Extreme supermodular set functions over five variables*

Milan Studený — Acad. Sci. Czech Rep.
Remco R. Bouckaert — Crystal Mountain Inf. Tech.
Tomáš Kočka — Univ. of Economics Prague

January 26, 2000

Abstract

The class of supermodular functions on the power set of a non-empty finite set N forms a cone. It can be viewed as the direct sum of a linear subspace and of a cone of standardized supermodular functions which has finitely many extreme rays. Every extreme ray can be described by a standardized integer-valued set function. We analyse the situation in the case when N has five elements (= variables). A computer program was used to obtain a catalogue of all classes of permutably equivalent extreme standardized supermodular functions on the power set of N. We consider several alternative ways of representation of these equivalence classes and use various characteristics to describe them. Moreover, two relevant hypotheses valid in case of four variables are disproved in case of five variables.

1 Introduction

The problem treated in this research report has two basic sources of motivation.

First, in [11] an attempt to develop a specific non-graphical approach to description of probabilistic conditional independence structures was made. In contrast to classic graphical methods this approach (for a brief overview see [10]) makes it possible to describe all probabilistic conditional independence structures induced by discrete distributions and by non-degenerate Gaussian distributions as well. Conditional independence structures can be described by special integer-valued functions on the power set of the set of variables called *structural imsets*. Moreover, the approach offers simple algebraic operations with integers as a tool for inference among conditional independence statements. However, the question of computer implementation of the inference mechanism leads to the problem of finding all extreme rays of a certain cone in multidimensional real vector space, namely of the cone of (standardized) *supermodular functions* on the power set of the set of variables.

Second, supermodular set functions play an important role in game theory where they are named either 'convex games' [5] or 'convex set functions' [4]. Distinguished position have then 'extreme convex set functions' which correspond to extreme rays of the above mentioned cone. In [4] the problem of characterization of extreme supermodular functions is raised and a certain necessary and sufficient condition for a supermodular function to be extreme is given. However, this condition is formulated in terms of specific representation of supermodular functions and it does not seem to be suitable for practical purposes of computer verification. Let us note that the question of characterization of all extreme supermodular functions may have relevance to theory of polymatroids, that is certain submodular functions [2].

*This research is supported by the grant MŠMT n. VS96008 and hopefully will be by GAČR n. 201/00/0471.

The task to find all extreme rays of the cone of standardized supermodular functions (on the power set of the set of variables) is easy in case of 3 variables. In case of 4 variables it is also manageable without help of a computer [7]; in fact, these function were already determined by S. A. Cook in 1965 (private communication to L. S. Shapley [5]). The primary aim of this paper is to make basic formal analysis in case of 5 variables. That is, to find all extreme standardized supermodular function on the power set of a set $N = \{a, b, c, d, e\}$. The secondary aim of this work is to make reasonably reduced catalogue of suitable representatives of these extreme supermodular functions. This catalogue can serve as a basis for later deeper analysis which may hopefully one time in future lead to proper intuitive understanding of the form and structure of these extreme supermodular functions. The third goal of our work is a computer program which realizes the corresponding inference mechanism for conditional independence statements over five variables mentioned in [10].

The report has the following structure. The next section contains definitions of basic concepts and recalls elementary facts. Section 3 describes our method of obtaining the basic catalogue of representatives of extreme supermodular functions using a computer program. Every item in the catalogue contains a certain representative of a class of permutably equivalent extreme standardized supermodular functions. A specific conjecture concerning the form of these functions valid in case of four variables [11] is refuted in Section 4. Several important characteristics of the above equivalence classes are introduced there as well. Other (alternative) ways of representation of extreme supermodular set functions are mentioned in Section 5. Section 6 is an overview of our internal notation which was agreed for the purpose of easy transition between theoretical concepts and their computer implementation. Section 7 describes the catalogue (realized in the form of a computer file). The Appendix illustrates pictorially the situation in case of 4 variables. The core of the report is self-contained. However, let us warn the reader that some parts (mainly additional remarks) can be fully understood only when the reader is aware of the content in the references.

2 Basic concepts

Let N be a finite non-empty set of *variables*. The symbol $\mathcal{P}(N)$ will denote the power set of N, that is the class of all its subsets $\{A\,;\, A \subseteq N\}$. The symbol $\mathbf{R}$ denotes the set of real numbers, the symbol $\mathbf{Q}$ the set of rational numbers, the symbol $\mathbf{Z}$ the set of integer numbers. *Imset* over N is an integer-valued function on the power set, that is an arbitrary function $i : \mathcal{P}(N) \to \mathbf{Z}$.

2.1 Supermodular function

A real function $m : \mathcal{P}(N) \to \mathbf{R}$ is called *supermodular* if

$$m(U \cup V) + m(U \cap V) \geq m(U) + m(V) \quad \text{for every } U, V \subseteq N\,.$$

Let us denote the class of supermodular functions by $\mathcal{M}$. It forms a cone, that is

$$\forall m_1, m_2 \in \mathcal{M} \;\; \forall \alpha, \beta \geq 0 \quad \alpha \cdot m_1 + \beta \cdot m_2 \in \mathcal{M}\,.$$

For the purpose of computer handling we need to characterize supermodular functions by means of minimal number of inequalities. This is done hereafter.

2.1.1 Elementary imsets

Elementary imset is an imset of a specific form. Given a triplet $\langle i,j|K\rangle$ where $K \subseteq N$ is a set of variables and $i,j \in N \setminus K$ are distinct variables ($i \neq j$) we introduce the corresponding *elementary imset* u over N, sometimes denoted by $u_{\langle i,j|K\rangle}$ as follows

$$u(K) \;=\; u(\{i,j\} \cup K) \;=\; +1, \quad u(\{i\} \cup K) \;=\; u(\{j\} \cup K) \;=\; -1, \quad u(S) \;=\; 0 \;\text{ for other } S \subseteq N\,.$$

The class of elementary imsets over N will be denoted by $\mathcal{E}(N)$. One can easily compute the number of distict elementary imsets, namely $n \cdot (n-1) \cdot 2^{n-3}$ where n is the number of elements of N.

Remark Thus, the imset $u_{\langle i,j|K\rangle}$ corresponds to the elementary conditional independence statement $\{i\} \perp\!\!\!\perp \{j\} \,|\, K$ over N. Note that (when one deals with probabilistic conditional independence structures) it is no reason to distinguish between $\{i\} \perp\!\!\!\perp \{j\} \,|\, K$ and $\{j\} \perp\!\!\!\perp \{i\} \,|\, K$ which is reflected by the fact that $u_{\langle i,j|K\rangle} = u_{\langle j,i|K\rangle}$.

2.1.2 Scalar product

Let $m : \mathcal{P}(N) \to \mathbf{R}$ and $u : \mathcal{P}(N) \to \mathbf{Z}$ be two function on the power set of N. Their *scalar product*, denoted by $\langle m,u\rangle$, is the real number defined by the formula

$$\langle m,u\rangle \;=\; \sum_{S \subseteq N} m(S) \cdot u(S)\,.$$

It is indeed the scalar product on the space $\mathbf{R}^{\mathcal{P}(N)}$.

2.1.3 Equivalent definition of supermodular function

It is quite easy to verify the following simple fact (see for example Lemma 2.7 in [11]).

Lemma 2.1 Function $m : \mathcal{P}(N) \to \mathbf{R}$ is supermodular if and only if

$$\forall u \in \mathcal{E}(N) \quad \langle m,u\rangle \geq 0\,.$$

2.1.4 Examples

Let us give a simple canonical example of a supermodular function. Given $A \subseteq N$ we introduce the function $m_A : \mathcal{P}(N) \to \mathbf{R}$ by the formula

$$m_A(S) = \begin{cases} 1 & \text{if } A \subseteq S \\ 0 & \text{otherwise} \end{cases} \quad \text{for every } S \subseteq N\,.$$

Evidently, m_A is supermodular for every $A \subseteq N$.

Remark Another important example of a supermodular function, the one we have particularly in mind, is the *multiinformation function* m_P induced by a probability distribution P over N. The value $m_P(S)$ for every non-empty $S \subseteq N$ is defined as the relative entropy of the marginal distribution P^S (= the marginal of P for S) with respect to the product of its one-dimensional marginals $\prod_{i \in S} P^{\{i\}}$ and $m_P(\emptyset) = 0$ by convention. This works properly (that is, it gives finite values) both in case of discrete distributions (= finite sample space for every $i \in N$) and in case of non-degenerate Gaussian distributions (= normal distributions with a regular covariance $N \times N$ matrix).

2.1.5 Submodular functions

A real function $l : \mathcal{P}(N) \to \mathbf{R}$ is called *submodular* if

$$l(U \cup V) + l(U \cap V) \le l(U) + l(V) \quad \text{for every } U, V \subseteq N\,,$$

or equivalently if $-l$ is supermodular. There exist functions which are simultaneously supermodular and submodular, for example the function $m_\emptyset$ and the functions $m_{\{i\}}$ for every $i \in N$.

2.2 Standardization

As mentioned earlier the aim of this work is to examine the cone of supermodular functions. The only problem which complicates easy view on this cone is that there exist functions which are simultaneously supermodular and submodular. However, this ambiguity can be avoided by a suitable standardization of supermodular functions done hereafter.

2.2.1 Contained linear subspace

Let us denote by $\mathcal{L}$ the class of those functions which are simultanously supermodular and submodular, that is $\mathcal{L} = \mathcal{M} \cap (-\mathcal{M}) = \{l \in \mathcal{M};\ -l \in \mathcal{M}\}$. It follows from Lemma 2.1 that

$$l \in \mathcal{L} \quad \Leftrightarrow \quad [\forall u \in \mathcal{E}(N) \quad \langle l, u \rangle = 0\,]\,.$$

Hence, $\mathcal{L}$ is a linear subspace of $\mathbf{R}^{\mathcal{P}(N)}$. As mentioned in Section 2.1.5 the functions $m_\emptyset$ and $m_{\{i\}}$, $i \in N$ belong to $\mathcal{L}$. They are evidently linearly independent. Actually, it is easy to show that they form a basis of $\mathcal{L}$ (see for example Lemma 2 [6]). Hence, one can summarize.

LEMMA 2.2 The set $\mathcal{L}$ is a linear subspace of $\mathbf{R}^{\mathcal{P}(N)}$ of dimension $n+1$ where n is the number of elements of N.

2.2.2 Strong equivalence of supermodular functions

We say that two real functions m_1 and m_2 on $\mathcal{P}(N)$ are *strongly equivalent* if

$$\forall u \in \mathcal{E}(N) \quad \langle m_1, u \rangle = \langle m_2, u \rangle\,.$$

Clearly, m_1 and m_2 are strongly equivalent if and only if $m_1 - m_2 \in \mathcal{L}$. By Lemma 2.1, whenever a function is strongly equivalent to a supermodular function then it is supermodular as well. From our point of view there is no reason to distinguish between strongly equivalent functions. Thus, to have a clear view on the situation one should choose one representative from every class of strongly equivalent supermodular functions in a systematic way.

2.2.3 Direct sum

To have also reasonable geometric insight one should simply do the choice 'linearly'. Thus, one should choose another linear subspaces $\mathcal{S}$ such that the direct sum $\mathcal{L} \oplus \mathcal{S}$ is the whole space $\mathbf{R}^{\mathcal{P}(N)}$. That means $\mathcal{L} \cap \mathcal{S}$ contains only the zero function and every function in $\mathbf{R}^{\mathcal{P}(N)}$ can be written as the sum $l + s$ where $l \in \mathcal{L}$ and $s \in \mathcal{S}$ (then l, s are determined uniquely). The fact $\mathbf{R}^{\mathcal{P}(N)} = \mathcal{L} \oplus \mathcal{S}$ then easily implies $\mathcal{M} = \mathcal{L} \oplus (\mathcal{M} \cap \mathcal{S})$. Hence, every $m \in \mathcal{M}$ is strongly equivalent to just one $s \in \mathcal{M} \cap \mathcal{S}$ and one can represent the corresponding equivalence class by s. We chose the following subspace as our standard:

$$\mathcal{S} \;=\; \{\, m : \mathcal{P}(N) \to \mathbf{R}\ ;\ m(\emptyset) = 0 \text{ and } m(\{i\}) = 0 \text{ for every } i \in N\,\}\,.$$

Thus, the result of our standardization is the cone of *standardized supermodular* functions $\mathcal{M} \cap \mathcal{S}$, denoted below by $\mathcal{M}_{st}$. Important fact is that it already does not contain a non-trivial linear subspace, that is $\mathcal{M}_{st} \cap (-\mathcal{M}_{st})$ contains only the zero function. Straightforward consequence of our choice (see for example Lemma 2.10 [8]) is the following pleasant property.

LEMMA 2.3 Every $m \in \mathcal{M}_{st}$ is a non-decreasing function, that is

$$m(S) \leq m(T) \quad \text{whenever } S \subseteq T \subseteq N\,.$$

Remark Our choice of the subspace $\mathcal{S}$ above is motivated by our wish to represent supermodular functions in multiinformation-like style. Indeed, every multiinformation function mentioned in Section 2.1.4 belongs to $\mathcal{M} \cap \mathcal{S}$. However, they are other possible ways of standardization. For example, one can consider the following subspace $\mathcal{R}$ instead of $\mathcal{S}$

$$\mathcal{R} \,=\, \{\, r : \mathcal{P}(N) \to \mathbf{R}\ ;\ r(S) = 0 \text{ whenever } |S| \geq |N| - 1\,\}$$

which leads to 'reverse' standardization by non-increasing functions. Another interesting option is to take simply the orthogonal complement $\mathcal{L}^{\perp}$ in place of $\mathcal{S}$.

2.2.4 Model equivalence of supermodular functions

We say that two real functions m_1 and m_2 on $\mathcal{P}(N)$ are *model equivalent* if

$$\forall\, u \in \mathcal{E}(N) \quad \langle m_1, u\rangle = 0 \;\text{ if and only if }\; \langle m_2, u\rangle = 0\,.$$

This concept is motivated by the approach from [11] where a certain model of conditional independence structure was ascribed to every supermodular function. Namely, given $m \in \mathcal{M}$ by the ('elementary' version of the) *induced independence model* I_m can be understood the class of those elementary conditional independence statements $\{i\} \perp\!\!\!\perp \{j\} \,|\, K$ for which $\langle m, u_{\langle i,j|K\rangle}\rangle = 0$. Thus, two supermodular functions are model equivalent if and only if they induce the same model.

Remark The multiinformation function m_P (induced by a probability distribution P over N) mentioned in Section 2.1.4 is important for the following reason. It encodes whole conditional independence structure induced by P since one has (for every elementary conditional independence statement) $\{i\} \perp\!\!\!\perp \{j\} \,|\, K\ [P]$ (i.e., $\{i\} \perp\!\!\!\perp \{j\} \,|\, K$ is valid in P) if and only if $\langle m_P, u_{\langle i,j|K\rangle}\rangle = 0$. Thus, the model induced by m_P is exactly the ('elementary' version of the) conditional independence model of P. One can consider the concept of *probabilistically representable* (standardized) supermodular function, that is such a function $m \in \mathcal{M}_{st}$ that there exists a probability distribution P over N and $\alpha > 0$ such that $m = \alpha \cdot m_P$. So, supposing $m \in \mathcal{M}_{st}$ is probabilistically representable, the induced independence model is indeed the conditional independence model induced by a probability distribution.

2.3 Extreme supermodular functions

A non-zero standardized supermodular function $m \in \mathcal{M}_{st}$ is called *extreme* if only segments in $\mathcal{M}_{st}$ having it as an internal point are segments consisting of multiples of m, that is $0 \neq m \in \mathcal{M}_{st}$ satisfies the following condition

$$\forall\, r, s \in \mathcal{M}_{st} \quad m = \frac{1}{2} \cdot r + \frac{1}{2} \cdot s \;\text{ implies }\; r = \alpha \cdot m \;\text{ for some } \alpha \geq 0\,.$$

Of course, zero function satisfies the above condition as well, but the definition requires that an extreme function must be non-zero. The point is that one can find a finite set of such functions characterizing the whole cone $\mathcal{M}_{st}$ (as the conical hull of this finite set). One can achieve uniqueness of such a set as a result of a certain type of normalization done hereafter. Note that the facts mentioned and used below have no immediate proof but they are certainly evident to everyone who is familiar with linear programming.

2.3.1 Extreme rays

Let us recall a few basic concepts and facts from the theory of convex cones in finite-dimensional real vector spaces (they will be applied to the case $\mathbf{R}^{\mathcal{P}(N)}$). By the *ray* of a cone $\mathcal{K}$ generated by $0 \neq x \in \mathcal{K}$ is understood the set $\{\alpha \cdot x\,;\, \alpha \geq 0\}$. A ray $\mathcal{R}$ is called extreme if $\forall u, v \in \mathcal{K}$ the fact $\frac{1}{2} \cdot u + \frac{1}{2} \cdot v \in \mathcal{R}$ implies $u \in \mathcal{R}$. Thus, a function from $\mathcal{M}_{st}$ is extreme in the sense of earlier definition if and only if it belongs to an extreme ray of $\mathcal{M}_{st}$.

The first basic fact is that every cone of the form $\mathcal{K} = \{m \in \mathbf{R}^{\mathcal{P}(N)}\,;\, \forall u \in \mathcal{F} \quad \langle m, u\rangle \geq 0\}$ where $\mathcal{F} \subseteq \mathbf{Q}^{\mathcal{P}(N)}$ is a finite set, can be equivalently described as the conical hull of a finite non-empty $\mathcal{G} \subseteq \mathbf{Q}^{\mathcal{P}(N)}$, that is $\mathcal{K} = \{\sum \alpha_v \cdot v\,;\, v \in \mathcal{G} \quad \alpha_v \geq 0\}$ (see Proposition 5b in [9]). It can be applied to the case of $\mathcal{M}_{st}$ since one can put

$$\mathcal{F} \,=\, \mathcal{E}(N) \cup \{\delta_\emptyset, -\delta_\emptyset\} \cup \{\delta_{\{i\}}\,;\, i \in N\} \cup \{-\delta_{\{i\}}\,;\, i \in N\}\,,$$

where

$$\delta_A(S) = \begin{cases} 1 & \text{if } A = S, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $\mathcal{M}_{st}$ can be obtained as the conical hull of a finite $\emptyset \neq \mathcal{G} \subseteq \mathbf{Q}^{\mathcal{P}(N)}$.

The second basic fact is that every such a cone $\mathcal{K}$ satisfying also $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$ has finitely many extreme rays and all of them are generated by elements of $\mathbf{Q}^{\mathcal{P}(N)}$ (see Consequence 5b in [9]). As mentioned earlier in Section 2.2.3 this is also the case of $\mathcal{M}_{st}$.

The third basic fact is that then for every finite set $\mathcal{H} \subset \mathcal{K} \neq \{0\}$ generating all extreme rays of $\mathcal{K}$ one has $\mathcal{K} = \{\sum \alpha_v \cdot v\,;\; v \in \mathcal{H} \quad \alpha_v \geq 0\}$ (see Proposition 4 in [9]). In case of the cone $\mathcal{M}_{st}$ it means that it suffices to choose just one element (representative) from every extreme ray.

2.3.2 Normalization

A natural way of the choice of a representative within every extreme ray of $\mathcal{M}_{st}$ is the following one. We say that an imset m over N is *normalized* if the collection of integers $\{m(S);\, S \subseteq N\}$ has no common prime divisor. Since we know that every extreme ray of $\mathcal{M}_{st}$ is generated by a rational-valued function, it has to contain a non-zero integer-valued function (= imset). Therefore it contains a normalized imset as well. Of course, it can contain at most one non-zero normalized imset. So, one can summarize the results as follows.

LEMMA 2.4 For every extreme $r \in \mathcal{M}_{st}$ there exist unique normalized imset m such that $r = \alpha \cdot m$ for $\alpha > 0$ ($m \in \mathcal{M}_{st}$ is extreme as well). The class of normalized extreme (standardized) supermodular imsets $\mathcal{M}_{ex}$ is finite and every $m \in \mathcal{M}_{st}$ can be written in the form

$$m \,=\, \sum_{v \in \mathcal{M}_{ex}} \alpha_v \cdot v \quad \text{where} \quad \alpha_v \geq 0\,.$$

Thus, the class of normalized integer-valued extreme standardized supermodular functions $\mathcal{M}_{ex}$ is uniquely determined finite class and we will call it the *skeleton* (terminology from [11]). Their elements are called *skeletal imsets*.

2.3.3 Interpretation of extreme supermodular functions

The reason why we are interested in extreme supermodular functions is that they play an important role in the approach from [11]. We have already mentioned in Section 2.2.4 that every supermodular function induces a certain model of conditional independence structure. Clearly, zero function induces the maximal conditional independence model (involving all elementary conditional independence statements). However, one can derive by combining the results from [11, 9] (Lemma 2.3, Consequence 2.4, Assertion 1.4) the following conlusion.

LEMMA 2.5 A standardized supermodular function m is extreme if and only if the induced independence model I_m is a submaximal independence model within the class of independence models induced by supermodular functions, in sense that the only distinct independence model containing I_m is the maximal independence model.

Remark Let us note that every independence model induced by a supermodular function is the intersection of the maximal independence model and a collection of submaximal independence models (possibly empty collection). This fact is a basis of our computer program which realizes the coresponding inference mechanism for conditional independence statements over 5 variables (see Section 7.4).

2.4 Permutations

Further simplification can be achieved as a result of considering of permutations of variables. These allow one to construct one extreme supermodular function from another one.

2.4.1 Permutation of variables

By a *permutation* of N is understood any bijective mapping $\pi : N \to N$. The class of all permutations of N forms a group: the result of composition of two permutations is a permutation, identical mapping is the null-element of the group, and every permutation π has an inverse permutation π^{-1}. Every permutation π can be extended to a bijective mapping $\boldsymbol{\pi} : \mathcal{P}(N) \to \mathcal{P}(N)$ as follows

$$\boldsymbol{\pi}(A) = \{\pi(i)\,;\, i \in A\} \quad \text{for every} \quad A \subseteq N\,.$$

Given a real function $m : \mathcal{P}(N) \to \mathbf{R}$ the formula

$$m_\pi(A) = m(\boldsymbol{\pi}(A)) \quad \text{for every} \quad A \subseteq N$$

then defines a permutated function m_π. The following observations are quite obvious.

LEMMA 2.6 Supposing $m : \mathcal{P}(N) \to \mathbf{R}$ is supermodular and π is a permutation of N, the function m_π is supermodular as well. Supposing m is standardized, the same holds for m_π. Supposing m is a normalized imset, m_π is also a normalized imset.

Thus, given a permutation π of N, the mapping $m \mapsto m_\pi$ is a bijective linear mapping which maps $\mathcal{M}_{st}$ onto $\mathcal{M}_{st}$. This fact makes it possible to derive that whenever $m \in \mathcal{M}_{st}$ is extreme them m_π is extreme as well. So, we can conlude the following fact.

LEMMA 2.7 Let π be a permutation of N, $m \in \mathcal{M}_{ex}$. Then $m_\pi \in \mathcal{M}_{ex}$ and the mapping $m \mapsto m_\pi$ is a bijective mapping onto $\mathcal{M}_{ex}$.

2.4.2 Permutable equivalence of supermodular functions

We say that two real functions m_1 and m_2 on $\mathcal{P}(N)$ are *permutably equivalent* if

$$\exists \text{ permutation } \pi \text{ on } N \text{ such that } \; m_1 \,=\, (m_2)_\pi \,.$$

It is evidently an equivalence relation. By Lemma 2.7, whenever a function is permutably equivalent to $m \in \mathcal{M}_{ex}$ then it belongs to $\mathcal{M}_{ex}$. Thus, it can be understood as equivalence relation on the skeleton $\mathcal{M}_{ex}$. Permutations enable one to do further reduction. Instead of keeping the whole list of elemets of the skeleton it suffices to choose one representative of each equivalence class of permutably equivalent skeletal imsets (according to a suitable criterion) and keep it on a reduced list.

Remark Conditional independence models mentioned in Section 2.2.4 can be permutated as well. Indeed, given a permutation π of N one can introduce the following bijective mapping

$$\mathit{\Pi} \,:\, \{i\} \perp\!\!\!\perp \{j\} \,|\, K \quad \mapsto \quad \boldsymbol{\pi}(\{i\}) \perp\!\!\!\perp \boldsymbol{\pi}(\{j\}) \,|\, \boldsymbol{\pi}(K)$$

between elementary conditional independence statements. Thus, every class of elementary conditional independence statements I can be permutated into the class of permutated statements $\mathit{\Pi}(I)$. It makes no problem to see that for every supermodular function m one has $I_{m_\pi} = \mathit{\Pi}_{-1}(I_m)$. Thus, permutations can be extended to independence models as well.

One can also show that permutation of a probabilistically representable supermodular function is a probabilistically representable supermodular function. Indeed, every probability distribution P over N, defined on the sample space $\prod_{i\in N} \mathbf{X}_i$, can be permutated to a probability distribution P_π, defined on the sample space $\prod_{i\in N} \mathbf{Y}_i \equiv \prod_{i\in N} \mathbf{X}_{\pi(i)}$, which is also a probability distribution over N. It follows from the definition of multiinformation function that $(m_P)_\pi = m_{P_\pi}$. In particular, permutation of a probabilistically representable element of $\mathcal{M}_{ex}$ is a probabilistically representable element of $\mathcal{M}_{ex}$. Therefore, probabilistic representability is an invariant property with respect to permutable equivalence.

2.5 Examples

Let us describe the situation in case $|N| < 5$. The trivial case of 2 variables is omitted.

2.5.1 Three variables

In case $N = \{a, b, c\}$ the class of elementary imsets $\mathcal{E}(N)$ has 6 elements. It is easy to find out that the skeleton $\mathcal{M}_{ex}$ has then 5 elements [6]. Skeletal imsets are shown in Figure 1 by means of special diagrams. The diagrams are in principle Hasse diagrams of the lattice of subsets N: the ovals in the diagrams correspond to subsets of N, links between two ovals are made when the symmetric difference of the represented sets is a singleton, and the integer number written in the oval is the value of the represented imset in the set represented by the oval. There are only 3 classes of permutably equivalent skeletal imsets over $N = \{a, b, c\}$. Rows in Figure 1 correspond to them.

2.5.2 Four variables

There exist 24 elementary imsets in case of 4 variables. The skeleton has 37 elements [7] and decomposes into 10 classes of permutably equivalent imsets (= types). Representatives of these types are depicted in the Appendix (Figures 6 - 15).

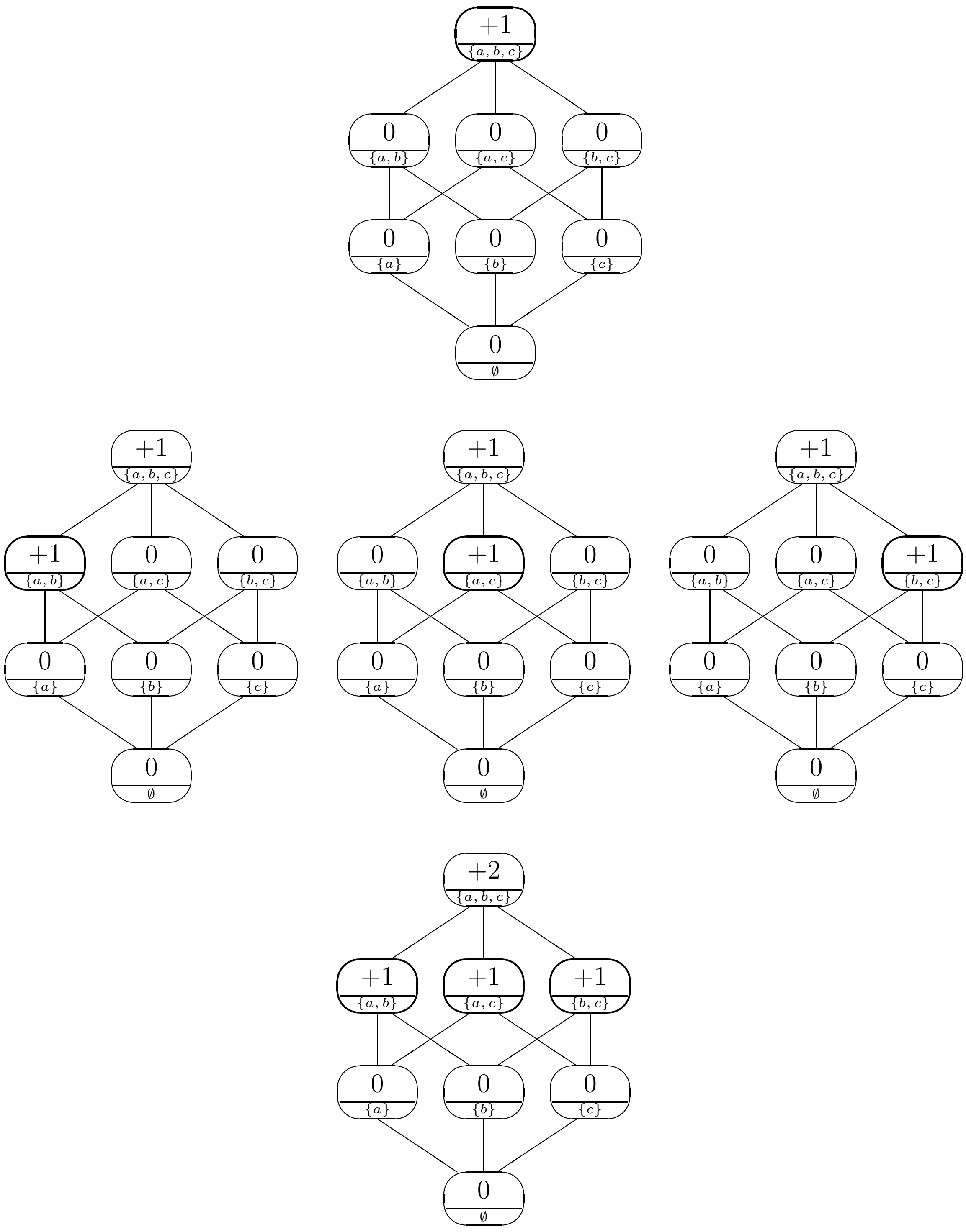


Figure 1: All skeletal imsets over $N = \{a, b, c\}$.

Remark Nine of ten types of skeletal imsets over $N = \{a, b, c, d\}$ consist of probabilistically representable supermodular functions [12]. Figure 14 gives an example of a skeletal imset which is not probabilistically representable by a discrete probability distribution.

3 Our method

3.1 Computation

A computer program PORTA [13] was used to obtain all extreme rays of $\mathcal{M}_{st}$ in case of 5 variables, that is $N = \{a, b, c, d, e\}$. PORTA is a collection of routines for analysing polytopes and polyhedra. Its procedure `traf` finds all extreme points of a bounded convex set defined by means of linear inequalities with rational coefficients. It uses special integer arithmetics so that the result is accurate (no rounding off during computation is made).

To adapt our problem for the procedure we have intersected $\mathcal{M}_{st}$ with the linear manifold

$$\mathcal{T} = \{ m : \mathcal{P}(N) \rightarrow \mathbf{R} \, ; \, m(N) = 1 \}$$

so that $\mathcal{M}_{st} \cap \mathcal{T}$ is a non-empty bounded convex set and every extreme ray of $\mathcal{M}_{st}$ contains just one point of $\mathcal{T}$. Then extreme points of $\mathcal{M}_{st} \cap \mathcal{T}$ correspond to extreme rays of $\mathcal{M}_{st}$. Owing to Lemma 2.1 $\mathcal{M}$ can be characterized by 80 inequalities with integer coefficients. Restriction to the linear subspace $\mathcal{S}$ (see Section 2.2.3) was done implicitly: every element of $\mathcal{S}$ was encoded into 26-dimensional real vector in such a way that every set $S \subseteq N$ with $|S| \geq 2$ was identified with a component of the vector. Restriction to the manifold $\mathcal{T}$ was realized by one extra equality restriction. Thus, an input text file describing 80 inequality and 1 equality restrictions in 26-dimensional in real vector space was written. Then, we have used the above mentioned procedure `traf` to get a 16 megabyte text output file. Note that computation for the five variables case took approximately eight hours while the four variables case took only a few seconds. This indicates that most likely it is not feasible to investigate the case of six or more variables using this method in future.

3.2 Normalization

The next step was to convert the above mentioned output text file into a suitable data file. Ideally, every obtained 26-dimensional real vector $\mathbf{x}$ should be modified as follows.

- Change $\mathbf{x}$ into a 32-dimensional real vector $\underline{\mathbf{x}}$ by incorporating additional zero components which are identified with sets $S \subseteq N$ satisfying $|S| \leq 1$.
- Multiply $\underline{\mathbf{x}}$ by a suitable rational constant so that the resulting 32-dimensional vector $\boldsymbol{x}$ has integer components with no common prime divisor (i.e. normalize it in sense of Section 2.3.2).

In such a way a big 117978×32 matrix describing the skeleton could be obtained. Its rows are skeletal imsets, the chosen coding of sets into components is described in Section 6.1.

3.3 Reduction

However, such a matrix is too big for an analysis without help of a computer. Fortunately, it can be reduced in the way described in Section 2.4.2. To have a proper criterion of choice of type representatives we considered weight $w(S)$ for every $S \subseteq N$:

$$w(S) = 2^{\delta(S,a)} \cdot 3^{\delta(S,b)} \cdot 4^{\delta(S,c)} \cdot 5^{\delta(S,d)} \cdot 6^{\delta(S,e)}$$

where

$$\delta(S,x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases} \qquad \text{for every } x \in \{a,b,c,d,e\}\,.$$

Thus, from every class of permutably equivalent skeletal imsets a representative m minimizing $\langle m, w\rangle$ was chosen, i.e.

$$m \,=\, \text{argmin}\,\{\sum_{S\subseteq N} m(S)\cdot w(S)\}\,.$$

Note that a similar principle of choice of type representative was used in the case of 4 variables (see the Appendix). The intuition behind is clear: the chosen representative m 'ascribes its mass' mainly to a, variable b has the second priority, after that c, d, and e follow (in this order).

Actually, we performed the above mentioned procedure before the normalization step described in Section 3.2. Indeed, this is possible because permutably equivalent vectors must be multiplied by the same constant during normalization. Thus, unnecessary repetition of normalization steps was avoided. In such a way a reduced 1319×32 matrix describing types of skeletal imsets was obtained. Moreover, the number of skeletal imsets belonging to each type was computed. This number can be considered as an auxiliary characteristic of the type.

3.4 Summary

Thus, in case of 5 variables one has 80 elementary imsets. The skeleton has 117978 imsets and decomposes into 1319 types. We have made a virtual catalogue of all types of skeletal imsets over $N = \{a,b,c,d,e\}$ (an analogue of the printed catalogue in case of 4 variables from the Appendix). The catalogue is described in Section 7.

4 Refuted conjectures

An analysis in case of 4 variables led to a naive conjecture how to obtain all skeletal imsets. Roughly said, the hypotesis was that every skeletal imset is determined by its support, i. e., the class of sets with a non-zero value. In such a case the skeleton over N could have been obtained by means of thorough analysis of ascending classes of subsets of N.

4.1 Minimal sets

A class $\mathcal{A} \subseteq \mathcal{P}(N)$ is called *ascending* if

$$S \subseteq T,\;\; S \in \mathcal{A} \;\;\Rightarrow\;\; T \in \mathcal{A}\,.$$

Every ascending class $\mathcal{A} \subseteq \{S \subseteq N\,;\, |S| \geq 2\}$ is uniquely determined by the class of its *minimal sets* with respect to inclusion

$$\mathcal{A}_{\min} \,=\, \{S \in \mathcal{A}\,; \forall T \in \mathcal{A} \;\; T \subseteq S \;\Rightarrow\; T = S\,\}\,,$$

which is typically smaller. Supposing $m \in \mathcal{M}_{st}$ its *support*

$$\mathcal{A}_m \,=\, \{S \subseteq N\,;\, m(S) > 0\,\}$$

is an ascending subclass of $\{S \subseteq N\,;\, |S| \geq 2\}$ by Lemma 2.3. Conversely, every such a class $\mathcal{A}$ induces a standardized supermodular imset $m_{\mathcal{A}}$ as follows. We put

$$\begin{aligned} m_{\mathcal{A}}(S) \,&=\, 0 \qquad && \text{for } S \subseteq N,\; S \notin \mathcal{A}, \\ m_{\mathcal{A}}(S) \,&=\, 1 \qquad && \text{for } S \in \mathcal{A}_{\min}, \end{aligned}$$

and determine the remaining values (for $S \in \mathcal{A} \setminus \mathcal{A}_{\min}$) gradually (first for S with $|S| = 2$, then for S with $|S| = 3$ etc.) by means of the formula

$$m_{\mathcal{A}}(S) \;=\; \max\{\, m_{\mathcal{A}}(S \setminus \{i\}) + m_{\mathcal{A}}(S \setminus \{j\}) - m_{\mathcal{A}}(S \setminus \{i,j\})\,;\;\; i,j \in S,\;\; i \neq j\,\}\,.$$

Indeed, the formulas above imply that for every $S \subseteq N$, $|S| \geq 2$ and every $u_{\langle i,j|K\rangle} \in \mathcal{E}(N)$ with $\{i,j\} \cup K = S$ one has

$$m_{\mathcal{A}}(S) = m_{\mathcal{A}}(\{i,j\} \cup K) \geq m_{\mathcal{A}}(\{i\} \cup K) + m_{\mathcal{A}}(\{j\} \cup K) - m_{\mathcal{A}}(K)\,.$$

Hence, by Lemma 2.1 $m_{\mathcal{A}} \in \mathcal{M}_{st}$ and $\mathcal{A}$ is the support of $m_{\mathcal{A}}$ (by Lemma 2.3). Thus, a certain class of standardized supermodular functions can be obtained in this way. An interesting fact is that in case $|N| \leq 4$ every $m \in \mathcal{M}_{ex}$ has this form. The reader can verify it directly (see Figure 1 and the Appendix where minimal sets of the support are emphasized). The conjecture was that every skeletal imset can be obtained in a similar way.

4.2 Incremental transformation

Let us reformulate the conjecture in a more elegant way. For every $S \subseteq N$ with $|S| \geq 2$ denote

$$\mathcal{E}^{\uparrow}_{S}(N) \;=\; \{u_{\langle i,j|K\rangle} \in \mathcal{E}(N)\,;\, \{i,j\} \cup K = S\}\,.$$

Given $m : \mathcal{P}(N) \to \mathbf{R}$ define the corresponding *incremental function* as follows:

$$\begin{array}{rcll} \Delta_m(S) &=& \min\{\, \langle m,u\rangle\,;\, u \in \mathcal{E}^{\uparrow}_{S}(N)\,\} & \text{for } |S| \geq 2,\\ \Delta_m(S) &=& 0 & \text{otherwise.} \end{array}$$

Clearly, by Lemma 2.1 m is supermodular if and only if Δ_m is non-negative. For every $S \subseteq N$, $|S| \geq 2$ one can rewrite $\Delta_m(S)$ in the form

$$m(S) - \max\{\, m(\{i\} \cup K) + m(\{j\} \cup K) - m(K)\,;\, u_{\langle i,j|K\rangle} \in \mathcal{E}^{\uparrow}_{S}(N)\,\}\,.$$

This gives an inverse formula for standardized $m \in \mathcal{S}$

$$m(S) \;=\; \Delta_m(S) + \max\{\, m(\{i\} \cup K) + m(\{j\} \cup K) - m(K)\,;\, u_{\langle i,j|K\rangle} \in \mathcal{E}^{\uparrow}_{S}(N)\,\} \;\; \text{for } S \subseteq N,$$

which has to be applied with increasing cardinality of S (maximum of empty collection is 0 by convention). Thus, $\Delta_m(S)$ is indeed the 'increment' of $m \in \mathcal{M}$ at S with respect to a natural lower estimate. Moreover, the formula above implies that $m \in \mathcal{M}_{st}$ is uniquely determined by Δ_m. Of course, supposing $\mathcal{A}$ is the support of $m \in \mathcal{M}_{st}$ one has

$$\begin{array}{rcll} \Delta_m(S) &=& 0 & \text{for } S \subseteq N,\; S \notin \mathcal{A},\\ \Delta_m(S) &=& m(S) & \text{for } S \in \mathcal{A}_{\min}. \end{array}$$

The conjecture mentioned in Section 4.1 can be reformulated as follows:

$$\forall m \in \mathcal{M}_{ex} \quad \begin{array}{ll} \Delta_m(S) = 1 & \text{for } \; S \in \mathcal{A}_{\min},\\ \Delta_m(S) = 0 & \text{for } \; S \notin \mathcal{A}_{\min}. \end{array}$$

Both conditions mentioned above need not hold in case $N = \{a,b,c,d,e\}$. An example that $m(S) = \Delta_m(S) > 1$ for a certain minimal set S is given in Figure 2. An example that the incremental function can be non-zero also for non-minimal sets is given in Figure 3.

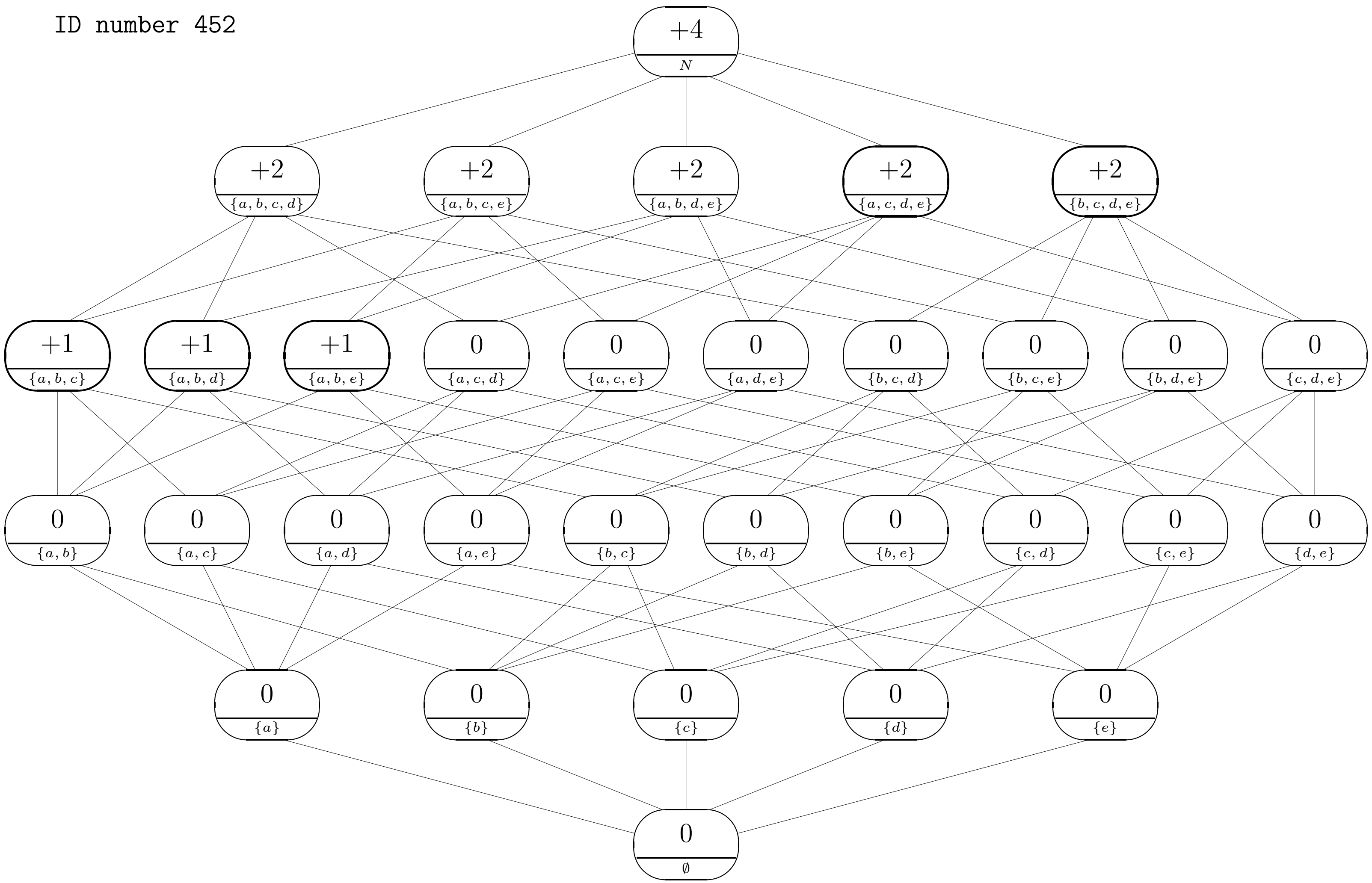


Figure 2: Skeletal inset m and a minimal set $S = \{b,c,d,e\}$ of $\mathcal{A}_m$ with $m(S) > 1$.

ID number 1264

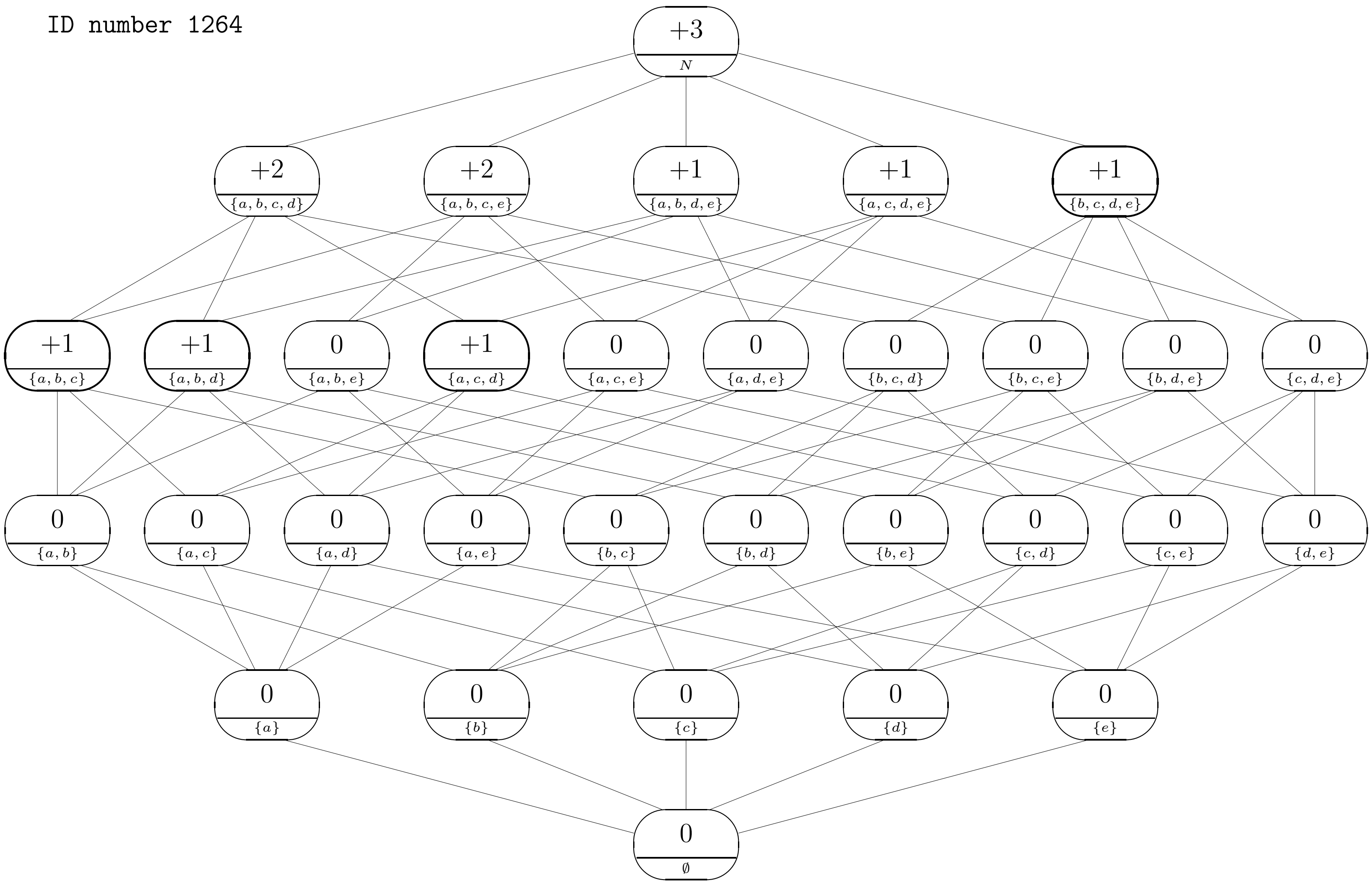


Figure 3: Skeletal imset m and a non-minimal set $S = \{a,b,c,e\}$ of $\mathcal{A}_m$ with $\Delta_m(S) = 1$.

4.3 Grade

Another important characteristic of $m \in \mathcal{M}_{st}$ is the record of its scalar products with elementary imsets

$$\mathsf{S}_m : \; u \in \mathcal{E}(N) \; \to \; \langle m, u \rangle \, .$$

It naturally encodes the induced independence model I_m mentioned in Section 2.2.4: the record has zeros in S_m at positions corresponding to induced valid conditional independence statements. Moreover, m can be reconstructed from S_m since it determines the incremental function Δ_m mentioned in Section 4.2. Another fact illustrating simplicity of cases 3 and 4 variables is that for every skeletal imset $m \in \mathcal{M}_{ex}$ the corresponding record of scalar products S_m consists of zeros and ones [6, 7]. However, this is not true in case of 5 variables as already demostrated in Figure 2. Thus, an interesting characteristic of a skeletal imset m can be its *grade*

$$g(m) \; = \; \max \{ \, \langle m, u \rangle \, ; \, u \in \mathcal{E}(N) \, \} \, .$$

It is easy to see that $g(m_\pi) = g(m)$ for every permutation π over N (see Section 2.4). Therefore, the grade can be considered as a characteristic of the type of skeletal imsets. The maximal possible grade over the set of variables N

$$\mathsf{G} = \max \{ \, g(m) \, ; \, m \in \mathcal{M}_{ex} \, \}$$

appears to be a quite important number. We have found that in case of 5 variables the maximal grade is 7. The corresponding example is given in Figure 4.

Remark The maximal grade plays crucial role in computer verification of 'facial' implication among elementary imsets [10] which enables one to derive 'automatically' formal properties of probabilistic conditional independence. For example, the following property

$$\left[\begin{array}{cc} \{c\} \perp\!\!\!\perp \{d\} \,|\, \{a,b,e\} & \{b\} \perp\!\!\!\perp \{e\} \,|\, \{a,c,d\} \\ \{a\} \perp\!\!\!\perp \{c\} \,|\, \{d,e\} & \{c\} \perp\!\!\!\perp \{e\} \,|\, \{d\} \\ \{d\} \perp\!\!\!\perp \{e\} \,|\, \{c\} & \{a\} \perp\!\!\!\perp \{d\} \,|\, \{e\} \\ \{b\} \perp\!\!\!\perp \{d\} \,|\, \{a\} & \{a\} \perp\!\!\!\perp \{b\} \,|\, \{c\} \end{array} \right] \;\Longleftrightarrow\; \left[\begin{array}{cc} \{d\} \perp\!\!\!\perp \{e\} \,|\, \{a,b,c\} & \{c\} \perp\!\!\!\perp \{e\} \,|\, \{a,b,d\} \\ \{a\} \perp\!\!\!\perp \{b\} \,|\, \{c,d,e\} & \{c\} \perp\!\!\!\perp \{d\} \,|\, \{b,e\} \\ \{b\} \perp\!\!\!\perp \{d\} \,|\, \{e\} & \{b\} \perp\!\!\!\perp \{e\} \,|\, \{c\} \\ \{a\} \perp\!\!\!\perp \{c\} \,|\, \{d\} & \{a\} \perp\!\!\!\perp \{d\} \,|\, \{c\} \\ \{b\} \perp\!\!\!\perp \{e\} \,|\, \{a\} & \end{array} \right]$$

is an easy consequence of the equality

$$\begin{aligned} u_{\langle c,d|abe \rangle} + 2 \cdot u_{\langle b,e|acd \rangle} + u_{\langle a,c|de \rangle} + u_{\langle c,e|d \rangle} + u_{\langle d,e|c \rangle} + u_{\langle a,d|e \rangle} + u_{\langle b,d|a \rangle} + u_{\langle a,b|c \rangle} \quad = \\ u_{\langle d,e|abc \rangle} + u_{\langle c,e|abd \rangle} + u_{\langle a,b|cde \rangle} + u_{\langle c,d|be \rangle} + u_{\langle b,d|e \rangle} + u_{\langle b,e|c \rangle} + u_{\langle a,c|d \rangle} + u_{\langle a,d|c \rangle} + u_{\langle b,e|a \rangle} \, . \end{aligned}$$

An interesting fact is that one cannot derive the above mentioned property of conditional independence as a consequence of a similar equality where the coefficient of the term $u_{\langle b,e|acd \rangle}$ is less than 2. On the other hand, one can show that the integer coefficients corresponding to elementary imsets can be at most the maximal grade G (this will be topic of another paper). Thus, in case of 4 variables the maximal coefficient is 1 while in the case of 5 variables the maximal coefficient is 7. An example of an equality with a coefficient 7 which cannot be diminished is here:

$$\begin{aligned} 7 \cdot u_{\langle c,e|abd \rangle} + 2 \cdot u_{\langle c,d|be \rangle} + 3 \cdot u_{\langle b,d|ce \rangle} + u_{\langle b,c|de \rangle} + u_{\langle a,e|bd \rangle} + 2 \cdot u_{\langle a,d|be \rangle} + 2 \cdot u_{\langle b,c|ae \rangle} \\ 4 \cdot u_{\langle a,c|bd \rangle} + 4 \cdot u_{\langle d,e|a \rangle} + u_{\langle a,b|c \rangle} + u_{\langle a,c|b \rangle} + u_{\langle b,c|a \rangle} + u_{\langle b,d|\emptyset \rangle} + u_{\langle a,e|\emptyset \rangle} \quad = \\ 3 \cdot u_{\langle d,e|abc \rangle} + 4 \cdot u_{\langle b,c|ade \rangle} + 6 \cdot u_{\langle c,e|bd \rangle} + 2 \cdot u_{\langle a,e|cd \rangle} + 2 \cdot u_{\langle a,d|ce \rangle} + 5 \cdot u_{\langle a,c|be \rangle} \\ 2 \cdot u_{\langle b,d|c \rangle} + 2 \cdot u_{\langle c,d|a \rangle} + u_{\langle a,b|e \rangle} + 2 \cdot u_{\langle b,d|a \rangle} + u_{\langle d,e|\emptyset \rangle} + u_{\langle c,e|\emptyset \rangle} \, . \end{aligned}$$

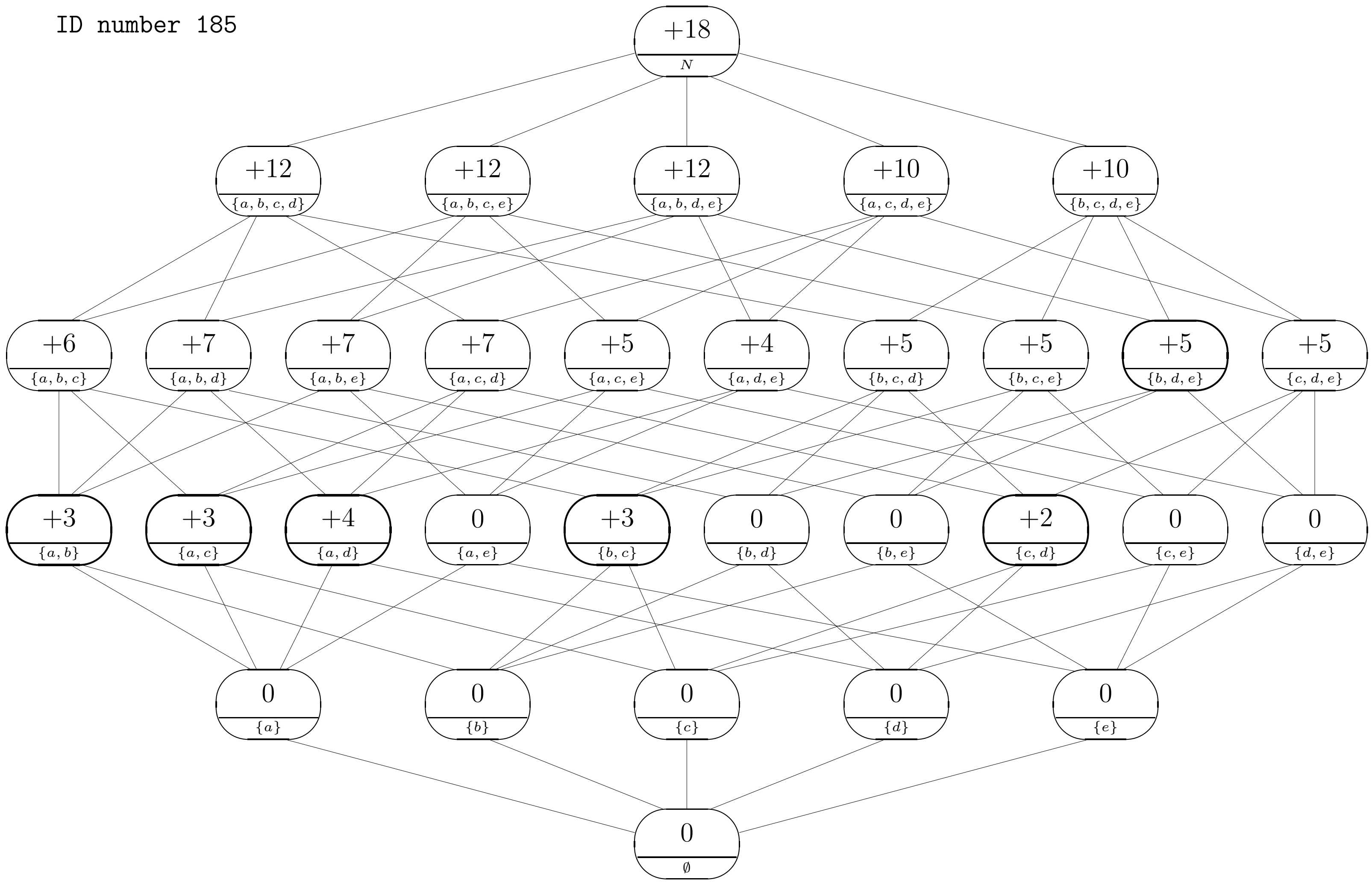


Figure 4: A skeletal imset over $N = \{a, b, c, d, e\}$ with maximal grade 7 attained at $u_{\langle a,b|e\rangle}$.

4.4 Sorting of types

A natural question how to sort types of skeletal imsets in an elegant way arises. Let us describe a certain method which works nicely in case of 4 variables. For every $l = 1, \ldots, |N| - 1$ we introduce the class of elementary imsets of level l:

$$\mathcal{E}_l(N) \;=\; \{\, u_{\langle i,j|K\rangle}\,;\, |K| = l-1 \,\}\,.$$

Clearly, $\mathcal{E}(N)$ decomposes into $\mathcal{E}_l(N)$, $l = 1, \ldots, |N| - 1$. Then every skeletal imset $m \in \mathcal{M}_{ex}$ can be characterized by an integer-valued $(|N|-1) \times (\mathsf{G}+1)$ matrix Z (where G is the maximal grade) with elements

$$z_{lk} \;=\; |\{\, u \in \mathcal{E}_l(N)\,;\, \langle m, u\rangle = k \,\}| \qquad l = 1, \ldots, |N|-1, \quad k = 0, \ldots, \mathsf{G}\,.$$

Thus, the l-th row in the *zip matrix* Z characterizes 'behaviour' of m on the l-th level. Again, it is easy to verify that permutably equivalent imsets have the same zip matrix. Thus, the matrix can be considered as a type characteristic.

For example, in case $N = \{a, b, c\}$ one has $\mathsf{G} = 1$ and zip matrices are 2×2 matrices. Zip matrices of skeletal imsets (see Figure 1) are then

$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}.$$

Evidently, they distinguish all 3 types. Similarly, in case of 4 variables one can sort all types of skeletal imsets easily since the corresponding 3×2 zip matrices distinguish all types (we leave it to the reader). However, the hope that zip matrices distinguish types also in case $N = \{a, b, c, d, e\}$ appeared to be abortive. Examples of two different types having the same 4×8 zip matrix are given in Figure 5.

5 Other ways of representation

There are many other ways of representation of extreme supermodular set functions. Some of them are described in this section.

5.1 Reverse representation

We have mentioned in Section 2.2.3 (Remark) that one can consider an alternative way of standardization of supermodular functions leading to non-increasing functions. That means, instead of the standard representative $m \in \mathcal{M}_{st} = \mathcal{M} \cap \mathcal{S}$ of a class of strongly equivalent supermodular functions one can choose another representative, the one which belongs to $\mathcal{M} \cap \mathcal{R}$ where

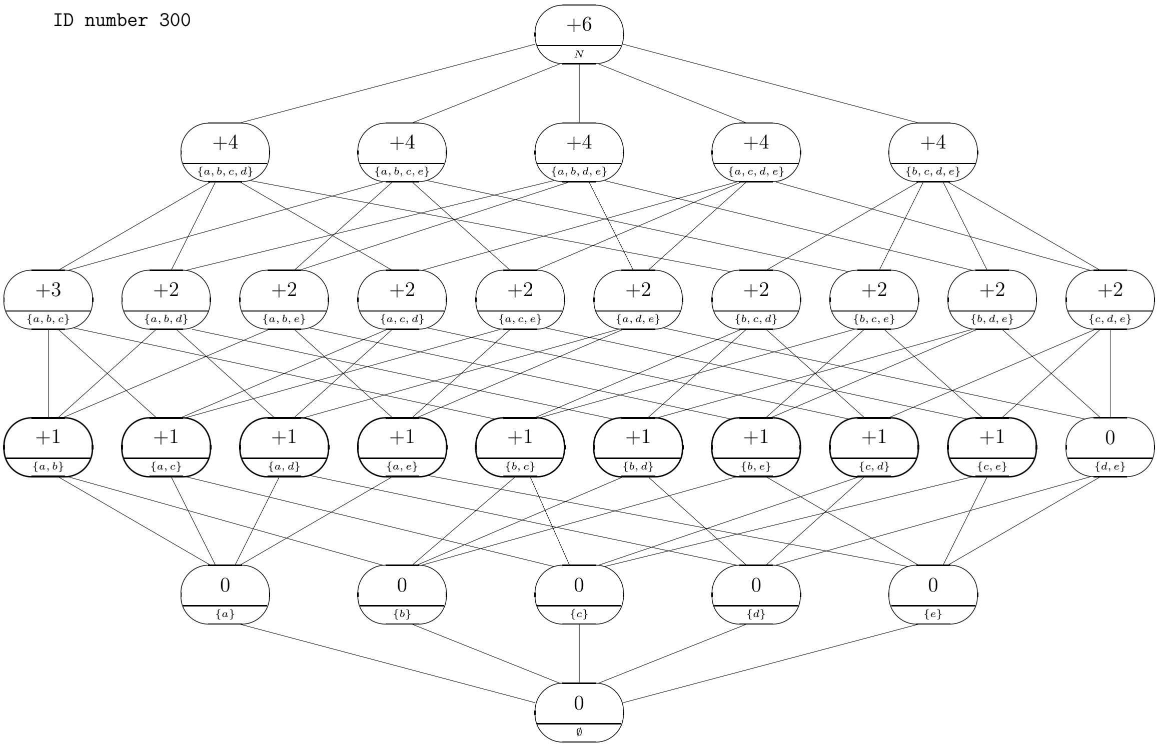
$$\mathcal{R} \;=\; \{\, r : \mathcal{P}(N) \to \mathbf{R}\,;\, r(S) = 0 \text{ whenever } |S| \geq |N| - 1 \,\}\,.$$

Of course, there is one-to-one correspondence between m and r. Indeed, to compute r on basis of m we first put

$$\nu(i) \;=\; m(N) - m(N \setminus \{i\}) \quad \text{for } i \in N, \qquad k \;=\; -m(N) + \sum_{i \in N} \nu(i)\,.$$

We do not know whether the terms $\nu(i)$ and k have certain interpretation, but analogous formulas appeared also in [4] on page 345. Then, r can be obtained from m as follows:

$$r(S) \;=\; m(S) + k - \sum_{i \in S} \nu(i) \quad \text{for } S \subseteq N\,.$$



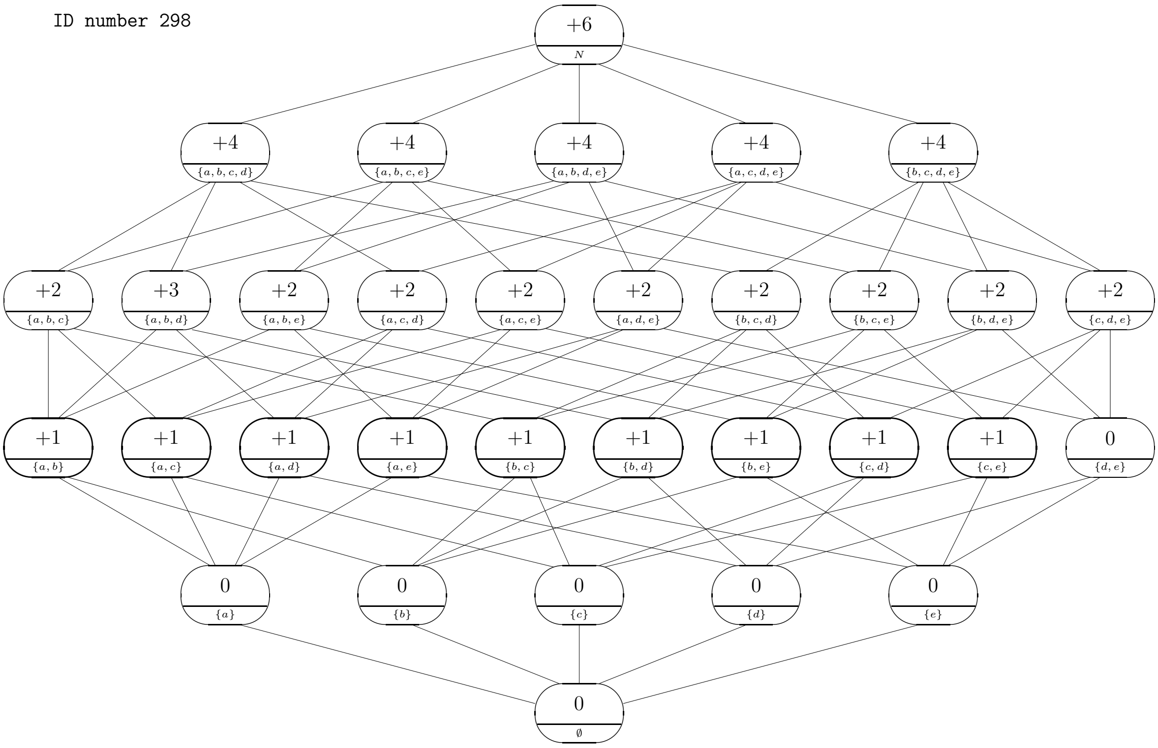


Figure 5: Two permutably non-equivalent skeletal imsets with the same zip matrix.

Clearly, r and m are strongly equivalent since $r - m \in \mathcal{L}$ (see Section 2.2.1). By construction $r(S) = 0$ whenever $|S| \geq |N| - 1$. Given $r \in \mathcal{M} \cap \mathcal{R}$, the formula above can serve as an inverse formula for m, it suffices to put

$$k \,=\, r(\emptyset), \qquad \nu(i) \,=\, k - r(\{i\}) \quad \text{for } i \in N\,.$$

5.1.1 Maximal sets

The choice of the way of standardization appears to be a matter of taste. Alternative choice of reverse standardization leads to completely 'reverse' point of view. Analogously to Lemma 2.3 one can see that every $r \in \mathcal{M} \cap \mathcal{R}$ is a non-increasing function, that is

$$r(S) \geq r(T) \quad \text{whenever } S \subseteq T \subseteq N\,.$$

In particular, the support

$$\mathcal{D} \,=\, \{S \subseteq N\,;\, r(S) > 0\,\}$$

of $r \in \mathcal{M} \cap \mathcal{R}$ is a descending subclass of $\{S \subseteq N\,;\, |S| \leq |N| - 2\,\}$. The class of its *maximal sets*

$$\mathcal{D}_{\max} \,=\, \{S \in \mathcal{D}\,;\, \forall\, T \in \mathcal{D} \;\; T \supseteq S \;\Rightarrow\; T = S\,\}$$

can be considered as a characteristic of a 'reverse' skeletal imset (an analogue of the class of minimal sets $\mathcal{A}_{\min}$ mentioned in Section 4.1).

5.1.2 Decremental transformation

The 'reverse' point of view leads to further possible way of representation. For every $S \subseteq N$ with $|S| \leq |N| - 2$ denote

$$\mathcal{E}^{\downarrow}_{S}(N) \,=\, \{u_{\langle i,j|K\rangle}\,;\, K = S\}\,.$$

Given $r : \mathcal{P}(N) \to \mathbf{R}$ define the corresponding *decremental function* as follows:

$$\delta_r(S) \,=\, \min\{\,\langle r, u\rangle\,;\, u \in \mathcal{E}^{\downarrow}_{S}(N)\,\} \quad \text{for } S \subseteq N\,,$$

where the minimum of empty collection is 0 by convention. One can repeat the arguments from Section 4.2 to show that there is one-to-one correspondence between r and δ_r.

5.1.3 Reflection

The 'reverse' point of view also enables one to recognize further symmetry within the skeleton. Let us consider the following bijective mappping $\boldsymbol{\iota} : \mathcal{P}(N) \to \mathcal{P}(N)$:

$$\boldsymbol{\iota}(A) \,=\, N \setminus A \quad \text{for every} \;\; A \subseteq N\,.$$

Then every real function $m : \mathcal{P}(N) \to \mathbf{R}$ can be transformed as follows

$$m_{\iota}(A) \,=\, m(\boldsymbol{\iota}(A)) \quad \text{for every} \;\; A \subseteq N\,.$$

Like in Section 2.4.1 one can verify that the mapping $m \mapsto m_{\iota}$ is bijective (self-inverse) linear mapping which maps $\mathcal{M} \cap \mathcal{S}$ onto $\mathcal{M} \cap \mathcal{R}$. In particular, extreme rays of $\mathcal{M} \cap \mathcal{S}$ are transformed into extreme rays of $\mathcal{M} \cap \mathcal{R}$. Therefore, every $m \in \mathcal{M}_{ex}$ is transformed into 'reversely' standardized extreme normalized imset $r_{\star} = m_{\iota} \in \mathcal{M} \cap \mathcal{R}$. This reverse representative $r_{\star}$ can be again linearly transformed (see above, Section 5.1) into the corresponding standard representative $m_{\star} \in \mathcal{M} \cap \mathcal{S}$ called the *reflection* of m. The previous arguments show that the reflection mapping $m \mapsto m_{\star}$ is a bijective mapping of $\mathcal{M}_{ex}$ onto $\mathcal{M}_{ex}$. For example, in case $N = \{a, b, c\}$ the reflection of the skeletal imset in the first row of Figure 1 is the imset in the third row of Figure 1 and conversely. The imsets in the second row are self-reflexive.

5.2 Polymatroid representation

Every supermodular function $m : \mathcal{P}(N) \to \mathbf{R}$ corresponds uniquely to a submodular function $-m$. This simple fact leads to further alternative way of representation of skeletal imsets. *Polymatroid* [3] is an non-decreasing submodular real function $h : \mathcal{P}(N) \to \mathbf{R}$ which satisfies $h(\emptyset) = 0$. *Rank* of the polymatroid is the value $h(N)$. Every $m \in \mathcal{M}_{st}$ can be identified with a polymatroid in the following way. We put

$$\nu(i) \;=\; m(N) - m(N \setminus \{i\}) \quad \text{for every } i \in N,$$

and introduce

$$h(S) \;=\; -m(S) + \sum_{i \in S} \nu(i) \quad \text{for } S \subseteq N\,.$$

It follows from formulas in Section 5.1 that

$$\forall\, S \subseteq N \qquad h(S) + r(S) \;=\; k\,,$$

where $k = -m(N) + \sum_{i \in N} \nu(i)$ is a constant. Hence, h is a polymatroid and there is one-to-one correspondence between h and the reverse representative r. Indeed, one has $h(N) = k = r(\emptyset)$.

Remark Given a discrete probability distribution P over N, the *entropy function* h_P induced by P is a function which ascribes the entropy of the marginal P^S to every $S \subseteq N$. It is always a polymatroid. Thus, the polymatroid way of representation is motivated by the wish to represent things in entropy-like style.

5.3 Möbius inversion

Another possible linear transformation applicable to standard supermodular function $m \in \mathcal{M}_{st}$ is the Möbius inversion. Indeed, one can put

$$t(S) \;=\; \sum_{T \subseteq S} (-1)^{|S \setminus T|} \cdot m(T) \quad \text{for every } S \subseteq N\,.$$

It is well known that the inverse formula is

$$m(S) \;=\; \sum_{T \subseteq S} t(T) \quad \text{for every } S \subseteq N\,,$$

and therefore the mapping $m \mapsto t$ is a bijective linear mapping.

Remark Note that in case $m = m_P$, where P is a probability distribution over N, the number $t_P(S) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} \cdot m_P(T)$ is known in information theory [1] as the *mutual information* of variables in S. Note that the mutual information of more than 2 variables can be negative (in contrast to mutual information of two variables).

6 Internal notation

This is an overview of our internal conventions in notation.

6.1 Coding of sets

We have chosen and fixed the following way of indentification of subsets of the set $N = \{a, b, c, d, e\}$ with integers $j \in \{1, \ldots, 32\}$. The set encoded by j will be denoted by $\mathsf{Set}(j)$.

1	$\emptyset$	9	$\{a,d\}$	17	$\{a,b,c\}$	25	$\{b,d,e\}$
2	$\{a\}$	10	$\{a,e\}$	18	$\{a,b,d\}$	26	$\{c,d,e\}$
3	$\{b\}$	11	$\{b,c\}$	19	$\{a,b,e\}$	27	$\{a,b,c,d\}$
4	$\{c\}$	12	$\{b,d\}$	20	$\{a,c,d\}$	28	$\{a,b,c,e\}$
5	$\{d\}$	13	$\{b,e\}$	21	$\{a,c,e\}$	29	$\{a,b,d,e\}$
6	$\{e\}$	14	$\{c,d\}$	22	$\{a,d,e\}$	30	$\{a,c,d,e\}$
7	$\{a,b\}$	15	$\{c,e\}$	23	$\{b,c,d\}$	31	$\{b,c,d,e\}$
8	$\{a,c\}$	16	$\{d,e\}$	24	$\{b,c,e\}$	32	$\{a,b,c,d,e\}$

6.2 Standard representation

Standard representatives of types of extreme supermodular functions on $\mathcal{P}(N)$ will be stored in a matrix

$$M = (m_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32\,.$$

Row m_i of this matrix describes the standard representative of the i-th type. One can also consider a 'non-reduced' matrix M^* whose rows are all standard representatives of extreme supermodular functions

$$M^* = (m^*_{i,j}) \quad i = 1, \ldots, 117978, \quad j = 1, \ldots 32\,.$$

6.3 Coding of elementary independence statements

We have fixed the following way of indentification of elementary independence statements over the set $N = \{a, b, c, d, e\}$ with integers $k \in \{1, \ldots, 80\}$.

1	$\{d\} \perp\!\!\!\perp \{e\} \mid \{a,b,c\}$	21	$\{a\} \perp\!\!\!\perp \{d\} \mid \{c,e\}$	41	$\{c\} \perp\!\!\!\perp \{d\} \mid \{e\}$	61	$\{c\} \perp\!\!\!\perp \{d\} \mid \{a\}$
2	$\{c\} \perp\!\!\!\perp \{e\} \mid \{a,b,d\}$	22	$\{a\} \perp\!\!\!\perp \{c\} \mid \{d,e\}$	42	$\{c\} \perp\!\!\!\perp \{e\} \mid \{d\}$	62	$\{a\} \perp\!\!\!\perp \{b\} \mid \{e\}$
3	$\{c\} \perp\!\!\!\perp \{d\} \mid \{a,b,e\}$	23	$\{d\} \perp\!\!\!\perp \{e\} \mid \{a,b\}$	43	$\{d\} \perp\!\!\!\perp \{e\} \mid \{c\}$	63	$\{a\} \perp\!\!\!\perp \{e\} \mid \{b\}$
4	$\{b\} \perp\!\!\!\perp \{e\} \mid \{a,c,d\}$	24	$\{b\} \perp\!\!\!\perp \{e\} \mid \{a,d\}$	44	$\{b\} \perp\!\!\!\perp \{d\} \mid \{e\}$	64	$\{b\} \perp\!\!\!\perp \{e\} \mid \{a\}$
5	$\{b\} \perp\!\!\!\perp \{d\} \mid \{a,c,e\}$	25	$\{b\} \perp\!\!\!\perp \{d\} \mid \{a,e\}$	45	$\{b\} \perp\!\!\!\perp \{e\} \mid \{d\}$	65	$\{a\} \perp\!\!\!\perp \{b\} \mid \{d\}$
6	$\{b\} \perp\!\!\!\perp \{c\} \mid \{a,d,e\}$	26	$\{a\} \perp\!\!\!\perp \{e\} \mid \{b,d\}$	46	$\{d\} \perp\!\!\!\perp \{e\} \mid \{b\}$	66	$\{a\} \perp\!\!\!\perp \{d\} \mid \{b\}$
7	$\{a\} \perp\!\!\!\perp \{e\} \mid \{b,c,d\}$	27	$\{a\} \perp\!\!\!\perp \{d\} \mid \{b,e\}$	47	$\{b\} \perp\!\!\!\perp \{c\} \mid \{e\}$	67	$\{b\} \perp\!\!\!\perp \{d\} \mid \{a\}$
8	$\{a\} \perp\!\!\!\perp \{d\} \mid \{b,c,e\}$	28	$\{a\} \perp\!\!\!\perp \{b\} \mid \{d,e\}$	48	$\{b\} \perp\!\!\!\perp \{e\} \mid \{c\}$	68	$\{a\} \perp\!\!\!\perp \{b\} \mid \{c\}$
9	$\{a\} \perp\!\!\!\perp \{c\} \mid \{b,d,e\}$	29	$\{c\} \perp\!\!\!\perp \{e\} \mid \{a,b\}$	49	$\{c\} \perp\!\!\!\perp \{e\} \mid \{b\}$	69	$\{a\} \perp\!\!\!\perp \{c\} \mid \{b\}$
10	$\{a\} \perp\!\!\!\perp \{b\} \mid \{c,d,e\}$	30	$\{b\} \perp\!\!\!\perp \{e\} \mid \{a,c\}$	50	$\{b\} \perp\!\!\!\perp \{c\} \mid \{d\}$	70	$\{b\} \perp\!\!\!\perp \{c\} \mid \{a\}$
11	$\{d\} \perp\!\!\!\perp \{e\} \mid \{b,c\}$	31	$\{b\} \perp\!\!\!\perp \{c\} \mid \{a,e\}$	51	$\{b\} \perp\!\!\!\perp \{d\} \mid \{c\}$	71	$\{d\} \perp\!\!\!\perp \{e\} \mid \emptyset$
12	$\{c\} \perp\!\!\!\perp \{e\} \mid \{b,d\}$	32	$\{a\} \perp\!\!\!\perp \{e\} \mid \{b,c\}$	52	$\{c\} \perp\!\!\!\perp \{d\} \mid \{b\}$	72	$\{c\} \perp\!\!\!\perp \{e\} \mid \emptyset$
13	$\{c\} \perp\!\!\!\perp \{d\} \mid \{b,e\}$	33	$\{a\} \perp\!\!\!\perp \{c\} \mid \{b,e\}$	53	$\{a\} \perp\!\!\!\perp \{d\} \mid \{e\}$	73	$\{c\} \perp\!\!\!\perp \{d\} \mid \emptyset$
14	$\{b\} \perp\!\!\!\perp \{e\} \mid \{c,d\}$	34	$\{a\} \perp\!\!\!\perp \{b\} \mid \{c,e\}$	54	$\{a\} \perp\!\!\!\perp \{e\} \mid \{d\}$	74	$\{b\} \perp\!\!\!\perp \{e\} \mid \emptyset$
15	$\{b\} \perp\!\!\!\perp \{d\} \mid \{c,e\}$	35	$\{c\} \perp\!\!\!\perp \{d\} \mid \{a,b\}$	55	$\{d\} \perp\!\!\!\perp \{e\} \mid \{a\}$	75	$\{b\} \perp\!\!\!\perp \{d\} \mid \emptyset$
16	$\{b\} \perp\!\!\!\perp \{c\} \mid \{d,e\}$	36	$\{b\} \perp\!\!\!\perp \{d\} \mid \{a,c\}$	56	$\{a\} \perp\!\!\!\perp \{c\} \mid \{e\}$	76	$\{b\} \perp\!\!\!\perp \{c\} \mid \emptyset$
17	$\{d\} \perp\!\!\!\perp \{e\} \mid \{a,c\}$	37	$\{b\} \perp\!\!\!\perp \{c\} \mid \{a,d\}$	57	$\{a\} \perp\!\!\!\perp \{e\} \mid \{c\}$	77	$\{a\} \perp\!\!\!\perp \{e\} \mid \emptyset$
18	$\{c\} \perp\!\!\!\perp \{e\} \mid \{a,d\}$	38	$\{a\} \perp\!\!\!\perp \{d\} \mid \{b,c\}$	58	$\{c\} \perp\!\!\!\perp \{e\} \mid \{a\}$	78	$\{a\} \perp\!\!\!\perp \{d\} \mid \emptyset$
19	$\{c\} \perp\!\!\!\perp \{d\} \mid \{a,e\}$	39	$\{a\} \perp\!\!\!\perp \{c\} \mid \{b,d\}$	59	$\{a\} \perp\!\!\!\perp \{c\} \mid \{d\}$	79	$\{a\} \perp\!\!\!\perp \{c\} \mid \emptyset$
20	$\{a\} \perp\!\!\!\perp \{e\} \mid \{c,d\}$	40	$\{a\} \perp\!\!\!\perp \{b\} \mid \{c,d\}$	60	$\{a\} \perp\!\!\!\perp \{d\} \mid \{c\}$	80	$\{a\} \perp\!\!\!\perp \{b\} \mid \emptyset$

6.4 Elementary imsets

Every elementary conditional independence statement corresponds to an elementary imset which is an integer-valued function on the power set of $N = \{a, b, c, d, e\}$. Taking into consideration coding of subsets of N mentioned in Section 6.1 elementary imsets can be described by a matrix

$$U = (u_{k,j}) \quad k = 1, \dots, 80, \quad j = 1, \dots 32 .$$

6.5 Lattice of subsets

The power set of N is ordered by inclusion and forms a distributive lattice. The corresponding Hasse diagram can be represented in the form of a matrix

$$L = (l_{j,J}) \quad j = 1, \dots, 32, \quad J = 1, \dots 32 .$$

composed of zeros and ones, where

$$l_{j,J} = \begin{cases} 1 & \text{if } \mathsf{Set}(J) \text{ contains } \mathsf{Set}(j) \text{ plus one extra element,} \\ 0 & \text{otherwise.} \end{cases}$$

6.6 Minimal sets

The classes of minimal subsets (see Section 4.1) will be represented also in the form of a zero-one matrix

$$M^{\min} = (m^{\min}_{i,j}) \quad i = 1, \dots, 1319, \quad j = 1, \dots 32,$$

whose elements are determined as follows

$$m^{\min}_{i,j} = \begin{cases} 1 & \text{if } \mathsf{Set}(j) \text{ is a mimimal set with non-zero value in } m_i, \\ 0 & \text{otherwise.} \end{cases}$$

6.7 Incremental (ascending) transformation

Standard representatives can be transformed by 'incremental' transformation (see Section 4.2). The corresponding matrix is then

$$A = (a_{i,j}) \quad i = 1, \dots, 1319, \quad j = 1, \dots 32 .$$

6.8 Table of scalar products

The record of scalar products (see Section 4.3) is described by a matrix

$$S = (s_{i,k}) \quad i = 1, \dots, 1319, \quad k = 1, \dots 80 .$$

which is nothing but matrix product of M and the transpose of U. Similarly, a non-reduced form (where M is replaced by M^*) is denoted analogously

$$S^* = (s^*_{i,k}) \quad i = 1, \dots, 117978, \quad k = 1, \dots 80 .$$

6.9 Grade

Maximum of every row in the matrix S is *grade* (see Section 4.3) of the coresponding type and is stored in a vertical vector

$$G = (g_i) \quad \text{with } g_i = \max_{k=1,\dots,80} s_{i,k} \text{ for } i = 1, \dots, 1319 .$$

6.10 Reverse representation

Reverse standardization (see Section 5.1) is stored in a matrix

$$R = (r_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32 \,.$$

Row r_i in R describes the reverse representative of the i-th type. The corresponding non-reduced matrix whose rows are all reverse representatives of extreme supermodular functions is then

$$R^* = (r^*_{i,j}) \quad i = 1, \ldots, 117978, \quad j = 1, \ldots 32 \,.$$

6.11 Decremental (descending) transformation

Reverse transformation can be transformed by 'decremental' transformation (see Section 5.1.2) which will be decribed by a matrix

$$D = (d_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32 \,.$$

6.12 Maximal sets

The classes of maximal subsets obtained on basis of reverse representation (see Section 5.1.1) will be represented in the form of a zero-one matrix

$$R^{\max} = (r^{\max}_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32$$

where

$$r^{\max}_{i,j} = \begin{cases} 1 & \text{if } \mathsf{Set}(j) \text{ is a maximal set with non-zero value in } r_i, \\ 0 & \text{otherwise.} \end{cases}$$

6.13 Polymatroid representation

Polymatroid representation (see Section 5.2) will be stored in the matrix

$$H = (h_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32 \,.$$

6.14 Mutual information

Mutual information representation (i. e. Möbius inversions of standard representatives, see Section 5.3) will be described by a matrix

$$T = (t_{i,j}) \quad i = 1, \ldots, 1319, \quad j = 1, \ldots 32 \,.$$

7 Catalogue

There are three versions of the catalogue described hereafter. However, in all three versions the types of skeletal imsets have the same identification (ID) number (a number between `1` and `1319`) and the same standard representative. The ID numbers were assigned to the types automatically by a computer during the choice of representatives described in Section 3.3. They have no special meaning, it is just an auxiliary tool for browsing through the catalogue (we have mentioned the ID numbers in our figures already).

All files mentioned below including the original input text file `fivevar.ieq` for PORTA mentioned in Section 3.1 can be found at

http://www.utia.cas.cz/user_data/studeny/fivevar.htm .

7.1 PC viewer

The first version of the catalogue is an original computer program for PC (under Windows) which makes it possible to visualise any list of (possibly alternative) representatives of types of skeletal imsets written in a special data file `*.prn`. The representatives are visualised in the form of diagrams similar to the figures used in this report. The program enables user to see simultaneously two (possibly different) representatives of the same type and to print those pictures. The choice of ways of representation can be done by importing (possibly two) data files. The order in the joint list of pictures depends on the order in the first imported data file `*.prn`. We have prepared the following data files:

`m.prn`	standard representatives of types of skeletal imsets (see Section 2.2.3),
`a.prn`	incremental transformation of the standard representatives (see Section 4.2),
`r.prn`	reverse representatives (see Section 5.1),
`h.prn`	polymatroid representatives computed by the formula from Section 5.2,
`t.prn`	Möbius inversions of the standard representatives (see Section 5.3).

7.2 WWW viewers

The second version is a series of five interactive catalogues (= online viewers) of representatives of types of skeletal imsets for clients of World Wide Web. Every viewer uses just one of five ways of representation mentioned in Section 7.1 and enables user to browse through the list of respective representatives visualised in a similar way. Moreover, one can optionally highlight minimal and maximal sets of the support (see Sections 4.1 and 5.1.1) in the way used in our figures. Finally, a special mode allows one to browse similarly through all permutations of a chosen fixed representative. So, a complete list of representatives of skeletal imsets is available for viewing.

7.3 Catalogue in EXCEL

The third version of the catalogue is a file in EXCEL which contains vector description of all five representatives of types of skeletal imsets mentioned above and the table of scalar products. Moreover, the file contains (in another list) several numerical characteristics of the type, namely

1. number of skeletal imsets belonging to the type (see Section 3.3),
2. grade of the type (see Section 4.3),
3. sum of values of the standard representative,
4. rank of the polymatroid (see Section 5.2),
5. zip matrix of the type (see Section 4.4).

Note that these lists are ordered according to the ID numbers. The characteristics are designed to enable user to reorder types according to specific (user's) criteria (command ORDER in menu DATA of EXCEL). This reordering automatically results in reordering of the list EXPORT_o.prn, which is the list of standard representatives of types of skeletal imsets ordered according to the new chosen criteria (not according to the ID number any more). This list can be exported into a file `*.prn` using the command SAVE_AS in menu FILE of EXCEL. Such a file can be used by the above mentioned PC viewer as the first imported file governing the order of types in a joint list.

7.4 Program realizing inference mechanism

The approach described in [10] leads to a certain method of inferring conditional independence statements from other conditional independence statements (*facial implication*). Effective computer implementation of this inference mechanism substantially depends on the knowledge of the skeleton. One of the goals of this work was to realize this inference mechanism in case of 5 variables. We have prepared such a computer program `verimpli.exe` for PC. The program gives an answer to the question whether a certain set of elementary conditional independence statements L facially implies another conditional independence statement v. The question is encoded in the form of a text input file `input.txt` which has the form

```
0 0 + 0 + + 0 ... 0 - + 0 + +
END
```

where the first line contains 80 symbols which correspond to elementary conditional independence statements ordered according to their codes from Section 6.3. The symbol `+` in the record indicates that the corresponding conditional independence statement belongs to L, the symbol `-` indicates v and the digit `0` indicates the remaining conditional independence statements. The program needs description of the skeleton in form of a data file `m.poi`. The command

```
verimpli input.txt m.poi
```

(where `input.txt` is the input file) written on a DOS command line can be used to execute the program. An experimental online version of this program is also available on web.

7.5 Program for decomposition

Another program `verstruc.exe` for PC allows one to find out whether an imset which is a combination of elementary imsets with integer coefficients has a non-zero non-negative multiple which is a combination of elementary imsets with non-negative integer coeffients (that is, whether it is a *structural imset* in sense of [10]). An input file describing the respective combination of elementary imsets has an analogous structure

```
0 0 1 0 5 -1 0 ... 0 -3 2 0 1 1
END
```

where the first line contains 80 integers, namely the coefficients of elementary imsets encoded in the way described in Section 6.3. The command

```
verstruc input.txt m.poi
```

written on a DOS command line is analogous to the previous case. The program can sometimes help one to decompose a structural imset, that is to write it as a combination of elementary imsets with non-negative integer coefficients. Equalities mentioned in Section 4.3 were obtained in this way.

8 Conclusions

The goals of the paper were roughly achieved. The basic results were achieved by a computer program PORTA. Computation took several hours. The runtime for four variables compared to five and the nature of the problem suggests that this approach is not feasible for finding extreme supermodular set functions over six or more variables. The result also suggests that the number of extreme supermodular set functions grows superexponentially with the number of variables. Several questions remain open, other questions emerged.

- Is there any suitable characterization of extreme supermodular functions which makes it possible to find the skeletal imsets directly on basis of 'combinatorial' properties?
- Is seems possible to realize facial implication [10] without knowing the skeleton, just on basis of knowledge of grade (see Section 4.3). How to compute the grade for every number of variables? Is such a method of verification of facial implication more efficient than the method based on knowledge of skeleton?
- How to recognize probabilistically representable and non-representable types of skeletal imsets on basis of their numerical characteristics? What are probabilistically representable supermodular functions over 5 variables? We were informed (private communication by F. Matúš) that a sufficient condition for probabilistical representability of a polymatroid h over N is that $\sum_{i\in N} h(\{i\}) \leq 7$.

Acknowledgment

We thank to the authors of PORTA program for development of this useful computational tool and to our colleague František Matúš for directing our attention to it. Our thanks are also devoted to the Fields Institute for Research in Mathematical Sciences, University of Toronto, Canada. The results reported in this paper were achieved in October 1999 while the authors were visitors there.

References

[1] T. M. Cover, J. A. Thomas: Elements of Information Theory, John Willey 1991.

[2] S. Fujishige: Submodular Functions and Optimization, Annals of Discrete Mathematics 47, North-Holland 1991.

[3] F. Matúš: Conditional independence structures examined via minors, Annals of Mathematics and Artificial Intelligence 21 (1997), pp. 99-128.

[4] J. Rosenmüller, H. G. Weidner: Extreme convex set functions with finite carrier: general theory, Discrete Mathematics 10 (1974), pp. 342-382.

[5] L. S. Shapley: Cores of convex games, International Journal of Game Theory 1 (1971/72), pp. 11-26.

[6] M. Studený: Convex set functions I., research report n. 1733, Institute of Information Theory and Automation, November 1991.

[7] M. Studený: Convex set functions II., research report n. 1734, Institute of Information Theory and Automation, November 1991.

[8] M. Studený: Multiinformation and conditional independence II., research report n. 1751, Institute of Information Theory and Automation, September 1992.

[9] M. Studený: Convex cones in finite-dimensional real vector spaces, Kybernetika 29 (1993), pp. 180-200.

[10] M. Studený: Structural semigraphoids, International Journal of General Systems 22 (1994), pp. 207-217.

[11] M. Studený: Description of conditional independence structures by means of faces and imsets, (a series of 3 papers) 1st part: introduction and basic concepts, 2nd part: basic theory, 3rd part: examples of use and appendices, International Journal of General Systems 23 (1994/1995), pp. 123-137, 201-219, 323-341.

[12] M. Studený, P. Boček: CI-models arising among 4 random variables, in Proceedings of WUPES'94, September 11-15, 1994, Třešt', Czech Republic, pp. 268-282.

[13] PORTA - a Polyhedron Representation Transformation Algorithm, written by T. Christof (University of Heidelberg), revised by A. Loebel and M. Stoer (ZIB Berlin), available at http://www.IWR.Uni-Heidelberg.de/iwr/comopt/soft/PORTA/readme.html .

Addresses

Milan Studený, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Prague, Czech Republic. (studeny@utia.cas.cz)

Remco R. Bouckaert, 86 Eskdale Road, Papakowhai, Wellington, New Zealand. (xtal_mntn@yahoo.com)

Tomáš Kočka, Laboratory of Intelligent Systems, University of Economics Prague, Ekonomická 957, 14800 Prague, Czech Republic. (kocka@vse.cz)

Appendix

All 10 types of skeletal imsets over $N = \{a, b, c, d\}$ are described by means of the following figures.

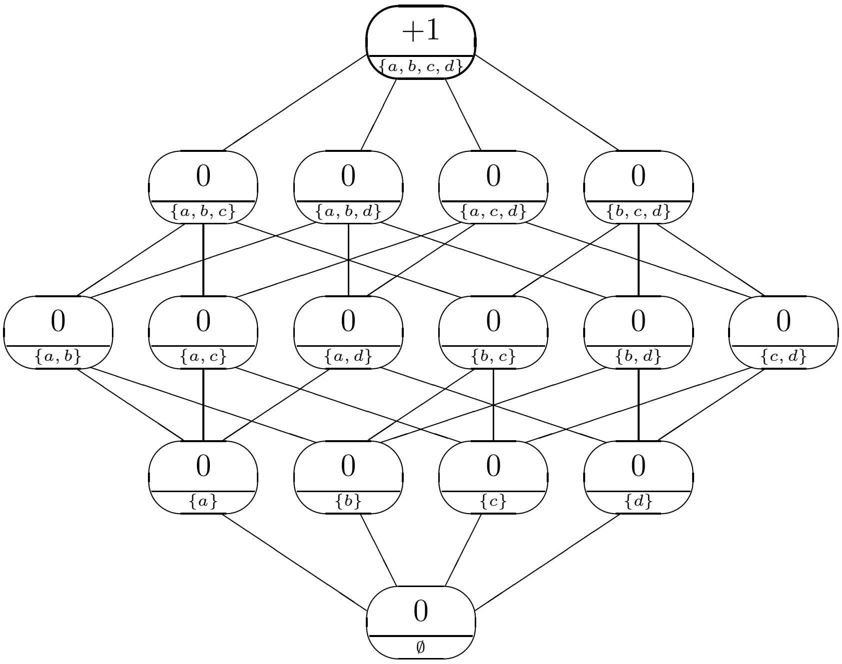


Figure 6: The first type of skeletal imsets over $N = \{a, b, c, d\}$ - 1 representative only.

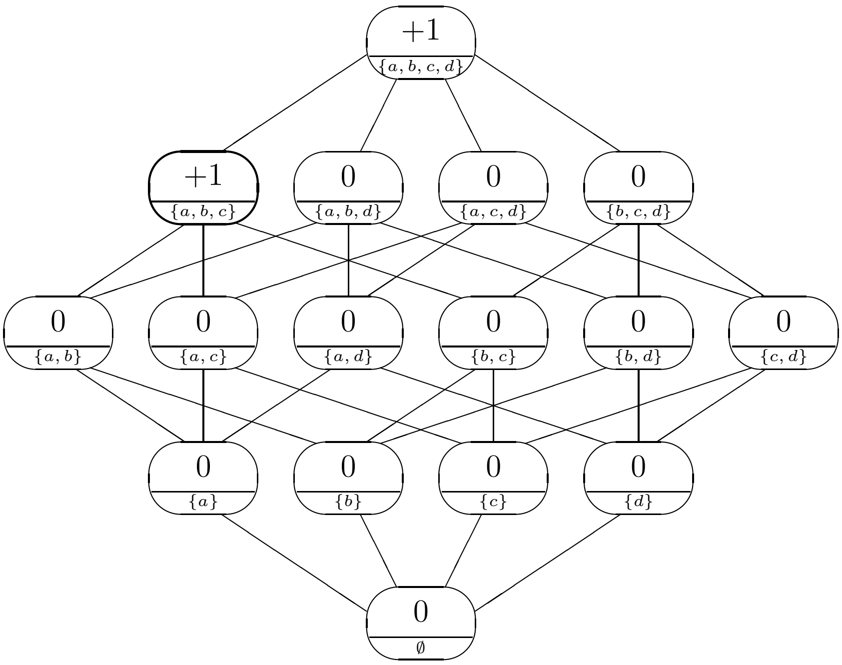


Figure 7: The second type of skeletal imsets - one of 4 possible representatives.

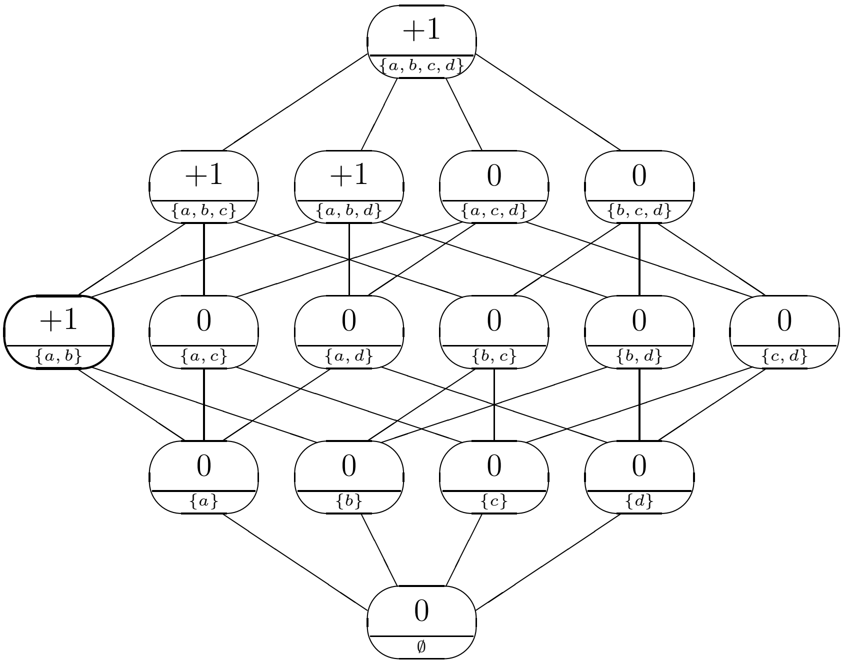


Figure 8: The third type of skeletal imsets - one of 6 possible representatives.

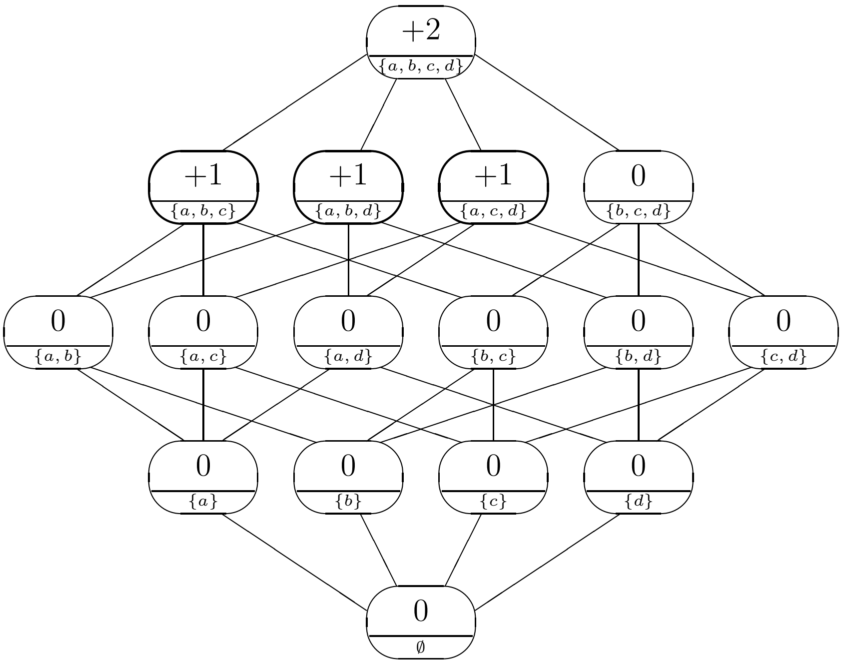


Figure 9: The fourth type of skeletal imsets - one of 4 possible representatives.

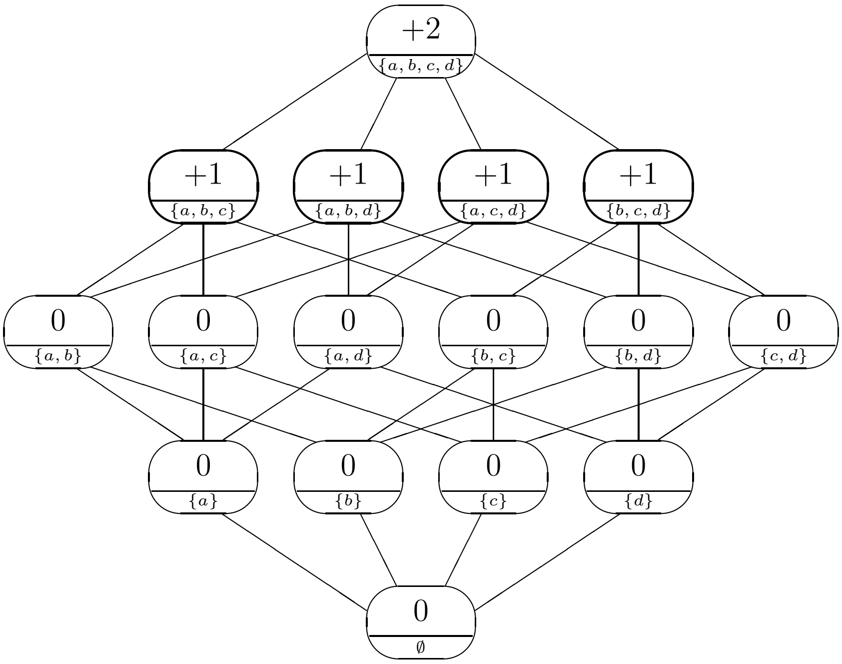


Figure 10: The fifth type of skeletal imsets over $N = \{a, b, c, d\}$ - only 1 representative.

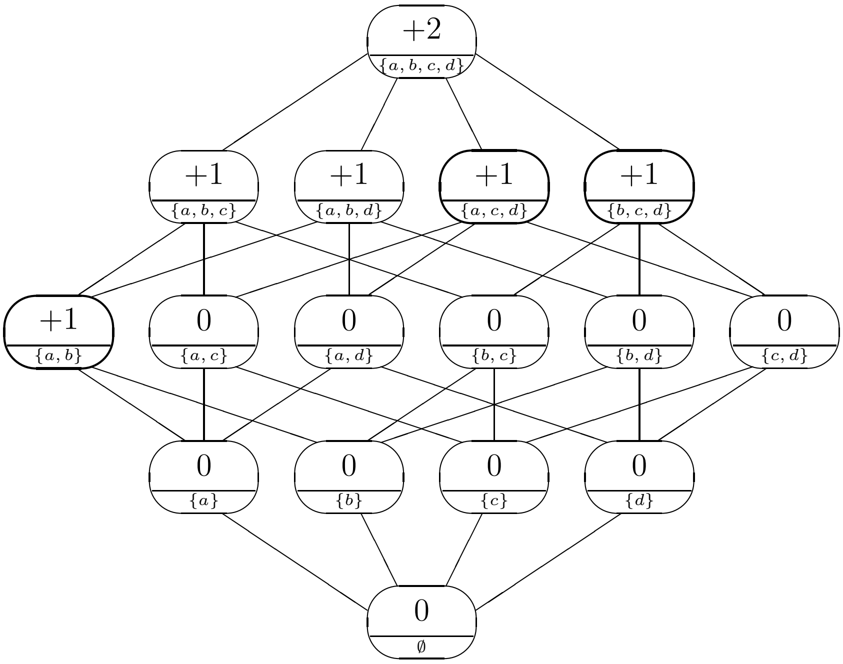


Figure 11: The sixth type of skeletal imsets - one of 6 possible representatives.

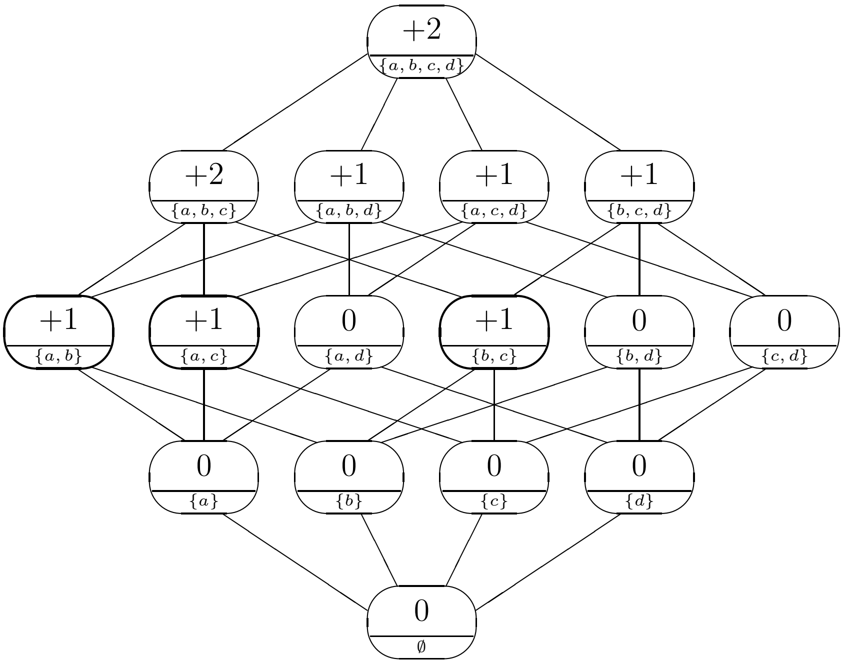


Figure 12: The seventh type of skeletal imsets - one of 4 possible representatives.

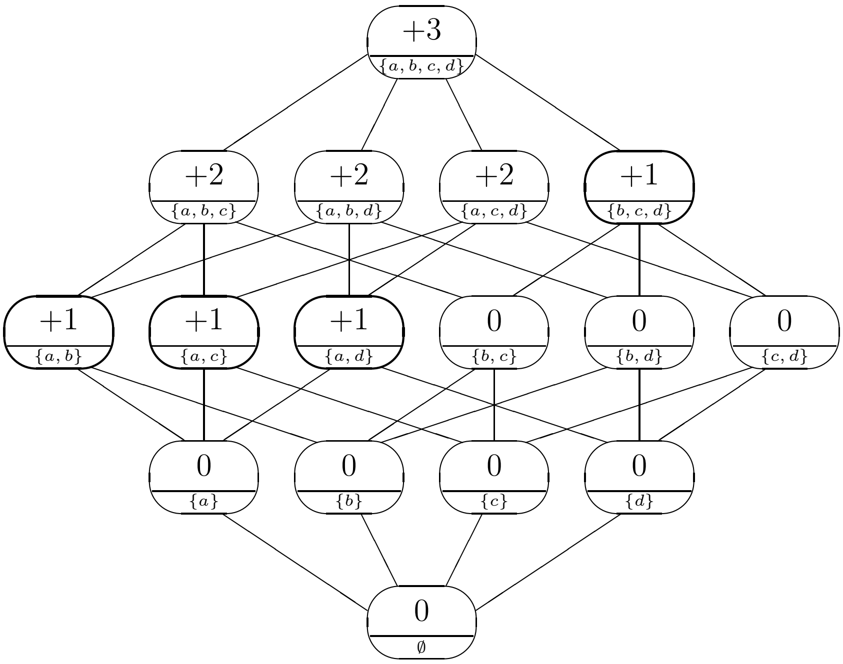


Figure 13: The eighth type of skeletal imsets - one of 4 possible representatives.

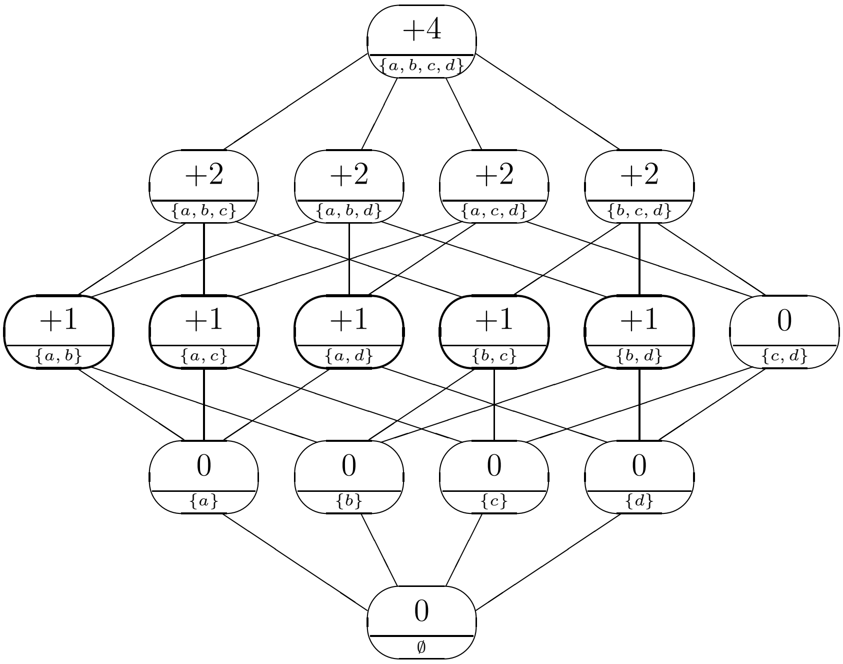


Figure 14: The ninth type of skeletal imsets - one of 6 possible representatives.

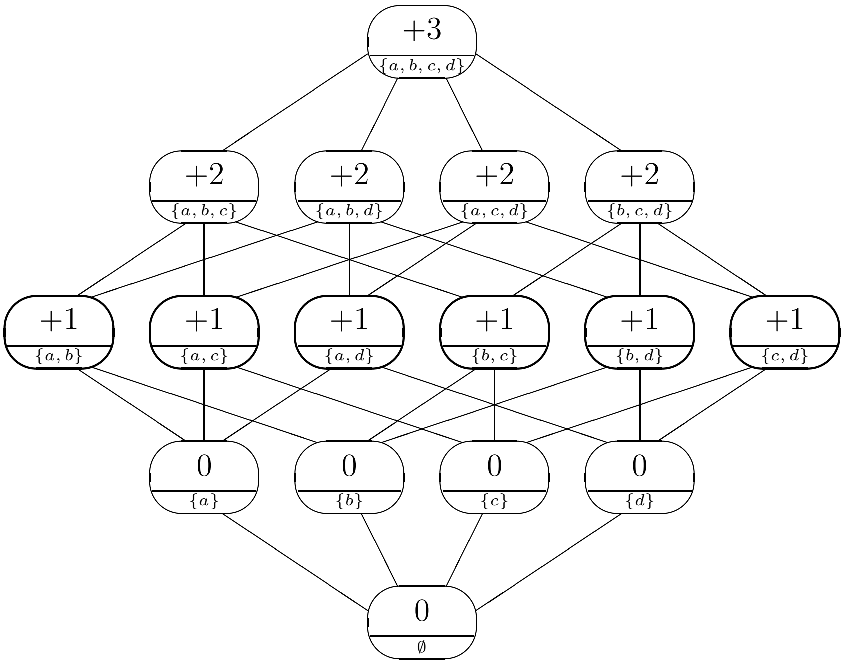


Figure 15: The tenth type of skeletal imsets over $N = \{a, b, c, d\}$ - only 1 representative.