dashed arcs of Figure 11 do not exist in fact. Notice that, as a subset of I which contains no circuit, J contains no circuit.

The dashed arcs with their two extremities inside X_1 cannot exist, otherwise there would exist a circuit in J. Now consider an arc (y_j, z_i) with $y_j \in Y$, $z_i \in Z$ (thus $i \neq j$) and with an extremity inside X_1 and the other inside X_2. As one extremity belongs to X_1, the arc (z_i, y_i) or the arc (z_j, y_j) exists in I, and thus in J since its weight is not equal to 0. Also, by construction of G, (y_i, z_j) is an arc of G, and thus of I ($W_2 \subseteq I$). Assume that (z_i, y_i) belongs to I (and thus to J); then y_i and z_i belong to X_1, while y_j and z_j belong to X_2. In this case, as I is transitive, the arcs (y_j, z_i), (z_i, y_i), (y_i, z_j) involve the existence of the arc (y_j, z_j), a contradiction with the belonging of y_j and z_j to X_2. Similarly, the dashed arcs with their two extremities inside X_2 cannot exist in I. Indeed, assume that such a pair of arcs (y_j, z_i) and (y_i, z_j) exist with $y_i \in Y \cap X_2$, $y_j \in Y \cap X_2$, $z_i \in Z \cap X_2$, and $z_j \in Z \cap X_2$ $(i \neq j)$ exist (notice that if one of these two arcs exists, the other one must exist too). As y_i, y_j, z_i, and z_j belong to X_2, the arcs (y_i, z_i) and (y_j, z_j) do not belong to I. But then the arcs (y_j, z_i) and (y_i, z_j) do not respect the definition of an interval order. So the look of J is as the one shown by Figure 11 without the dashed arcs.

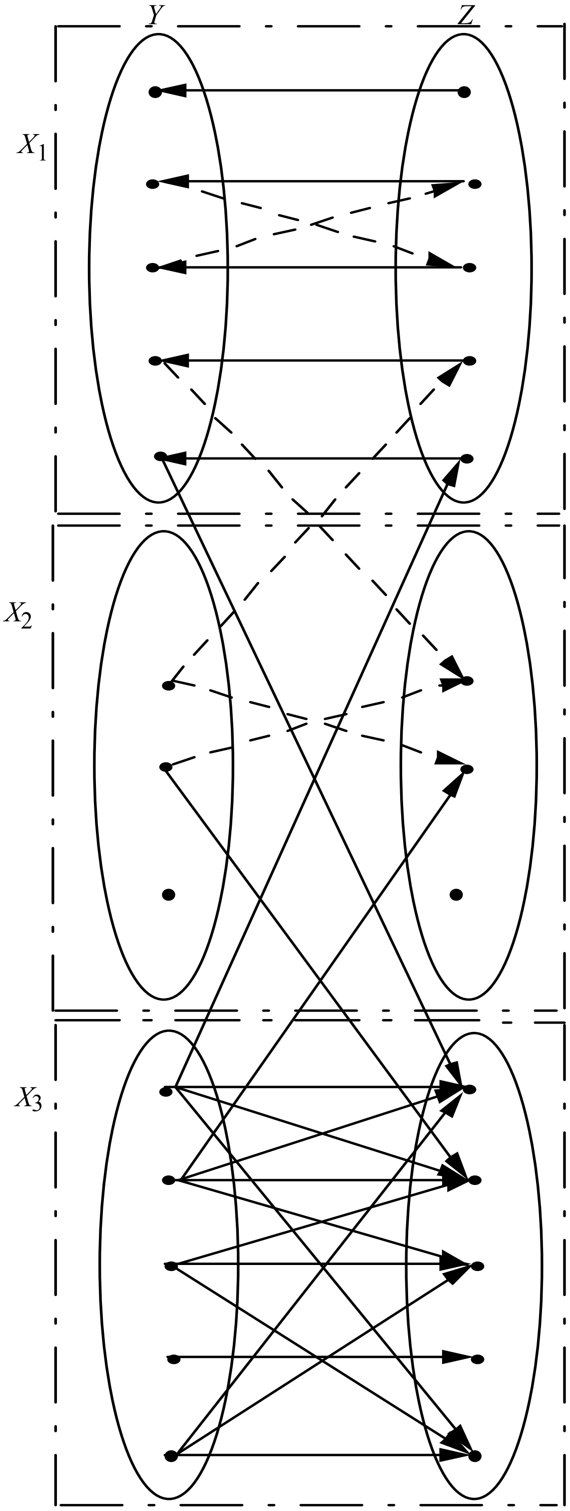

Figure 11. The graph induced by J for $Q_0(\mathcal{L}, \mathcal{Z})$ when $\mathcal{Z}$ is equal to $\mathcal{I}$ or $\mathcal{S}$.

Now, set $J' = J \cup \{(z, y) \in W_1$ for $y \in X_2$ and $z \in X_2\}$ (with respect to Figure 11, we add all the horizontal arcs from right to left with their two extremities in X_2). As the vertices of X_2 are linked only with vertices of X_3, it is easy to see that J' is still without any circuit. As the weights of these arcs are positive, we get: $\sum_{(x,y) \in J'} m_{xy} \geq \sum_{(x,y) \in J} m_{xy} = \sum_{(x,y) \in I} m_{xy} \geq K$. As J' is without circuit, by Lemma 5 b, we may extend J' into a linear order L. As we had $W_2 \subseteq I$ and as, for any index k with $1 \leq k \leq n$, y_i and z_i are already linked by an arc belonging to J', all the arcs that we add in order to define L from J' have a weight equal to 0. Hence $\sum_{(x,y) \in L} m_{xy} = \sum_{(x,y) \in J'} m_{xy} \geq K$. The end of the proof is exactly the same as in Theorem 15, and we do not duplicate it here: from L we define a subset W' which shows that the instance (H, h) of BFAS′ admits the answer « yes », which completes the proof for the subcase $\mathcal{Z} \in \{\mathcal{I}, \mathcal{S}\}$.

- 2nd subcase: $\mathcal{Z} = \mathcal{C}$

As for the previous case, we are going to prove that, if the answer admitted by the instance (G, K) is « yes », then we can build a linear order which gives this answer « yes ». Then the conclusion will be the same as above.

So, assume that there exists a subset C of U_X such that (X, C) is a complete preorder with $\sum_{(x,y) \in C} m_{xy} \geq K$. As above, this inequality involves that C contains all the elements of W_2 and no arc (x, y) such that (y, x) would belong to W_2 (details are left to the reader). Let D be the set made of the arcs of C with a non-zero weight: $D = C - \{(x, y) \in C$ with $m_{xy} = 0\}$. Moreover, gather the vertices of X into the following three sets:

* $X_1 = \{y_k \in Y, z_k \in Z$ such that $(z_k, y_k) \in D$ and $(y_k, z_k) \notin D\}$

* $X_2 = \{y_k \in Y, z_k \in Z$ such that $(y_k, z_k) \in D$ and $(z_k, y_k) \notin D\}$

* $X_3 = \{y_k \in Y, z_k \in Z$ such that $(y_k, z_k) \in D$ and $(z_k, y_k) \in D\}$.

The look of the graph induced by D is given by Figure 12. We are going to show that the dashed arcs of Figure 12 do not exist in fact.

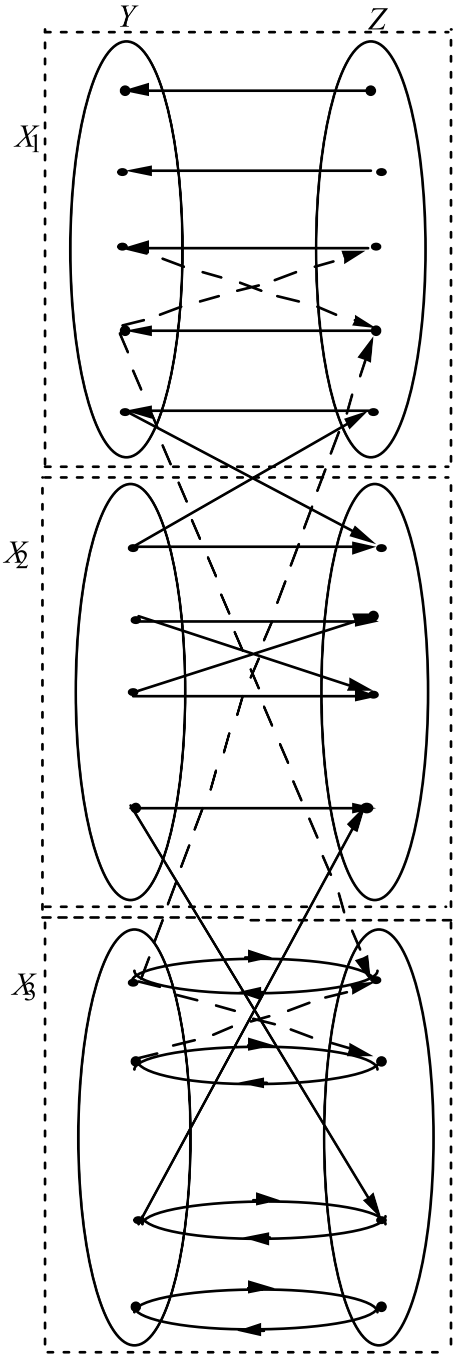

Figure 12. The graph induced by D for $Q_0(\mathcal{L}, \mathcal{C})$.

Indeed, let (y_j, z_i) be such an arc with $z_i \in Z$ and $y_j \in Y$. Then its weight is not equal to 0, and it is the same for (y_i, z_j), which thus belongs to D. If we assume that the four vertices y_i, y_j, z_i, and z_j belong to $Z_1 \cup Z_3$, then the arcs (z_i, y_i) and (z_j, y_j) belong to D and, by transitivity, the arcs (z_j, y_i) and (z_i, y_j) also belong to D, what is impossible (see above). So, the look of the graph induced by D is the one depicted by Figure 12 without the dashed arcs.

The next step consists in showing that we may extract a set D' of arcs from D such that D' is without circuit while its weight $\sum_{(x,y)\in D'} m_{xy}$ is still greater than or equal to K. For this, let D' be defined by $D' = D - \{(y_i, z_i) \text{ for } y_i \in X_3, z_i \in X_3\}$ (in other words, with respect to Figure 12, we remove the – almost – horizontal arcs inside X_3 and oriented from left to right). As the removed arcs have a negative weight, we get: $\sum_{(x,y)\in D'} m_{xy} \geq \sum_{(x,y)\in D} m_{xy} \geq K$. Moreover, D' is without circuit. Indeed, consider any circuit in D, which is transitive. Such a circuit must contain an arc of the form (z_i, y_i) with $y_i \in Y$ and $z_i \in Z$ (since it is the only way to go from Z to Y in the graph induced by D). Because of the transitivity of D applied to the considered circuit, (y_i, z_i) must also be an arc of D, and so y_i and z_i must belong to X_3. So the removal from D of the arcs (y_i, z_i) with $y_i \in X_3$, $z_i \in X_3$ leaves a graph (induced by D') without any circuit.

We may now conclude. As D' is without any circuit and by Lemma 5 b, we may complete it into a linear order L by adding appropriate arcs. As D' already contains W_2 and, for $1 \leq i \leq n$, exactly one of the two arcs (y_i, z_i) or (z_i, y_i), the extra arcs have a weight equal to 0. So, we get: $\sum_{(x,y)\in L} m_{xy} \geq \sum_{(x,y)\in D'} m_{xy} \geq K$. Then it is sufficient to apply the same argument as in Theorem 15 to show the existence of a subset W' of $W_1 \cup W_2$ which gives the answer « yes » to the instance (H, h) of BFAS′, which completes the proof for the subcase $\mathcal{Z} = \mathcal{C}$. ❑

Corollary 21. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$ and for $f = \Omega(n^2/\log n)$, with f taking even values, $P_f(\mathcal{L}, \mathcal{Z})$ is NP-hard. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$ and for $f = \Omega(n)$ with f taking even values, $P_f(\mathcal{T}, \mathcal{Z})$ is NP-hard.

Remark. If we transpose the proof of Theorem 20 in terms of preferences, we build a profile of linear orders such that there exists an optimal interval order, or an optimal semiorder, or a complete order which is in fact a linear order. An interesting question would be to know whether it is always the case, for any profile of linear orders.

Proofs similar to the previous ones (and not given here) lead to the following results:

Theorem 22. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$, the problems $Q_1(\mathcal{L}, \mathcal{Z})$ are NP-complete. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$ and for $f = \Omega(n^3/\log n)$, with f taking odd values, $P_f(\mathcal{L}, \mathcal{Z})$ is NP-hard. For $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{S}\}$ and for $f = \Omega(n^2)$, with f taking odd values, $P_f(\mathcal{T}, \mathcal{Z})$ is NP-hard.

To study the complexity of the problems $P_f(\mathcal{Y}, \mathcal{Q})$ and $P_f(\mathcal{Y}, \mathcal{W})$, we first prove a lemma. In order to state it, we recall a previous notation. For any set $\mathcal{Z}$ of binary relations defined by some properties, we define $\mathcal{Z}^a$ as the set of preferences which are the asymmetric part of a preference belonging to $\mathcal{Z}$. In particular, we have $\mathcal{C}^a = \mathcal{W}$, $\mathcal{Q}^a = \mathcal{S}$, and $\mathcal{P}^a = \mathcal{O}$.

Lemma 23. For $\mathcal{Y} \in \{\mathcal{L}, \mathcal{T}\}$ and for any set $\mathcal{Z}$ and any function f, $P_f(\mathcal{Y}, \mathcal{Z})$ and $P_f(\mathcal{Y}, \mathcal{Z}^a)$ have the same complexity.

Proof. The result comes from the fact that we have $\Delta(\Pi, Z) = \Delta(\Pi, Z^a)$, for any profile Π of linear orders or of tournaments and any element Z of $\mathcal{Z}$. ❑

Corollary 24.

- For $f = \Omega(n/\log n)$, with f taking even values, $P_f(\mathcal{L}, \mathcal{P})$ is NP-hard. For $f = \Omega(n^2/\log n)$, with f taking even values, $P_f(\mathcal{L}, \mathcal{Q})$ and $P_f(\mathcal{L}, \mathcal{W})$ are NP-hard.
- For $f = \Omega(n^2/\log n)$, with f taking odd values, $P_f(\mathcal{L}, \mathcal{P})$ is NP-hard. For $f = \Omega(n^3/\log n)$, with f taking odd values, $P_f(\mathcal{L}, \mathcal{Q})$ and $P_f(\mathcal{L}, \mathcal{W})$ are NP-hard.
- For $m \geq 2$ with m even, $P_f(\mathcal{T}, \mathcal{P})$ is NP-hard. For $f = \Omega(n)$ with f taking even values, $P_f(\mathcal{T}, \mathcal{Q})$ and $P_f(\mathcal{T}, \mathcal{W})$ are NP-hard.
- For $f = \Omega(n)$ with f taking odd values, $P_f(\mathcal{T}, \mathcal{P})$ is NP-hard. For $f = \Omega(n^2)$ with f taking odd values, $P_f(\mathcal{T}, \mathcal{Q})$ and $P_f(\mathcal{T}, \mathcal{W})$ are NP-hard.
- For any even $m \geq 2$, the problems $P_m(\mathcal{R}, \mathcal{P})$ and $P_m(\mathcal{R}, \mathcal{Q})$ are NP-hard.
- For $f = \Omega(n)$ with f taking odd values, $P_f(\mathcal{R}, \mathcal{P})$ is NP-hard. For any odd $m \geq 3$, $P_m(\mathcal{R}, \mathcal{Q})$ is NP-hard.

Proof. For $\mathcal{Y} = \mathcal{L}$ or $\mathcal{Y} = \mathcal{T}$, these results come as a consequence of Theorem 19, Corollary 21 and Theorem 22 and from the application of Lemma 23 to $\mathcal{Z} = \mathcal{O}$, $\mathcal{Z} = \mathcal{Q}$ and to $\mathcal{Z} = \mathcal{C}$. For $\mathcal{Y} = \mathcal{R}$, this comes from Lemma 23, Corollary 7 and Corollary 13 for $\mathcal{Z} = \mathcal{Q}$ or $\mathcal{Z} = \mathcal{C}$, or from the complexity of $P_f(\mathcal{T}, \mathcal{P})$ for $\mathcal{Z} = \mathcal{Q}$ (by considering $\mathcal{T}$ as included into $\mathcal{R}$). ❑

The last result of this section deals with any set $\mathcal{Y}$ containing $\mathcal{L}$.

Theorem 25. For $f = \Omega(n/\log n)$, with f taking even values, for any set $\mathcal{Y}$ with $\mathcal{L} \subseteq \mathcal{Y}$, for $\mathcal{Z} \in \{\mathcal{A}, \mathcal{L}, \mathcal{O}, \mathcal{P}\}$, $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard. For $f = \Omega(n^2/\log n)$, with f taking odd values, for any set $\mathcal{Y}$ with $\mathcal{L} \subseteq \mathcal{Y}$, for any set $\mathcal{Z} \in \{\mathcal{A}, \mathcal{L}, \mathcal{O}, \mathcal{P}\}$, $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard. For $f = \Omega(n^2/\log n)$, with f taking even values, for any set $\mathcal{Y}$ with $\mathcal{L} \subseteq \mathcal{Y}$, for any set $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{Q}, \mathcal{S}, \mathcal{W}\}$, $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard. For $f = \Omega(n^3/\log n)$, with f taking odd values, for any set $\mathcal{Y}$ with $\mathcal{L} \subseteq \mathcal{Y}$, for any set $\mathcal{Z} \in \{\mathcal{C}, \mathcal{I}, \mathcal{Q}, \mathcal{S}, \mathcal{W}\}$, $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard.

Proof. The previous results give the statement of Theorem 25 for $\mathcal{Y} = \mathcal{L}$. For $\mathcal{L} \subset \mathcal{Y}$, it is sufficient to consider any instance of the NP-hard problem $P_f(\mathcal{L}, \mathcal{Z})$ as an instance of $P_f(\mathcal{Y}, \mathcal{Z})$. This transformation (the identity !) is obviously polynomial and keeps the answer. Hence the result. ❑

In particular, we may apply Theorem 25 when $\mathcal{Y}$ is any one of the sets $\mathcal{A}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{L}$, $\mathcal{O}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, or $\mathcal{W}$, but also to « mixed » profiles belonging to any union of two or more sets $\mathcal{A}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{L}$, $\mathcal{O}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, or $\mathcal{W}$, for instance to profiles which may contain tournaments, preorders, and interval orders simultaneously...

5 Conclusion

The previous section was devoted to NP-hard problems. There are also some problems $P_f(\mathcal{Y}, \mathcal{Z})$ which are polynomial. It is trivially the case for $P_f(\mathcal{Y}, \mathcal{R})$ and for $P_f(\mathcal{Y}, \mathcal{T})$, for any set $\mathcal{Y}$ and any function f. Indeed, if we consider the associated problems $Q_0(\mathcal{Y}, \mathcal{R})$, $Q_1(\mathcal{Y}, \mathcal{R})$, $Q_0(\mathcal{Y}, \mathcal{T})$, or $Q_1(\mathcal{Y}, \mathcal{T})$, it is easy to see that an optimal solution consists in keeping all the arcs of G with a positive weight for $\mathcal{Z} = \mathcal{R}$, or in keeping, for each pair of arcs (x, x') and (x', x), the arc with the greatest weight for $\mathcal{Z} = \mathcal{T}$. Another interesting polynomial case is the one of unimodular orders; in this case, the aggregation of unimodular orders into a unimodular order is polynomial (see D. Black (1948)).

Anyway, it seems that properties like the lack of circuits or transitivity usually lead to NP-hard problems. The previous complexity results illustrate this trend. We may summarize them by the tables of Figure 13 (m even) and Figure 14 (m odd). In these tables, « NPH » means that the considered problem $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard. In such a case, we indicate the range of a lower bound of the number m of relations inside the profile which ensures that $P_f(\mathcal{Y}, \mathcal{Z})$ is NP-hard; for instance, $m = \Omega(n)$ with m odd for $P_f(\mathcal{T}, \mathcal{L})$ means that $P_f(\mathcal{T}, \mathcal{L})$ is NP-hard if the range of the odd number m of tournaments of the profile is at least n. As a general result, remember that the NP-hardness of $P_f(\mathcal{Y}, \mathcal{Z})$ involves the one of $P_{f+2}(\mathcal{Y}, \mathcal{Z})$. To my knowledge, when not trivial, the complexity for lower values of m is not known. The letter « P » means that $P_f(\mathcal{Y}, \mathcal{Z})$ is (trivially) polynomial. Remember also that all the results displayed in the tables of Figures 13 and 14 remain the same if we add the reflexivity or the irreflexivity to the considered types of relations.

From this table, it appears that some cases are still unsolved, when m is low. One such interesting case is the problem stated by P. Slater (1961), i.e. $P_1(\mathcal{T}, \mathcal{L})$ for which Π is reduced to one tournament while the median relation must be a linear order. In spite of repeated efforts, its complexity remains open...

Median relation ($\mathcal{Z}$)	$\Pi \in \mathcal{R}^m$ ($\mathcal{Y} = \mathcal{R}$)	$\Pi \in \mathcal{T}^m$ ($\mathcal{Y} = \mathcal{T}$)	$\Pi \in \mathcal{Y}^m$ with $\mathcal{L} \subseteq \mathcal{Y}$
binary relation ($\mathcal{R}$)	P	P	P
tournament ($\mathcal{T}$)	P	P	P
acyclic relation ($\mathcal{A}$)	NPH, $m \geq 2$	NPH, $m \geq 2$	NPH, $m = \Omega(n / \log n)$
complete preorder ($\mathcal{C}$)	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
interval order ($\mathcal{I}$)	NPH, $m \geq 2$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
linear order ($\mathcal{L}$)	NPH, $m \geq 2$	NPH, $m \geq 2$	NPH, $m = \Omega(n / \log n)$
partial order ($\mathcal{O}$)	NPH, $m \geq 2$	NPH, $m \geq 2$	NPH, $m = \Omega(n / \log n)$
preorder ($\mathcal{P}$)	NPH, $m \geq 2$	NPH, $m \geq 2$	NPH, $m = \Omega(n / \log n)$
quasi-order ($\mathcal{Q}$)	NPH, $m \geq 2$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
semiorders ($\mathcal{S}$)	NPH, $m \geq 2$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
weak order ($\mathcal{W}$)	NPH, $m \geq 2$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$

Figure 13. The complexity results for m even.

Median relation ($\mathcal{Z}$)	$\Pi \in \mathcal{R}^m$ ($\mathcal{Y} = \mathcal{R}$)	$\Pi \in \mathcal{T}^m$ ($\mathcal{Y} = \mathcal{T}$)	$\Pi \in \mathcal{Y}^m$ with $\mathcal{L} \subseteq \mathcal{Y}$
binary relation ($\mathcal{R}$)	P	P	P
tournament ($\mathcal{T}$)	P	P	P
acyclic relation ($\mathcal{A}$)	NPH, $m \geq 1$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
complete preorder ($\mathcal{C}$)	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^3 / \log n)$
interval order ($\mathcal{I}$)	NPH, $m \geq 3$	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^3 / \log n)$
linear order ($\mathcal{L}$)	NPH, $m \geq 1$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
partial order ($\mathcal{O}$)	NPH, $m \geq 3$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
preorder ($\mathcal{P}$)	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n)$	NPH, $m = \Omega(n^2 / \log n)$
quasi-order ($\mathcal{Q}$)	NPH, $m \geq 3$	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^3 / \log n)$
semiorders ($\mathcal{S}$)	NPH, $m \geq 3$	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^3 / \log n)$
weak order ($\mathcal{W}$)	NPH, $m \geq 3$	NPH, $m = \Omega(n^2)$	NPH, $m = \Omega(n^3 / \log n)$

Figure 14. The complexity results for *m* odd.

References

D. Arditti (1984), Un nouvel algorithme de recherche d’un ordre induit par des comparaisons par paires”, *in Data analysis and informatics III*, E. Diday *et al*. (eds.), North Holland, Amsterdam, 323-343.

J.-P. Barthélemy (1979), Caractérisations axiomatiques de la distance de la différence symétrique entre des relations binaires, *Mathématiques et Sciences humaines* 67, 85-113.

J.-P. Barthélemy, G. Cohen, A. Lobstein, *Algorithmic Complexity and Communication Problems*, *UCL Press*, London, 1996.

J.-P. Barthélemy, C. Flament, B. Monjardet (1989) : Ordered sets and social sciences, *in Ordered Sets*, I. Rival (ed.), D. Reidel, 721-758.

J.-P. Barthélemy, O. Hudry, G. Isaak, F. S. Roberts, B. Tesman (1995), The reversing number of a digraph, *Discrete Applied Mathematics* 60, 39-76.

J.-P. Barthélemy, B. Leclerc (1995), The median procedure for partitions, *in Partitioning data sets*, I.J. Cox, P. Hansen and B. Julesz (eds), *DIMACS Series in*

Discrete Mathematics and Theoretical Computer Science 19, Amer. Math. Soc., Providence, RI, 3-34.

J.-P. Barthélemy, B. Leclerc, B. Monjardet (1986): On the use of ordered sets in problems of comparison and consensus of classifications, *Journal of Classification* 3, 187-224.

J.-P. Barthélemy, B. Monjardet (1981), The median procedure in cluster analysis and social choice theory, *Mathematical Social Sciences* 1, 235 - 267.

J.-P. Barthélemy, B. Monjardet (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.

J.J. Bartholdi III, C.A. Tovey, M.A. Trick (1989), Voting schemes for which it can be difficult to tell who won the election, *Social Choice and Welfare* 6, 157-165.

D. Black (1948), On the rationale of group decision-making, *Journal of Political Economy* 56, 23-34.

K.P. Bogart (1973), Preference structures I: distance between transitive preference relations, *Journal of Mathematical Sociology* 3, 49-67.

K.P. Bogart (1975), Preference structures II: distances between asymmetric relations, *SIAM Journal of Applied Mathematics* 29 (2), 254-262.

J.S. de Cani (1969), Maximum likelihood paired comparison ranking by linear programming, *Biometrika* 3, 537-545.

I. Charon, A. Guénoche, O. Hudry, F. Woirgard (1997), New results on the computation of median orders, *Discrete Mathematics* 165-166, 139-154.

B. Debord (1987), Caractérisation des matrices de préférences nettes et méthodes d'agrégation associées, *Mathématiques et Sciences humaines* 97, 5-17.

P. Erdös, L. Moser (1964), On the representation of directed graphs as unions of orderings, Magyar Tud. Akad. Mat. Kutato Int. Közl. 9, 125-132.

P.C. Fishburn (1985), *Interval orders and interval graphs, a study of partially ordered sets*, John Wiley, New York.

M.R. Garey, D.S. Johnson (1979), *Computers and intractability, a guide to the theory of NP-completeness*, Freeman, New York.

A. Guénoche, B. Vandeputte-Riboud, J.-B. Denis (1994), Selecting varieties using a series of trials and a combinatorial ordering method, *Agronomie* 14, 363-375.

L. Hubert (1976), Seriation using asymmetric proximity measures, *Br. J. Math. Statist. Psychol*. 29, 32-52.

O. Hudry (1989), *Recherche d'ordres médians : complexité, algorithmique et problèmes combinatoires*, thèse de l'ENST, Paris.

O. Hudry, B. Leclerc, B. Monjardet, J.-P. Barthélemy (2005), Médianes métriques et laticielles, in *Concepts et méthodes pour l'aide à la décision*, D. Dubois, M. Pirlot, D. Bouyssou and H. Prade (éds.), Hermès, Paris.

M. Jünger (1985), *Polyhedral combinatorics and the acyclic subdigraph problem*, Heldermann Verlag, Berlin.

R.M. Karp (1972), Reducibility among combinatorial problems, *in* R.E. Miller and J.W. Tatcher (eds.), *Complexity of computer computations*, New York, Plenum Press, 85-103.

M. Krivanek, J. Moravek (1986), NP-hard problems in hierarchical-tree clustering, *Acta Informatica* 23, 311-323.

D. McGarvey (1953), A theorem on the construction of voting paradoxes, *Econometrica* 21, 608-610.

B. Monjardet (1979), Relations à éloignement minimum de relations binaires, note bibliographique, *Mathématiques et Sciences humaines* 67, 115-122.

B. Monjardet (1990), Sur diverses formes de la « règle de Condorcet » d'agrégation des préférences, *Math. Inf. Sci. hum* 111, 61-71.

J.B. Orlin (1988), unpublished manuscript.

G. Reinelt (1985), *The linear ordering problem: algorithms and applications*, Research and Exposition in Mathematics 8, Heldermann Verlag, Berlin, 1985.

P. Slater (1961), Inconsistencies in a schedule of paired comparisons, *Biometrika* 48, 303-312.

R. Stearns (1959), The voting problem, *Amer. Math. Monthly* 66, 761-763.

A.W. Tucker (1960), On directed graphs and integer programs, *1960 Symposium on Combinatorial Problems*, Princeton University, USA.

Y. Wakabayashi (1986), *Aggregation of binary relations: algorithmic and polyhedral investigations*, PhD thesis, Augsbourg.

Y. Wakabayashi (1998), The Complexity of Computing Medians of Relations, *Resenhas* 3 (3), 323-349.

D.H. Younger (1963), Minimum feedback arc sets for a directed graph, *IEEE Trans. of the profes. tech. group in circuit theory* 10 (2), 238-245.